Audit firm scope and auditor independence.

R. Doogar*
doogar@uiuc.edu

Department of Accountancy,
College of Commerce,
University of Illinois
284 Commerce West,
1206 South Sixth Street
Champaign, IL 61820.

First draft: October 1999.
This version: December 29, 2000.

* I wish to acknowledge, without implicating in any way, helpful discussions with Andy Bailey, Paul Beck, George Deltas, Rob Easley, Bob Halperin, Van Johnson, Young Kwon, Tom Omer, Mark Peecher, Dave Ricchiute, Marjorie Shelley and Ira Solomon.

Audit firm scope and auditor independence.

Abstract

I model production and organization costs for a large multi-service audit firm that enjoys economies of scope in the production of auditing and consulting services and examine the optimal choice of audit firm scope in a competitive market for mandated audit and voluntarily purchased consulting services. I show that when audits are mandated, firms produce more of both the audit and non-audit services than investors prefer. This happens because mandating audits removes the link between the utility of the audit to the investor and the audit fee. In effect, the firm is not penalized sufficiently for producing "too much" consulting relative to what investors would prefer and therefore chooses to invest too much in the factor that induces economies of scope between the two tasks.

Audit firm scope and auditor independence.

1. Introduction

On June 27, 2000 the SEC published its proposed revisions to the rules on auditor independence (SEC 2000a and b).[1] Of all the proposed revisions, the proposal to ban the provision of certain types of non-audit services to audit clients provoked the greatest controversy.[2] Under regulatory pressure, two of the Big Five audit-and-consulting firms decided to divest themselves of parts of their consulting operations (MacDonald 2000, Brick 2000).[3] After several rounds of combative public hearings,[4] much lobbying activity in Congress on behalf of the firms and threats of litigation by some Big Five firms, in November 2000, the SEC modified its original proposals on non-audit services to audit clients from an outright ban to requiring disclosure of fees for non-audit services paid to auditors.[5] Throughout much of the debate the SEC and the profession presented apparently irreconcilable views of the world: What the SEC felt was inappropriate pursuit of consulting revenues from audit clients, the accounting firms and the profession felt was merely due reward for developing an intimate knowledge of the client's business and industry. What the SEC saw as reasonable measures to forestall potentially catastrophic failures of public confidence in auditor independence (its proposed limitations on providing non-audit services to audit clients), the profession and the firms saw as needless regulatory intervention grounded in baseless premises.[6]

Both the question of the optimal scope of the audit firm and the role of pubic perceptions that figured so prominently in these debates, though not new in the policy debate over auditor independence, appear to have been largely ignored in prior research on audit firm scope and auditor independence.[7] In this study I outline a modeling framework within which such issues can studied rigorously and show how the diametrically opposed perspectives of the SEC and the audit profession can be reconciled within an institutionally descriptive model of the audit market and its informational and legal environment.

Any satisfactory economic analysis of the scope question must start with a representation of the audit firm that includes both a model of audit and non-audit production and an examination of organization costs. I generalize a two-factor Cobb-Douglas production function that has been used to

explain both aggregate and firm specific audit market shares (Doogar 1994, Doogar and Easley 1998; Doogar, Easley and Ricchiute 2000) to include a role for knowledge capital which is claimed to be the source of economies of scope between auditing and consulting. Based on this production technology and a discussion of organization costs, I derive a cost function for a multi-service firm. I then examine the choice of firm scope under market competition when audits are mandated but consulting services are voluntarily purchased.

In my model, while knowledge capital creates strong economies of scope and incentives for the firm to grow larger, organization costs provide a countervailing tension in the opposite direction. Both coordination costs, which I model as the costs of delay, and internal contracting costs of aligning incentives within the firm, increase with firm size. Moreover, due to information voids and imperfect contracting, investors' belief in auditor independence decreases as the firm's practice grows larger since the potential pool of rents available to clients to discipline the auditor (impair the auditor's independence) increases. The optimal size of the firm is determined by the tradeoff between the incremental gains from knowledge efficiencies against the incremental costs of organization as the firm grows larger.

The principal conclusions are as follows. First, since auditor independence is hard to evaluate and since firms can always settle lawsuits without admitting guilt or wrongdoing, rational investors may, by pure chance, conclude that auditors lack independence even when this conclusion is, in fact, false. The demand that a regulator be able to produce a "smoking gun" in order to undertake prophylactic measures ignores the difficulty of establishing independence failures and the role the litigation settlement process plays in the economy. The possibility of information cascades leading to a run on investor confidence in audits offers a plausible rationale for preventive (or, in the SEC's terminology, "prophylactic") measures.

Second, *whenever* auditing is mandated and auditing and consulting enjoy economies of scope in production costs, audit firms invest too much in knowledge capital and therefore do too much auditing

and consulting relative to investor preferences. The intuition behind these results is simple but striking. When investors have concerns about auditors also providing non-audit services, the value of the audit, and hence the audit fee, declines with the extent of non-audit services provided. If auditing is mandatory, the audit fee is not capped above by the value of the audit to the investor and an increase in the level of consulting does not directly affect the auditor's revenue. This leads the auditor to ignore investor preferences and become "too large" (invest too much, produce "too much" auditing and consulting relative to investor preferences).

If the gap between the optimal size of the firm determined with and without the costs of investors' concerns being accounted for is large, regulatory intervention may be called for simply because markets in this case do not suffice to provide the auditor with the "right" signal. However, regulatory intervention may not be necessary if firms find it optimal to voluntarily spin off consulting services once the size of the consulting practice reaches a threshold. My analysis identifies conditions when spinoffs become optimal. Interestingly, since the firm's knowledge capital cannot be divested even when the consulting practice is sold off (the same knowledge base is used for auditing and consulting), after a spinoff, the audit firm in effect functions as a natural business incubator, creating new uses for its knowledge capital and spinning off new practice areas as they reach critical mass.

The rest of the paper is organized as follows: Section 2 reviews prior related research. Section 3 presents a model of a multi-service firm with both internal and external organization costs and describes the market setting. Section 4 establishes the main result on divergent investor and auditor preferences over the firm's optimal scope. Section 5 summarizes and concludes.

2. Prior Related Research

The theory of multi-production and economies of scope has spawned a vast literature that is beyond the scope of this study to review. The approach I adopt in this study is inspired in large part by the type of models used in the study of contestable markets (Baumol et. al. 1988). While this approach has been criticized on account of its neglect of entry and exit costs and informational frictions, this

approach does provide a useful way to organize our thinking about the ways in which market forces may (or fail to) discipline firm conduct (Spence 1983, Brock 1983).

I depart from strict contestability analysis by incorporating some contracting frictions into the analysis and also consider the role of mandated demand in some markets on the supply of services to all markets. In general economic models assume that demand is voluntary and hence demand curves are downward-sloping in price. A key feature of audit markets is that audits are mandatory for publicly traded (SEC) clients: for them, the demand curve is not downward sloping in price over any effective range. I incorporate this unique aspect of audit markets into my analysis.

The transactions-cost theory of the firm emphasizes that economies in multiproduction i.e. the production of multiple outputs by the same firm, *cannot* be explained by purely technological features, e.g. economies of scale in joint production of multiple products using common facilities, alone. In this perspective, analysis of organizational costs is central to any successful explanation of the boundaries of the firm (Teece 1980, 1982). A recent variant of transactions cost theory, the intermediation theory of the firm, suggests that internal costs of organization organizes the discussion of transactions costs around the themes of internal and external costs of contracting (Spulber 1999). This distinction, I find, offers a natural framework for discussing the organization costs of an audit firm.

To survive, firms must efficiently coordinate activities across tasks and between agents.[8] There is a large economic literature on the costs of coordination that I believe auditing research has yet to make good use of. I borrow from this literature notions such as optimal hierarchies and minimum delay costs that help construct a more complete and satisfactory understanding of the costs of organizing a large multi-service audit firm. Internal contracting costs arising from imperfect control of agents' actions has also largely been ignored in prior research on auditor independence.[9] Yet these costs can often pose significant risks for large audit firms. A recent AICPA-sponsored research study notes that "Misdeeds by one individual can impair the reputation of an entire firm so monitoring of individual behavior is warranted and routinely done. It is difficult to tell directly however whether this monitoring is sufficient

to curb systemic independence violations."(Antle et. al. 1997: 23). Acemoglu (1994) presents an interesting model of auditor-auditee collusion in a hierarchy where the auditor conceals information because of her "career concerns" which arise from the expectation of future rents.[10] Finally, a careful consideration of investors' information environments points to links with the literature on information cascades and herding behavior. Since even a brief review of these diverse literatures would be intolerably lengthy, I defer discussion of the relevant studies to the appropriate model development sections.

Simunic (1984), Beck, Frecka and Solomon (1988a) and Antle and Demski (1991) model varying aspects of the joint production of audit and non-audit services.[11] Simunic (1984) examines how economies of scope in production might affect the demand for audit and non-audit services but does not examine the implications of multiproduction for auditor independence. This study is similar in spirit to Simunic's work as it focuses on the economics of multiproduction. A maintained assumption in Simunic's study (and in Antle and Demski) is that independence is not compromised by the provision of non-audit services.[12] Recently, however, the explosion in audit firms' consulting practices has raised questions about the validity of this assumption.[13] Independence concerns are central to current regulatory initiatives (Levitt, 2000) and to my analysis. Unlike Simunic and other prior researchers who assume there is no independence problem due to multiproduction, I explicitly consider the links between independence and market competition.

Beck et. al. use a game-theoretic approach to show how rents from repeated consulting engagements can impair auditor independence. I also assume that both activities generate rents that are the root of the independence problem . However, I use a classical production-function approach and develop audit cost functions endogenously from economic primitives. This permits me to generalize the analysis to multi-client settings and examine the effects of market competition on the choice of firm scope.

3. Model

I assume that the audit and consulting markets are competitive so that firms compete for clients in prices and that a price equilibrium obtains in which fees are set by the costs of each firm's nearest rival. This process can be modeled formally as a game in which firms bid for clients till no client can find a lower cost service provider.[14] The market is said to have cleared when a) every client has found an auditor (and if it has a non-zero consulting demand, a consultant as well) such that the client cannot save money by switching to a different service provider and b) no auditor has an unprofitable client. After the market clears, each firm that has any clients then chooses a cost-minimizing production plan to serve its clients, services are rendered and fees paid. Restricting production to occur after market clearing decouples pricing from production decisions and simplifies the analysis.

