THE JOURNAL OF FINANCE • VOL. XLVIII, NO. 2 • JUNE 1993

Limitation of Liability and the Ownership Structure of the Firm

ANDREW WINTON*

ABSTRACT

This paper models the optimal choice of shareholder liability. If investors want managers to be monitored, the monitors should be residual claimants (shareholders), and monitoring and firm value will increase as shareholders commit more of their wealth to the firm. When liquidating wealth is costly, contingent liability dominates direct investment as a wealth commitment device; however, if wealth is unobservable, under this regime only relatively poor investors will hold shares in equilibrium. This may be prevented at a cost by verifying shareholder wealth and restricting stock transfers. Comparative statics on various liability regimes are used to motivate actual contractual arrangements.

MUCH HAS BEEN WRITTEN on designing and using financial securities to control agency conflicts among a firm's investors, or between managers and investors.[1] This literature has largely focused on the allocation of the firm's cash flows and control rights. Less attention has been paid to a related set of issues: how owners (investors with legal control of the firm) choose their financial commitment to the firm and to other investors, their ability to trade their ownership rights, and their actual control over management. More briefly, what is the firm's optimal ownership structure?[2]

Many researchers implicitly assume one form of ownership, namely that of the large modern corporation. This usually has the following features: (i) the firm is owned by a large number of shareholders, (ii) shareholders bear no

* Kellogg Graduate School of Management, Northwestern University. I am grateful to Franklin Allen, Chitru Fernando, Gary Gorton, Bruce Grundy, Sreenivas Kamma, and Morris Mendelson for advice and comments, and to seminar participants at the 1991 Hilliard-Lyons Symposium on the Design of Contracts, Institutions, and Markets and the following universities: British Columbia, Chicago, Illinois, MIT, Ohio State, Northwestern, Rochester, and Washington. Part of this research was made possible by a Unisys Fellowship.

[1] Two recent surveys are by Allen and Winton (1992), who focus on security design, and Harris and Raviv (1991), who focus on the optimal mix of debt and equity.

[2] Previous work in this area includes Fama and Jensen (1983a, 1983b, 1985), who discuss ways in which the choice of ownership structure is influenced by tradeoffs between risk sharing, agency costs, and the nature of the organization's activities. While they address issues such as whether ownership and control should be separated, who should have residual claims on the organization, and how residual claimants may exercise control over agents, I assume a certain amount of organizational form and focus more narrowly on the form and liquidity of ownership claims for a firm. However, my discussion of professional partnerships largely parallels theirs.

liability (beyond initial investments in their shares) for the firm's obligations, and (iii) shares are freely transferable. However, this structure is far from unique in practice. In this paper I model the circumstances under which this "typical" form is optimal, and those under which regimes making use of additional shareholder liability and/or trading restrictions dominate. Some past and present examples of these other regimes now follow.

Until the middle of the 19th century, unlimited shareholder liability for company debts was in fact the rule in Britain. Even when limited liability was allowed on a general basis, the new regime allowed shareholders to be liable for a fixed amount over and above their initial investment. A survey of limited liability corporations formed in Britain from 1863 to 1866 reveals that average paid-in capital was less than 10 percent of the total amount for which shareholders were liable.[3] Double liability was the rule for nationally chartered banks in the United States from the mid 19th century until the early 20th century, and at least one British bank did not eliminate "uncalled liability" until 1957.[4] Finally, additional liability is still used by corporations in Australia and New Zealand, although the liability must be due and payable on a set date in the future if the firm's shares are to be publicly listed.[5]

In the present day, the well-known syndicates of Lloyd's of London are unlimited liability partnerships which play a significant role in certain areas of the insurance industry.[6] Partnerships with unlimited liability are not uncommon in the United States, particularly in fields such as accounting, law, and medicine. Key shareholders of closely held corporations often give personal guarantees for the firm's debt.[7] Finally, the notion of zero liability has been overridden on occasion in the U.S. courts, where shareholders have been forced to pay damages over and above their initial investment.[8]

As with zero liability, the ability to transfer shares freely is not as universal as one might think. Most of the more recent examples of additional

[3] Cf. Hunt (1936), p. 150.

[4] This was Martins Bank. Its uncalled liability then totaled £16.4 million, compared with paid-in capital of £4.3 million and total advances (loans) of £91.5 million; cf. Chandler (1964), pp. 517, 519. Uncalled liability was a feature of all the large British banks well into the 20th century; cf. Sykes (1926), p. 88.

[5] Cf. Bruce, McKern, Pollard, and Skully (1986), pp. 86–91. These shares should not be confused with contributing ("no liability") shares, issued primarily by mining companies both in Australia and the United States. The additional liability on contributing shares is optional; owners can refuse to pay, in which case they forfeit their shares. Brennan and Dunlop (1991) model this type of contract as a response to managerial overinvestment.

[6] Cf. Hodgson (1984).

[7] Cf. Greve (1987), pp. 7–8, and Halpern, Trebilcock, and Turnbull (1980).

[8] This principle, known as "piercing the corporate veil," typically involves corporations of two types: closely held ones, and subsidiaries of publicly held firms. Here, concentration of ownership suggests that owners can be expected to monitor, and therefore bear some responsibility for losses; conversely, the threat of such liability will also give owners more incentive to monitor. See Easterbrook and Fischel (1985).

liability shares cited above also restrict the trading of shares; generally, these shares cannot be sold anonymously in public markets, and in some cases they may not be sold at all without the approval of other shareholders. This suggests a link between limitation of liability and the ability to trade. However, in 19th-century Britain, the move towards publicly traded firms actually preceded the general adoption of limited liability (much less zero liability) by several decades, so the relationship is not perfect.

Thus, a model that analyzes the efficient choice of ownership structure must explain a wide range of outcomes. However, previous work on this subject has focused on the drawbacks of additional liability without describing any advantages that might lead to its use. Jensen and Meckling (1976) point out that, under unlimited liability, shareholders will wish to monitor each other's wealth so as to estimate their own liability, which will be quite costly if shares are widely held.[9] Halpern, Trebilcock, and Turnbull (1980), Easterbrook and Fischel (1985), and Woodward (1985) discuss the adverse selection effects associated with additional liability which are derived in Section III of this paper.[10]

The difficulty with these arguments is that the drawbacks they describe can be overcome by having shareholders invest their wealth in the firm from the start. Thus, it is not clear why *any* form of contingent liability (limited or not) has been used. By contrast, I model an explicit advantage that contingent liability has over initial investment, and I focus on the choice between these two methods of wealth commitment.

Because the relative costs of *contingent* liability (unlimited or otherwise) are linked to lack of information about investors' wealth rather than to risk aversion, I simplify the analysis by using a risk-neutral setting. A firm with two possible levels of output is operated by a manager, whose ability or costly effort affects the distribution of returns. If ability or effort is not freely observable, there is a role for costly evaluation or monitoring by the firm's outside investors.

I assume that the bulk of one's wealth is typically illiquid (real estate, human capital, and so forth). As a result, ceteris paribus, it is cheaper to finance the firm with small investments from large numbers of investors.

[9] This argument is used by Carr and Mathewson (1988) to support their claim that laws mandating unlimited liability serve only as a barrier to entry: that is, incumbents with limited liability induce regulators to impose unlimited liability on potential entrants. However, by exogenously fixing managerial performance, they leave no role for the shareholder bonding arguments made in this paper, ruling out any benefit to additional liability from the start. Thus, their model cannot explain the situation in law or medicine, where unlimited liability was initially imposed on *all* firms.

[10] Of these, Woodward's discussion is closest to this paper. She describes how additional liability creates incentives for wealthier shareholders to sell their shares to poorer investors. However, she does not formally model the monitoring role of shareholders, nor does she consider liquidation costs.

However, monitoring or evaluation is best given to a relatively small number of investors so as to keep free rider problems and other costs of cooperation at a minimum. Since monitoring benefits the firm as a whole, but the effort involved is unobservable, there will be moral hazard problems both among the monitors and between them and the other investors. Giving the monitors more of the firm's risk gives them more incentive to monitor so as to reduce the probability of poor firm returns. Thus, an arrangement where several large investors (shareholders) own the firm and take a residual position relative to the other investors (debtholders) will dominate. I show that the shareholders will in fact take on as much risk as their wealth allows, and that they will prefer to do this through contingent liability rather than additional investment: the first arrangement avoids liquidation costs in at least some states of the world, whereas the second incurs such costs with certainty.

