

August 13, 2001  
**Preliminary**

**A Theory of Organizations  
To Supersede the Theory of the Firm <sup>1</sup>**

**Yoram Barzel**

**Introduction**

The theory of organizations that I formulate in this paper has three central themes—enforcement, measurement and the claim that the firm is not a useful concept and should be replaced with organizations. Agreements within and across organizations must be enforced, and enforcement issues suffuse agreement

---

<sup>1</sup> I wish to thank Tim Dittmer, Kristen Foss, Nicolai Foss, Chris Hall, Gary Libecap and Armelle Maze for their comments.

A more extensive paper under the same title, available from my web site cover the themes: 1. Enforcement. 2. Measurement. 3. Developing a theory of organizations. This version, prepared for the ISNIE presentation contains the Enforcement section, as well as the introduction, the conclusions and the full references, but not the discussion of Measurement and of the desirability of replacing the firm with organizations.

formation and execution. I focus on the difference between state enforcement and enforcement by long-term relations. State enforcement is at the interface of organizations, whereas enforcement by long-term relations characterizes, among others, within-organization interactions.

Costly information, and its operational counterpart, costly measurement, is a basic ingredient in the analysis of institutions. Adopted here is a measurement costs framework that unifies and augments the analysis of the leading transaction costs models of the firm. I contend, however, that in spite of its popularity, the firm itself is not a useful concept, and that we should rather focus on organizations.

In Barzel (2002) I argue that the state, with control over violence, is the efficient enforcer of easy to-delineate (i. e., to measure) transactions. These, governed by contract, are classified as being “in the market.” I also argue that except for caveat emptor transactions, contracts are used to enforce only parts of transactions, and that long-term relations are used to enforce other parts of the same transactions. The comparative advantage of the latter is in enforcing components that are less explicitly specified. This difference is significant, among other things, for the study of vertical integration and for explaining the radical change in the character of the pre industrial revolution business enterprise to the one that emerged subsequent to it. More generally, I attempt to explain which enforcement method will be utilized in particular cases or, in different words,

which activities will be undertaken in the market and which (mostly) within organizations.

The definitions of organization and its scope that I adopt are as follows:

**Definitions:** *An organization is a nexus of the agreements and parts of agreements guaranteed by centralized equity capital and enforced without the state's assistance.*

*The scope of the organization is the ratio of its guaranteeing capital to (some measure) of its expected guarantee payments.<sup>2</sup>*

The measurement cost framework enables us to consolidate the existing transaction costs models of the firm (now understood as organizations). I show that this framework can unify the models by Alchian and Demsetz (1972), Williamson (1975), Klein, Crawford and Alchian (1978), and Barzel (1982). It also encompasses the role of guarantee capital (Barzel and Suen (1995) and Barzel (1997)). The need for such capital helps determine the size and boundaries of organizations. An implication that carries through *all* these models is that as the cost of measuring commodity attributes declines, more activities will be carried out in the market and fewer within the firm (or within organizations).

---

<sup>2</sup> Insurance firms are characterized by a low scope relative to the amount of guaranteeing capital they have.

Although this paper is in the tradition of the transaction costs theory of the firm, the firm is a too diffuse a target for explanation. Some writers on the firm, especially in the law and economics literature, equate firms to for-profit corporations, subject to corporate law.<sup>3</sup> This is sensible in that it is clear what falls under the definition. More importantly, the rather common assumption that the sole objective of the firm is profit maximization seems reasonable here (though it is easy to think of real life caveats). Students of the firm, however, tend to hold a more expansive view of what constitute firms. Viewing profits as owners' *sole* objective ceases to be useful in the analysis of ownership forms such as partnerships, coops, or condominiums. In residential coops, for instance, an apartment that is offered for sale does not necessarily go to the highest bidder. Moreover, by invoking this assumption we forego the opportunity to ask why are these organizational forms needed; why aren't they all organized as for-profit corporations? When owners maximize with respect to two or more objectives, we need to know what these are and to form our models correspondingly.

In any case, what is *the firm*? Economists seem to clearly know what firms are. This "knowledge," however, evaporates once one is required to define or describe them. To broaden the coverage of the term beyond the corporation, we need to know what falls under it, but no useful stopping rule suggests itself for

---

<sup>3</sup> One example is Rock and Wachter (2001).

doing so. An operational theory of the firm requires dividing organizations into two mutually exclusive sets, one consisting of firms, and the other of non-firms. At the present state of knowledge such separation is unlikely to emerge; the existing transaction-cost based theories of the firm are not amenable for such separation.<sup>4</sup> For instance, we observe employers employing teams of workers, paying them wages for their effort, thus seemingly conforming to the Alchian and Demsetz model. Such a production mode, however, is not unique to for-profit corporations. Team production also takes place, among others, in not for profit organizations and even in government. The notion of team production may help determine what takes place within organizations and what takes place outside them, but as it is not unique to any narrow concept of firms, it is not helpful in delineating them.