Markets are assumed to be contestable so that entry is free but since there are large firms already in the market, entry may not be profitable. I also assume that the market is large enough that a multi-firm equilibrium obtains so that each firm, no matter how large, is small enough relative to the market to act as a price taker. This assumption has the following force: When demand for a service is voluntary, as is the case for consulting services, fees are set by the minimum of a) the value of the service to the client and b) the costs of the second-lowest provider operating in the market. When demand for a service is mandated, as is the case for auditing services, the value of the audit to the client is irrelevant to the pricing of the audit services and audit fees are set by the costs of the second-lowest provider operating in the market.

For ease of exposition the development of the firm's cost function is organized as follows. In Section 3.1, I present a model of audit and consulting service production using human capital and knowledge capital as inputs. Next, in 3.2, I discuss internal contracting costs focusing largely on coordination (equivalently, delay) costs and on control (equivalently, agency) costs. In Section 3.3, I consider external contracting costs of setting up relationships and the costs of reputational externalities

due to investors' imperfect information about auditor independence. Section 3.4 provides a discussion and some remarks.

3.1. Production

For expositional simplicity, consider a professional services firm which provides only two services, labeled A (auditing) and C (consulting), to a number of clients. Since the production of services involve mostly human capital and knowhow, little generality is lost by assuming that this firm uses only three inputs to produce both services for all its clients. Two of these inputs represent two distinct types of labor, experienced labor which we can think of as seasoned professionals with titles such as manager, partner etc. in a large professional services firm while the other type of labor represents inexperienced labor such as fresh accounting undergraduates.[15] The third input represents the investment made by the firm to understand its clients' business and the industries in which the clients operate. This investment can take the form of tacit information, libraries of relevant information, expert systems available to aid audit and consulting staff, routines that have been customized to particular industries or types of transactions and all other investments that create knowledge or expertise that can be used to perform either task (e.g. training).

One commonly heard explanation for the extension of audit firms' practices into consulting services is that the knowledge capital investments firms must make in order to perform good audits enable them to provide consulting services at low incremental cost (e.g. Antle et. al. 1997, Section 5.2). Following this logic, I specify the model of production so as to generate *strong* economies of scale in auditing and consulting, this is, I assume that *every* dollar of investment in knowledge capital confers benefits on *both* tasks for *every* client.[16]

Let a_i and c_i represent the joint demand of A and C by the $i^{th.}$ client. Assume also that the firm's production technology for the two products has the form

$$a_i = A(K) f_{a_i}^{\alpha} l_{a_i}^{(1-\alpha)}; \qquad c_i = C(K) f_{c_i}^{\beta} l_{c_i}^{(1-\beta)} \tag{1}$$

where K is the firm's investment in knowledge capital, f_a and l_a are, respectively, the quanta of experienced and inexperienced labor used on an audit assignment and f_c and l_c the amounts of labor used on the consulting task. $A(K)$ and $C(K)$ represent the contribution of knowledge capital in producing a and c. It is sensible in the context of a professional service firm for knowledge capital and labor to be gross substitutes in production, so assume that $A(\cdot)$ and $C(\cdot)$ are strictly increasing in K.

The firm is assumed to be a price-taker for all inputs. Let the price vector which clears the input market be [1, r, w], so that capital is simply measured in monetary units while the prices of units of f and l are r and w respectively. $TDC(A,C)$, the total direct labor cost of serving a clientele $\{A,C\}$, is:

$$TDC(a_i,c_i) = K + r\sum_i (f_{a_i} + f_{c_i}) + w\sum_i (l_{a_i} + l_{c_i}) \qquad \textbf{(2)}$$

After the market clears and each client selects a provider, every firm that has a client selects the optimal input vector $I^{**}=[K^{**}, f_a^{**}+f_c^{**}, l_a^{**}+l_c^{**}]$ given its equilibrium clientele. Then we have:

Theorem 1:[17] Given the production technology in **(1)**, investment in capital is non-decreasing in the number of clients and the number of outputs produced by the firm. ■

The proof of Theorem 1 relies heavily on the strict concavity of the functions A and C, i.e. on the assumption that the increasing returns to scale to knowledge capital in both audit and consulting activities tail off as the investment in knowledge capital increases. If either map is non-concave, i.e. if investing in information technology produces accelerating returns to scale on either activity, there may be regions where the multi-product firm may not invest more than a single product firm.[18] However, we always have:

Theorem 2: A multiproduct firm never invests less than a single product firm. ■

Since investments in knowledge capital are likely to generate successively smaller incremental returns to scale, I focus on the more likely case where A and C are strictly concave in K, i.e. increases in knowledge capital produce successively smaller increases in returns to scale in each activity. Restricting both A and C to being strictly concave yields the following version of Theorem 2:

Theorem 2A: A multiproduct firm with both *A* and *C* strictly concave, invests at least as much as, but not more than twice as much as, the sum of the investments made by two comparable stand-alone firms.■

Two interesting conclusions follow from Theorem 2A. First, the theorem highlights the range of indeterminacy of investment when the researcher does not know the functions *A* and *C*, but it shows that the upper bound on the relative scale of investments made by a multi-product firm is a factor of two. The second interesting observation is that the larger the a firm's investment in knowledge capital, the lower its use of labor and ceteris paribus, the lower its variable costs of production. In price competition with other firms, a firm with larger investments in knowledge capital can compete more aggressively in two ways: first, by lowering its own cost structure the firm can increase the "gap" between its costs and the costs of the second-lowest-cost producer, which increases the rents that the incumbent can earn. Second, by lowering its cost structure a firm may be able to attract clients of other firms and, if the demand for some services is voluntary, lower costs can increase the size of the market.[19]

Under the assumption that labor inputs display at least constant returns to scale (the exponents of *f* and *l* sum to at least one), it can be shown that this technology exhibits increasing returns to scale and satisfies the necessary conditions for a natural monopoly (Baumol et. al., Theorem 8C4).[20] In other words, given the assumption of constant returns to scale in labor *and absent any of the countervailing costs to be discussed in Sections 3.2 and 3.3*, only one firm can remain active in the audit and consulting market. In order to state this result formally, it is useful to make precise the definition of a natural monopoly.

Definition (Baumol et. al., Definition 8D1): *An entrant to an industry is said to be a full supplier vis-a-vis a (former) monopolist if it undertakes to supply a proper subset S of the monopolist's outputs N, charging for each product a price no higher than the monopolist's and offers to sell whatever quantity of each product in S is demanded by the market at the relevant prices. The relevant prices are those set by the entrant for the goods in S and those set by the monopolist for the remaining goods in {N-S}.* ■

Remark: The definition applies to any firm that can enter *some* subset of markets of a potential multi-product monopolist and undercut the monopolist while still making a profit. Given this definition of a natural monopoly, Theorem 3 sets out the formal proposition of interest.

Theorem 3: *An audit production technology of the form **(1)** with A(K) and C(K) concave leads to a sustainable multi-product industry structure that is a natural monopoly*. ■

Theorem 2 shows that the non-exclusive nature of the firm's investment in knowledge capital does the job of entry barriers for free in a single product market. Theorem 3 shows that this works for all markets as well. The essence of the proof is that given natural monopoly in auditing and the global subadditivity of costs across any arbitrary combination of levels of *A* and *C*, no stand-alone consulting firm can compete with the incumbent in the audit market who wishes to extend the scope of the audit firm to the output *C*.

In sum, the production function in **(1)** is designed to ensure that a unit of investment in knowledge capital lowers labor costs on all clients *and* all products.[21] The complementarities induced by knowledge capital render the cost function strictly sub-additive[22] and absent any countervailing tensions, the only sustainable equilibrium is a single producer with some form of pricing scheme that recovers no more than the efficient cost of production.[23] While I have formally derived this result for a two-service firm, with suitable changes in notation, the same results hold for an arbitrary number of services.

I now turn to a discussion of organization costs. Spulber (1989, 1999) suggests that both internal and external contracting costs jointly determine the boundaries of the firm. *Internal contracting costs* are, for Spulber, the costs of contracting with employees, while *external contracting costs* are the costs of contracting with outsiders (customers, vendors). In the audit setting, the internal/external contracting dichotomy provides a natural and convenient way to organize the discussion of organization costs. I also consider the costs of coordination as part of my discussion of internal contracting costs in Section 3.2, since coordination costs determine the structure of the employment relationship which in turn affects internal contracting costs. External contracting is discussed in section 3.3.

3.2 Internal Contracting Costs.

Teece (1980) argues that for multi-product enterprise to be efficient the firm must enjoy "the recurrent use of proprietary knowhow or the common and recurrent use of a specialized and indivisible physical asset." Knowledge capital functions as such an asset for the large professional services provider described in section 3.1. However, Lloyd (1983), Spulber (1989) and others argue that for a single *firm* to sell multiple outputs (as opposed to producing the outputs using the indivisible or recurrent common factor), there must also be some uncertainty in demand or contracting frictions that preclude franchising or resale contracts for the multiple products.[24] The preferred organizational form will be one that minimizes the sum of these costs (Williamson 1975, see also Williamson and Masten, 1998, and Carroll and Teece, 1999, for recent overviews).

Two major sources of contracting frictions in large firms result from the communication and coordination problems and incentive problems that arise when the activities and information of multiple individuals have to be combined. Communication and coordination issues are the central concern of formal models of teams and hierarchies (Marschak and Radner 1972) and a large literature on organizing optimal hierarchies that satisfy some formal criterion of minimizing communication and coordination costs has developed over the past decade (Aoki 1990, Cremer 1990, Radner 1993 and 1996, Bolton and Dewatripont 1994, Meyer 1994, Friedman and Oren 1995, Prat 1997, Jehiel 1999). However these models generally ignore control costs. In section 3.2.1, I use these models to build a cost function for coordination costs within the firm. In section 3.2.2, I address control costs or the costs of frictions arising from incentive problems (see Prendergast 1999 and Gibbons 1999 for overviews).