This result assumes that each shareholder's wealth is freely observable to all other investors. If this is not the case, the adverse selection problems described in the earlier literature come into play. The higher a shareholder's wealth, the lower the value of his shares, because more wealth means a larger effective share of the firm's liabilities. If trading in shares is allowed, this result will lead to an exodus of wealthy investors and an inflow of poorer ones, which will weaken the bond on the shareholders and reduce equilibrium monitoring and firm value. While this may be overcome by verifying each shareholder's wealth and/or restricting the transfer of shares, such solutions are likely to be costly. Which liability regime is most efficient will depend on the relative magnitudes of verification and liquidation costs, as well as the number of shareholders and their individual wealth levels.

The remainder of the paper is organized as follows. Section I describes the basic model and assumptions. Section II analyzes optimal ownership structure when investor wealth is freely observable. Section III introduces unobservable wealth and examines the relative performance of firms under various ownership regimes as the underlying parameters of the model change. Section IV interprets a number of modern and historical ownership structures in light of the comparative statics results of the previous section. Section V summarizes the paper and discusses possible extensions.

I. The Model

There is a technology (firm) which lasts two periods, requiring a fixed input of I units in the first period and returning a random amount of output in the second. Firm output is A with probability p and B with probability $1 - p$, where p depends on a manager and $A > I > B$. While output is freely observable, p is not. All agents (manager and investors alike) have utility that is time additive with a discount factor of one. The period utility function is $x - e$, where x is consumption and e is effort; thus agents are risk neutral.

There are two interpretations of the model, depending on whether p is a function of the manager's effort or of her innate ability; to simplify exposition,

I will focus on the first.[11] In this interpretation, the manager controls p with costly effort, where the cost can represent the disutility of diverting time away from personal (rather than investors') goals as well as aversion to work per se. Investors have the choice of making the manager's compensation depend on realized firm output or else monitoring her effort at some cost and writing an effort-based contract.

If investors observe the manager's effort choice, they can write a contract that depends directly on this, inducing her to choose the first-best level of p, p_f, at which total expected returns net of the manager's effort cost are maximized. If they do not observe her effort, they must rely on output based incentives. Optimally, the manager will be paid more in the good state than in the bad, and if her wealth is sufficient to let her bear all the firm's losses in the bad state, the first-best allocation of effort and expected output can be achieved (Harris and Raviv (1979)). If her wealth is limited, achieving p_f requires that the manager receive more than her reservation return; as a result, investors will typically choose a contract that leads to a probability of the good state, p_c, that is lower than p_f. For simplicity, assume that $p_c \cdot (A - B) + B - I$ is positive, so that the firm is profitable even with this second-best contract.

If an outsider wishes to observe (henceforth, "monitor") the manager's effort, he can do so with a probability of success q at an effort cost $C(q)$, where $C(0)$ equals zero and C is strictly increasing and weakly convex.[12] I make three more assumptions about this monitoring technology. First, monitoring gives perfect knowledge of the manager's effort when it succeeds and is otherwise completely uninformative. Second, monitors can verify any successful observations in a court of law. Third, the manager can condition her effort on the success or failure of monitoring (i.e., she knows whether or not she is being observed). These assumptions are chosen for simplicity; qualitatively, all that is needed is that monitoring is sequentially rational, that the signal received is verifiable in court and increases in precision with the amount of effort the monitor expends, and that a higher monitoring effort causes the

[11] In the adverse selection interpretation, investors must choose a manager whose hidden ability determines the probability of the good state. They can evaluate the manager's ability at a cost, or they can write output-based contracts that attempt to separate out the highest ability type. Hagerty and Siegel (1988) show that, ignoring the manager's wealth constraint, the optimal separating contract corresponding to each type in this adverse selection setting will be identical to the optimal contract required to elicit the equivalent level of effort in the case of moral hazard. Limited managerial wealth will require that separating contracts pay higher types more than their reservation return, and investors may instead choose contracts that result in some pooling. As a result, costly evaluation may be Pareto improving; however, proper incentives must be provided for the evaluator, which is essentially the same problem as that of providing shareholders with incentives to monitor managerial effort.

[12] The assumption of convexity is for ease of exposition. If there were nonconvexities, monitors could convexify their cost by randomizing the level of effort they used. Thus the case of linear costs cq can be reinterpreted as a fixed cost of monitoring c combined with randomized choice by the monitor.

manager to choose a higher effort level herself.[13] Finally, I assume that multiple monitors receive independent signals, and that the effort put into monitoring is not freely observable.[14]

For simplicity, I will ignore the manager's wealth and compensation, so that A and B are unaffected by changes in p. The probability that output equals A is p_f if one or more agents successfully monitor and p_c otherwise. This means that if n agents monitor with probabilities of success $\{q_i\}$, the expected probability of A is $P[Q(q_1, \ldots q_n)] = Q(\cdot) \cdot \Delta p + p_c$, where $Q(\cdot) = 1 - \Pi_{i=1}^{n}(1 - q_i)$ is the probability that at least one agent monitors successfully, and $\Delta p = p_f - p_c$.

Because successful monitoring improves overall firm performance, there is an externality between those investors who monitor and those who do not. Since the effort used to monitor is not freely observable, monitors cannot credibly contract to do any level of monitoring without either being monitored by others or receiving output-linked compensation. In this paper, I ignore hierarchies in which some investors monitor the manager and are monitored in turn, and focus on the case where those who monitor receive incentives based on firm output. As with the manager, an investor's wealth will limit how much of the firm's risk he can bear and thus how much effort he will put into monitoring.

Suppose investors fall into two types based on wealth. Large investors have most of their wealth invested in relatively illiquid assets. Specifically, there is a loss of k for each unit of wealth liquidated in the first period (whether for consumption or investment), where k is between 0 and 1.[15] Small investors are individually negligible but numerous, so that their total wealth is enough to totally finance the firm; their wealth is held in liquid form. That small investors have only liquid assets can be explained by indivisibilities in the size of illiquid assets or in the information or other costs involved in their acquisition. An alternative and more realistic interpretation is that the bulk of any investor's wealth is illiquid, so that the distinction between large and small is a convenience for distinguishing between those who commit only their liquid wealth to the project and those who commit more.

Illiquidity can arise from a number of frictions. Information asymmetries and search costs may make it costly to sell assets or use them as collateral for loans; examples include shares in small firms, real estate, and used cars.

[13] If monitoring were done ex post, sequential rationality would require that the manager be penalized for bad signals, and the manager would randomize her effort; the assumption that the manager knows whether or not she is being observed allows me to abstract from issues of managerial compensation.

[14] The assumption of independence is not necessary; even with correlated signals, there are equilibria in which monitors independently randomize their choices of effort. Whether such equilibria are dominated by those in which only one agent monitors will depend on the shape of the cost function $C(q)$.

[15] Although many liquidation costs might be present in both periods, this would not affect the model's qualitative results. This will be discussed in the next section.

Investors may have special advantages relating to certain assets, such as skills or experience in running a business, or idiosyncratic tastes with regard to homes and other material possessions.[16] Diamond and Dybvig (1983) suggest that long-term production technologies may physically involve a certain amount of illiquidity. Since much personal wealth is still invested in real estate, human capital, and small businesses, illiquidity is a factor even in the modern United States.

Several reasons may account for why investors hold the bulk of their wealth in such illiquid assets. Some wealth is inherited, in which case its liquidity is not a matter of choice. Much wealth takes the form of future labor earnings, so that loans against the future may be costly to obtain because of moral hazard concerns; a similar argument may apply to the case of small privately held businesses. Other illiquid assets may be freely acquired because of special benefits, the most obvious example being housing. Finally, this paper suggests that, if one's other assets are fairly illiquid, there may be incentives to adopt firm ownership forms that restrict trading; thus, asset illiquidity can be self-reinforcing.[17]

Returning to the subject of who monitors the manager, it should be clear that large investors will always dominate small investors at this task. Because their wealth is negligible, small investors can only absorb a negligible share of the firm's risk, leaving them without sufficient incentives to monitor (see Corollary 1). By contrast, large investors can acquire a sizeable share of the firm and its risk, giving themselves more incentive to monitor. Monitoring investors will optimally absorb as much of the firm's risk as they can, so they are residual claimants; this and the fact that they are responsible for evaluating and monitoring the manager makes these investors the effective owners of the firm in the legal sense.

II. Observable Investor Wealth

Assume n large investors ("shareholders") own the rights to the firm's technology, having either invented, inherited, or purchased it.[18] Since running the technology is always profitable, ownership of the firm is a valuable asset. In addition, each shareholder has "outside" wealth (i.e., wealth excluding the value of owning the firm) w, which is illiquid. For simplicity, assume

[16] Lacker (1990) uses this to help motivate the optimality of collateralized debt as a lending contract; this is in fact a form of contingent liability.