I propose to focus instead on explaining organizations. The rationales underlying the existing models of the firm may contribute to the explanation of organizations. I argue that it is useful to group organizations according to their ownership structure. The basic hypothesis here is that each ownership structure has a comparative advantage in reducing capture costs in an identifiable class of activities. Our task then is to predict which activities belong to which ownership

---

<sup>4</sup> The model of the firm by Hart and associates, as Holmstrom points out, is subject to the same problem. In a separate paper I critically scrutinize that model (as well as the traditional, text book theory of the firm).

structures. Such a theory should allow us to determine which activities will take place within which organization and when would each type expand or shrink.

## **Enforcement**

### *Introduction*

In this section I attempt to determine which exchanges will be transacted by contract and carried out in the market, and which ones will be carried out within organizations.<sup>5</sup> The state imposes certain restrictions (e. g., quid pro quo, or sometimes the use of official units of measurement) that contracts must conform to. Aside from these, the transactors can stipulate in their contract whatever they wish. We expect contractors to stipulate those attributes that are most readily measurable relative to their values. As a rule, the state does not participate in writing contracts. Therefore, information about what the parties actually agreed to is entirely based on the contracts themselves. To be effective, then, contracts must clearly stipulate the agreed upon terms. For this reason, transactors tend to use explicit, and where possible standardized measures of the attributes they stipulate in their contracts. In turn, contracts tend to be impersonal. One immediate

---

<sup>5</sup> Rock and Wachter (2001) are concerned with the same issue. They hypothesize that in the case of for profit corporations, the courts step in when norms and self-governing arrangements are insufficient.

implication is that the more valuable is exchange among strangers provided they are subjects of the same state the more effective is the exchange by contract.

In the common-stock corporation with no limitation on share transfers, the owners need not know each other, and thus, as a rule, are not subject to long-term relations. Long-term relations cannot govern such organizations; they can be erected *only* by contract, subject to state jurisdiction. Where contract enforcement cannot be relied on, the range of ownership forms is expected to be rather limited. This, for instance, may explain why organizations in much of Southeast Asia are family-dominated. Moreover, where owners are not free to transfer their shares, it is unlikely that organizations with a large number of owners could function. This applies not only to conventional for-profit firms, but also to organizations such as guilds or labor unions. Any of these organizations could have been maintained provided that owners are restricted to tight knit groups, or alternatively, where the organizers themselves have access to the use of violence, as is the case in some labor unions. But then such organizations themselves acquire some of the characteristics of the state.

Non-contractual agreements are less formal and less clearly spelled out than contractual ones. Lon-term relations have a comparative advantage in enforcing transactions that are less explicitly specified and that rely more heavily on the transactors' and the enforcers' reputations. For example, we observe that in the

marriage relationship the state enforces the distribution of marital property in divorce, but not love and affection, and not even the quality of cooking. Agreements that are enforced by a third party enforcers using long term relations tend to lack formal documentation. The parties must apprise the enforcer of their intentions when they make the agreement; otherwise the enforcer would not know well enough what to enforce. From this angle it is seen that the ease of measurement and with that the ease of delineation determines, in part, whether transactions will be in the market, enforced by the state, or within organization, enforced by other means.

Prediction: As attributes become easier to measure, their enforcement will shift from long-term relations to the state. This will happen, for example, when new commodity standards are formed. A movement between forward contracts and futures contracts could test the proposition.

In the rest of this section I first clarify some terms. I then discuss the enforcement power of the state vis a vis that of long term relations. Next I indicate what is the nature of the activities over which each of these enforcement forms has a comparative advantage in enforcing. I conclude by bringing out some new notions about vertical integration.

### *Terms and definitions*

Enforcement within organizations (or firms), as is well recognized, differs from enforcement in the market. Not much attention, however, has been given to the nature of the two enforcement forms. In this section I compare enforcement by the use, or threat of use of violence, which is the realm of the state with that resulting from long-term relations that organizations use.<sup>6</sup>

To proceed, certain terms have to be clarified and transaction costs defined. I make a distinction between agreements and contracts. *Agreements* encompass entire relationships. *Contracts* are agreements or part of agreements that the state enforces.<sup>7</sup> In addition, by my definition, all agreements governed by contract are ‘*in the market.*’ I also distinguish between legal (property) rights and economic (property) rights. *Legal rights* are individuals’ rights that the state helps enforce.

