

**EQUITY AS A GUARANTEE:
A Contribution to the Theory of the Firm**

By

Yoram Barzel and Wing Suen¹

August 1997

¹ We wish to thank Douglas Allen, John Hause, Levis Kochin, Keith Leffler and Dean Lueck for their comments.

**EQUITY AS A GUARANTEE:
A Contribution to the Theory of the Firm**

By

Yoram Barzel and Wing Suen

ABSTRACT

In order to maximize the value of a transaction, transactors, backed by guaranteeing-capital, allocate the bulk of the income variability produced by their transaction to the party that will gain most by affecting the outcome of the transaction. Individuals who are short of guaranteeing-capital can gain by collaborating with capital-rich individuals who will, however, require them to become employees. The gains are the result of reducing moral hazard, and not from risk-sharing. Both the use of guaranteeing-capital and the supervision of employees who perform similar functions are subject to economies of scale. We define the firm by the extent to which its equity capital guarantees outcomes. The firm tends to employ similarly skilled workers who are short of capital. Independent workers who sell their services to firms will tend to join these firms when their output becomes more costly to guarantee; conversely, firms will tend to shrink as the variability associated with activities previously performed by their employees declines.

I. Introduction

The economic literature offers (at least) three prominent explanations for the existence of firms. The first and earliest is offered by Coase, who views the firm as a means to economize on the use of prices. The second is offered by Alchian and Demsetz, who hypothesize that the firm is a means to realize the economies of team production, and that members of the team are bound to the firm by a nexus of contracts. Such a firm consists of a residual claimant-supervisor who is at its center and of others whose output is difficult to observe and who produce the firm's output under a fixed-wage contract, supervised by the residual claimant. The third explanation is offered by Williamson as well as by Klein, Crawford and Alchian, who argue that vertical integration, which contributes to the creation of firms, is useful to prevent the capture of specialized capital. Barzel's (1982) argument on measurement suggests that as an intermediate output becomes more costly to measure, it is less likely to be sold in the market and is likely to remain within the firm. This force toward vertical integration is quite similar, but not identical, to Williamson's and to Klein, Crawford and Alchian's argument for vertical integration.

By themselves, these explanations satisfactorily account neither for the scope of the firm nor for its boundaries. There appears to be no empirical work which demonstrates that the size and structure of actual firms conform to the three models. Casual empiricism does not provide support to the notion that these are the main explanations of the existence of actual firms. For instance, it is difficult to see how the size and structure of both Safeway and of the corner grocery store can be attributed to forming prices, to team production, to specific capital or to the measurement of intermediate output. Regarding the boundaries of the firm, Cheung points out the ambiguity in what constitutes a firm. He argues persuasively that if we cannot classify activities as firm or non-firm activities, then there can be no operational theory of the firm.

We propose that the guarantee function of equity capital is subject to scale economy that might help explain both the size and the scope of firms. Moreover, the guarantee function provides a means of defining the boundaries of the firm. The view we take is that guarantees and the capital backing them are at the heart of any organization, particularly of the firm. The outcome of the transaction among individuals (say, when they are engaged in team production or in an exchange relationship) is subject to variability. The contract governing a transaction implicitly allocates the variability of the transaction-outcome to the transacting parties, and therefore affects their behavior. In the absence of the guarantee problem, the optimal choice of contracts can be analyzed in terms of the optimal allocation of variability to minimize moral hazard losses. There is no reason why such optimal contracts should center around “the firm,” and Cheung’s argument applies. However, the bearing of variability is subject to the guarantee problem—individuals with limited wealth, even if they are risk neutral, cannot bear the consequence of adverse outcomes. The main component of our model is in the assembly of equity capital which transactors use to guarantee their activities. We argue that economies of scale in the use of guarantee capital produce “firms” which are essentially constellations of equity capital surrounding the contracts which it guarantees. We classify a sale fully guaranteed by a firm’s equity capital as being within the guaranteeing firm, and a sale partly guaranteed by equity capital as being only partly within that firm.¹ As a rule, no party fully guarantees its transactions, but most guarantee their transactions in part. Transactors who are unable or unwilling to guarantee their actions may collaborate with others who assume the bulk of the exchanges’ variability, placing these exchanges (largely) within the guarantors’ firms.²

¹ Liability plays a central role in “The Problem of Social Cost.” Liability problems arise because of variability; similarly, guarantees arise because of variability. This paper, then, can be traced more directly to the “Problem of Social Cost” than to “The Nature of the Firm” (see Barzel and Kochin).

² By our definition, to the extent that a firm’s activities are insured, they lie within the insurance companies.

At times, if not always, our model competes with Alchian and Demsetz's model.³ According to our argument, workers who cannot guarantee their own work may profitably contract with guarantors. In order to induce the appropriate performance, the latter would need to constrain the former. The wage contract is the method through which the constraints are exercised. So, although a nexus of contracts emerges here too, the reason for its emergence is entirely different.⁴ Finally, the measurement argument, as well as the asset specificity theory, are not mutually exclusive with the argument presented here. To the extent that each is valid, each may contribute to the theory of the firm.⁵ The paper is also related to Eswaran and Kotwal's (1989) theory of why capitalists are the bosses, although we are putting it in a broader framework.

We define the firm in terms of the totality of activities that it (partially) guarantees. The definition is operational and it yields firms that differ from each other in size and in the nature of their contracts with other firms. If the theory has merit, then, it must be able to explain at least some of these differences. Since the definition allows for measuring firm size, Cheung's criticism does not apply to it. Moreover, our model is capable of addressing questions that cannot arise under Cheung's proposition that only contractual forms be studied.

A road map may be helpful in reading this paper. In Section II, we discuss the principle behind the optimal allocation of variability. To illustrate this principle, we use it to examine the firm's purchase of fire insurance and its use of bond underwriters. Sections III and IV form the core of our paper. In Section III we discuss the problem that arises when the

³ In Knight's model too, the firm is a nexus of contracts between entrepreneur and workers. His entrepreneur assumes risk and pays fixed wages to workers. In Alchian and Demsetz's model and in ours preferences toward risk play no role in explaining the firm.