[17] If investors face preference shocks as in Diamond and Dybvig (1983), they will desire to keep some wealth in liquid form. If shares are even somewhat illiquid, this would be another reason for committing some wealth to the firm on a contingent basis rather than investing it up front. Of course, the possibility of such shocks might affect the wealth on hand to meet calls on contingent liability, further complicating analysis.

[18] Modeling the purchase of the technology would not change any of the model's qualitative results; a proof is available on request.

shareholders have no other liabilities outstanding that could lead to claims against their shares in the firm.[19]

As owners, the shareholders are collectively responsible for raising additional funds from themselves or other investors and for monitoring management. Assuming that additional monitors are not needed, funds raised from other investors are best raised from "small" investors, who can provide wealth in the first period without liquidation costs.[20] Since the negligible value of a small investor's wealth commitment to the firm means that he or she would never choose to monitor, and since risk sharing is not a concern, the original shareholders will prefer to issue a mix of securities which leaves themselves with as much of the firm's risk as possible. Although I will call the issued security debt, it could just as easily be considered to be a mix of riskless debt and zero liability equity.[21]

In order to maximize the net value of their shares, the shareholders face the following decisions. They must first decide the amount s of their outside wealth that they each will liquidate and invest in the firm (net investment is $s(1 - k)$), and how much additional wealth l they each will allow to be seized in the event that the firm defaults on its debt. Once s and l have been chosen and observed by other investors, the shareholders choose the current value D of outside debt which they wish to raise, along with the face value D' they will repay. Each shareholder i then privately decides how much monitoring q_i he will do. All decisions but the last are publicly observable and thus can be made collectively, as assumed. In addition, I have focused on symmetric wealth commitments for simplicity.

Defining $S = ns$ and $L = nl$, a number of conditions must hold. First, shareholders cannot commit more than their total wealth to the firm: $S + L \leq nw$. Also, the total amount raised in the first period must be enough to fund the firm: $S(1 - k) + D \geq I$. This will hold with equality, because observable investment rules out fraud and investors have no other preference for consumption today instead of tomorrow. Next, the expected payoff to creditors given the monitoring probabilities $\{q_i\}$ that will be chosen must weakly

[19] Although outside wealth could include additional liability shares of another firm, the possibility of claims against shareholders by creditors of the other firm would complicate analysis. Other possible components of shareholders' outside wealth include human capital, sole proprietorships, personally held real estate, and zero liability shares of small (hence illiquid) firms.

[20] Since large investors face liquidation costs when investing, the only reason to approach one would be to gain another monitor with a substantial wealth commitment to the firm—i.e., another large shareholder.

[21] Because the adverse selection problems discussed in Section III would make additional liability for small shareholders meaningless, any equity issued to them should have zero liability. If the original shareholders issued *only* additional equity to small investors, they would share output B in the bad state with the new shareholders on a pro rata basis; their risk would not be maximized, and any additional liability on their part would never come into play. This is why at least some of the issued securities should be senior to equity. Of course, factors such as risk aversion or costs of financial distress would tend to favor issuing more equity and fewer senior claims.

exceed their reservation return D. Finally, given S, L, and D', $\{q_i\}$ must make up a Nash equilibrium among the shareholders; that is, each chooses q_i to maximize his own return holding the other shareholders' choices constant. While many equilibria are possible, I will again focus on symmetric outcomes, so that $q_i \equiv q$.[22]

To solve the shareholders' problem of maximizing their returns, I begin with their choice of monitoring probabilities once the financing variables S, L, and D' are set. By symmetry, an individual shareholder's objective function is an equal share of total monetary returns to shareholders (firm profits less shareholders' initial investment S and payments to debtholders), less the individual monitoring cost $C(q_i)$. Payments to debtholders are the smaller of the debt's face value D' and the sum of firm output and shareholder liability L. The case where $D' > A$ can be ruled out, since in this case the shareholders would be better off not financing the firm in the first place. Thus, a shareholder's net return is given by

$$[P(Q(\cdot)) \cdot [A - D'] - [1 - P(Q(\cdot))] \cdot \min\{L, D' - B\} - S] \cdot n^{-1} - C(q_i), \tag{1}$$

which he maximizes over q_i in the interval [0, 1]. Let μ_1 and μ_0 be the Kuhn-Tucker multipliers associated with the upper and lower constraints on q_i. Given $q_i = q$ for all $j \neq i$, the necessary and sufficient condition for optimal q_i is

$$n^{-1}\Delta p \cdot (1 - q)^{n-1} \cdot [A - D' + \min\{L, D' - B\}] - C'(q_i) + \mu_0 - \mu_1 = 0. \tag{2}$$

An immediate result is

LEMMA 1: *In the problem described by equation* (1), *a unique symmetric outcome exists, with q given implicitly and uniquely by*

$$C'(q) = [A - D' + \min\{L, D' - B\}] \cdot (\Delta p/n) \cdot (1 - q)^{n-1} \tag{3}$$

if $(\Delta p/n)$ *times the term in braces is greater than* $C'(0)$, *and* $q = 0$ *otherwise.*[23]

Thus monitoring with certainty ($q = 1$) is never a symmetric equilibrium if there is more than one shareholder: monitoring has no benefit if someone else is sure to do it. The Implicit Function Theorem shows that q is decreasing in D' and n and increasing in L. It can also be shown that if q is positive, it will be strictly less than the value obtained by choosing all q_i's simultaneously and identically to maximize shareholder profits; this difference represents the free rider effect. This cooperative value is in turn no higher than the q

[22] Since shareholders are symmetric, asymmetric equilibria could best be justified by randomly assigning equilibrium monitoring probabilities to shareholders. The qualitative effects of increases in shareholders' wealth or numbers would be unchanged from the symmetric case analyzed here.

[23] Unless otherwise noted, all proofs are in the Appendix.

that would be chosen in maximizing firm returns net of monitoring costs, because shareholders don't always bear all of the risk of the firm. Since shareholder profits are strictly concave in q, it follows that changing the debt structure so as to bond shareholders to increase their individual monitoring without increasing the expected payout to debtholders will increase shareholder returns.

Moving back to the optimal choice of S, L, and D', notice that the same values maximize individual and total shareholders' returns. Letting $q(S, L, D') \equiv q$ denote the solution from Lemma 1, and using the fact that $D = I - (1 - k)S$, the shareholders' problem can be written as

$$\underset{S, L, D'}{\text{Max}} \; P(Q) \cdot [A - D'] - [1 - P(Q)] \cdot \min\{L, D' - B\} - S - nC(q) \quad (4)$$

such that

$$P(Q) \cdot D' + [1 - P(Q)] \cdot \min\{B + L, D'\} - I + S \cdot (1 - k) \geq 0, \quad (5)$$

(debtholders receive at least their reservation return),

$$nw - S - L \geq 0 \quad (6)$$

(shareholder commitments do not exceed their total wealth), and

$$S \geq 0, L \geq 0, D' \geq 0. \quad (7)$$

PROPOSITION 1: *In the solution to equations* (4) *to* (7), (*i*) *constraint* (5) *is always binding*; (*ii*) $S \equiv 0$; (*iii*) *if* $C'(0) < \Delta p(A - B)/n$, *there is an interval of values for* w, $(w_*, (I - B)/n]$, *with* $w_* \geq 0$, *such that within the interval, the constraint* $nw - L \geq 0$ *is binding and the optimal* q *increases with* w, *and outside the interval the constraint is not binding and* q *does not vary with* w; *and* (*iv*) *if* $C'(0) \geq \Delta p(A - B)/n$, *then the constraint is never binding and* $q \equiv 0$, *i.e., monitoring never occurs.*

Parts (i) and (ii) are fairly straightforward: debtholders receive their reservation return, leaving shareholders with the firm's entire net present value; because the firm can obtain funds costlessly from a large number of small investors, it is optimal to have no initial investment by shareholders, keeping liquidation costs to a minimum.[24] Parts (iii) and (iv) rely on the fact that $C'(0)$ is the smallest possible marginal cost of monitoring, while $\Delta p(A - B)/n$ is the greatest marginal benefit additional monitoring can have for any one shareholder.[25] When the condition in (iv) holds, monitoring never pays—the

[24] If there are no liquidation costs ($k = 0$), it is easy to show that shareholders are indifferent between committing their wealth through contingent guarantees or through initial investment. Because contingent guarantees have drawbacks detailed in the next section, shareholders would have incentive to invest as much as possible of their wealth initially, even in an unlimited liability regime. Since the historical record does not bear this out, this supports the idea that illiquidity of wealth has been a factor in practice. (In addition to the examples noted in the introduction, Scottish banks with unlimited liability typically had initial paid-in capital that was small in comparison with subsequent calls for funds; cf. Munn (1981), pp. 103–106.)