---

<sup>6</sup> In Barzel (2002) I compare the two enforcement modes and dwell on the restrictions that subjects of rule of law states impose on the violence-using enforcer.

<sup>7</sup> Masten, among others, also defines contracts as agreements that the state enforces.

*Economic rights* are what people can do with their commodities or assets.<sup>8</sup> This view of what economic rights are conforms to Alchian's (1965, 1987) and Cheung's (1970) notion of property rights.

Allen (1991) provides a most attractive definition of transaction costs (while, like Alchian and Cheung, using the term "property rights" for what is called here "economic rights").<sup>9</sup>

**Definition:** "*Transaction costs are the resources used to establish and maintain property rights*" [boldface in original].<sup>10</sup> Elaborating, he states that "They include

---

<sup>8</sup> The concept of economic rights is strictly positivistic, and has nothing to do with "entitlements." What individuals maximize are their economic rights.

<sup>9</sup> See also Allen (2000).

<sup>10</sup> The substance of this definition is identical with Jensen and Meckling's definition of "agency cost." Under the definition of transaction costs adopted here the costs of excess measurement as well as the costs of such things as locks or of not wearing expensive jewelry for fear of theft are transaction costs. Jensen and Meckling (1976) deal with a narrower subject, and it may escape notice that these costs also fall under their definition of agency costs. Since the parties incurring such costs are not always "principals" and "agents," the term "agency cost" does not convey the full breadth of the concept. Holmstrom and Milgrom (1994) notion

the resources used to protect and capture (appropriate without permission) property rights, plus any deadweight costs that result from any potential or real protecting and capturing.”

As defined here, then, transaction costs are the “dual” to property rights.<sup>11</sup>

*Some aspects of state enforcement*

The violence using enforcer, i.e., the state, does not depend on long-term relations for enforcement power, and is more likely to confiscate than are those using long term relations. Seemingly to abet the threat of confiscation, subjects of rule-of-law states require the state to use objective and explicit criteria for adjudicating contract-disputes (Barzel 2002). The more successful they are, the higher the quality of the safeguards, and the more attractive is state enforcement. We expect contractors to write contracts that are relatively clearly delineated in order to take advantage of state enforcement. Two additional factors affect the attractiveness of state-enforcement: 1. How well delineated the law is (e. g., more so for traditional transactions than for new types). 2. The probity of judges, as well as the level of corruption of the judicial system.

---

of agency cost is the same as that of Jensen and Meckling. It seems to me that they handcuffed themselves by the narrowness that the term connotes.

<sup>11</sup> Neither Williamson (1975) nor Grossman and Hart (1986) are aware that each concept is the dual of the other.

Consider now caveat emptor transactions. These are consummated at the time of the agreement, and lack built-in continuation. They can accommodate trade among anonymous transactors only under state enforcement. What the state enforces here consists of ensuring that each gives the other something of value. In spite of its minimal level, such enforcement tends to alleviate the last period problem, and of course plain theft, which is an extreme form of the last period problem.

Transactors attempt to determine the levels of the attributes of the commodities exchanged in caveat emptor transactions must occur before they are consummated. Certain transactions would not take place under such terms because adequate determination (measurement) is too expensive. The parties, however, may use the state to secure and enforce the levels of various attributes of the transaction. Alternatively, as discussed later, the parties may use long term relations for that purpose. Now, as shown in Barzel (2002), the state enforces by itself only caveat emptor transactions. The state enforces all other transactions, especially those that span a period of time, at most only in part. Long-term relations are employed to enforce some of the other attributes of these transactions. Transactors, then, must form long-term relations to be able to engage in non caveat emptor exchange.

*Enforcement by long-term relations*

Seemingly, the literature does not fully recognize, and definitely does not much explore transactors' need to use long-term relations for the enforcement of all of their interactions (other than those by caveat emptor). There are two well-known conditions for transactions to be self-enforced. One is that each transactor expects the relationship to extend into the indefinite future. The other is that each perceives that at any point in the life of the relationship the other's gain from renegeing is less than the present value of the loss in opportunities for future exchange.

One immediate implication of these conditions is that in the absence of state enforcement, strangers cannot effect exchange even if each has a good the other desires and they would have agreed on the terms of exchange. The reason is that neither expects the other to perform and thus neither of them views the exchange as worthwhile. Rather, if they actually meet, the stronger of the two will simply walk away with the other's good. It may appear that breaking down a lumpy, one shot transaction, into many small ones to be traded successively may create conditions for exchange, since a person not paying for a batch would be deprived from additional units. But because of the finite duration of the interaction, the exchange will not be effected. Whoever runs out of his commodity first will not get paid for the last batch because he cannot penalize the other for non-payment. But then he would not offer his last batch, and similarly for the succession of all earlier

batches. Thus, as is well known, the transaction will unravel. Among transactions with positive joint net value that do not meet the indefinite life condition necessary for self-enforcement are lending transactions and sporadic exchanges. The latter may include the intermittent exchange of assets that are not economical to turn into the continuing exchange of the assets' services.