3.2.1 Coordination Costs

I adopt the view of a firm as a network of decision-makers allocating resources to various tasks. Minimizing coordination costs for a network has firm has two principal components. First there is the process of distributing knowledge capital effectively throughout the firm for its maximal deployment in production of services -- this can be thought of as setting up a hierarchy of decision-makers to minimize

the actual costs of constructing the hierarchy (i.e. the costs of hiring and using the decision-makers). This question can be further broken down into questions like "How many people do we need? At what levels of the organization should they be placed? What should reporting relationships look like?" Second is the cost of coordinating in the sense of getting decision makers who may have local information to agree on a common plan of action. Here it is useful to think of the decision-makers for each assignment polling each other about their utilities from a proposed plan and using some set of efficient algorithms to use a price-like mechanism to select the optimal allocation of firm resources across all the firm's projects (Friedman and Oren 1995). The key cost here is the cost of delay in the sense of the time taken to obtain a decision.

For the first set of costs, it is useful to turn to the literature on organizational hierarchies (Radner 1993 and 1996, Bolton and Dewatripont 1994, Meyer 1994, Prat 1997, Jehiel 1999). These studies develop a theory of optimal organization structures absent incentive problems. Their results suggest two principal conclusions: First, these studies suggest that the optimal organization structure is one that arranges the processors in a hierarchy where the size of the hierarchy and the number of levels is a function of the capabilities of the processors (Prat 1997) and the likelihood of information loss (Jehiel 1999). The optimal mix of processors to be used is a function of their relative costs. The hierarchical model of the firm predicted by this literature is similar to the class of organization structures derived from the Cobb-Douglas production technology used in section 3.1 (Beckmann 1983, Doogar 1994). Thus it may be argued that for our present purposes, once the labor allocations in section 3.1 have been chosen optimally, the appropriate organization structure has been chosen given the relative capacities of the two labor inputs (respectively, $\alpha/(1-\alpha)$ for auditing and $\beta/(1-\beta)$ for consulting) and their costs (r, w, the respective input prices).

Second, this literature examines the delay costs associated with decision-making in networks. The *iron law of delay* (Radner 1996) sets a lower bound on processing times for information. To process N units of information arriving every T units of time, the minimum processing time is $1+\log_2 N$ and the

minimum number of processors required is $(1+\log_2 N)(N/2T)$. Thus as the amount of information to be processed increases, either the delay time increases or the number of people involved increases.[25]

When resources have to be allocated over a network, Friedman and Orren (1995) conjecture that the computational complexity of a price mechanism, described as the time T required to find a solution, is increasing in the number of tasks and people to be assigned and the level of accuracy and timeliness desired.[26] Since delay or error in disseminating information is potentially costly to non-audit clients and may lead to mis-estimating client assertions in the audit context, it is sensible to assume that delay costs are non-linear, and more tractably, strictly convex in the amount of delay. I therefore assume that coordination costs are convex in the size of the firm's workforce and the size of the firm's clientele.

Let $\eta(f,l,\{A\},\{C\})$ be the total organizational and coordination costs for a firm with total labor demands f and l and total audit and non-audit clienteles $\{A\}$ and $\{C\}$ respectively. By arguments just presented, $\eta(\cdot)$ increases faster with increases in the size of the firm.

3.2.2 Control Costs

Control costs are the costs of inducing goal congruity in employees whose payoffs may not entirely coincide with the payoffs to the entire aggregate called the firm. Myerson's Revelation Principle (Myerson 1979) which says that an optimal contract cannot pay any agent less than the maximum utility they could achieve by lying. Therefore optimal contracts merely mean the principal can extract the maximal surplus possible given the fundamental information asymmetry prevailing.[27]

The literature on agencies suggests that when agent environments are highly correlated, the principal can reduce the costs of overcoming agent moral hazard, possibly quite significantly, through either tournaments or relative performance evaluation schemes (see Prendergast 1999 for a review). Thus, for routine tasks performed by lower ranks of the firm hierarchy, monitoring and ensuring goal congruity may be relatively less expensive. On the other hand, the audit task is complex, involves considerable judgement and output is hard to see. Consulting tasks involve considerable discretion and in both cases objectives and outcomes are determined by superiors. Very little is understood about

optimal contracting under such settings (Prendergast 1999: 58), but recent work suggests that optimal contracts, even with agents lower in the hierarchy, may be less effective in providing incentives than had been thought to be the case (Gibbons 1999, Lazear 2000, Baker 2000, Prendergast 2000). Consequently even when the firm writes second-best contracts, these contracts may not induce as much goal-congruence as expected and reliance on psychological reinforcement through notions of culture and through professionalization might be more effective than incentive contracting (Kreps 1990 and 1998, Gibbons 1999).[28] Acemoglu (1994) highlights how career concerns may impair individual auditors' independence. Thus, for an audit firm, to a first approximation, I assume that costs of internal contracting will increase both with the number of levels in the hierarchy as well as the size of hierarchy. They may also increase as the set of service professionals the firm has to manage grows more diverse. Let $\xi(f,l)$ denote the costs of contracting. By the foregoing arguments, $\xi(.)$ can be assumed to be increasing convex in the numbers of partners and staff.

Based on the discussion in 3.2.1 and 3.2.2, the internal contracting costs of a large professional services firm can be written as, to a first approximation, the sum of its coordination and control costs, i.e. $\Xi = \eta(f,l,\{A\},\{C\}) + \xi(f,l)$ each term of which is convex in its arguments.

3.3 External Contracting Costs

A firm's *external contracting costs* include both *relationship specific* costs and *reputation* costs. Relationship specific costs include the startup and discovery costs of establishing the limits of engagements (more costly for consulting than auditing, at least for SECPS clients), learning the exact nature and scope of activities that will be required and allocating responsibility for tasks between client and auditor (internal auditor, consultant), as well as costs of establishing and maintaining relationship-specific mechanisms for communication and dispute resolution. Reputation costs are the costs of entering into incomplete contracts with buyers of the services. I discuss each component in turn.

3.3.1 Relationship-specific costs

For some types of services, startup and discovery costs may be lower (installing an accounts payable system) and for others considerably higher (installing an ERP system, executive searches and strategic consulting). Similarly the dispute resolution costs may be low for a less critical service but very high for a strategic task. For instance, anecdotal evidence suggests that the post-1993 explosion of consulting revenues for Big Six audit firms is in large measure due to the demand for their expertise in information systems consulting for both audit and non-audit clients. It is therefore credible that for these types of services, firms will develop standardized contracting routines and offer clients menus of choices, significantly reducing startup and discovery costs. On the other hand, due to the crucial nature of these tasks for the efficiency of the client's business, relationship maintenance costs may be high.

The foregoing arguments suggest that external contracting costs, $\zeta(\cdot)$, are likely to be a function of the number of distinct practice specialities (i.e. activities requiring a critical mass of distinct skill inputs to successfully develop routines), the level of standardization of services (decreasing as a service becomes standardized, increasing with increasing customization), and, the criticality of service s to the client (and increasing in criticality).[29] Since I model only one additional activity, I assume that ζ, the external contracting costs is simply a (large) recurrent fixed cost that has to be paid each period.[30]

3.3.2 Reputation Costs and Information Externalities

The key external contracting costs that audit firms face is the cost of assuring investors and other financial statement users that independence was not compromised. Ceteris paribus, decreases in investor confidence in the auditor's independence increase the cost of capital and decrease the set of investment that come to market, i.e. slow down economic growth. It is not therefore surprising that current regulatory guidelines suggest that perception is given as important a role as substance in determining whether auditor conduct is appropriate (e.g. "Auditors must not only be independent, they must be seen to be independent." quoted in SEC 2000). On the other hand, the AICPA, professional firms, bodies as the Public Oversight Board and some academics dispute the importance attached to perceptions by

regulators.[31] They argue that perceptions, being inherently subjective, are not a good basis for policy formulation. An underlying theme to this argument is that markets can be expected to discipline the mix of audit and non-audit services firms chose, given the value investors attach to each task (Antle et. al. 1997). This view ignores important information voids present in the audit setting.

Auditor independence problems arise from the existence of future rents that client management can use as a bargaining chip to impair the auditor's independence in the event the auditor disagrees with management's financial reporting practices.[32] Absent rents, there is no independence problem. In Beck et. al. (1988), rents accrue to the audit firm. However, the analysis in 3.2 shows that sometimes employee rents, rather than rents earned by the firm may be the source of independence failures. Banning consulting services will not necessarily affect the component of the independence problem that relates to audit employee misconduct. This effect, while minor, has some important implications for investor beliefs. For one thing it implies that even in one-shot audit engagements, as long as there is imperfect internal contracting, rational investors will never assign probability one to the event that the auditor was, in fact, independent. Investors' expectations of the likelihood of independence failures will likely increase with the magnitude of rents from repeat engagements and also depend on their perceptions of firms' activities and promotion and compensation practices.[33]

On the other hand, from an investors' perspective, *an unambiguous determination of the auditor's independence or lack thereof is virtually impossible.* Independence is, to use a cliche, a state of mind, and is hard to establish ex-post. Second, when sued, audit firms can, and quite often do, settle lawsuits and seal the proceedings from public scrutiny, leading to a lack of clear public information about auditor conduct. However the settlement of lawsuits, at least in the United States, is not tantamount to an admission of guilt because the randomness of jury trials and massive discovery costs can often force innocent defendants to settle lawsuits. Consequently, once a lawsuit has been settled, the investor does not know whether a guilty auditor settled the lawsuit in order to "hush up" an independence failure or an innocent auditor settled to avoid substantial litigation costs.