⁴ We note two major problems that the Alchian and Demsetz model has. First, the economies of team production seem to be confined to teams *much* smaller than is the workforce of many observed firms. Second, in the shareholder public corporation the monitor is *not* the main claimant to the residual of team production. Managers-monitors receive salaries and shareholders do very little monitoring.

⁵ Regarding the role of the firm in economizing on the use of price, our model can explain why prices are formed within firms. For the sake of brevity we simply assert that had price formation been a market activity, the required guarantee would make the arrangement uneconomical and thus it is performed, as a rule, within firms. This is a connection between our model and Coase's.

capital individuals have is insufficient to back up their guarantees. Because of the economies of scale in the use of guarantee capital, people have an incentive to pool their capital and organize their transactions around the pool. This—we argue in Section IV—is the essence behind the organization called the “firm.” Section V then uses our theory of the firm to discuss some old topics of the field—separation of ownership and control, entrepreneurship, and integration. In Section VI, we provide a rudimentary discussion of some issues involved in using reputation-capital as a guarantor and the last section summarizes our argument.

II. The Guaranteeing Role of Equity Capital

A fundamental proposition regarding the action of parties to a contract in the presence of variability is as follows: *In order to maximize the value of exchange, the exchange contract should assign to a party more of the variability of the joint action when that party's ability to affect the outcome rises.* This principle can be formalized in a variety of models. In the literature on product liability rules (e.g., Cooper and Ross [1986]; Kambhu [1982]), there is a tradeoff between pre-sale care by the manufacturer and post-sale care by the customer. Optimal contracting calls for allocating a greater share of the liability to the party with greater control of the outcome, as this would minimize the cost due to moral hazard. In the choice of contracts involving team production, Eswaran and Kotwal (1985) and Barzel and Suen (1992) argue that it is optimal to assign a larger share of the (variable) output to the party whose effort has a greater marginal effect on output. Similarly, to minimize under-investment in specific assets, Grossman and Hart (1986) and Hart and Moore (1990) show that property rights should reside in the party whose investment in the final outcome is greater. The organizational implications of this principle, however, has not been explored fully. Our main concern is how the bearing of variability—which is the same as having a claim on the residual and providing a guarantee for it—delineates the function and scope of a firm.

We define the firm by its guarantee capital and by the scope of its guarantees. The scope of the firm comprises *the set of contracts whose variability is contractually guaranteed by common equity capital*. The firm, then, is a nexus of outcome guarantees.

The usefulness of this definition is evident when considering, among other things, the franchise operation. As an example, take McDonald's. Are the franchisees part of McDonald's or are they separate firms? Since each franchisee invests in his outlet, he has some guarantee capital. He bears, for example, the effect of variability in the quality of his employees; he reaps the benefits of hiring better ones and suffers a reduction in the value of his enterprise when retaining poor ones. By our definition, then, each franchisee is a firm. The variability in the operation of a franchisee's outlet borne by the franchisor, however, is part of the franchisor's firm; if the franchisee pays the franchisor a fraction of the franchise revenue, the franchisor then bears part of the franchise revenue variability. The two entities are intertwined, but the definition indicates how to allocate the variability between them.

Parties to a transaction may guarantee to one another shares in the outcome of that transaction. The guarantees by firms (and other organizations) of given sources of variability, however, are not completely separable, and therefore are not mutually exclusive. The sharing of variability by two or more organizations may generate net benefits even though it is prohibitively costly to fully separate the shares. For this reason, property rights to guaranteed actions are not fully defined, and the Pareto outcome is unattainable.

Since people will spend resources to capture the not-fully-assigned guarantees, guarantors in general, and firms in particular, are expected to impose constraints on the guaranteed in order to reduce the capture expenditures. Whereas no attempt is made here to measure the firm, its decision to expand or to reduce the level of its guarantees is, by the definition here, a decision to respectively expand or shrink. In the following discussion, the organization we mostly focus on will be the publicly held corporation.

Typically, each party to a contract assumes part of the outcome-variability and, therefore, most of the transactions of a publicly held corporation are partly in the firm. The

share of outcome-variability a firm assumes, however, varies greatly across transactions. Note that when owners of factors who choose to guarantee their own actions sell their services to firms, the transactions largely occur outside the firms being considered. The primary conclusion of this section is that in terms of the inputs used to produce the output it sells, a firm that transacts mostly with factor owners who guarantee their actions would emerge as a nearly empty shell, consisting of little other than equity capital. Sears is one example of a firm that comes close to this characterization. Most of the products it sells are produced and guaranteed by other firms. Sears itself, as distinct from its suppliers, guarantees only a small fraction of each of its transactions with its customers. Factors leading to the formation of full-bodied firms will be introduced in the next section.

Owners of productive resources transact in order to exploit their diverse skills and physical assets. Besides the conventional production skills, owners are also expected to acquire skills for managing their assets. Individuals vary greatly as to the management skills of their own assets. At one end of the spectrum are owners of specialized resources who, not surprisingly, are better informed about the nature of their output than their potential buyers are. Being better informed they are able to pass low-quality output off as high-quality, if it is not checked. At the other end there often are numerous shareholders, i.e., owners of capital, who acquire a diverse set of intermediate inputs from specialists to produce a product. The individual capital owners have neither the incentive to acquire the skills necessary to supervise those inputs which collaborate with their capital, nor the incentive to become knowledgeable about other aspects of the enterprise.

Buyers who do not possess the necessary skills to cheaply evaluate their purchases may require that the price, along with other terms of exchange, be adjusted commensurably with the success of their purchase. The seller who could otherwise easily gain at the buyer's expense is expected to bear the effect of his or her action by guaranteeing the output. An advantage of the guarantee arrangement over prior inspection arises from the fact that output quality tends eventually to be revealed (at low cost). This allows avoiding the acquisition of skills for

inspection and the actual inspection at the time of purchase. Most sales, indeed, are subject to some guarantee.