Reputation may be built up and become a substitute for the indefinite life of relationships. If each of two would-be transactors has some reputation, when one wishes to get what the other has he will attempt to demonstrate that his losses from renegeing will be large. The more difficult it is to determine the value of what one receives, the more difficult it is to figure out if trade is worth while, and thus to reach agreement about it. In addition, the more the execution of the exchange puts one in a position to take or to steal, the less desired he is as an exchange partner. Recall that there is no state here to deter theft. The main deterrent is the loss of reputation that one suffers when revealed to be a thief.

The higher the level or quality of one's reputation, the higher the likelihood that others will seek him as an exchange partner. Reputational quality is not fixed; people can affect its level by their action. In a multi-installment exchange, for instance, as exchanges are carried out, each party is likely to encounter reasons to renege. Each time one gives the other the benefit of the doubt by deciding not to renege he engages in a reputation enhancing investment. The value of one's

reputation also depends on others' perceptions; among other things, the larger the number of potential trading partners who become aware of such behavior, the higher the return from it. The value of the reputation also increases with its durability. Thus, for instance, a given amount of reputational capital tends to be more valuable to younger people than to older ones.

State enforcement may allow exchanges to take place where reputation is insufficient or where it is absent. Much of the time, however, it suffices to assign to the state a partial role. In the example of the transaction that is broken down into a number of successive steps, it suffices to enforce the last installment for executing the entire transaction. We expect, among other things, that the easy-to-measure attributes of transactions will be enforced by the state and that other attributes will be enforced by long-term relations.

### *The enforcement of the employment relationship*

The employment relationship, though usually called the "employment contract," is governed only partly by contract. The contract typically stipulates attributes such as the wage rate and the length of paid vacation. It does not, however, spell out with any degree of completeness what the employer can demand of the employee, nor what the employer is expected to do for the employee. Indeed, under a contract that would have spelled out *all* such attributes, the employer would have had no leeway for making *any* master-servant type

demands on the employee. Thus the relationship would not become what is usually considered to be an employment relationship.

As the state does not enforce entire employment relationships, long-term relations must be deployed to enforce additional within-organization attributes. What is the nature of the enforcement power embedded in the long-term relations here and how potent is it? Note right away that the relationship is mutual. If either party has no reputation, they would be unable to form an employment relationship between them.

I proceed by first discussing a two person relationship, assuming that one will be employer and one employee, and then turn to briefly suggest and discuss a general reason why there would be such a relationship. Consider an employer and an employee. Each expect the other to provide some valued services not stipulated in their contract, and each would require the other to possess a sufficient brand name to back his (implicit) promise. Suppose they already possess brand names, say acquired from their high school acquaintance.<sup>12</sup> Each has some prior

---

<sup>12</sup> The term “day worker” characterizes a worker possessing the least amount of reputation and is still employable. It seems to imply that the worker’s reputation is insufficient to support a longer duration relationship. Such workers typically supply routine services that resemble what for commodities is sometimes called “commoditized.”

expectations of the other's (as well as of his own) performance. As they interact, they presumably encounter opportunities to operate at the initially expected level, at a higher level, or at a lower level. If the employer acts at least as well as expected, the employee is likely to continue the relationship. On the other hand, if the employee perceives underperformance, he is more likely to quit. When others observe a "quit," as when the separation occurs during a high employment period, they will lower the level of reputation they assign to the employer. However, when others observe a "dismissal," as when the separation occurs during a low employment period, they will lower the level of reputation they assign to the employee. A period without separation indicates that *both* are acting at or above their initial reputation level. Others, then, will assign both of them a higher level of reputation.<sup>13</sup> One implication here is that women that temporarily leave the labor force to bear children and then return to their old employers lose not only labor force experience but also the opportunity to enhance their brand name to the degree that the longer employment duration would have generated. The loss in brand

---

<sup>13</sup> It is well recognized that employees must be paid a premium to prevent them from reverting to the "competitive" labor force. The required symmetry seems to be overlooked, however. The employer must receive labor services valued more than the competitive wage.

name would be still greater if (absent regulation to the contrary) the old employers do not invite them back.