The literature on information cascades (Banerjee 1992, 1993; Bikhchandani et. al., 1992) shows that in cases where there is no strong public signal about the underlying source of uncertainty, rational actors who observe only other people's actions but not their information, can easily fall into a pattern of herd behavior in which they ignore their private information and act as if the public signal were correct.[34] Given that both guilty and innocent auditors settle lawsuits, the prior belief that the auditor's independence was, in fact, compromised, cannot, based on the evidence of a settlement alone, ever be revised to eliminate all residual uncertainty about the auditor's independence.[35] Additional signals such as the presence or absence of SEC investigations or news articles that can be observed by investors are likely to be noisy as well in such cases.[36] Since there is no mechanism by which an honest auditor's private information can credibly be shared with investors, the information asymmetry in this problem is incorrigible. In such cases, despite lack of a "smoking gun" the regulator may have to act simply because of the risk that investor behavior may get locked into the undesirable equilibrium in which everyone acts as if auditors were not independent (Levitt 2000b).[37] This is even more likely to be the case when private signals are weak.[38]

The consequences of information voids will be further exacerbated by any reputational externalities in which suspicions about one firm cast a doubt on the value of the services of all other firms. There is some evidence that the large audit firms (Big Six, Big Five), form an *inter-se*-undifferentiated source of supply in the minds of users (Novak 1998). Consequently, when there is a highly publicized failure of a major audit firm's client, the public discussion tends to tar all firms with a fairly broad brush.[39] The undifferentiated image of the large audit firms in the minds of consumers and the business press creates an externality in the audit market in that it leads to all firms' reputations being "punished" for problems that occur at one firm's clients.

For simplicity, I ignore all interactions and simply assume that the likelihood of cascades increases with v, the total amount of consulting services purchased by audit clients. Let $P(v)$ denote the net social costs of misplaced persistent skepticism about auditor independence. This is the *externality*

imposed on society by audit firms providing non-audit services to audit clients. Thus the external contracting frictions give rise to two very distinct sets of costs: while relationship-specific costs ζ are borne by the firm, the reputation costs, $P(v)$ are borne by investors.

3.4 Summary

The total out-of-pocket costs, TC, of a firm providing a set of services $\{A,C\}$ is given by:

$$TC = TDC(A,C) + \Xi(f,l,A,C) + \zeta \qquad \textbf{(3)}$$

Section 3.1 showed that TDC displays strong economies of scope. Section 3.2 established arguments why $\Xi(f,l,A,C)$, the firm's total internal contracting cost i.e. the cost of coordination and control (motivation), is likely to be increasing and convex in the number of decision-makers (f), resources (l) and tasks (A,C), i.e. in the size of the firm's personnel and clientele. Section 3.3 motivated the existence of relationship specific costs ζ that are borne by the firm.[40] The next section examines the choice of optimal investment in knowledge capital made by a firm that has a cost function TC.

4. Market Clearing and Investment Choice.

To determine the optimal level of investment in knowledge capital, which will, as suggested by the discussion following Theorem 2A, determine the optimal scope of the firm, it is necessary to consider how product market competition affects the investment decision. Under price competition, a firm's ability to win clients depends only on its relative costs which in this model are function of the level of investment K. Revenues are set by the costs of its competitors or the values of the service to the client and therefore are not directly a control variable for the auditor. The firm's optimal level of investment is determined by the profit maximizing choice of the input vector $I=[K, f=f_a+f_c, l=l_a+l_c]$. It follows from this that when audits are mandated, even if the value from auditing drops below the audit costs of the most efficient auditor in the market, clients will have to buy audits at the price of the second-most-efficient auditor in the market, not at the value of the audit to the investor. In effect, once audits are mandated, audit prices depend on the cost structures of the firms rather than the value of the audit to the

investor. The managers who buy the consulting services on the other hand will still be willing to pay up to the minimum of other firms' bids or their own utility for those services. However due to the competitive conditions in the audit market, there is no way for investors' diminished value for the audit produced in conjunction with consulting to be reflected in the audit fees. Consequently, *P(v)* has to be borne by investors and audit fees cannot be used to discipline auditors. I show below that in such a setting, auditors always produce too much consulting relative to what investors would prefer.

First, though, it is useful to consider the choices of *K*, *f*, and *l* made by a firm that minimizes TC rather than TDC.[41] Due to the addition of organization costs, the optimal choice of K^{**} is now given by (see Appendix):

$$1 = r\left(\frac{f_{a_i}^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right) + \Xi_f\left(\frac{f_{a_i}^{**}}{\gamma}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\delta}\frac{C'(K^{**})}{C(K^{**})}\right) \qquad \textbf{(4)}$$

Ξ_f is positive and $A'(K^{**})/A(K^{**})$ and $C'(K^{**})/C(K^{**})$ are both single valued and monotone decreasing in *K*, so the value of K^{**} which solves **(4)**, i.e., the problem with organizational costs is larger than the K^{**} that solves (A.1) i.e. the problem without organization costs. Also, *df/dK* is negative, so the firm with organizational costs uses less labor of each type relative to a firm with no organizational costs. Since both *A(K)* and *C(K)* are concave, this implies rising marginal costs.

Investors will, however, prefer that firm size and scope be chosen to minimize the total social costs of production, *SC*, i.e., all the costs in **(3)** as well as the costs of all externalities:

$$SC = TC + \mathrm{P}(\nu) \qquad \textbf{(5)}$$

where $\nu = \sum_{i \ni a_i \neq 0 \wedge c_i \neq 0} c_i$ is the amount of consulting purchased by audit clients. Assume that *P(v)* is convex in *v*. By arguments similar to those used to solve **(4)**, it is straightforward to show that *dP(v)/dK* is negative. We then have the following result:

Theorem 4: Investors would prefer the firm to invest less than the firm finds privately optimal. ■

As the firm invests less in knowledge capital, *K*, its cost advantage relative to stand-alone firms that do not combine auditing and consulting decreases and it does less of both tasks. While smaller firms are less efficient so that society would be better off with larger firms in this model, a single activity firm also has lower internal contracting costs, does not have to pay the integration cost ζ and imposes fewer externalities on investors. This tension suggests an interesting prediction for spinoffs of consulting practices by accounting firms. When the sum of internal contracting costs, Ξ, increases more rapidly as a result of growth in the firm's size than do the cost reductions from incremental investments in knowledge capital, a firm may find it optimal to spin off its consulting practice once that practice grows sufficiently large. Spinning off the consulting practice gives rise to a stand-alone firm that does not bear the costs of carrying the audit staff in addition to the consulting staff i.e. does not have to pay the integration costs of ζ. Consequently, the firm will find it optimal to split the consulting practice into a standalone firm.[42] However, due to the fungibility of knowledge capital in **(1)**, any knowledge capital retained by the parent firm for its audit task inevitably allows it to compete for some consulting clients. For instance, Arthur Andersen found itself in direct competition with its spun-off consulting division, Andersen Consulting, primarily because of the complementarities in the skills and knowledge bases required for both audit and certain information-technology consulting tasks (Petersen, 1999).

It remains an empirical question whether the magnitude of organization costs suffices to induce spinoff at levels that are close to investor-preferred levels. If the endogenous divestiture threshold is quite close to the solution to **(5)**, no intervention may be necessary. If the endogenous divestiture threshold is "far" away from the socially optimal level (i.e. the solution to **(5)**), the privately chosen level of diversification may be too high. Regulatory intervention offers one mechanism to reduce audit firm scope. Making audits voluntary might better allow investors' assessments of potential independence impairment due to either audit incumbency or consulting rents to be reflected in audit demand and fees.

5. Summary and Conclusions

I construct a model of service production in which knowledge capital generates economies of scope and organizational costs increase with firm size. I show that investors' and auditors' preferences over the optimal scope of the firm diverge because of a number of contracting frictions: inelastic demand for auditing, incentive problems within the audit firm, client-specific quasi-rents that may be threatened if the auditor disagrees with the client and information voids resulting from a lack of strong public signals about the true level of auditor independence. Prior research has treated only subsets of these factors and thus been unable to offer much insight into the underlying causes of the regulatory problem.

In contrast, I show how the litigation settlement mechanism and inelastic demand for auditing play crucial roles in creating a divergence between investor and auditor preferences. My analysis shows how the ability to settle suits out of court combined with the difficulty of establishing auditor independence creates a setting where rational investors might be led to conclude that auditors most likely lack independence even when they are in fact independent. I also provide an economic rationale for the differences between the SEC's investor perspective which sees the amount of consulting that firms do as excessive and the firms' view of the same level of consulting as the natural fruit of their investment in understanding their clients. The insights generated from the analysis suggest that the framework I adopt is both tractable and fruitful.

References

Abdel-khalik, R. 1990. The jointness of audit fees and demand for MAS: A self-selection analysis. *Contemporary Accounting Research*, (Spring): 295-322.

Acemoglu, D. 1994. Monitoring and Collusion: "Carrots" Versus "Sticks" in the Control of Auditors. Working Paper 94-9, Department of Economics, MIT .

Association of Certified Fraud Examiners. (ACFE). *Cooking the Books.* video.

AICPA, American Institute of Certified Public Accountants. 1996. *Statement on Consulting Services.* AICPA, New York.

Antle R., and J. Demski. 1991. Contracting Frictions, Regulation, and the Structure of CPA Firms. *Journal of Accounting Research*, Supplement, 29:1-30.

Antle, R., P.A. Griffin, D.J. Teece and O.E. Williamson. 1997. An economic analysis of auditor independence for a multi-client, multi-service public accounting firm. A report prepared on behalf of the AICPA in connection with the presentation to the Independence Standards Board of "Serving the public interest: A new conceptual framework for auditor independence." download from: http://ftp.aicpa.org/public/download/members/div/secps/isb/0117194.doc.

Antle, R. 1999. Accounting firms, the accounting industry and accounting research. *The British Accounting Review*, 31: 1-13.

Antle, R. 2000. Testimony at the First United States Securities & Exchange Commission Hearing on Auditor Independence, July 26. Archived at http://www.sec.gov/rules/extra/audmin.htm.

Aoki, M. 1990. The participatory generation of information rents and the theory of the firm. In Aoki, M., B. Gustafsson and E.E. Williams ed. *The Firm as a Nexus of Treaties.* Sage, New York.

Baker, G. 2000. The use of performance measures in incentive contracting. *American Economic Review*, 90(2): 451-20.

Balachandran, B.V. and R.T.S. Ramakrishnan. 1987. A theory of audit partnerships: Audit firm size and fees. *Journal of Accounting Research*, 25(1): 111-26.

Banerjee, A.V. 1992. The economics of rumours. *Review of Economic Studies*, 60(2): 309-27.

Banerjee, A.V. 1993. A simple model of herd behavior. *Quarterly Journal of Economics*, 107: 797-818.