Shareholders whose firms purchase *fully guaranteed* goods have no reason to supervise the sellers' production effort and would not seek master-servant relationships; instead, they engage in across-firm transactions. Such guaranteed purchases constitute transactions between well-informed, specialized input owners and less informed buyers. The guarantee that specialized input owners provide to the buyers of their products or of their advice consists of these owners assuming the bulk of the variability of the transaction outcome. The next few paragraphs explain the sense in which certain transactions constitute the guaranteed advice sold to firms. Two examples, typically given a different explanation than ours, illustrate the argument: the firm's purchase of fire insurance and the service it obtains from bond underwriters.

The ability to protect against fire is one useful skill which shareholders or even managers of most shareholding firms are unlikely to possess. Shareholders and their managers would, therefore, like to purchase the advice of fire prevention specialists. Lacking expertise, however, they are unable to cheaply assess the quality of the service (asymmetry of information). The quality of the advice obviously affects the incidence and severity of fires. According to our basic proposition, the specialists are expected to assume the bulk of the variability generated by their services. In this case, assuming the variability means providing fire insurance. The buyers of fire insurance may be highly diversified, or simply risk neutral, but they still seek the insurance; what they actually obtain is guaranteed advice. Since the behavior of shareholders or of their employees also affects the losses from fire, the insurer is not expected to bear all the expenses resulting from fire damage; the shareholders are expected to bear part of them. Still, the fire insurer typically bears the bulk of the fire-damage expenses and is the primary *claimant*, i.e., owner, of the insured income-variability associated with fire and its prevention. The fire prevention service that a firm obtains from another organization

that guarantees its service is (largely) produced within the fire-insurer firm. (See Goldberg [1990].)

One other service that shareholders may wish to obtain is investment advice. Here too they are ignorant; they are generally not knowledgeable about the proper scope of the operations of their own firms, and they will attempt to acquire advice from specialists. Different specialists provide advice regarding various aspects of the scope of the firm's operations. We will argue that bond underwriters, in cooperation with buyers of the bonds, are one class of such specialists. They give advice about lines of activity to move into or to steer away from. We will focus on the guarantee of the advice they provide.

Bond underwriters' advice is implicit, and is contained in the bonds' covenants. They guarantee their advice by buying the entire issue of bonds. New bonds are generally assigned the lowest seniority among debts and are paid off last in case of bankruptcy; their claim on the firm's assets stands just above the firm's equity. Such bonds are secure only as long as the value of the equity remains positive. Bond-underwriters usually buy the bonds from issuers in order to sell them to the public. They gain as the net price at which they can sell the bonds increases, and therefore their profit margin will decline as would-be buyers' expectation of bankruptcy of the issuing firm rises. In order to obtain the highest price for the bonds net of their own costs, the underwriters must acquire expertise in the borrower's line of business such that they are able to impose restrictions or covenants that steer the borrowing firm toward profitable investments and away from losing propositions. The underwriters also attempt to impress upon prospective bondholders the high probability that the firm will stay solvent until the bonds are redeemed. The choice of covenants requires specific knowledge; a run-of-the-mill investment for one firm is exotic for another. Competition between potential underwriters allows the borrowing firm to choose the most desired combination of the financial costs of the bonds and the associated restrictions. By issuing low priority bonds, then, the firm in fact purchases guaranteed advice regarding its new operations.

The two particular examples that have just been discussed show how owners of equity-capital can acquire expert advice on matters about which they know less than their transacting partners. These equity-capital owners succeed in limiting their potential losses due to ignorance because the advice they obtain is guaranteed. Since owners of all factors, human and non-human, are expected to be experts in their particular fields, all are expected to guarantee their performance. If they do guarantee their performance, they will, then, operate largely outside the firms to which they sell their services. By itself, the application of the allocation-of-variability principle to the collaboration between owners of equity capital and owners of other resources who are able to guarantee their actions suggests that such collaboration need not be within the capitalists' firm.

The ability of resource owners to operate outside the firm with which they transact requires guarantees, and guarantees require capital. Factor owners who lack sufficient guaranteeing capital may find it advantageous to cooperate with owners of capital, and this cooperation is a major function of the equity capital of firms.

III. Equity Capital and the Firm's Nexus of Contracts

In this section we analyze the relationship between the amount of guarantee-capital backing a product and the liability associated with the product. Our primary conclusion is that those workers who command less capital than is needed to guarantee their action stand to gain by becoming employees.⁶

As already stated, efficiency dictates that a party should bear more of the variability of a transaction when its gain from affecting the outcome increases. The means available to individuals, however, are not always adequate to bear the effect of their actions.⁷ Individuals

⁶ While our discussion is stated in terms of product liability, the principle involved is also applicable to other types of transactions such as the undertaking of risky investment.

⁷ Capital takes many forms, including those of human capital and brand-name capital. The use of human capital as collateral is constrained by anti-slavery laws. The capital-value of brand name, or of reputation may also be used as collateral. In such a case, however, the guaranteed do not directly gain what their providers lose. We will discuss the use of such capital in Section VI below.

who are less able to back their actions are expected to generate a higher liability loss and the *net* value of their output will be lower. To illustrate the relationship between liability loss and guarantee capital, suppose that the rate of defective products depends on how labor is applied. Let the total value of damage, X , be distributed with the c.d.f. $F(x;a)$ on the support $[0,\bar{x}]$, where a is the level of labor effort. A higher level of effort reduces damage in the sense of first order stochastic dominance, i.e., $F_a(x;a) \geq 0$ for all x and a . The efficient labor effort is given by the level a that minimizes the sum of labor cost, $c(a)$, and expected damage, $E[X|a]$. The marginal benefit of an increase in effort is given by $-\partial E[X|a]/\partial a$, which is equal to $\int_0^{\bar{x}} F_a(x;a)dx$. If, relative to the per-unit labor cost, the damage⁸ that particular defects are likely to cause is low (defects may be frequent but are minor), workers can readily guarantee their work. They are expected, on this score, to become claimants to the appropriate residual and to produce the optimal product. As the potential damage from a defect rises, the ability of a capital-poor worker to cover an early run of defects declines. When buyers of the product do not expect the worker to be able to bear the entire damage he or she may cause, the worker's inducement to produce products free of defects declines, too. For example, if the worker's guarantee capital K is less than the maximum damage \bar{x} , the expected damage payment is only $\Pr[X \leq K]E[X|a; X \leq K] + \Pr[X > K]K$. For this worker, the marginal benefit of an increase in effort is then only

$$\int_0^K F_a(x;a)dx,$$

which is less than the full marginal benefit. (See also Sappington [1983]; Barzel and Suen [1992].) Were the worker to operate independently and to be the (nominal) residual claimant, he or she would fail to produce the optimal product. The worker's net earnings would also