How are the employer-employee relationships enforced? Casual observation may suggest that the value of the brand names that different individuals have is often low compared with what it needs to enforce to secure the desired behavior. Transactors, however, may have a choice of the form of the agreement they adopt. Changing the agreement form from a market to an employment relationship involves a radical change of incentives that the stipulations of the agreement induce. This change seems to be the main force for inducing the parties to act in conformity with each other's desire, and correspondingly affects the guarantees that others look for.

Consider the 'make or buy' choice. A transactor may acquire a product, the 'buy' decision, thus essentially operating in the market. Or he may acquire inputs, the 'make' decision, thus essentially operating within an organization. In both the 'buy' and the 'make' agreements, the contract governing the transaction is "incomplete." Left out of the agreement are those aspects that are too costly to measure and enforce. Forming and enforcing agreements is costly, and their incompleteness opens the door for capture activities.<sup>14</sup> One input is that of labor

---

<sup>14</sup> In their discussion of incomplete contracts Grossman and Hart (1986) assume that owners are able to maintain *full* control over all uncontracted attributes. This

services, constituting the employment agreement within the ‘make’ decision. Among the left out terms in the agreement are the exact tasks workers are expected to perform, their promotion as well as various aspects of working conditions which must be self-enforced (Rock and Wachter).

Transactors will adopt the form they expect to generate the higher net gain (i. e., the one subject to lower transaction cost). We have to determine, then, under what conditions will one form of agreement be preferred over the other. An agreement regarding inputs may be preferred to one regarding the output when, for instance, evaluating the inputs is cheaper than evaluating the product itself. Although this difference may seem innocuous, it is actually of great significance.

Suppose the commodity to be exchanged is automobile breaks. When breaks fail, the resulting damage may be rather large. If determining the quality of the breaks at the time of purchase is difficult the buyer may require a guarantee. The reputation of an independent producer selling his devices, however, is unlikely to suffice for such purpose. A pecuniary guarantee, which is likely to be enforced by the state, may serve here well but neither is the seller’s wealth likely to suffice for

---

implies that owners own both economic and legal rights over these attributes. They fail to recognize that attributes that are not (fully) delineated such as the use of office copier are prone to theft and other forms of capture, especially by employees and by other contracting partners.

such a guarantee. If the seller cannot be fully penalized for careless behavior, we expect him to be more careless than is “efficient” and to produce a product that will cause “excessive” damage. He is not likely to be “careful enough” in his drive to maximize the value of his output. One may attempt to revert to a third party to provide the guarantee. This party, however, may charge an unacceptable high premium unless the seller agrees to be constrained; if not constrained, he has no incentive to avoid morally hazardous behavior. In some cases modest constraints on the seller would not suffice. One radical constraining method is to turn the seller into the guarantor’s employee.

An employee obviously has to be induced to perform; he has to have a reputation to lose in case the employer lays him off because of low level of performance. Now turning a person from operating independently to becoming an employee does not make the product less prone to cause damage. The change is in the terms of the agreement under which the two types of operations take place. Consider again the potential for faulty breaks. Under the employment agreement the need to guarantee them falls on the employer. When an independent producer becomes an employee, however, his incentives are radically altered. Being rewarded by a wage, and given the incentives that the wage contract brings about, he is unlikely to extend himself as he gains little from working hard. But because of its indirect effect on product quality, this is the behavior that the employer

wishes to induce. When he does not work hard he is less likely to be careless and produce faulty products that can harm their users.<sup>15</sup> By the same token, a truck driver transporting delicate china who is paid by the hour is inclined to shirk by “taking it easy,” which in this case actually plays into his employer’s hand. His shirking is harnessed towards the desired outcome.

In terms of reputation, what the employee has to guaranty is diligent and careful work. For that a modest amount of reputation seems adequate. Thus by turning the worker from operating independently to becoming employee, the amount of reputation needed to guarantee performance is radically reduced.

On his part the worker has to believe that the employer is motivated to perform as implicitly promised. The employer’s reputation must be adequate to guarantee such things as good ambience in the work place. The work place may also pose the danger of injury. If the work involved exposes the worker to serious injury, he may view the employer’s reputation as an inadequate to induce the appropriate behavior. There is no reason to expect the reputation of an individual employer to be adequate to guarantee his own action more than the worker’s guaranteeing his. The employer’s problem can be resolved, however, if the need

---

<sup>15</sup> This parallels Holmstrom and Milgrom’s statement (1994, p. 989) that “an optimal incentive structure may require the elimination or muting of incentives which in a market relationship would be too strong.”

for injury compensation can be documented, as that part of the agreement can be included in the employment contract. The employer then will have to demonstrate that he is in command of the pecuniary capital. Here he can take advantage of the scale economies to insurance by insuring many workers whose potential injuries are not highly correlated. He also insures the product. Obviously, he has to amass the capital necessary for that purpose.<sup>16</sup>