Baumol, W.J., J.C. Panzar and R.D. Willig. 1988. *Contestable Markets and the Theory of Industry Structure.* New York..

Beck, P., T. Frecka and I. Solomon. 1988a. A model of the market for MAS and audit services: Knowledge spillovers and auditor-auditee bonding. *Journal of Accounting Literature*, 7: 50-64.

Beck, P., T. Frecka and I. Solomon. 1988b. An empirical analysis of the relationship between MAS involvement and auditor tenure: Implications for auditor independence. *Journal of Accounting Literature,* 7: 65-84.

Beckmann, M.J. 1983. *Tinbergen Lectures on Organization Theory.* Springer-Verlag, New York.

Biggs, J. 2000. Auditors and consultants shouldn't be too close. *Wall Street Journal*, June 13.

Bikhchandani S., D. Hirshleifer and I. Welch. 1992. A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy*, 100(5): 992-1026.

Bikhchandani S., D. Hirshleifer and I. Welch. 1998. Learning from the behavior of others: conformity, fads, and informational cascades. *Journal of Economic Perspectives,* 12(3): 151-70.

Bolton, P. and M. Dewatripont. 1994. The firm as a communication network. *Quarterly Journal of Economics*, 109(4): 809-39.

Brick, M. 2000. SEC proposes changes in accounting rules. *New York Times*, June 27.

Brock, W.A. 1983. Review of "*Contestable Markets and the Theory of Industry Structure.*" *Journal of Political Economy.*

Burns, S. 2000a. SEC to seek comment on auditor independence rules. *Wall Street Journal*, June 27.

Burns, S. 2000b. Audit firms to square off vs. SEC on independence. *Wall Street Journal*, July 10.

Carroll, G.R. and D.J. Teece, ed.. 1999. *Firms, markets, and hierarchies: The transaction cost economics perspective*. Oxford University Press, 1999.

Coase, R.H. 1937. The nature of the firm. *Economics*, New Series 4: 386-405.

Cremer, J. 1990. Common knowledge and the coordination of economic activities. In Aoki, M., B. Gustafsson and O.E. Williams ed. *The Firm as a Nexus of Treaties.* Sage.

DeAngelo, L. 1981a. Auditor independence, 'low balling', and disclosure regulation. *Journal of Accounting and Economics*, 3: 113-127.

DeAngelo, L. 1981b. Auditor size and audit quality. *Journal of Accounting and Economics*, 3:183-199.

Doogar, R. 1994. Competition in the audit market: An industry equilibrium analysis. In *Three Analytical Essays in Accounting,* Ph.D. Thesis, The Pennsylvania State University.

Doogar R., and R.F. Easley 1998. Concentration without differentiation: A new look at the determinants of audit market concentration. *Journal of Accounting and Economics*. 28(3): 235-253.

Doogar, R, R.F. Easley and D.N Ricchiute. 2000. Switching costs, leverage and audit firm mergers. Working paper.

Economist. 1997. Double entries. December 17.

Economist. 1998. Too few auditors. September 15.

Economist. 2000. Auditory discomfort. January 15.

Friedman, E.J. and S.S. Oren. 1995. The complexity of resource allocation and price mechanisms under bounded rationality. *Economic Theory*, 6: 225-250.

Gibbons 1999. Incentives in organizations. *Journal of Economic Perspectives*, 12(4): 133-150

Goldwasser, D. 1999. The tasks awaiting the ISB. *Accounting Today*, June 13(10): 7.

Jehiel, P. 1999. Information aggregation and communication in organizations. *Management Science,* 45(May): 659-669.

Jenkins, H. 2000. Abused accountants say 'take this job and . . .' *Wall Street Journal*, January 19.

Kreps, D. 1990. Corporate culture and economic theory. In J.E. Alt and K.A. Shepsle eds., *Perspectives on Positive Political Economy*, 1990, pp. 90-143.

Kreps, D. 1998. Intrinsic Motivation and Extrinsic Incentives. *American Economic Review,* 87: 359-64.

Lazear, E.P. 2000. The power of incentives. *American Economic Review* 90(2): 410-15.

Lloyd, P.J. 1983. Why do firms produce multiple outputs? *Journal of Economic Behavior and Organizations*, 4(1): 41-51.

Levitt, T.J. 2000. Renewing the Covenant With Investors. Speech at the New York University Center for Law and Business, May 10. download from: http://www.sec.gov/news/speeches/spch370.htm.

Levitt, T.J. 2000b. Comments during the First Hearing on the Proposed Modification of the SEC's Auditor sIndependence Rules. July 26.

Macdonald, E. 1999. PricewaterhouseCoopers Plans Steps Assuring Auditor Objectivity. *Wall Street Journal*, November 18.

MacDonald, E. 2000. PricewaterhouseCoopers to Divide Into Two or More, Under Pressure. *Wall Street Journal*, February 18.

Marschak, J. and R. Radner. 1972. *Economic Theory of Teams.* Yale University Press, New Haven.

Meyer, M.A. 1994. The dynamics of learning with team production: Implications for task management. *Quarterly Journal of Economics*, 109: 1157-84.

Myerson, R. 1979. Incentive compatibility and the bargaining problem. *Econometrica* 47(1): 61-73.

Novak, G., 1998, Do clients care about the big 6 mega-mergers?, Accounting Today (March 16) 7.

O'Malley, S. 2000. Remarks made at the concurrent session titled *Public Oversight Board Panel on Audit Effectiveness: Report and Recommendations*. American Accounting Association Meetings, Philadelphia, August 14.

Palmrose, Z.V. 1986. The effect of nonaudit services on the pricing of audit services: Further evidence. *Journal of Accounting Research*, 24(Autumn): 405-11.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37(1):7-63.

Prendergast, C. 2000. What trade-off of risk and incentives? *American Economic Review* 90(2): 421-5.

Petersen, M. 1998. How the Andersens Turned Into the Bickersons. *The New York Times* (March 15), Section 3, 1.

Prat, A. 1997. Hierarchies of Processors with Endogenous Capacity. *Journal of Economic Theory*, 77: 214-222.

Radner, R. 1993. The organization of decentralized information processing. *Econometrica*, 61: 1009-46.

Radner, R. 1996. Bounded rationality, indeterminacy and the theory of the firm. *The Economic Journal*, 106(September): 1360-1373.

Ramsey, F. 1927. A contribution to the theory of taxation. *Economic Journal*, 37(March): 47-61.

Rasmussen. 1989. *Games and Information*. Blackwell, New York.

Scheiner, J.H. 1984. An empirical assessment of the impact of SEC non-audit service disclosure requirements on independent auditors and their clients. *Journal of Accounting Research*, 22(Autumn): 789-97.

Scheiner, J.H. and J. Kiger. 1982. An empirical assessment of auditor involvement in non-audit services. *Journal of Accounting Research*, 20(Autumn Pt II): 482-96.

Schroeder, M. 2000. SEC probes Andersen for conflict of interest. *Wall Street Journal* August 25.

SEC, Securities and Exchange Commission. 2000a. Fact Sheet: The Commission's Proposal To Modernize the Rules Governing the Independence of the Accounting Profession. Published at http://www.sec.gov/news/extra/audfact.htm.

Securities and Exchange Commission (SEC). 2000b. Proposed Rule: Revision of the Commission's Auditor Independence Requirements. Published at http://www.sec.gov/rules/proposed/34-42994.htm.

Simunic, D.A. 1984. Auditing, consulting and auditor independence. *Journal of Accounting Research*, 22(Autumn): 679-702.

Spence, M.A. 1983. Contestable Markets and the theory of Industry Structure: A review article. *Journal of Economic Literature*, 21(September): 981-990.

Spulber, D.F. 1999. *Market microstructure: intermediaries and the theory of the firm.* Cambridge University Press, New York.

Spulber, D.F. 1989. *Regulation and Markets*. MIT Press, Cambridge, Mass.

Teece, D.J. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organizations*, 1(3): 223-247.

Teece, D.J. 1982. Towards an economic theory of the multi-product firm. *Journal of Economic Behavior and Organizations*, 3(1): 39-63.

Telberg, R. 1999. Auditing the auditors' auditors. *Accounting Today* 13(17): 6.

Williamson, O.E. 1975. *Markets and Hierarchies*. Free Press.

Williamson, O.E. and S.E. Masten. 1999. *The Economics of Transaction Costs*. Elgar.

Appendix

Lemma: Derivation of optimal solutions shown in (A.1)

The firm minimizes **(2)** subject to **(1)**. Use **(1)** to eliminate l_a and l_c from **(2)** to obtain an unconstrained minimization program:

$$\underset{K,f_{a_i}}{Min}\ K + r\sum_i (f_{a_i}+f_{c_i}) + w\sum_i \left[\left(\frac{a_i}{A(K)}\right)^{\gamma+1} f_{a_i}^{-\gamma} + \left(\frac{c_i}{C(K)}\right)^{\delta+1} f_{c_i}^{-\delta}\right]$$

The first order condition w.r.t f_{a_i} and f_{c_i} is:

$$r = \gamma w\left(\frac{a_i}{A(K^{**})}\right)^{\gamma+1} f_{a_i}^{**-(\gamma+1)} = \delta w\left(\frac{c_i}{C(K^{**})}\right)^{\delta+1} f_{c_i}^{**-(\delta+1)}$$

from which one may solve for $f_{a_i}, f_{c_i}, l_{a_i}$ and l_{c_i}. The first order condition wrt K yields:

$$1 = (\gamma+1)\, w\left(\frac{a_i}{A(K^{**})}\right)^{\gamma+1} f_{a_i}^{**-\gamma}\frac{A'(K^{**})}{A(K^{**})} + (\delta+1)\, w\left(\frac{c_i}{C(K^{**})}\right)^{\delta+1} f_{c_i}^{**-\delta}\frac{A'(K^{**})}{A(K^{**})}$$

$$= \frac{\gamma+1}{\gamma} r f_{a_i}^{**}\frac{A'(K^{**})}{A(K^{**})} + \frac{\delta+1}{\delta} r f_{c_i}^{**}\frac{C'(K^{**})}{C(K^{**})} = r\left(\frac{f_{a_i}^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right)$$