⁸ The term "damage" is used as shorthand for the difference between the realized value of the service or product and its promised mean value.

fall, since the price buyers will pay is expected to fall *by more than* the amount which the worker would save by producing a lower-quality product.⁹

It may seem that in order to insure their product to their customers, capital-poor workers will purchase guarantees for their products in the insurance market. An unconstrained market transaction, however, is not viable here. The suppliers of guarantees can gain from guaranteeing only if the guaranteed are induced to produce a product which is less damaging than the one which they would produce on their own. As long as the guarantor does not constrain the actions of the producers, they can still gain by producing a low-quality product; they retain the cost reduction associated with the lower quality and do not bear the burden of damage payment. The price, or premium, such guarantee suppliers would have to charge, therefore, would exceed the value of the service, and there would be no takers for their product. Pure insurance, then, will not do here. The guarantor, however, may instead employ the producers, constrain their behavior, and guarantee their work.

The wage contract tends to induce low levels of the damaged product. Workers will gain less from carelessness and haste if paid per unit of time than if paid by output. Workers' *capacity* for causing damage is not altered if they become employees, but the changed contract radically alters their *incentives*. Pay by time, however, reduces the incentive for *any* action. The employer then will impose those restrictions which serve to reduce shirking, but at the same time also direct effort so as to reduce damaging actions.¹⁰

Workers may encounter guarantee problems not only because of variability in their product, but also because of the costs they may impose on fellow workers and other resource

⁹ Our assertion that individuals who lack guarantee capital are expected to engage in socially damaging behavior is not new. Indeed, it is central to the medieval institution of "frankpledge." Thus Warren, in writing about pre-Norman England, states that "every...man had to demonstrate...that he had sufficient pledges (probably 12) who would be answerable for his appearance in court in response to accusations, and who would themselves discharge any financial penalty laid upon him if he defaulted in the redress of a wrong he had committed....Failure to find pledges deprived a man of 'law-worthiness'...in resisting an accusation." (p. 41)

¹⁰ Within-firm transactions are often considered to be motivated by risk-aversion. The wage contracts described here redistribute risk: The owners of equity capital are risk takers since they guarantee others' actions, and those guaranteed are buying risk reduction. It is evident, however, that such an exchange of risk enhances specialization and results in greater efficiency. Risk-neutral transactors, then, are expected to engage in such exchanges.

owners. This problem seems particularly acute when team production enhances productivity. First, workers, by the nature of their work, can easily harm each other—by using power tools or explosives. Second, and more commonly, they can impose costs in the form of reduced productivity. Carelessness on the part of one worker can disrupt production of the whole team; and small clerical errors can have major financial implications. Third, when production follows a sequence of stages, shoddiness of work in one stage can affect other stages (see Kramer 1993). Here, too, efficiency requires individuals to bear the effect of their actions, but few individuals are wealthy enough to compensate others for severe injuries or for major work disruption. Were workers to remain independent, they would not take sufficient precautions, and their net output would suffer in consequence, lowering the value of their workplace. Capitalists can employ capital-poor workers and in addition to guaranteeing their products can take charge of the working environment and guarantee the workers against injuries and disruptions.¹¹

IV. Scale and Guaranteeing-Capital

Guarantees are subject to both economies and diseconomies of scale. Suppose that initially numerous individuals are operating independently and guaranteeing their own actions. Besides the gain through exchange when the individuals differ in their command over capital, the gains also depend on a number of additional factors, among which are: (1) the ease of supervising the recipients of capital, (2) economies of scale to the amount of capital, and (3) diseconomies in the assembly of the capital. We now consider these factors in greater detail.

Capitalists who cooperate with workers must either supervise the workers themselves or use specialized supervisors. The payoff to supervision is higher the more workers a supervisor can supervise, and the returns from investing in skill in supervision (or in any other skill) is highest when it is exploited on a full-time basis. Moreover, the opportunity to compare

¹¹ The employer has to guarantee the wage payment at the agreed-upon rate within the employment duration against defaulting and laying off workers when the agreed-upon wage exceeds the market wage.

similarly skilled workers with each other facilitates the supervision of these workers which, in turn, enhances scale economies in supervision. Capitalists, therefore, will tend to guarantee the actions of many workers who perform similar tasks. This may explain why fire-prevention experts who do not guarantee their own advice are usually employed and their work guaranteed by capitalists specializing in insurance; not by the capitalists who seek fire-prevention advice.¹² Market size, and the lack of guarantee capital, however, constrain the scope of the supervising unit.

We will now turn to features of guarantees and guarantee-capital. Guarantee-capital is not physical capital, but rather the right over it. Effecting a guarantee requires the transfer of rights from the guarantor to the guaranteed. The productiveness of guarantee-capital is a function of its transferability and is distinct from the productiveness of physical capital. The lower the cost of transfer, the more valuable it is for guarantee purposes. The ease of transfer enhances the guarantee function.

A basic feature of guarantees is the random nature of the timing and of the size of guarantee payments. Guarantee-capital serves as a “standby,” and it is called into use only in case of an “emergency,” when certain random outcomes are realized.¹³ Because of its standby character, such capital may serve several guarantee prospects at once. This feature of guarantee-capital gives rise to economies of scale through its use: as the number of guarantee prospects increases, the amount of guarantee capital required to attain a given guarantee level increases too but at a lower rate.