[25] This occurs when (a) no other shareholders monitor, and (b) shareholders gain the entire benefit of increased monitoring at the margin, which means that the debt is fully insured.

cost always outweighs the benefit to the monitor. Otherwise, it is possible for monitoring to be worthwhile, but the monitors must have sufficient incentive. Thus, there may be a minimum level of wealth commitment (possibly zero) before monitoring becomes incentive compatible. Once the total wealth commitment renders the debt riskless, additional wealth commitments do not add value.

The conclusion that shareholders set initial investment to zero would hold even if liquidation costs applied in the second as well as the first period. This is because the deadweight losses imposed by initial investment occur in all states, while those incurred through contingent liability only occur in default states. However, the presence of liquidation costs in both periods might cause shareholders to commit less than their total wealth to the firm, since at some point the deadweight losses from liquidation could outweigh the gains from further bonding. Since the shareholders' choice of commitment device would be unchanged, I have opted for the simpler framework.

The effects of changing the number of shareholders are somewhat ambiguous. On the one hand, given individual monitoring q, increases in n increase total monitoring Q; the convexity of $C(\cdot)$ may even allow a larger number of shareholders to reduce q, keep Q constant, and reduce total monitoring costs. On the other hand, as n increases, each shareholder's stake in the firm decreases, decreasing his incentive to monitor—the free rider effect. The following corollaries provide conditions on the slope and curvature of $C(\cdot)$ such that free rider effects eventually overcome any "economies of scale" in monitoring.

COROLLARY 1: *If* $C'(0) > 0$, *there exists* n' *such that, for all* $n \geq n'$, *the equilibrium choice of* q *is* 0, *and firm value is constant at* $V_0 \equiv p_c(A - B) + B - I$.

COROLLARY 2: *If there is an* $\alpha < 1$ *such that* $qC''(q)/C'(q) \leq \alpha$ *for all* $q \in [0, 1]$, *then for* $n \geq \max\{(I - B)/w, (1 - \alpha)^{-1}\}$, *monitoring and firm value decrease with* n.

COROLLARY 3: *If* $C(q) = cq$, $c > 0$, *then whenever the optimal monitoring level* q *is positive,* q *and firm value will decrease with* n.

The intuition for Corollary 1 is that, if the marginal cost of monitoring is strictly positive, increasing the number n of shareholders eventually causes the per shareholder benefit of any marginal monitoring to be outweighed by its cost, and no monitoring occurs. If the relative curvature of the monitoring cost function is not too large, Corollary 2 suggests that "economies of scale" in monitoring do not increase fast enough to offset the reduction in per shareholder benefit as n increases, so that total monitoring and firm value decrease. Finally, Corollary 3 specializes the last result to the case where monitoring costs are linear. With no convexity in costs, it is always most efficient to have only one investor monitor—so long as that investor has sufficient wealth at stake to make monitoring incentive compatible. Further

increases in n simply reduce the effectiveness of any monitoring that takes place, since there is some costly "doubling up" of effort.

I use Corollary 3 in Section III to simplify the discussion of optimal ownership regimes. For other cost functions, the relation between firm value and n may be more interesting, with interior maxima before free rider effects take hold. Calculations of this sort would be especially important for analyzing the relative efficiency of asymmetric and hierarchical equilibria—an issue of particular interest if the shareholders that monitor management the most are viewed as the firm's directors.

III. Unobservable Shareholder Wealth

A. Adverse Selection Effects

So far, it has been assumed that each shareholder's external wealth is freely observed by all other agents. If this is not true, a share contract that uses additional liability will face a "lemons" market problem (Akerlof (1970)). With additional liability, any shareholder's liability is effectively limited by his unobservable wealth; thus shareholders with lower wealth have less incentive to monitor than those with higher wealth. If share price reflects expected monitoring based on the range of possible wealth levels, the poorest shareholders get a free ride at the richest shareholders' expense, and the latter will prefer to sell their shares to (poorer) outside investors. This behavior reduces monitoring and firm value, offsetting some or all of the advantages contingent liability has over initial investment.

Suppose all shareholders start with outside wealth $w > 0$, and the chosen regime is liability L equal to nw with no initial investment S. Immediately before monitoring takes place, each shareholder undergoes a mean zero wealth shock. Shocks are identically and independently distributed, and cause wealth to assume a value in the interval $[w^L, w^H]$, with $w^L < w < w^H$. The distribution of shocks can be continuous, discrete, or mixed without changing the general nature of the following results.

Let $\boldsymbol{w}_i$ be the (random) value of shareholder i's outside wealth after the shock. Defining $\boldsymbol{Z}_i(w_i)$ as i's residual liability given that $\boldsymbol{w}_i$ equals w_i, this will be a random variable whose value depends on the wealth of each of the other shareholders. Liability is assumed to be "joint and several"; i.e., total liability is spread evenly until one or more shareholders' wealth is exhausted, after which the process is repeated among remaining shareholders until either (i) the total liability is met, or (ii) total shareholder wealth is exhausted. It follows that, for any configuration of shareholder wealth, increases in w_i strictly increase the liability i faces if and only if w_i is less than the larger of $(D' - B)/n$ or $D' - B - \Sigma_{j \neq i} \min\{w_i, w_j\}$.[26] Since w^L is the smallest wealth level that can occur, it follows that

[26] "Joint and several" liability is the rule for unlimited liability partnerships. By contrast, under "several" liability each shareholder is responsible for his share of the firm's debts and no more, and liability strictly increases with w_i i.f.f. $w_i < (D' - B)/n$. There are lemon problems in either case.

LEMMA 2: *If* $w^L < (D' - B)/n$, *then* $E[\boldsymbol{Z}_i(w_i)]$ *is strictly increasing in* w_i *for* w_i *in* $[w^L, D' - B - (n-1)w^L)$, *and constant for* $w_i \geq D' - B - (n-1)w^L$. *If* $w^L \geq (D' - B)/n$, *then* $E[\boldsymbol{Z}_i(w_i)]$ *is constant in* w_i.

In other words, so long as the smallest possible realization of total shareholders' wealth is not sufficient to cover the firm's debts, a shareholder's expected liability increases with his wealth until he is capable of fully making up any shortfall himself. If each shareholder's wealth always covers his share of the firm's debts, this externality will not arise.

If the shareholders' monitoring choices were unaffected by the shock, it could easily be shown that the value of i's share would be decreasing in his realized wealth w_i: the value of i's share once q_i and Q are fixed is decreasing in i's liability when firm output is low, and this is increasing in w_i by Lemma 2. Changing a shareholder's expected liability in the low output state changes his marginal benefit from monitoring, so monitoring choices will be affected by wealth shocks. Since each shareholder only observes his own shock, and shocks are independent, i's monitoring choice q_i will be a function of w_i alone. Thus the (random) value of i's share given wealth w_i and monitoring choice q_i is

$$\boldsymbol{v}_i(w_i, q_i) \equiv P[Q(\cdot)] \cdot \left[\frac{A - D'}{n} + \boldsymbol{Z}_i(w_i)\right] - \boldsymbol{Z}_i(w_i) - C(q_i), \qquad (8)$$

where $Q(\cdot)$ has arguments $(q_i, \boldsymbol{q}_{-i})$, and $\boldsymbol{q}_j \equiv q_j(\boldsymbol{w}_j)$ for j not equal to i. The next lemma shows that expected share value will still be decreasing in wealth.

LEMMA 3: *Suppose shareholders choose their monitoring probabilities conditional only on the knowledge of their own wealth. Then the expected value of* i*'s share conditional on wealth* w_i *is decreasing in* w_i. *It is strictly decreasing if* $w^L < (D' - B)/n$ *and* $w_i < D' - B - (n-1)w^L$, *and constant otherwise.*

Since wealth affects the value of shares, there may be incentives to trade following wealth shocks. To keep matters simple, I focus again on symmetric Nash equilibria (those in which shareholders who have the same level of wealth choose the same level of monitoring). Let the equilibrium expected value of shareholder i's share conditional on his wealth level w_i be denoted $E[\boldsymbol{v}_i(w_i)]$.

Suppose shareholders can sell their shares to other large investors immediately after the wealth shocks. If the price v equals the unconditional expected value of a share in the firm, it follows that shareholders with high wealth will wish to sell: $E[\boldsymbol{v}_i(w_i)] < v$ for w_i sufficiently close to w^H, since $E[\boldsymbol{v}_i(w_i)]$ is decreasing in w_i. Similarly, shareholders with sufficiently low wealth will strictly prefer to hold their shares. Outside investors with wealth close to $w^L + v/(1-k)$ will prefer to buy shares if liquidation costs are not too large: once they do, their outside wealth will be close to w^L, so the value of their shares will exceed v and may exceed the purchase cost $v/(1-k)$. Outside investors with wealth close to $w^H + v/(1-k)$ will prefer not to buy.