Proof of Theorem 1: Since production occurs after the market has cleared, it suffices to consider only cost minimizing input vectors so that the optimal input vector I^{**} minimizes **(2)** subject to **(1)**. I^{**}, being an input vector, takes values in $\mathbb{R}^3_+$, the three-fold product space of the non-negative half line, $\mathbb{R}_+$. Let $\gamma=\alpha/(1-\alpha)$ and $\delta=\beta/(1-\beta)$. For a single client firm with a client demand $\{a,c\}$ for the two services, **(1)** always holds with equality and the preceding **Lemma** shows that following equilibrium relationships hold:

$$f_a^{**} = \left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{a}{A(K^{**})}; \qquad f_c^{**} = \left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\frac{c}{C(K^{**})}$$

$$l_a^{**} = \left(\frac{r}{\gamma w}\right)^{\frac{\gamma}{\gamma+1}}\frac{a}{A(K^{**})}; \qquad l_c^{**} = \left(\frac{r}{\delta w}\right)^{\frac{\delta}{\delta+1}}\frac{c}{C(K^{**})} \tag{A.1}$$

$$1 = r\left(\frac{f_a^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})}+\frac{f_c^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right)$$

Separability in the labor inputs for each task in **(1)** implies the ratios f_a^{**}/l_a^{**} and f_c^{**}/l_c^{**} i.e. the labor input proportions on each task are independent of the size of the other task and of the level of capital investment (K^{**}). The level of capital investment, K^{**} depends on both tasks: the last equation in (A.1) shows that the firm chooses K^{**} such that the marginal benefit of spending is equal to the cost savings from the decrease in the demand for labor across both tasks. Recasting this last equation in the form of

equation (A.2) below shows that the optimal level of K is a function of wage rates and other exogenous technological parameters alone. Substituting the solutions for f_a^{**} and f_c^{**} into the last equation in (A.1) yields:

$$1 = r\left[\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{A'(K^{**})}{A(K^{**})^2} + \frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\frac{C'(K^{**})}{C(K^{**})^2}\right] \tag{A.2}$$

Setting $a=c$, $A(K)=C(K)$, $\alpha=\beta$ reduces the model to one of a multi-client audit firm studied in prior work with the additional feature that the firm in this model has a strongly positive return to investment in knowledge capital.[43] Let the solution to the case c=0 be $I^*=[K^*, f_a^*, l_a^*, 0, 0]$. Substituting this into the last expression in (A.1) and comparing to the case c>0 yields

$$r\left[\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{A'(K^{*})}{A(K^{*})^2}\right] = r\left[\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{A'(K^{**})}{A(K^{**})^2} + \frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\frac{C'(K^{**})}{C(K^{**})^2}\right] = 1$$

$$\Rightarrow \frac{A'(K^{*})}{A(K^{*})^2} - \frac{A'(K^{**})}{A(K^{**})^2} = \psi\frac{C'(K^{**})}{C(K^{**})^2}, \qquad where\ \psi = \frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\left[\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\right]^{-1} > 0 \tag{A.3}$$

When $A(K)$ is a concave (increasing) function of K, $A'(K)/A(K)$ is everywhere positive and decreasing in K and as long as $C(\cdot)$ is an increasing function of K, it follows from (A.3) that $K^{**} > K^*$.[44] By exactly the same arguments in the case where $C(\cdot)$ is concave increasing in K, it follows that joint production increases the firm's level of investment relative to a firm the produces only c as well. Jointly, these two results imply that the demand for knowledge capital is non-decreasing in the number of clients demanding each service. ■

Proof of Theorem 2: Note that the first two equations in (A.1) and $K^{**}<K^*$ imply that $f_a^{**}>f_a^*$ and $l_a^{**}>l_a^*$ and $f_c^{**}>f_c^{**}(A(K^*))$ and $l_c^{**}>l_c^{**}(A(K^*))$ where $f_c^{**}(A(K^*))$ and $l_c^{**}(A(K^*))$ are the optimal labor demands for task c when the firm sets $K=K^*$. In other words, since from (A.1) , cost minimization for a given level of investment implies that firms buy labor inputs in fixed proportions, independent of the level of investment in K, but when the level of capital is higher, firms buy proportionately less of each type of labor for each job. It thus follows from (A.1) that by investing less a joint product firm has to buy more auditing labor than it did before and *more* consulting labor than it would have had to had it not decreased investment. This cannot be a cost minimizing choice. The firm could do better simply by not decreasing investment below the level of the undiversified firm. However, if both A and C are strictly concave functions of K, the multi-product firm will do better by increasing investment relative to the levels of the single product firms. Moreover, substituting the solutions for $f_a^{**}, f_c^{**}, l_a^{**}$ and l_c^{**} into **(2)** yields:

$$\begin{aligned} DC &= K^{**}(A,C) + \left[r\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}} + w\left(\frac{r}{\gamma w}\right)^{\frac{\gamma}{\gamma+1}}\right]\frac{\sum_i a_i}{A(K^{**}(A,C))} + \left[r\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}} + w\left(\frac{r}{\delta w}\right)^{\frac{\delta}{\delta+1}}\right]\frac{\sum_i c_i}{C(K^{**}(A,C)} \\ &= K^{**}(A,C) + (\gamma+1)\, w\left(\frac{r}{\gamma w}\right)^{\frac{\gamma}{\gamma+1}}\frac{A}{A(K^{**}(A,C))} + (\delta+1)\, w\left(\frac{r}{\delta w}\right)^{\frac{\delta}{\delta+1}}\frac{C}{C(K^{**}(A,C))} \end{aligned} \tag{A.4}$$

The cost of production is linear in the size of the clientele of each type of service but since the choice of K^{**} is an increasing function of both A and C, the total auditing and consulting clienteles, the average cost function is strictly decreasing. ■

Proof of Theorem 2A: Let the optimal investments by these two "specialist" firms be k_a^* and k_c^* and that by the multi-product firm be k^{**}. Then it follows from rewriting equation (A.2) that:

$$\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{A'(k_a^*)}{A(k_a^*)^2}+\frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\frac{C'(k_c^*)}{C(k_c^*)^2} = 2\left[\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}\frac{A'(K^{**})}{A(K^{**})^2}+\frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}\frac{C'(K^{**})}{C(K^{**})^2}\right] \quad \text{(A.5)}$$

Let

$$\mu = \frac{\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{1-\alpha}}{\frac{a}{\alpha}\left(\frac{\gamma w}{r}\right)^{\frac{1}{\gamma+1}}+\frac{c}{\beta}\left(\frac{\delta w}{r}\right)^{\frac{1}{\delta+1}}},\ \phi_1(k) = \frac{A'(k_a^*)}{A(k_a^*)^2}\ \textit{and}\ \phi_2(k) = \frac{C'(k_c^*)}{C(k_c^*)^2} \quad \text{(A.6)}$$

When A and C are concave increasing in k, ϕ_i, $i=1,2$ decreases in k and $\phi_i(k)>2\phi_i(2k)$. Then for k_a^*, $k_c^*>0$, (A.5) implies

$$\mu\phi_1(k_a^*+k_c^*)+(1-\mu)\phi_2(k_a^*+k_c^*) < \mu\phi_1(k_a^*)+(1-\mu)\phi_2(k_c^*) = 2\left(\mu\phi_1(K^{**})+(1-\mu)\phi_2(K^{**})\right) \quad \text{(A.7)}$$

so that $K^{**}/2 < k_a^*+k_c^*$. Thus a multiproduct firm will not invest less than the total investment of two specialist firms optimally serving the same clients, but it will not invest more than twice as much either.[45]

Proof of Theorem 3: The proof requires that the production technology lead to a natural monopoly in at least *one* market and that no full supplier of consulting alone can compete with a full supplier of auditing and consulting. Natural monopoly in *each* market is trivial from (A.3). The second part follows from noting that input vector $\{\max(k_a^*, k_c^*), f_a^*+f_c^*, l_a^*+l_c^*\}$ is less costly than the cost of the sum of the stand-alone input vector $\{\max(k_a^*+k_c^*), f_a^*+f_c^*, l_a^*+l_c^*\}$. If the input vector $\{k^{**}, f_a^{**}+f_c^{**}, l_a^{**}+l_c^{**}\}$ did not cost less than $\{\max(k_a^*+k_c^*), f_a^*+f_c^*, l_a^*+l_c^*\}$, the firm could simply invest $\max(k_a^*, k_c^*)$, but from the discussion immediately following (A.3) the multi-product firm finds it optimal to invest more than $\max(k_a^*+k_c^*)$. Hence the input vector $\{k^{**}, f_a^{**}+f_c^{**}, l_a^{**}+l_c^{**}\}$ must cost less than $\{\max(k_a^*+k_c^*), f_a^*+f_c^*, l_a^*+l_c^*\}$. ■

Derivation of Equation (4): Now the firm minimizes **(3)** subject to **(1)**. Eliminate l_a and l_c from **(3)** and solve the unconstrained problem to obtain the first order condition wrt f_{a_i} and f_{c_i} as:

$$r + \Xi_f = \gamma w\left(\frac{a_i}{A(K^{**})}\right)^{\gamma+1} f_{a_i}^{**-(\gamma+1)} = \delta w\left(\frac{c_i}{C(K^{**})}\right)^{\delta+1} f_{c_i}^{**-(\delta+1)}$$

which can be solved for $f_{a_i}, f_{c_i}, l_{a_i}$ and l_{c_i} as well. Also note that $\frac{df}{dK} = \frac{f_a}{\alpha}\frac{A'(K)}{A(K)} + \frac{f_c}{\beta}\frac{C'(K)}{C(K)}$. The new first order condition w.r.t K is:

$$
\begin{aligned}
1 &= (\gamma+1)\,w\left(\frac{a_i}{A(K^{**})}\right)^{\gamma+1} f_{a_i}^{**-\gamma}\frac{A'(K^{**})}{A(K^{**})} + (\delta+1)\,w\left(\frac{c_i}{C(K^{**})}\right)^{\delta+1} f_{c_i}^{**-\delta}\frac{A'(K^{**})}{A(K^{**})} + \Xi_f\frac{df}{dK} \\
&= \left(r+\Xi_f\right)\left(\frac{f_{a_i}^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right) - \Xi_f\left(f_{a_i}\frac{A'(K^{**})}{A(K^{**})} + f_{c_i}\frac{C'(K^{**})}{C(K^{**})}\right) \\
&= r\left(\frac{f_{a_i}^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right) + \Xi_f\left(\frac{f_{a_i}^{**}}{\gamma}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\delta}\frac{C'(K^{**})}{C(K^{**})}\right)
\end{aligned}
\tag{A.8}
$$