Let the random variable X be the potential liability of a producer. This producer will be able to follow through on his guarantee as long as $X < K$. Thus the effectiveness of his guarantee is increasing in $\Pr[X < K]$. If this probability is low, the guarantee is empty and the

¹² Alchian and Demsetz (p. 784) argue that the central provision of power and the production line reduced the cost of supervision and thus made team production cheaper. The same factors also enhance the scale economies in supervision. According to the hypothesis here, then, many similarly skilled workers may be employed in a firm even when they are not producing in “teams.”

¹³ Klein and Leffler (1981) are also concerned with guarantees. In the case they consider, however, the guarantee can be effected only if the capital backing it is sunk.

producer would have little incentive to exercise care in avoiding damage. Efficiency would be impaired.

How does the probability that the producer will honor his guarantees depend on the number of guaranteed prospects? Let the mean and variance of X be μ and σ^2 respectively. From the one-sided Chebyshev Inequality (see, e.g., Ross [1989]),

$$\Pr[X < K] \leq \frac{\sigma^2}{\sigma^2 + (\mu - K)^2}.$$

Rearranging the terms, one can see that if

$$K = \mu + \sigma \sqrt{\frac{p}{1-p}},$$

then the probability that the guarantee will be made good is at least p . In other words the capital required to guarantee a “security level” of p is given by the above expression. The guaranteeing capital is linearly related to the standard deviation of the liability in the above expression. If the producer combines his operation with another producer whose potential liability is Y , the standard deviation of the combined liability is $\sigma(X+Y)$. Unless X and Y are perfectly positively correlated, $\sigma(X+Y)$ is always less than $\sigma(X) + \sigma(Y)$. This means that producers can save on the amount of guaranteeing capital if they pool their activities. Alternatively, the combined guaranteeing capital would give a higher “security level” (and hence greater efficiency) than what is obtained if producers are independent operators.

If the expected value of claims is zero, that is, if guaranteeing is a fair gamble, it may appear that a risk-neutral person would agree to participate in the guarantee at no charge. Guarantee-capital, however, is not free. Suppose that claims are settled in cash. When paying a claim, the guarantor’s cash holding will decline, and when no claims are made, his cash holding will increase as he receives the insurance premiums. People, however, hold cash in the first place because of its *ease of transferability*, and there is a trade-off between transferability and productivity. People sacrifice the latter when they enhance the former. In case of claims, a guarantor must be able to transfer assets of value to the guaranteed. If he does that by holding cash, he sacrifices productivity—i.e., the return he could earn on the

capital. Suppose he chooses to transfer productive assets instead? Since any commodity specimen is unique, transfer, like trade, consumes resources.¹⁴ The more idiosyncratic the assets, the greater the cost of transfer. As the guarantee claims increase, then, for each one dollar value received by the guaranteed, the value lost to the guarantor exceeds one dollar by an increasing amount since the latter switches from transferring the less idiosyncratic assets to the more idiosyncratic ones. Therefore, even if the guarantor is risk neutral he will act as if he were risk averse.

Guarantee-capital, then, is costly, and the amount of it held is seldom, if ever, sufficient to cover *all* possible guarantee obligations at once. Because of the random nature of guarantee claims, it is not possible to assign to claimants clear priorities over the limited amount of guarantee-capital, and individuals' rights over the guaranteed action are incomplete. Therefore, such prospects are, in part, in the public domain, and individuals will expend resources in order to claim them. The guaranteed may attempt to advance the actual claims in order to increase the chance that when they make claims, enough guarantee capital will still be available to cover them. The "rush to collect" from firms on the verge of bankruptcy is a dramatic illustration of such action (as is a "run on the banks").

The better the delineation of claims over guarantee-capital, the less resources will be spent in claiming it, and the more it is valued. Owners of guarantee-capital can enhance its value by assembling prospects which are least positively correlated with each other.¹⁵ The correlation among guarantee claims can be reduced in various ways. For example, if the value of a commodity which is subject to substantial supply fluctuations is guaranteed, then guaranteeing complements to it reduces the overall variability to which the guarantee capital is subject. Another way to reduce the correlation among potential claims is to adopt production methods that are less likely to generate long runs of defects. A subtle but more direct option

¹⁴ The less "liquid" assets are, the higher the cost of transfer. "Illiquid" assets are simply costly to transfer.

¹⁵ Banks attempt to solve the same problem when structuring the riskiness of their loan portfolio.

which reduces the incentive to rush, and which, therefore, enhances the rights to the guarantee-capital, is to make the maximum payment which each prospect may receive an increasing function of time (as banks often do). Finally, the value of the guarantee capital can be enhanced through the maintenance of the priority order of damage claims. Such order, although unlikely to ever be perfect, may still be adversely affected by mixing together what previously were reasonably well ordered distinct sets of claims.¹⁶

When guaranteeing-capital is supplied by single individuals, its amount depends on the individuals' personal wealth. Individuals, of course, may cooperate with each other in order to assemble more capital than single individuals are able or willing to supply, and thus benefit from the scale economies of guaranteeing-capital. The cooperation in the assembly of capital, however, encounters its own guarantee problem which counters the economies of scale in the use of the capital.

What forms can cooperation among suppliers of guarantee-capital take? Suppose that previously each individual was using the capital to back his own production. If the cooperation takes the form of allowing each to use the combined capital for guaranteeing without restrictions, each then can reduce his own cost by producing more damaging output. As suppliers of capital, the individuals will try to prevent its abuse. They may, for example, restrict production to only one narrowly specified commodity in order to reduce the cost of monitoring. As the number of the suppliers of capital increases, however, their incentive to directly monitor each of the producers declines. The public shareholders corporations offers a way to resolve the problem.

Rather than having individual capitalists guarantee each other, pooling capital as a form of cooperation among capitalists, as publicly held corporations do, has several attractive features. The most basic of these is that capital owners cannot free ride on each other's guarantee capital. There are two other features to this arrangement. First, guarantee prospects

¹⁶ As conditions change, some priority orders may become more ambiguous. Firms may then divest themselves of some activities in order to generate clearer priorities. Other priorities may become easier to combine, and such activities may then be merged within firms.

are easy to coordinate. And second, the guaranteed find it easier to estimate the total amount of capital backing guarantees, which reduces the resource cost of evaluating the quality of the guarantee.