Even if the initial market price does not equal the unconditional expected value of a share in the firm, the preceding analysis shows that there will be gains to trade between relatively rich current shareholders and relatively poor (but still large) outside investors. Thus, adverse selection on both sides of the market will cause the average outside wealth of shareholders to drop as wealthier types sell out to poorer ones. If wealth was not illiquid, the equilibrium price would reflect the value of the firm when all shareholders had the lowest residual outside wealth consistent with having bought a share at its true value; wealth illiquidity imposes a transactions cost which will mitigate this process to some extent.

Some of these gains in share value come at the expense of bondholders, since the shareholders' effective liability is lower than the agreed upon nw. This will be foreseen when the debt is issued, causing its face value to be higher than it would be if the effective liability were nw. This reduces monitoring and firm value. In equilibrium, the shareholders absorb the entire loss caused by adverse selection on wealth.

To prevent this, the firm's owners may pass rules forbidding the sale of shares without approval by all shareholders, or require residual liability on the part of the seller for some period of time after the sale. In fact, such rules have been and are still used in many unlimited liability settings.[27] There will typically be costs to enforcing these restrictions or monitoring shareholder wealth at the time of sale, and shares will be illiquid, which will lower their value to the extent liquidity is desirable in a security.

Suppose instead that wealth shocks occur before the firm is formed. It is easy to show that, in a contingent liability regime with full wealth commitment, only shareholders with relatively low wealth will acquire shares (unless $nw^L > I - B$, in which case any collection of n shareholders can completely guarantee the firm's debt with no variation in their individual liabilities). Shareholders will have incentive to use costly wealth verification at the outset, and trading restrictions will still be needed to prevent shareholders from selling out to poorer outside investors. In general, the effective commitment of shareholders under contingent liability will be weakened if free trading in shares is allowed.

[27] Bylaws requiring shareholder or director approval for sales of stock seem to have been common among the Scottish banking "copartneries" (Munn (1981), pp. 159–160). In 1834, an Act of Parliament mandated residual liability for three years following the transfer of shares of certain companies. This provision was repealed in 1837 (Hunt (1936), p. 57). In the present day, an unlimited partner typically cannot sell his share in the firm without the other partners' approval, and investors in Lloyd's remain liable for any losses on policies written during their membership in a syndicate, even if they later resign.

These restrictions do not solve another problem associated with wealth shocks. If nw^H is greater than $I - B$, expected liability is in fact lower than nw even when liability is unlimited: the debtholders can never collect more than $I - B$, so states where most shareholders are poorer are not fully offset by states where most are richer.

B. Comparison of Different Regimes

The preceding subsection showed that contingent liability may have costs as well as benefits if shareholder wealth is unobservable. The next step is to compare the performance of various liability regimes in this setting. To simplify matters, I assume that the cost of monitoring $C(q)$ equals cq, wealth shocks are distributed uniformly over $[w^L, w^H]$ and occur *before* the firm is formed, and there are infinitely many large investors, so that for any $w \in [w^L, w^H]$, n investors with approximately that much outside wealth (after buying ownership in the firm) can always be found.[28] With linear monitoring costs, Corollary 3 suggests the optimal number of shareholders will typically be close to one; however, I will treat n as exogenous, because other factors (risk aversion, liquidity, and different cost functions) favor having more shareholders.

Several regimes are possible. One (Regime H) involves verifying shareholders' wealth and using the optimal structure from Section II: shareholders will have the highest possible outside wealth w^H and matching contingent liability.[29] Suppose there is an effort cost ϕ for verifying each shareholder's outside wealth to other investors in the firm, and ϕ is continuous and nondecreasing in n.[30]

Another option (Regime L) is to "ignore" the adverse selection problem, using contingent liability without restrictions. Since wealth is unverified, maximum effective liability per shareholder is w^L.

A third alternative (Regime I) involves initial investment w^I by each shareholder. Because any wealth levels between w^I and w^H are possible, allowing more liability would not be credible: there would be a tendency for only investors with wealth of exactly w^I to invest. w^I will be chosen to maximize the net value of the firm to the shareholders, subject to $w^I \leq w^H$.

Finally, the large investors could bow out entirely, selling ownership of the firm to many small investors (Regime 0). Free rider effects will prevent any monitoring, and the firm's net value will be V_0, which equals $p_c \cdot [A - B] + B - I$. The new owners will fund the firm's investment by issuing claims to more small investors; since no monitoring will occur in any case, it does not matter whether these additional claims are debt or equity.

Let $q(w, n)$ be equilibrium individual monitoring given n shareholders with additional wealth commitment w each, and let $Q(w, n)$ be the corre-

[28] Strictly speaking, the shock should affect total wealth including the amount required to buy ownership of the firm. However, as noted before, this would not change the results of the model, so I focus on the comparative statics of remaining outside wealth.

[29] If $w^H > (I - B)/n$, outside wealth and contingent liability need only be $(I - B)/n$.

[30] At one extreme, ϕ is constant if each shareholder's wealth could be publicly verified at fixed cost. At the other extreme, Carr and Mathewson (1988) assume verification is private information, which in the limit would result in total cost $n(n - 1)\phi$ for shareholders *alone, plus* $n\phi$ per bondholder. However, free rider issues would arise, particularly among bondholders (who are assumed fairly numerous), and the actual expected cost would probably be some intermediate amount increasing in the number of shareholders. In addition, ϕ could be increasing in the amount of wealth verified, but this would not change most of the qualitative results of the model.

Table I

Alternative Ownership Regimes

This table summarizes the various ownership/liability regimes discussed in Section III. $q(w, n)$ is the individual monitoring probability when n shareholders each commit w net wealth to the firm. $k \in (0, 1)$ is the cost of liquidating a unit of wealth invested outside the firm. A shareholder's outside wealth (net of initial share in the firm) is at least w^L and at most w^H, where $0 < w^L < w^H$.

Regime	Description	Monitoring Probability
H	Each shareholder has contingent liability w^H; his wealth is publicly verified to be w^H at a cost of ϕ.	$q(w^H, n)$
L	Each shareholder has contingent liability equal to the minimum wealth level w^L.	$q(w^L, n)$
I	Each shareholder liquidates wealth $w^I \leq w^H$ at cost kw^I. The net proceeds are invested in the firm, and there is no contingent liability.	$q((1 - k)w^I, n)$
0	Shareholders sell the firm to many small liquid investors; no monitoring takes place.	0

sponding total probability of monitoring. For initial investment s and additional liability l, equilibrium q equals $q(s(1 - k) + l, n)$. Table I summarizes the various regimes.

For the first three regimes, net firm value V_{regime} is given by

$$V_H = P[Q(w^H, n)] \cdot [A - B] + B - I - ncq(w^H, n) - n\phi, \tag{9}$$

$$V_L = P[Q(w^L, n)] \cdot [A - B] + B - I - ncq(w^L, n), \tag{10}$$

$$V_I = P[Q(w^I, n)] \cdot [A - B] + B - I - ncq(w^I, n) - nkw^I. \tag{11}$$

Define $R(w, n) = P[Q(w, n)] \cdot [A - B] - ncq(w, n)$. Then necessary and sufficient conditions for the dominance of various regimes are

$$V_I \geq V_L: \quad R(w^I(1 - k), n) - R(w^L, n) \geq nkw^I \tag{12}$$

$$V_I \geq V_H: \quad R(w^I(1 - k), n) - R(w^H, n) \geq nkw^I - n\phi, \tag{13}$$

$$V_H \geq V_L: \quad R(w^H, n) - R(w^L, n) \geq n\phi. \tag{14}$$

By Proposition 1, $R(w, n)$ is weakly increasing in w. Thus, a necessary condition for (12) is $w^I(1 - k) \geq w^L$. Also, since $w^H \geq w^I(1 - k)$, (13) implies that $\phi(n) \geq kw^I$, with strict inequality if $q(w^H, n) > 0$. Finally, since both Regimes I and H involve deadweight losses, they cannot dominate unless w^H and $w^I(1 - k)$ are high enough to result in strictly positive monitoring. As a result, there will be a minimum level of wealth w^H below which both of these regimes are unviable.

Formal comparative statics are as follows. If w^L increases, V_L increases, while V_H, V_I, and V_0 are unchanged. If w^H increases, V_L and V_0 are unaffected, V_H increases, and, because the constraint on w^I is weakened, V_I may also increase.[31] Thus increases in minimum wealth make unrestricted contingent liability more attractive, while increases in maximum wealth make costly wealth commitment regimes more attractive.