### **Additional enforcement features of the employment relationship**

In his critique of Alchian and Demsetz (1972), Holmstrom (1999, p. 80) states: “Any contract between the owner-monitor and the workers could just as well be carried out in the market as within a firm.” I agree, and this, indeed, seems to apply to any relationship that might take place either in the market or within a firm. Holmstrom, however, seems to ignore the possibility that the costs of the two forms of organizing relationships tend to differ. He acknowledges (footnote 20) that the “argument is weaker if we think of the contract as implicitly enforced, say through repeated interaction,” but he does not follow through on this observation.

What is the nature of the difference between independent and employed workers, or between production that is carried out “in the market” and production that is carried out “within an organization?” As Holmstrom implies, the effort a

---

<sup>16</sup> I address some of these issues in greater detail in the section on measurement.

worker exerts is equally observable to the would-be market buyers of his product as it is to employers. For three reasons the cost associated with observing workers and using the information within firms is less than the corresponding cost when they are observed in the market.

1. Market transactions are enforced by the coercive power of the state whereas firms enforce transactions that take place within them. Consider the enforcement of the level of effort. Effort has to be explicitly documented if it is to be adjudicated by the courts. Operating within the firm is different. An employer may rely on his informal judgement of the effort. He may reward a worker for “intangibles,” a term often used in team sports, which I interpret as observed behavior that is difficult to quantify and to document. The “employment at will doctrine” is the other side of this coin. Under this doctrine, an employer need not be able to demonstrate to the courts why he chose to terminate an employee.<sup>17</sup> The

---

<sup>17</sup> As state earlier, the state too could use informal observations for adjudication. Because the violence user can more easily abuse his power than can users of long-term relations, rule of law states are restricted to use only explicitly documented evidence.

greater the cost advantage of an implicit measurement relative to an explicit one, the higher the chance the transaction will be conducted within the firm.<sup>18</sup>

2. An employer who pays wages must observe employees' effort. When employing several workers he may place them next to each other to observe them together at the same time. Market buyers are less likely to operate that way; it seems "unmarket like" to constrain several sellers to operate side by side. The costs of the two forms would be the same only if the "market" arrangement would be subjected to the same restrictions (not state enforced) are imposed on the employment relationship. But then such "market transactions" are turned, as the IRS would agree, into employment relationships.<sup>19</sup>

---

<sup>18</sup> Suppose it is desired to induce harmonious relations among workers. Enforcement by the courts requires explicitly documenting the inputs to the relationship and the level of harmony actually attained, which seems to be exceedingly expensive. Therefore, the creation of such relationships may be left to the employer's less formal methods.

<sup>19</sup> Buyers of plumbing services or of medical services tend to be present when the service is performed, and thus are in position to supervise the provider. These services are often charged for by the hour. Such relationships seem to straddle the boundary between firm transactions and market transactions.

3. Workers interact when one transmits a product to the other, share machines, or move in the same space. They function best as independent workers if it is easy to explicitly delineate their rights. Suppose, however, that the quality of the specimens that one worker transmits to the other varies, or that they intermittently use the same machines or move in the same space. If they operate independently, each is likely to attempt to gain at the other's expense. Alternatively, a single employer may employ both individuals. The employer delineates the relationship between the two.

The instructions he gives and the rules he imposes constitute a redelineation of rights and he becomes a third party enforcer of these. The employer is expected to impose such rules and give such instructions so as to reduce the capture, and more importantly, reduce the incentive for capture. For instance, he may restrict the amount of time available for picking and choosing, and reduce the reward for choosing only high quality specimens. Or he may provide mutually exclusive territories to salespeople. The employer here induces a cooperative equilibrium. Independent transactors find such arrangements difficult to make, and Nash equilibrium is the more likely outcome here. I hypothesize that organizations are formed for the express purpose of creating rights that are more economically enforced by non-state means than by state means

These considerations yield the two following predictions:

1. *When the state's comparative advantage in enforcement has increased, (declined) more (fewer) stipulations will be agreed upon contractually.*
2. *When the state's comparative advantage in enforcement has increased (declined), the scope of the state will expand (shrink).*

### **Enforcement and Vertical integration.**

Definition: **Vertical integration** is a state wherein a residual claimant of a firm, or of some other organization, has taken on himself the combined variability of two or more vertical operations.