Pooling the capital of a number of individual capitalists, however, has one major drawback. The smaller the individuals' share in the total capital, the less incentive they have to spend resources on its management. They can employ managers to do the job for them, and they can tie their reward to performance. But the incentives of the managers who do not provide all the residual-bearing capital are bound to diverge from those of the suppliers of capital. This well known problem of the "separation of ownership and control" results in an advantage to pools where at least one individual contributes a significant share of the total capital. This factor, however, favors relatively small scale operations, thus countering the scale economies in the use of the guarantee-capital.

V. Entrepreneurship, Equity Capital, Management, and Integration

In dispersed-ownership corporations shareholders whose ownership-shares are small have neither the incentive to acquire skills necessary to manage their corporations nor to actually spend resources on managing them. Our basic proposition states that variability should be borne by the parties with the most control over it. A large part of the variability in the stock value of a firm comes from systematic (economy-wide) risks over which nobody has direct control. It is then befitting that such systematic risks are borne by passive diversified shareholders who have little control over the management of the firm. If small shareholders are not expected to provide entrepreneurship, who will produce these services?¹⁷ What is the system of rewards which generates the entrepreneurial services actually produced? and how are they guaranteed?

¹⁷ Large (undiversified) shareholders, and various independent suppliers bear the non-systematic risks associated with the firm. The large shareholders are the claimants to the *firm's* variability (as opposed to economy-wide variability), and they do have an interest in the operation of the firm. Few of the largest firms, however, have large individual shareholders.

Entrepreneurship is not a routine activity. The difficulties incurred by its exchange stem from the lack of a yardstick to measure its contribution. Not being routine, entrepreneurial output is subject to large variability. Its consumers, including cooperating input owners and the dispersed shareholders, would want it guaranteed. Dispersed shareholders may buy the services from entrepreneurs who possess the capital to guarantee their action and who operate independently. Alternatively, they may obtain the services from employed managers. The level and the nature of the services obtained from employed managers is limited, however, because such managers do not fully guarantee their actions.

Consider first entrepreneurial services supplied by independent operators. A whole variety of firms, often owned by experts, supply management services to others. Among these services are those supplied by labor-management firms, by accounting firms, and by investment bankers. To illustrate the case of purchased managerial services, consider some of the services that investment bankers provide and how they are guaranteed. Investment bankers specialize, in part, in linking budding entrepreneurs with suppliers of equity capital. Syndicates of investment bankers assess the value of new enterprises and suggest a price at which to sell the stock to the public. In a parallel manner to the bond underwriters discussed in Section II, the syndicates guarantee to the stock issuer major aspects of their advice on the value of the enterprise by purchasing all the shares to be offered to the public at a price agreed upon in advance.¹⁸ They also convey the information to potential investors, and guarantee to them certain attributes of their product, for instance that the enterprise claims are not fraudulent.

Investment bankers perform another major task: they prepare the charter of the enterprise, called the Articles of Incorporation. The charter specifies, among other things, the means available to shareholders for controlling their managers. Since syndicates formed by investment bankers are, in part, residual claimants to the difference between the price they receive and the price they pay for new issues, the investment bankers are induced to write

¹⁸ Corporate raiders who engaged in predicting the value of enterprises also guarantee their prediction. They commit their own capital to the enterprises they deem underpriced.

charters that maximize, net of their own costs, the value of the enterprises to the new shareholders. Investment bankers, then, perform an entrepreneurial function which they sell to other firms; it is narrow, but clear-cut, and they guarantee its quality.

Not all managerial services can be obtained from the outside; in some cases the expected variability in outcome exceeds experts' means to adequately guarantee their advice. Some such services may be provided by employed managers.¹⁹ The problem of securing entrepreneurial services from managers is complex. A glance at existing remuneration schemes helps us to grasp the problem. Existing remuneration schemes consist of slow moving wages, of various performance bonuses, and of ownership of corporate stock and of stock options. It is sometimes inferred that efficiency will be enhanced if managers' reward is made to correspond more closely to the performance of the firms they manage (Jensen and Murphy 1990). The total value of individual corporations, however, tends to swing widely. According to the view presented here, managers tend to become others' employees precisely because their wealth is insufficient to bear sizable corporate losses. Were they the claimants of a large fraction of corporate gains, but not of losses, they would have been induced to undertake bad gambles. Their actual managerial reward seems to be structured such that their incentive to take gambles of any kind is small.²⁰

Turning to "true" entrepreneurs, consider a person who has a promising business idea, but lacks the requisite capital to fully develop it. Fixed interest lenders would furnish the necessary funds were they to receive a guarantee that the idea was sound and that it would be properly developed. Such a guarantee is unlikely to be forthcoming; the idea-man lacks the capital necessary to provide such a guarantee, otherwise he would not need the lender in the first place. Investors may agree to finance the venture if they obtain both the ownership rights

¹⁹ Some services will simply not be provided in that case. This suggests that enterprises managed by their owners will differ from dispersed ownership enterprises in their "character"—i.e., in the kind of entrepreneurial services *produced*.

²⁰ The other side of the coin of managers' pay structures is that shareholders gain from exceptionally good management and lose from bad management. These are gains and losses which, on the margin, are unrelated to shareholders' contributions (although giving them the aura of "owners").

to the residual of the enterprise and the right to constrain the developer's action. Making the developer their employee is part of the constraining process. The developer, however, is not expected to become an employee immediately after coming up with the idea. Capitalists would be willing to consider financing the enterprise only after the idea acquires a relatively concrete form and the operation of the firm becomes sufficiently routine so that control of management is not exceedingly expensive.