Since monitoring costs are linear, Corollary 3 holds, and increases in the number of shareholders n decrease the value of the three regimes that involve shareholder monitoring. Because they involve deadweight losses, Regimes H and I will eventually be dominated by L (although it is possible that, for such n, Regime L results in no monitoring and has the same value as the atomized ownership of Regime 0). V_H is likely to decrease more than V_I, since ϕ increases in n, while kw^I need not.[32]

Changes in the cost parameters k and ϕ have straightforward effects. Higher k decreases V_I without affecting the other regimes: for any choice of w^I, deadweight losses are higher and net investment is lower, reducing optimal q. Similarly, increases in ϕ directly reduce V_H but do not affect the other regimes.

These comparative statics results are summarized in Table II.

The effects of changes in other parameters on the relations among the regimes are difficult to assess, due to the nonlinear nature of the monitoring problem's solution. There will be a range over which changes that increase the value of monitoring and of the firm (e.g., increases in p_f, decreases in c, and so forth) increase the net benefit of Regimes H and I, both of which use costly measures to bolster shareholders' monitoring incentives. However, such changes will eventually give shareholders incentive to monitor without any additional bonding, and in the limit costly regimes will be unnecessary.

IV. Applications

The comparative statics results from the preceding section can now be used to study actual business practices, past and present. Five examples are discussed: the evolution of the corporate form in Britain, the insurance syndicates of Lloyd's of London, modern partnerships in the United States, leveraged buyouts, and standby letters of credit.

[31] However, large increases in w^H tend to improve Regime H relative to Regime I. Either w^I attains an interior maximum and stops increasing with w^H, so that further increases in w^H improve Regime H's relative position, or w^I continues to increase. In this last case, deadweight losses due to liquidation grow; once w^I exceeds ϕ/k, Regime H must dominate Regime I.

[32] Once more, these results take n as exogenous. In a more complex model where n is determined endogenously, a given firm might have different numbers of shareholders under different regimes. For example, since Regime H is most likely to dominate when n is small, the optimal arrangement under this regime might well have fewer shareholders than the optimal arrangement under Regime L.

Table II

Effects of Parameter Increases on Firm Value

This table summarizes the comparative statics effects of increases in parameters on firm value under each of the four ownership/liability regimes discussed in Section III. A brief description of parameters and regimes follows; in the table itself, parentheses indicate relative tendencies.

Parameters

w^L = shareholder's minimum possible outside wealth (net of initial share in the firm).
w^H = shareholder's maximum possible outside wealth (net of initial share in the firm).
n = number of shareholders.
k = cost of liquidating one unit of outside wealth.
ϕ = cost of verifying a shareholder's wealth.

Regimes

H = each shareholder has contingent liability w^H, shareholder wealth is verified to be w^H.
L = each shareholder has contingent liability at minimum wealth level w^L.
I = each shareholder invests $w^I \leq w^H$ more in the firm (before costs); no contingent liability.
0 = firm is sold to many small liquid investors; no monitoring takes place.

Effect of Parameter Increase on Firm Value, by Ownership Regime

	Parameter				
Regime	w^L	w^H	n	k	ϕ
H	none	increase	decrease (more)	none	decrease
L	increase	none	decrease (less)	none	none
I	none	increase	decrease	decrease	none
0	none	none	none	none	none

A. Evolution of the Modern Corporation in Britain

All the regimes described in the preceding section have been used in Britain over the last two centuries.[33,34] At the start of the Industrial Revolution, firms tended to be small in financial and geographical scope. In this setting, opportunities for selling firm shares or other assets were limited, while firm owners' wealth was relatively less costly to verify, since creditors were familiar with owners on a day-to-day basis. Thus ϕ was low, k was high, restrictions on trading were unnecessary, and unlimited liability (Regime H) was the rule. As firms expanded, public posting of shareholders' names and assets became common, and restrictions on the sale of shares existed both legally and in corporate charters.

[33] This example draws heavily on Sykes (1926), Hunt (1936), and Munn (1981).

[34] The regulatory environment in the United States has been more restrictive in that firms have generally only been able to choose between unlimited liability (partnership) and zero liability (corporation). However, the imposition of double liability on nationally chartered banks noted earlier might have been an attempt to make use of an intermediate ownership structure for efficiency.

Pressure to increase the liquidity of shares seems to have been a major force in the move to limited liability. Limiting maximum liability per share limited the range of possible effective liabilities per share, reducing the size of the adverse selection problem and thus the need for restrictions on trading.[35] However, as shares traded more widely and minimum shareholder wealth w^L fell, the benefits of having additional liability without restricting shareholder wealth or trading in shares declined. Increases in the number of shares may have enhanced their liquidity, but at the cost of reduced monitoring due to free rider effects. In the limit, Regime L would effectively merge with Regime 0 (atomized ownership with no liability), and liquidity could be enhanced by dropping the additional liability feature, which would only discourage wealthy investors from holding shares.

Also contributing towards dominance of atomized ownership was the growth in the average size of the firm relative to the wealth of most investors. In my model, this change can be represented by multiplying A, B, I, and the cost of monitoring c by some $\alpha > 1$ while holding shareholder wealth constant, which is formally equivalent to dividing wealth by α and keeping other parameters unchanged. A reduction in value would follow by Proposition 1, and expanding the number of shareholders would only exacerbate this through the free rider effect. Thus fewer firms would be able to take advantage of regimes specifying costly shareholder commitment, even if increased monitoring was potentially more valuable than the desire to trade shares easily.

For firms that continued to restrict ownership to the largest investors, forces linked to population growth and geographic integration favored initial investment over contingent liability. First, these changes made economies more "anonymous," increasing the cost ϕ of verifying an investor's wealth to other shareholders and the firm's creditors.[36] Similarly, increases in the breadth and depth of markets for assets (including shares themselves) made them more liquid, reducing k; also, more liquid assets (stocks and bonds) made up a greater fraction of wealth as industry and government grew. Finally, as financial intermediaries increased in size and efficiency, raising liquid funds through a personal loan against assets became a cheap alternative to selling assets outright. While a loan secured by illiquid assets would still require verification of the collateral's value, this would probably be cheaper when the only counterparty was a bank rather than a number of debtholders and shareholders.[37]

An exception to this would have been the intermediaries themselves, whose raison d'etre involved gathering liquid funds from many individuals and providing scale economies linked to monitoring or risk sharing. Having bank

[35] Although in this model only the type with just enough wealth to purchase shares would do so, various complications—risk aversion, varying degrees of wealth observability, more complicated liquidity preferences or costs—might well lead to the intermediate results seen historically.

[36] This would also tend to increase the cost of collecting shareholder's guarantees in the event of bankruptcy, which would hurt unrestricted contingent liability (Regime L) as well.

[37] As Diamond (1984) notes, the bank has to maintain its own incentive compatibility at some cost, so increased bank efficiency is essential to the last argument.

or insurance company shareholders back their firm with contingent liabilities may well have been more efficient than having shareholders invest in the firm with funds raised through direct loans from individuals. In fact, banks and insurance firms were some of the earliest joint stock (unlimited liability) companies in the United Kingdom, banks seem to have been among the last firms to remove contingent liability from their corporate structure, and the next example shows that at least one major group of insurers still uses contingent liability for shareholders.

B. Lloyd's of London

Lloyd's is a modern example of unlimited liability syndicates performing a significant role in the international insurance market. Several facts should be noted. First, a syndicate has a very small number of managers: typically only one active underwriter is responsible for almost all the risk taking of the syndicate. The number of Names (outside partners) varies between several dozen and a thousand. Second, potential losses tend to be high and of uncertain probability, Lloyd's having made a reputation for exploring new, high-risk areas of insurance. A skilled and motivated underwriter can avoid the winner's curse problem that a less diligent or less knowledgeable underwriter will face.

Thus, in terms of my model, c is relatively low, $A - B$ is high, and Δp is high, all of which should make shareholder evaluation of the underwriter feasible if wealth commitment is sufficiently high. The Committee of Lloyd's has also developed a system for assessing Names' wealth and holding part of it as collateral, reducing ϕ. Thus, it is not surprising that Lloyd's continues to opt for the nontraded contingent liability regime, providing its members with the greatest incentive to monitor.[38]

By contrast, normal insurance companies typically deal in more standardized products (low Δp) with economies of scale; this leads to large organizations which are more difficult to monitor (high c). Thus there should be little gain to having owners use their wealth as a bond. Not surprisingly, their ownership form (mutual or stock ownership) is diffuse and does not involve contingent liability.