The distinction between state enforcement and other enforcement forms is at the heart of vertical integration as it is defined here. Although the definition is idiosyncratic, it will be seen to be quite useful. Understanding the nature of the enforcement of operations that take place within organizations sheds new light on the notion of vertical integration. In general, some of the interactions within an organization are enforced within them but some, though seemingly falling within the operations of organizations, are enforced by the state, and thus are actually in the market.

What does it mean that the state enforces agreements regarding activities that take place within an organization? Consider, say, Ford Motor Company buying batteries that the battery maker contractually guarantees to car buyers. The

battery maker, then, bears the variability in battery quality. It is true that the carmaker bought these batteries and installed them in his cars. But he acts here as his customers' agent and does not bear the effect of variability in the quality of the batteries. The batteries, therefore, are not a part of his vertical structure even while they passed through his plant.<sup>20</sup>

The merger of two vertical firms creates a single vertical firm out of them. I argue, however, that merger may actually *reduce* the level of vertical integration as it is defined here. Suppose that B, C, D, and E are vertical links in the production process. B, for instance, may be a wholesale buyer of fresh vegetables. C and D provide two successive links in transporting the vegetables to E, a supermarket. To proceed I make the following assumptions: 1. Freshness is a valued attribute. 2. Freshness is easy to inspect when B prepares the vegetables for shipment as well as when they are delivered to the supermarket. 3. Under the most economical packaging method, freshness is expensive to inspect when vegetables are in transit. 4. Shippers' costs are lower when skimping on maintaining freshness.

---

<sup>20</sup> In reality things are more complicated. Ford may make engines that draw too heavily on the battery and be able to shift some of the cost to the battery maker. On its part, the battery maker may make weak batteries resulting in cars not starting well, and customer may believe it is a car rather than a battery problem.

Suppose that while shippers C and D operate independently of each other, B employs the low-cost packaging method and guarantees freshness to E. In that case, C and D can gain by skimping on maintaining freshness since each can blame the other for the loss of freshness. The guarantee, then, encourages C and D to spend too little on maintaining freshness. Therefore, B's guarantee of freshness to E will not survive for long.

As long as C and D continue to operate independently, freshness would be properly maintained only if B *does not* guarantee freshness to E. In the absence of the guarantee, E will pay D an amount commensurate with the freshness of the vegetables, and similarly D will pay C only if he is satisfied with what he receives. This arrangement, however, requires the use of the more transparent and more expensive packaging method that accommodates low-cost inspection in transit.

B would be willing to guarantee freshness to E provided that C and D merge to become C+D and the merged firm agrees to see to it that freshness would not be harmed by improper shipping methods. Such an agreement is likely here since, given the assumptions, the cost is low of ascertaining whether or not the agreement was kept. If C+D agree to maintain freshness, it presumably will simply instruct its employees to take the necessary steps to do so. The merger of C and D, then, allows B to leapfrog both and contractually guarantee freshness to E. The

level of integration is actually *lowered* then. Indeed, this may well be the purpose of the merger.<sup>21</sup>

I now turn to another point regarding the relation between vertical integration and enforcement. Consider the effect of a change in the cost of legal delineation. For example, this cost would fall, as discussed in the next section, when the cost of measuring commodity attributes falls, or when new commodity standards are formed. The ease of legally enforcing different stipulations varies. It depends, among other things, on what there is to enforce, which allows us to test the model here. I predict that a fall in this cost will induce firms to lower the degree of vertical integration, and to increase the rate at which they spin off some of their operations. This prediction is sharper than Williamson's (1975) and Klein, Crawford, and Alchian's (1978) model might yield. To them vertical integration is binary; it is either present or absent. Two firms will merge, and thus integrate presumably when the quasi rents (which, to say the least, is not readily measurable) in their interface are "significant."

Consider another question: Are a franchisor and his franchisees integrated? Echoing Cheung (1983), this question cannot be satisfactorily answered if we just try to determine whether or not they "belong" in the same firm or whether they

---

<sup>21</sup> The illustration in the text demonstrates that integration is not necessarily the result of quasi-rent capture problems.

constitute separate firms. The question can be answered by using my definition of vertical integration. This is because the definition is designed to handle the notion that the *level* of integration is not constant and depends on the fraction of the exchanges between the franchiser and the franchisee that are enforced within the operation. The existing notion of vertical integration cannot cope with such changes.

## **Conclusions**

This paper proposes a theory of organizations. It contains three major ingredients: 1. The importance of the enforcement of agreements in general, and the distinction between enforcement by the state and enforcement by long-term relations. 2. The adoption of a measurement cost framework to unify and augment various strands in the literature on the firm. 3. The claim that, on the one hand, the attempt to formulate a theory of the firm is futile as no useful definition of the firm seems to be available and thus far pinning it down has not been successful. On the other hand, the demonstration that it is possible to formulate an operational theory of organizations.