Proof of Theorem 4: Investors would prefer the firm choose inputs to minimize **(5)** subject to **(1)**. Following the now familiar methods yields the first order condition wrt *K* which bears comparison to **(4)**

$$
\begin{aligned}
&= (\gamma+1)\,w\left(\frac{a_i}{A(K^{**})}\right)^{\gamma+1} f_{a_i}^{**-\gamma}\frac{A'(K^{**})}{A(K^{**})} + (\delta+1)\,w\left(\frac{c_i}{C(K^{**})}\right)^{\delta+1} f_{c_i}^{**-\delta}\frac{A'(K^{**})}{A(K^{**})} + \Xi_f\frac{df}{dK} + \mathrm{P_v}(\nu)\frac{d\nu}{dK} \\
&= r\left(\frac{f_{a_i}^{**}}{\alpha}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\beta}\frac{C'(K^{**})}{C(K^{**})}\right) + \Xi_f\left(\frac{f_{a_i}^{**}}{\gamma}\frac{A'(K^{**})}{A(K^{**})} + \frac{f_{c_i}^{**}}{\delta}\frac{C'(K^{**})}{C(K^{**})}\right) + \mathrm{P_v}(\nu)\,\theta\,\frac{C'(K)}{[C(K)]^2}
\end{aligned}
\tag{A.9}
$$

where θ is a positive term that does not involve *K*. By the logic used to argue that the solution to **(4)** is smaller than the solution to (A.1), it follows that the solution to **(5)** is smaller than the solution to **(4)**.

Endnotes

1. The Executive Summary of the SEC rule change proposal lists, among other factors: the widespread concern large audit firm mergers may leave too few competitors in the market for audits of large public clients (see also Economist, 1997, 1998), the public fracas over profits and control between the largely audit-oriented practice of Arthur Andersen and its erstwhile consulting arm, Andersen Consulting (cf. Petersen, 1998), the explosive growth of consulting operations that might lead to independence problems (cf. MacDonald 1999) and the widespread and highly negative publicity (cf. Economist, 2000) attendant on the discovery of widespread (but albeit arguably inconsequential -- see Jenkins 2000) violations of existing independence rules by partners of PricewaterhouseCoopers.

2. The American Institute of Certified Public Accountants (in both AICPA 1996 and Antle et. al. 1997) as well as some observers (Antle 1999) have expressed doubts about the utility of practice restriction proposals. The most recent SEC scope restriction proposals have been characterized as "directly harmful" by some observers (Burns 2000a) and as "on a collision course" with the recommendations of the AICPA's Panel on Audit Effectiveness (Burns 2000b). However, scope restrictions were also supported by some observers and investors (Telberg 1999, Goldwasser 1999, Biggs 2000).

3. The views of the Big Five firms on the question of divestiture (or more precisely, the scope of divestiture) of their consulting practices varies widely. Brick for instance reports "Some officials of (the) three accounting firms say that would bring an undesirable restructuring of their industry by forcing firms to split off their valuable consulting businesses. The other two large accounting firms have or are ready to split off their consulting units."

4. See, for instance, the exchange between the SEC Commissioners and representatives of the AICPA and major accounting firms at the second public hearings held in New York in September 2000 (the proceedings are archived at the SEC website at http://www.sec.gov).

5. Some business executives have stated that the disclosures required in the November 2000 rules makes it too risky for them to buy non-audit services from their auditors. In any event, these rules are subject to Congressional approval and will take effect only 60 days after publication in the Congressional Register for 60 days. The recent changes in political control of the U.S. presidency and Chairman Arthur Levitt's announcement (New York Times, December 21, 2000) that he will step down early in 2001, some observers suggest, leave open the possibility of reversal or rejection of these rules either by the next administration or by Congress.

6. One author of an AICPA-commissioned consulting report on auditor independence argued for instance that perceptions ought not constitute an appropriate basis for regulatory intervention on the grounds that for a long time the public perception was that the earth was flat. Less striking representations of the "*What* perception problem?" theme run through the testimony of many AICPA and audit firm representatives and the employees and academic consultants to these firms who testified.

7. In fact, except for Beck et. al. (1988a&b), prior studies of the joint production of auditing and consulting (e.g. Simunic 1984, Antle and Demski 1991) generally assume that there is no independence problem. See the next section for a more detailed discussion.

8. The guiding principle of the transactions cost literature is Coase's (1937) seminal insight that markets affect choice of organizational form just as much as they discipline firms' production and pricing choices.

9. For instance, in the first set of papers to draw a link between large audit firms, auditor independence and audit quality (albeit in a audit-only setting), DeAngelo (1981a and b) uses the term "auditor" and "he" interchangeably. This approach implicitly assumes that the behavior of audit firm employees is consistent with the welfare of the firm as a whole. Antle and Demski (1991) assume the auditor is a single person as well. Balachandran and Ramakrishnan (1987) is a rare exception in that it is the only other study I am aware of that conducts an economic analysis of the *internal* organization of the audit firm. Their analysis is conducted in a principal-agent setting, but despite the varying formalisms, the central concerns of our studies overlap to some extent.

10. Evidence from past frauds suggests that precisely these kinds of concerns are exploited by unscrupulous auditees (cf. the former ZZZ Best CEO and some-time convict Barry Minkow's comments in the ACFE video *Cooking the Books*).

11. Scheiner and Kiger (1982), Scheiner (1984), Simunic (1984), Palmrose (1986), Beck, Frecka and Solomon (1988b), Abdel-Khalik (1990) and numerous other studies present empirical evidence for the presence or absence of economies of scale, scope, knowledge spillovers and contracting costs.

12. The primary focus of these studies is to explain or explore the benefits of bundling auditing and consulting services. For instance, Simunic (1984) assumes that audit quantity and not audit quality is the appropriate choice variable and warns that his findings "do(es) not imply that joint performance of AMS and auditing is necessarily desirable." Similarly, Antle and Demski construct a model that explains multiproduction as a way for the client to negotiate a better fee for both audit and consulting tasks. They write (1991: 3): "Another intentional omission is auditor independence. ... The provision of consulting by CPA firms suggests that the benefits outweigh the costs; but our focus is on the nature of the benefits, not on the weighing of the benefits against the costs." Weighing costs versus benefits is central to my analysis.

13. From 1977 to 1999, consulting revenues as a share of total audit firm revenues rose from 12 percent to 50 percent while the share of auditing revenues dropped from 70 percent to 30 percent. Antle et. al. (1997) present evidence that consulting revenues really exploded in the post-1993 period. Since 1993, auditing revenues have been growing by 9 percent per year on average compared to consulting revenue growth of 27 percent per year (Levitt 2000). The rapid growth of audit firms' consulting practices makes independence *the* central issue in the current regulatory debate

14. In a similar model which examines the market for auditing alone, Doogar (1994) proves such an equilibrium exists. Doogar and Easley (1998) show that a version of the model in which the firm produces only audits, predicts Big Six and non-Big-Six audit market shares reasonably well. This study generalizes the single-product audit firm considered in these papers to a multi-product setting.

15. One can imagine a process by which over time, the second factor will eventually mature into the first factor, increasing the supply of the first factor (and decreasing the supply of the first). However, because of retirement or natural attrition, the supply of the first factor will also decrease over time. Meanwhile a fresh supply of accounting undergraduates will replenish the labor pool of the second type. In a steady state economy, one can imagine this process continuing over time. A formal analysis of this process is however, beyond the scope of this study.

16. This assumption, while strong, also ensures that the results of my analysis cannot be challenged on the grounds that still stronger economies of scale than what I have modeled would lead to different conclusions. Of course, if economies of scale are in practice weaker than what I have modeled, as might

well be the case, then the value of joint production would be even smaller and there would be even less rationale for combining large amounts of both services in a common provider.

17. All proofs (and derivations not in the body of the paper) may be found in the Appendix.

18. Consider for instance the technology $A(k)=b_0+b_1e^{ak}$. The function $g(K)=A'(K)A(K)^{-2}$ is non-monotonic with a maximum at $k=a^{-1}ln(b_1/b_0)$, being everywhere increasing to the left of k and everywhere decreasing to the right. Then, for *c* small enough the two-product firm may choose $K^{**}=K^{*}$, the choice of the audit-only firm.

19. Since fees for voluntary services are set by the minimum of the value of the service to the customer and other firms' costs of producing the service, lowering costs allows the firm to serve customers who would otherwise either have gone to other firms or might not have purchased the service at all.

20. It is obvious from **(1)** that when $A(K)$ is an increasing function of K and the exponents of the other factors sum to one, a doubling of all inputs (K,f,l) more than doubles output of both *A* and *C*. I should note that if labor displays diminishing returns then the net effect of increasing returns to knowledge and decreasing returns to labor inputs may well be indeterminate or decreasing. However as long as labor displays at least constant returns to scale, this technology as a whole also displays IRTS.

21. This point can be seen even more clearly if we take the production technology to be the 3-input Cobb Douglas production function $q=K^{\gamma}f^{\alpha}l^{1-\alpha-\gamma}$, $0<\alpha,\gamma<\alpha+\gamma<1$. Then, $A(K)=K^{\gamma}$ which is strictly increasing and concave. The classical 3-input Cobb-Douglas technology displays no economies of scale for each output, yet the fact that k also contributes to the cost of all other outputs (whether of A or C) introduces strong subadditivity and therefore everywhere decreasing average cost for a multi-product firm.

22. See, especially, Proposition 2A1, Baumol et. al., p. 18. Subadditivity of the cost function follows from parts (i) and (ii) of their proposition.

23. In this market, marginal cost-based incremental prices will not recover all costs of production and are therefore not sustainable. Among the sustainable pricing schemes proposed in the literature, an interesting one is Ramsey pricing in which markups above the variable cost of production decrease in the elasticity of demand for the product (Ramsey 1927, Baumol et. al. 1988). This would imply that the monopolist would employ higher markups on auditing and lower markups on consulting.