C. Unlimited Partnerships in the United States

Such arrangements seem to be concentrated among smaller firms and in certain professions, most notably medicine, law, accounting, and consulting.[39]

[38] Allegations of fraud at Lloyd's in the early 1980s involved the largest underwriter, Ian Posgate, who wrote for several syndicates totaling some 6000 Names; see Hodgson (1984). With such large syndicates, even unlimited liability may not be enough to overcome free rider problems. In fact, similar charges have been brought forward more recently, and Lloyd's is considering limiting the Names' liability (*Economist*, June 20th, 1992, pp. 71–72).

[39] While regulation forces partnerships on some of the professions, this varies by state; see Carr and Mathewson (1988). In any event, one might argue that economic considerations have some impact on the laws; certainly, there would be pressure to change the laws if partnerships were grossly inefficient.

For small firms, the situation is like that described for early British firms: monitoring is less costly, significant capital can be provided by a small group of investors (reducing free riding), and the costs of verifying the latter's wealth are not excessive. The professions, on the other hand, combine a large potential for moral hazard or adverse selection (e.g., medical malpractice) with a requirement of specialized knowledge on the part of monitors: Δp is high, and c is high unless one is also in the field. Thus, a professional's colleagues are best suited to monitor him or her, but much of their wealth is in human capital. The partnership form gives them maximum incentive to monitor without requiring large initial investment. However, large firms will still develop free rider problems, exacerbated by the need to monitor more and more partners. This is an alternative explanation for Carr and Mathewson's (1988) empirical finding that in states where law firms are permitted to use the corporate (limited liability) form, those firms that choose limited liability tend to be larger than those that do not.

D. Leveraged Buyouts

Many buyouts turn publicly held firms into privately held ones with more concentrated ownership and higher debt-equity ratios. While the firm remains a corporation, the "sponsors" (dealmakers) usually have a partnership form, with limited partners providing funding and general partners providing monitoring or even managerial skills.[40] Both the general partners and the managers are frequently asked for personal guarantees against the firm's debt (though these are typically limited rather than unlimited in nature). Since the number of shareholders is fairly small, verification and trading restriction costs are limited. To the extent shareholder wealth is illiquid (real estate, human capital), the regime of verified contingent liability (Regime H) may be most efficient. Thus, it is not surprising to see greater use of privately held equity and contingent liability in this setting.

E. Standby Letters of Credit

These are often used to secure commercial paper. In return for a fee, the bank agrees to reimburse the holders of the firm's paper should the firm default. Since the bank takes a junior position relative to the holders of the commercial paper, it has the greatest risk of all creditors and thus the highest incentive to monitor the borrower. While a direct loan funded with large CDs also gives the bank a junior claim, a letter of credit permits the bank to commit funds on a contingent basis, with a lower probability of incurring any costs associated with having assets "on the books." Examples of

[40] See Jensen (1989). Insofar as the general partners of the sponsor have monitoring expertise (low c in my notation), switching from Regime 0 to either Regime H or I makes sense; the sponsors' monitoring is more efficient than that of a typical large investor.

such costs include any transaction costs of raising funds, the cost of maintaining adequate capital against balance sheet assets in order to maintain the bank's charter and deposit insurance, and the reserve tax and FDIC premium paid on deposits.[41]

V. Concluding Remarks

There are a number of ways in which a firm's ownership structure can give owners the incentive to monitor management, each method having its own costs. Contingent liability forces a choice between adverse selection and costly restrictions on trading, while initial investment and no additional liability involves liquidation costs. The relative tradeoffs between these regimes as underlying parameters change can be used to motivate the evolution of the corporate ownership form in Britain, and are consistent with the stylized features of those present-day firms which have chosen alternatives to the "typical" zero liability corporate ownership structure. Related arguments may help to explain the increased use of standby letters of credit by U.S. banks during the 1980s.

A number of factors relevant to ownership structure choice were beyond the scope of this paper. Perhaps the most salient example is the demand for stock market liquidity. Although this paper has addressed the impact of such liquidity on monitoring under a contingent liability regime, Huddart (1992) shows that the ability to trade anonymously may also have a negative effect on large shareholder monitoring under a zero liability regime.[42] On the other hand, increased liquidity (possibly linked to wider ownership of shares) has

[41] Regulatory costs probably outweigh transaction costs in the modern financial world. Recently, regulators have begun to require some capital to be held against contingent liabilities as well, and banks need no longer pay reserves on time deposits. This should reduce the attractiveness of standby letters of credit, ceteris paribus.

[42] The idea is that, if the share price reflected monitoring by a large investor, he could anonymously sell his shares to a number of small investors and receive this value without incurring the effort. Since the monitoring does not take place, this is not an equilibrium, and so the ability of the initial investment regime to enforce monitoring incentives tends to break down.

Shleifer and Vishny (1986) show that inability to trade anonymously rules out such selling by a large shareholder but makes it suboptimal for him to acquire his position in the first place. However, they suggest that firms may reward those (large) investors who monitor by choosing actions which these investors prefer and presumably cannot get elsewhere (the example they give is paying dividends to please large corporate shareholders that get the dividend received deduction). This acts as a means of preventing the shareholder from selling his shares on the open market and abandoning the role of monitor. Note that the only reason the shareholder will in fact monitor if forced to hold shares is that he shares directly in the improvement in firm performance caused by monitoring. Thus, the arguments made in the current paper are basically unaffected by Shleifer and Vishny's proposal: once an investor is cast in the role of monitor, he should be given a junior claim (such as stock) in order to enhance his incentives to monitor; the cheapest way of committing his wealth will be through contingent liability; and, in a contingent liability regime, large shareholders will have incentive to sell their shares to other large investors who also prefer the firm's specialized actions but have lower wealth.

benefits as well; these need to be modeled more explicitly, so that the tradeoffs between liquidity and monitoring can be more fully examined.

The presence of risk aversion would also favor wider ownership so as to allow greater diversification and risk sharing. In addition, a large investor might be better off diversifying his wealth and allowing only contingent liability against his portfolio rather than investing everything in the firm, although this would place additional risk on the bondholders. Finally, if investors have decreasing absolute risk aversion, adverse selection caused by contingent liability would be lessened: poorer investors would now place less value on shares of given risk, offsetting their lower effective liability.

Taxes are yet another important influence on the choice of ownership structure in the United States today.[43] Nevertheless, the ownership forms described here were present before corporate income taxes were a factor, and the current tax advantages of partnerships do not explain why contingent guarantees are used in closely held corporations. Tax concerns seem likely to complement this paper's analysis rather than substitute for it.

As mentioned in Section II, boards of directors could be incorporated into this model by having a group of shareholders perform monitoring of management while the rest of the shareholders monitor that group. This would tend to reduce free rider problems while retaining aggregate shareholder wealth as insurance for debtholders. The use of additional liability for all shareholders may then be a simple means of insuring additional liability for directors.

Finally, other forms of moral hazard or adverse selection might have interesting implications for ownership structure choice. For example, shareholders have some incentive to take on additional risk at the debtholders' expense. If shareholders have a larger wealth commitment to the firm, such actions should be less attractive. This would provide another argument for greater shareholder bonding, and thus indirectly for the use of contingent liability.

Appendix

Proof of Lemma 1: The solution to (3) is unique if it exists, since the RHS is decreasing in q while the LHS is strictly increasing. If $(\Delta p/n) \cdot [A - D' + \min\{L, D' - B\}]$ exceeds $C'(0)$, the RHS exceeds the LHS when q is 0. However, when q is 1, the RHS equals 0 while the LHS equals $C'(1) > 0$, and so a solution exists. Q.E.D.

Proof of Proposition 1: (i) Suppose (5) is not binding. As noted in the text, the equilibrium q given S, L, and D' is weakly decreasing in D'. A sufficient decrease in D' will cause the constraint to be binding. Since q is now greater than before but still less than the level that maximizes $P(Q) \cdot (A - B) -$

[43] See, for example, Scholes and Wolfson (1992), Chapter 4, and Mackie-Mason and Gordon (1991).

$nC(q)$, and the expected payment to bondholders is smaller, shareholders are better off.

(ii) Since (5) is binding, the objective function becomes

$$P(Q) \cdot [A - B] + B - I - kS - ncq, \tag{A1}$$

where q is given by Lemma 1. In that lemma, $A - D' + \min\{L, D' - B\}$ now equals

$$A - B + \frac{\min\{0, S(1-k) + L + B - I\}}{P(Q)}. \tag{A2}$$

Increases in $S(1-k) + L$ weakly decrease D' and hence increase the equilibrium level of q, improving shareholder value. Since $S(1-k) + L$ is maximized subject to the wealth constraint by $S = 0$, $L = nw$, and $S > 0$ directly reduces the objective function, shareholder returns are maximized when S equals 0.