The capital suppliers, i.e., the shareholders, do not become the owners of the enterprise through some divine power, nor even by virtue of wealth. They become owners because the associated division-of-variability and constraint patterns maximize the value of the enterprise. These capitalists, of course, will make the stipulations necessary to induce the employed developer to act consistently with the maximization of the capitalists' net gain from the enterprise. On his part, the developer will attempt to raise capital at the lowest possible cost, offering to financial supporters a rate of return near the market rate of interest, plus a premium, also near the market rate, for assuming the guaranteeing function. The developer will then retain the value of his contribution to the enterprise net of the cost of outside financing. Because of the difficulty of guaranteeing his actions, the developer, who initially owned the business idea, is likely to relinquish both residual claimancy and control to the financial supporters and agrees to be constrained by them.²¹

Guarantee problems are likely to arise not only between the entrepreneur and the financial backers, but also between the entrepreneur and the input-owners that are unique to his or her production. Consider an entrepreneur who makes a new product that requires some novel specialized inputs as well. The variability in the value of these inputs per unit of the input service depends largely on the value of the final product. Cooperating input owners would be unwilling to invest in providing novel inputs unless there is some assurance of the commercial success of the new product. Lacking such an assurance, the entrepreneur can

²¹ The last two paragraphs echo Jensen and Meckling's (1976) analysis of the financial structure of the firm. Considering shareholders as "principals" and their employees as "agents," however, obscures the roles the two types play.

secure the needed inputs only by guaranteeing payments to the cooperating input owners and financing the investment in their new skills. Thus firms making new products tend to start out as vertically integrated entities, with the entrepreneur who first conceived of the business idea bearing most of the variability. As the product proves to be successful and the market expands, the need for an explicit guarantee to induce investment diminishes. Moreover, as the variability in the marketability of the new product is resolved, the variability in the quality of cooperating inputs becomes relatively more important. Cooperating input owners can guarantee the quality of their input services by assuming independent status; selling the input services to the entrepreneur and bearing their variability. Our theory, then, predicts vertical disintegration as new industries mature.²²

VI. Guarantee by Reputational-Capital

In the preceding sections we analyzed the guarantee function of the equity of the shareholder corporations. In this section we apply our model to a different kind of capital—reputational-capital. We show that when such capital is used, an increased market size is conducive to the formation of horizontally integrated firms employing many similarly specialized employees.

The sale of many goods and services is backed not by guaranteeing capital but by the reputation or brand name of the seller. Examples are services of self-employed plumbers, shoe repairmen, doctors, and lawyers, and, to a degree, the sale of many low-valued commodities such as candy bars or soap. In sales backed by reputation, or brand name, the seller loses reputational-capital when providing a poor service, but the buyer is not compensated for the poor service.

The use of reputation is advantageous when the buyer cannot demonstrate a loss easily. For example, a theater-goer may believe that the actors did not act as well as promised. However, it is very difficult to measure the quality of acting objectively, and had the seller

²² Stigler (1951) also predicts such a pattern, but his explanation is technological.

promised “satisfaction guaranteed or your money back,” it would have been impossible to sort out those demanding refund even though they were in fact satisfied by the performance. If, on the other hand, all the seller stands to lose is reputation by, for instance, foregoing repeat sales, satisfied buyers will gain nothing from punishing the seller.

Even in cases when damage can be demonstrated, the guaranteeing capital does not serve a useful purpose when the damage in each individual case is too small to justify activating the guaranteeing mechanism. Whereas the sellers may offer refunds and may indeed pay when consumers complain that the product was stale or otherwise defective, few consumers would bother to make a claim. Since consumers are aware they can be easily exploited then, the seller must demonstrate that he stands to lose if he provides low quality output, as he does when he has reputational-capital.²³

Consider now the formation of reputational capital. Buyers have some expectations regarding the quality of the goods they receive, perhaps in the form of the distribution of qualities. Whenever a buyer consumes a specimen of the commodity, he will determine where in the distribution it lies, and add to his or her own compilation of the seller’s cumulative performance. A bad specimen will lower the evaluation of a “well tested” seller only slightly, but would substantially lower the evaluation of a seller with only few accumulated observations. A new seller can establish a reputation by consistently selling high quality specimens. Establishing a reputation is an act of investment when such specimens are sold at, or below, the prevailing market price.

As is well known, establishing reputation is a function of such factors as the variability of the product, its value, frequency of purchase and ease of exchanging information among buyers. The degree of specializing tends to interact with these factors. We will now focus on the degree of specializing which is a fundamental determinant of the ease of establishing reputation.

²³ The two forms of capital are likely to be used together. Buyers of candy bars may require a real guarantee against poisoning.

People who spend their entire market time serving small communities cannot be highly specialized. A physician in such a community is likely to be a general practitioner, not a brain surgeon. Specializing is likely to increase with the size of the community or market. As a worker becomes more specialized, he or she serves a larger number of buyers, each spending a smaller amount on the purchase. Reputation, then, becomes more difficult to establish. Patients of general practitioners use their services often and can assess their performance more easily than patients of brain surgeons can assess the latter's performance.

The ease of establishing reputation depends on one more factor—the volume of output that forms the basis of the reputational record; the larger it is, the more accurate, and thus the easier to compile, is the record. The larger volume of output of a given commodity is likely to be purchased by more individuals rather than through larger purchases by given individuals. Still, reputation is cheaper to establish when the volume is large because of the information exchanged among buyers and because visibility seems easier to attain with larger volume.

An independent producer may increase the volume of his or her output, but obviously only within narrow limits. More output can be produced primarily by increasing the number of the producers of the output. The individual producers of the output can enhance their reputation by pooling their output only if they maintain a rigid regime of product quality. Otherwise the larger output will produce no reputational gain; if the producers remain independent they will each gain from saving on their costs by producing low quality output. The economies of scale in forming reputational-capital, then, require that the individual producers agree to be constrained, and, most likely, become employees. So in a manner parallel to the case of individuals who do not have enough guaranteeing-capital, becoming employees in horizontally integrated firms is also a method of resolving the difficulty of forming reputational capital. Therefore, in cases in which reputational or brand name capital is important and where individuals are highly specialized, firms employing individuals with similar skills are expected to emerge.²⁴

²⁴ It is ironic that as individuals become more specialized, and thus serve more consumers, they then gain from forming horizontally integrated firms that serve yet proportionately more consumers.