24. The auditing and consulting market qualifies on both counts. The human capital of audit and consulting firms is the firm's key asset and is rarely franchised or sub-contracted to other firms. In fact, if much of the human capital consists of tacit knowledge, then economies of scope for the *firm* arise precisely because it is very costly to communicate this tacit knowledge. Further, in a world of changing needs, while the demand for auditing may well remain predictable, demand shocks for consulting tasks alone would suffice to justify an audit firm's choice to provide additional services.

25. Holding the complexity of the organization constant, the minimum number of nodes required is roughly of the order of a polynomial in the logarithm of the amount information to be processed and the reciprocal of the delay time and accuracy desired (Radner 1993 and 1996, Friedman and Oren 1995).

26. More formally, the time to solution depends on 1) the probability Δ with which a solution of accuracy ε is desired, 2) the number, *n*, of tasks (processors to which resources should be allocated) available 3) the number, *r*, of resources to be allocated and 4) the computation times involved in a single

real number computation by a processor, *X*, in communication of a number, *I*, and in finding the utility of the solution, *Ω*, as follows:

$$T = \mathrm{O}\left(\left(r^3\left(\log\frac{nr}{\varepsilon\Delta}\right)\left[X\left(1+\frac{\log 1/\varepsilon}{n}+\frac{r^2}{n}+r^4\log(n/\varepsilon)\right)+Ir^2+\Omega r^2\log(n/\varepsilon)\right]\right)\right) \qquad \textbf{(11)}$$

where *O* indicates an order of magnitude relationship between the variables on the right hand side of the equation and the processing time, *T*. Externalities add some further delay to the process but the delay is still of the same order of magnitude. The RHS of **(11)** is concave in *n* and convex in *r* for reasonable values of *n* and *r*. Thus for a firm allocating *r* personnel over *n* client-task combinations, some with externalities (e.g. clients who acquire both audit and non-audit services) and some without externalities (e.g. clients who buy only one service), the time required to find the optimal allocation that satisfies the desired risk and accuracy parameters has the properties discussed in the text.

27. Rasmussen (1989:162) describes the import of the Revelation Principle as "the mother who agrees never to punish her daughter if she tells her everything will never hear any untruths." This is why agency contracts are often described as *second-best*.

28. Moreover, research in the cognitive psychology of auditing has indicated a number of puzzling if not troublesome patterns of individual information processing and decision-making that may indicate that even under reasonably correlated environments and with optimal contracts the firm may fail to achieve goal congruence because of information processing limitations. Moving up the hierarchy, as agents' environments become less correlated and as the demand for decision-making autonomy grows, the firm's ability to reduce the agent's slack falls even more. Against this enhanced scope for opportunistic behavior is the fact that the agent has more personal wealth invested in the firm (though not necessarily a larger proportion than the junior employees -- junior employees have more career concerns with no more exit opportunities than senior employees).

29. The particular variables of interest will vary from service to service but since the proposed scope restrictions in the areas of systems and internal auditing are likely to be the most controversial, I choose to focus on aspects of systems consulting for salience to the reader.

30. More generally, for a firm providing multiple non-audit services, this function may not be differentiable at a finite number of points (i.e. it may be a step function).

31. For instance the AICPA Chair, Robert Elliott, has argued that isolated instances of conjectured independence compromises should not be construed as indicative of a substantive problem (Schroeder 2000). Similarly, the head of the POB Panel on Audit Effectiveness recently remarked that if only a handful of newspaper headlines could be eliminated from the scene, the controversy over auditor independence would probably fade away (O'Malley 2000). Other observers have argued that regulatory intervention should discount public perceptions and delay intervention pending more attention to the value added by the joint production (Antle 2000).

32. The analysis in section 3.1 shows where these rents come from: the "gap" between the incumbent's costs and all other firms' costs is a perpetual source of rent for the incumbent. Note however, that future audit tasks generate rents as much as do future consulting tasks. Hence consulting rents at worst, merely exacerbate the rent problem, but they are by no means the sole source of the independence problem.

33. Practices such as promoting "rainmakers", compensating partners based on the amount of new business brought in or the profitability of the consulting operations or even the notion of having relationship managers who can enhance the client's one-stop shopping experience can raise investor doubts. Curiously, while the SEC's policy towards security issuance is disclosure-based ("sunshine is the best disinfectant") as opposed to merit-based (only "good" securities can be traded), its policy towards auditors is merit-based in that only qualified firms can perform audits of SEC registrants. No public disclosure is required of the auditors' finances, investments and compensation practices. Thus beyond a handful of summary measures of activity (revenues, partner and staff strengths, number of offices etc.), investors probably know very little about the finances and operations of the guardians of public disclosure.

34. Bikhchandani et. al. (1998:153) formulate the following interesting problem "Consider two scenarios. In both each individual starts with some private information, obtains information from predecessors, and then decides on a particular action. In the *observable actions* scenario, individuals can observe the actions but not the signals of their predecessors. (We) compare this to a benchmark *observable signals* scenario in which individuals can observe both the actions and signals of predecessors." They then show (ibid, 154) that the two scenarios are fundamentally different: "Because actions reflect information, it is tempting to infer that in only the actions of predecessors are observable, the public information set will also gradually improve ... we now show that a scenario of observable actions is actually quite different from a scenario of observable signals. In the observable actions case, individuals often converge on the same wrong action – that is the choice that yields a lower payoff. Furthermore behavior is idiosyncratic, in the sense that the error-prone choices of a few early individuals determine the choices of all successors."

35. Mapping the audit litigation and settlement scenario into the example used by Bikhchandani et. al., one may show that if, early on, just a few observers obtain a sequence of noisy private signals that indicate that independence may in fact have been compromised, and act on that information, e.g., by publishing critical commentary, filing complaints to the SEC or just passing on their information, it will be rational for investors later in the sequence to act as if the *mistaken* belief that independence was in fact correct, even if their private signals tell them that independence was probably *not* compromised. This paradoxical (and problematic) situation can persist for a long time unless a strong, clearly contradictory, signal is received. Actually, in such a setting it is far more likely that a client-auditor separation will occur and no strong contradictory information *can* emerge so that the subsequent actions of all rational individuals will reflect the *incorrect* belief that independence *was* compromised in that instance.

36. For instance, unless the SEC finds strong evidence of wrongdoing it may not prosecute or take other publicly visible actions. Press reports are inherently noisy since they are, unlike evidence produced in courts or directly from the files of clients and auditors, usually based on subjective and fragmentary information.

37. The case of a firm with a large number of clients is even worse. Even absent a single actual audit failure or compromised independence, if the parameters of the economy make out-of-court settlement of litigation economically optimal in some fraction of the cases, then the larger the firm's portfolio, the more often it gets sued and settles and the larger the likelihood of information cascades about the reputation of the firm. These costs may be worsened by the interaction between increasing contracting costs as the number of firm members increases and as the firm's clientele increases.

38. Bikhchandani et. al. (1998, 155) point out that when "private signals are very noisy; specifically the probability that the signal is correct is p-0.51 ... there is an approximately 75 percent chance that a(n) ... cascade forms after the first two individuals!"

39. For instance, in Schroeder (2000), the discussion of audit failure allegations at Waste Management, an Arthur Andersen client also included a discussion of problems at Pinnacle, a Price Waterhouse client and of a 1996 lawsuit involving Phar-Mor, a Coopers and Lybrand client and a negatively defensive portrayal of the role of the AICPA. On balance the evidence from various studies of fee premia as well as the tenor of discussions in the business press suggests that client failures of highly visible audit firms lead to negative perceptions about all audit firms.

40. I could assume that $\zeta(C)$ is a function of the amount of consulting work which would only strengthen my main result that firm size is no longer unbounded, i.e. the market is no longer is a natural monopoly. Assuming ζ constant allows me to drop a partial derivative in the analysis and simplify the exposition.

41. After the market has cleared and auditors and consultants chosen by all clients active in a given period, in the short run, each firm chooses the amount of staff it will hire. In the long run however, each firm must also choose its level of partners and investment in knowledge capital as well as the scope of the services it will offer.

42. This point can be demonstrated more formally as follows: As the firm invests less in K, the demand for labor inputs increases and the costs of complexity rise. In addition a stand-alone firm does not have to bear the cost ζ of integrating the two activities within the firm. Thus a multi-product firm has the cost structure $TC=TDC(A,C)+\Xi(f,l,A,C)+\zeta$ while two standalone firms will have the cost structure $TDC(A)+TDC(C)+\Xi(f,l,A)+\Xi(f,l,C)$. The cost difference between the multi-product firm and two stand-alone firms may be written as

$$\mu = \left[TDC(A,C) - \left(TDC(A) + TDC(C)\right)\right] + \left[\Xi(f,l,A,C) - \left(\Xi(f,l,A) + \Xi(f,lC)\right)\right] + \zeta$$

Due to economies of scope, the first term w ill be negative, due to convexity of organization costs, the second term will be positive for consulting practices that are sufficiently large and the third term is positive. The firm will prefer to organize as two stand-alone firms when $\mu > 0$. However, once the parent firm has shed its consulting practice, the fact that the residual investment of *K it has retained for its audit task* will continue to have applications to consulting and the firm will have incentives to provide consulting services once more, in effect competing with the spinoff unit.

43. Formally, setting $A(K)=K^0$ so that $A(K)$ is identically 1 yields the special cases studied in prior work.

44. $A(K)$ is increasing in K and concavity ensures that $A'(K)$ is decreasing in K.

45. Can we establish tighter bounds? For instance, would be nice to be able to show $K^{**}<k_a^*+k_c^*$. In general no: even assuming A''' and C''' positive (which would suffice to make ϕ_i convex decreasing in K) would not do the trick and further restrictions appear to be quite ad-hoc and suspect. The reason why this intuition does not work in general is insightful about the economies of scope modeled here: as the firm does both jobs, the return to *each* unit of K is higher, so it may increase investment beyond the sum-of-the-stand-alone-levels. In other words, due to the strong complementarities in the demand for the common input K, the multi-task and/or product firm may well invest more in knowledge production than the sum of the investments made by single-client or single product firms. Thus the model confers a natural advantage in knowledge production to larger firms.