(iii) The restated objective function (A1) gives shareholders the entire net present value of the firm, but q is constrained by Lemma 1 and (A2). Since increases in L weakly increase q and firm value V, increases in total wealth loosen the constraint on L and weakly increase firm value. Part (iii)'s condition on $A - B$ guarantees that equilibrium q is positive for L sufficiently large, giving w_*; further increases in committed wealth strictly increase q and firm value until debtors are fully insured, when $w = (I - B)/n$.

(iv) Since $A - D' + \min\{L, D' - B\} \leq A - B$, Lemma 1 and part (iv)'s condition on $A - B$ imply that equilibrium q is never positive. Q.E.D.

Proof of Corollary 1: While this follows from Proposition 1 (iv), a tighter bound can be found. Using the results of Proposition 1, substitute into (3) and multiply both sides by $(n/\Delta p) \cdot (1-q)^{1-n}$ to obtain the following equilibrium condition for $q > 0$:

$$(n/\Delta p) \cdot C'(q) \cdot (1-q)^{1-n} = \left[A - B - \frac{\max\{0, I - B - nw\}}{p_f - \Delta p \cdot (1-q)^n}\right]. \tag{A3}$$

The RHS is bounded above by $A - B - \max\{0, [I - B - nw]\}/p_f$, while the LHS is bounded below by $nC'(0)/\Delta p$; if $C'(0)$ is greater than 0, sufficiently large n implies the LHS is everywhere greater than the RHS, no solution exists, and $q = 0$. Q.E.D.

Proof of Corollary 2: If $nw \geq I - B$, (A3) becomes $(n/\Delta p) \cdot C'(q) \cdot (1-q)^{1-n} = A - B$. Since the LHS of this is increasing in q, the solution is unique. From the Implicit Function Theorem, dq/dn equals

$$\frac{-1 + n \cdot ln(1-q)}{n \cdot \dfrac{C''(q)}{C'(q)} + n(n-1) \cdot (1-q)^{-1}} \tag{A4}$$

which is negative. Because firm value V equals $P(Q) \cdot [A - B] + B - I - nC(q)$, and $P(Q) = p_f - \Delta p(1 - q)^n$, it can be shown that dV/dn equals

$$\left[\frac{-(n-1) - n \cdot \dfrac{C''(q)}{C'(q)} \cdot (1-q) \cdot ln(1-q)}{\dfrac{C''(q)}{C'(q)} + (n-1) \cdot (1-q)^{-1}} \right] \cdot C'(q) - C(q) \quad \text{(A5)}$$

Since $-(1 - q) \cdot ln(1 - q) < q$ for $q < 1$, the numerator of the bracketed term and thus (A5) are negative if $qC''(q)/C'(q) < (n - 1)/n$. Since $qC''(q)/C'(q) < \alpha$ and $n \geq (1 - \alpha)^{-1}$, this holds. Q.E.D.

Proof of Corollary 3: Corollary 2 shows that the desired result holds for $nw \geq I - B$. Since $C(q) = cq$, the LHS of (A3) is convex and increasing in q, and the RHS is concave and increasing for $nw < I - B$. Thus there are at most two values that solve (A3). Shareholders will choose the highest of these, and the slope of the LHS will be greater or equal to that of the RHS at that point by the second order condition for a maximum. Defining $F \equiv$ LHS $-$ RHS and applying the Implicit Function Theorem, it follows that dq/dn is negative wherever the theorem holds. Next, firm value equals $P(Q)[A - B] + B - I - ncq$; taking the total derivative with respect to n, a large amount of tedious algebra reveals that dV/dn is less than zero. Q.E.D.

Proof of Lemma 3: Given that other shareholders' choose their monitoring probabilities as functions of their own wealth only, shareholder i chooses q_i to maximize $E[\mathbf{v}_i(w_i, q_i)]$, where the expectation is taken over $\boldsymbol{w}_j$ for all $j \neq i$. Recall that liability $\boldsymbol{Z}_i(w_i)$ is decreasing in w_i. From (8) it follows that, for any $q \in [0, 1]$, $\mathbf{v}_i(w', q) - \mathbf{v}_i(w'', q) \geq 0$ if $w' \leq w''$, so $E[\mathbf{v}_i(w', q)] - E[\mathbf{v}_i(w'', q)] \geq 0$ if $w' \leq w''$. Now suppose q'' is an optimal monitoring choice for i when i's wealth is w'', and q' is optimal when i's wealth is w'. If $w'' \geq w'$, then

$$E[\mathbf{v}_i(w', q')] - E[\mathbf{v}_i(w'', q'')] \geq E[\mathbf{v}_i(w', q'')] - E[\mathbf{v}_i(w'', q'')] \geq 0 \quad \text{(A6)}$$

where the first inequality follows from the optimality of q', and the second follows from the argument immediately preceding (A6). Thus expected share value is nonincreasing in a shareholder's own wealth. The necessary first-order condition for optimal q_i can be derived from (8); it is such that optimal q_i is non-decreasing in $\boldsymbol{Z}_i(w_i)$ and only depends on w_i through $\boldsymbol{Z}_i(w_i)$. This together with Lemma 2 completes the proof. Q.E.D.

REFERENCES

Allen, F., and A. Winton, 1992, Corporate financial structure, incentives, and optimal contracting, Working paper no. 146, Department of Finance, Northwestern University.

Akerlof, G., 1970, The market for lemons: Qualitative uncertainty and the market mechanism, *Quarterly Journal of Economics* 84, 488–500.

Brennan, M., and I. Dunlop, 1991, Contributing shares, Unpublished manuscript, University of California at Los Angeles.

Bruce, R., B. McKern, I. Pollard, and M. Skully, 1986, *Handbook of Australian Corporate Finance*, 2nd ed. (University of Hawaii Press, Honolulu).

Carr, J., and G. F. Mathewson, 1988, Unlimited liability as a barrier to entry, *Journal of Political Economy* 96, 766–784.

Chandler, G., 1964, *Four Centuries of Banking*, Vol. I (B. T. Batsford Ltd., London).

Diamond, D., 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.

——— and P. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *Journal of Political Economy* 91, 401–419.

Easterbrook, F., and D. Fischel, 1985, Limited liability and the corporation, *University of Chicago Law Review* 52, 89–117.

Fama, E., and M. Jensen, 1983a, Separation of ownership and control, *Journal of Law and Economics* 26, 301–325.

———, 1983b, Agency problems and residual claims, *Journal of Law and Economics* 26, 327–349.

———, 1985, Organizational forms and investment decisions, *Journal of Financial Economics* 14, 101–119.

Greve, J., 1987, *How to Do a Leveraged Buyout*, 3rd ed. (Business Publications, Inc., San Diego, Calif.).

Hagerty, K., and D. Siegel, 1988, On the observational equivalence of managerial contracts under conditions of moral hazard and self selection, *Quarterly Journal of Economics* 103, 425–428.

Halpern, P., M. Trebilcock, and S. Turnbull, 1980, An economic analysis of limited liability in corporation law, *University of Toronto Law Journal* 30, 117–150.

Harris, M., and A. Raviv, 1979, Optimal incentive contracts with imperfect information, *Journal of Economic Theory* 20, 231–259.

———, 1991, The theory of capital structure, *Journal of Finance* 46, 297–355.

Hodgson, G., 1984, *Lloyd's of London* (Viking Press, New York).

Huddart, S., 1992, The effect of a large shareholder on corporate value, *Management Science*, Forthcoming.

Hunt, B. 1936, *The Development of the Business Corporation in England, 1800–1867* (Harvard University Press, Cambridge, Mass.).

Jensen, M., 1989, Eclipse of the public corporation, *Harvard Business Review* 67:5, 61–74.

——— and W. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, 305–360.

Lacker, J., 1990, Collateralized debt as the optimal contract, Working paper, Federal Reserve Bank of Richmond.

Mackie-Mason, J., and R. Gordon, 1991, Taxes and the choice of organizational form, Unpublished manuscript, University of Michigan.

Munn, C., 1981, *The Scottish Provincial Banking Companies 1747–1864* (John Donald Publishers Ltd., Edinburgh).

Scholes, M., and M. Wolfson, 1992, *Taxes and Business Strategy: A Planning Approach* (Prentice Hall, Inc., Englewood Cliffs, N.J.).

Shleifer, A., and R. Vishny, 1986, Large shareholders and corporate control, *Journal of Political Economy* 94, 461–488.

Sykes, J., 1926, *The Amalgamation Movement in English Banking, 1825–1924* (P. S. King & Son, Ltd., London).

Woodward, S., 1985, Limited liability in the theory of the firm, *Zeitschrift für die gesamte Staatswissenschaft (Journal of Institutional and Theoretical Economics)* 141, 601–611.