VII. Summary

Transactions are subject to variability, and their outcome can be affected by both transacting parties. The allocation of variability which will maximize a transaction's value must follow the principle that the greater a party's gain from affecting the outcome, the greater is the share of its variability he or she should assume. The gain from affecting the outcome depends on contract stipulations, while the ability to assume variability depends on the amount of wealth which backs the parties' actions.

Firm owners deal with various other resource owners whose cooperation is needed in order to produce the desired product. The dispersed firm owners, however, are not induced to acquire the skills necessary to run their firm. Therefore, they are more ignorant than the owners of cooperative factors in the latter's specialties. In order not to be cheated, firm owners will demand that cooperating factors guarantee their contributions. Owners of some cooperative factors supply their own guarantee-capital, and transactions subject to such guarantees are largely outside the firm. When buying guaranteed output, firm owners will get a refund if its quality is sub-par; therefore they have no incentive to supervise those factors who participate in producing such output. Shareholders also collaborate with those owners of factors whose capital is not adequate to fully guarantee their own actions. The shareholders supply, then, the guarantee capital and assume much of the effect of the collaborative variability. Those with capital will constrain those without capital, especially by making them employees, even if they are risk-neutral. Because employees are compensated by fixed pay, their incentive to produce defective products is greatly reduced, but they must be supervised if they are to be induced to produce the right amount of output at the right level of quality. Such a relationship is part of the firm. The firm is defined by the totality of its guarantees, i.e., by the set of contracts guaranteed by the pooled equity-capital, and it is characterized by the restrictions imposed on those resource owners whose actions it guarantees.

Guaranteeing is subject to economies of scale. Combining two units of guarantee-capital tends to more than double the guarantee level. Such a combination is usually achieved by creating a single pool of capital. As the number of individuals supplying capital to the pool increases, however, their incentive to manage it falls. Some managerial services are acquired in the market, others are obtained by employing managers. Managers may perform, as a part of their function, some entrepreneurial tasks. Independent innovators, however, seem to be a major source for providing entrepreneurial services. These latter individuals tend to initiate the development of ideas. Projects which require high guarantee levels must be set up initially so that their value can be easily assessed; entrepreneurs will then tend to sell out to those willing to provide the necessary guarantee capital. When managers take over such operations, much of the entrepreneurial work will have already been accomplished.

REFERENCES

- Alchian, Armen A. and Harold Demsetz, "Production, Information Costs, and Economic Organization," *AER*, December 1972.
- Barzel, Yoram, "Measurement Cost and the Organization of Markets," *J. of Law and Economics*, April 1982.
- Barzel, Yoram, "The Entrepreneurs' Reward for Self-Policing," *Economic Inquiry*, January, 1987, 103-16.
- Barzel, Yoram and Wing Suen, "The Allocation and Guarantee of Outcome Variability," Working Paper, 1992.
- Barzel, Yoram and Wing Suen, "Moral Hazard, Monitoring Cost, and the Choice of Contracts," Working Paper, 1992.
- Barzel, Yoram and Levis Kochin, "Ronald Coase on the Nature of Social Cost as a Key to the Problem of the Firm," *The Scandinavian J. of Economics*, Vol. 94, No. 1, 1992.
- Cheung, Steven N.S., "The Contractual Nature of the Firm," *J. of Law and Economics*, April 1983, 1-21.
- Coase, Ronald H., "The Nature of the Firm," *Economic Journal*, 1937.
- Coase, Ronald H., "The Problem of Social Cost," *J. of Law and Economics*, 1960.
- Cooper, Russell and Thomas W. Ross, "Product Warranties and Double Moral Hazard," *Rand Journal of Economics*, Winter 1988, 19, 495-515.
- Eswaran, Mukesh and Ashok Kotwal, "A Theory of Contractual Structure in Agriculture," *AER*, June 1985, 75, 352-367.
- Eswaran, Mukesh and Ashok Kotwal, "Why Are Capitalists the Bosses?" *Economic Journal*, March 1989, Vol. 99, 162-176.
- Goldberg, Victor P., *Production Functions, Transaction Costs and the New Institutionalism*, in George R. Feiwel, McMillan 1985.
- Goldberg, Victor P, "Aversion to Risk Aversion in the New Institutional Economics," *Journal of Institutional and Theoretical Economics*, March 1990, Vol. 146, 216-222.
- Grossman, Sanford J. and Oliver D. Hart, "The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, August 1986, 94, 691-719.
- Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *JPE*, December 1990, Vol. 98, 1119-1158.
- Jensen, Michael C. and William H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 1976, No. 3, 305-360.

- Jensen, Michael C. and Kevin J. Murphy. "Performance, Pay and Top-Management Incentives." *Journal of Political Economy*, April 1990, 98, 225-264.
- Kambhu, John, "Optimal Product Quality Under Asymmetric Information and Moral Hazard," *Bell Journal of Economics*, Autumn 1982, 13, 483-492.
- Klein, Benjamin and Keith B. Leffler, "The Role of Market Forces in Assuring Contractual Performance," *JPE*, August 1981, 615-641.
- Klein, Benjamin, Crawford and Armen A. Alchian, "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *J. of Law and Economics*, October 1978.
- Knight, Frank H., *Risk, Uncertainty and Profit*, Boston: Houghton and Mifflin, 1921.
- Kramer, Michael, "The O-Ring Theory of Economic Development," *QJE*, August 1993, 108, 551-575.
- Ross, Sheldon, *A First Course in Probability*, 3rd ed., New York: Macmillan 1989.
- Sappington, David, "Limited Liability Contracts Between Principal and Agent," *Journal of Economic Theory*, February 1983, 29, 1-21.
- Simon, Herbert A., *Models of Man*, New York: John Wiley and Sons, 1957.
- Stigler, George J. "The Division of Labor is Limited by the Extent of the Market." *Journal of Political Economy* 59, June 1951.
- Warren, W.L. *The Governance of Norman and Angevine England 1086-1272*, Stanford, 1987.
- Williamson, Oliver E., *Markets and Hierarchies: Analysis of Anti-Trust Implications*, New York, 1975.