To underscore the first two ingredients consider another aspect of the formation of charitable organizations. In the United States today a small donor can find a multitude of such organizations from which to choose the ones to donate to.

Two basic features of the United States today are the high standard of accounting practices, and the accountability of fiduciary managers. When applied to charitable institutions, and especially charitable corporations, their performance can be made transparent, and thus easy to measure as a result of the former. The latter means that the level of state enforcement of managers' duties is high. In the absence of these two ingredients, people would be reluctant to donate to charitable corporations. Indeed, in most places and in most eras such corporations have not been widely observed. Instead, churches, ethnic groups and local and worker associations managed such tasks. Large would-be donors often shunned charitable organizations altogether and instead personally directed their charitable activities.<sup>22</sup>

The unifying aspect of the measurement framework is reflected in that it generates implications that apply generally, independently of particular firm models. The core dual implications are: 1. As the cost of measuring products declines (increases), we expect the scope of organizations to fall (increase). 2. As

---

<sup>22</sup> In the early days of the zionist movement in Palestine, Baron Rothschild personally financed much of the settlement activity. Later on, the zionist movement, organized as a representative and thus transparent polity was able to collect donations from a large number of individuals, eclipsing Rothschild charitable endeavor.

the cost of measuring labor inputs declines (increases), we expect the scope of organizations to increase (decline).

The focus on enforcement, and on the difference between state enforcement and enforcement by long-term relations proves to be a major ingredient in determining what will fall within organizations and what will fall across them. For instance, as the common law becomes more entrenched, we predict that a larger fraction of all interactions will be between organizations. A fall in the probity of judges will pull in the opposite direction.

Finally, whereas we cannot pin the firm down (is the Red Cross a firm?) the ownership structures of different organizations allows us to predict the lines of activities in which they have comparative advantage, and such a predictions may be readily tested.

## **References**

Alchian, Armen A., "Some Economics of Property Rights," *Il Politico*, (1965), Vol. 30, 816-29.

Alchian, Armen A. and Harold Demsetz, "Production, Information Costs, and Economic Organization," *American Economic Review* 62, (1972), No. 5: 777-95.

Alchian, Armen A, and Susan Woodward, "The Firm is Dead; Long Live the Firm," *Journal of Economic Literature* 26, (1988), No.1: 65-79.

Allen, Douglas W., "What are Transaction Costs", *Research in Law and Economics*, Vol. 14, 1991, 1-18.

Allen, Douglas W., "Transaction Costs," in Encyclopedia of law and Economics, Bouckaert, Boudewijn and De Geest, Gerrit (eds.), The Encyclopedia of Law and Economics, vol. I, 893-926, Cheltenham, Edward Elgar (2000).

Alsberg, Carl L., "Protein Content: A neglected Factor in Wheat Grades," 2 *Wheat Studies* of the Food Research Institute, 163-76 (1926).

Bajari, Patrick, and Steven Tadelis, "Incentives versus Transaction Costs: A Theory of Procurement Contracts," mimeo, October 2000.

Barzel, Yoram, "Measurement Cost and the Organization of Markets," *JLE*, April 1982.

Barzel, Yoram, "Transaction Costs: Are they Just Costs?" *Journal of Institutional and Theoretical Economics*, Vol. 141, 4-16, (1985).

Barzel, Yoram, *Economic Analysis of Property Rights*, Cambridge University Press, First Edition, 1989, Second Edition, 1997.

Barzel, Yoram, 2002, *A Theory of the State: Economic Rights, Legal Rights, and the Scope of the State*, forthcoming, Cambridge University Press.

Barzel, Yoram and Wing Suen, "Equity as a Guarantee: A Contribution to the Theory of the Firm," mimeo, University of Washington, 1995.

Cheung, Steven N.S., "The Structure of the Contract and the Theory of Non-Exclusive Resources," *Journal of Law and Economics*, Vol. 13, pp. 49-70 (1970).

Cheung, Steven N. S., 1983 "The Contractual nature of the Firm," *Journal of Law and Economics*, April, Vol. 26, pp. 1-22.

Coase, Ronald H. (1937) "The Nature of the Firm," *Economica* 4, no. 3: 386-405.

Coase, Ronald H. (1960), "The Problem of Social Cost," *Journal of Law and Economics* (October).

Coase, Ronald H. (2000). "The acquisition of the Fisher body by general Motors," *Journal of Law and Economics* (April) 43: 15-32.

Demsetz, Harold, 1988. "A Framework for the Study of Ownership," in *Ownership, Control, and the Firm*. Oxford: Basil Blackwell.

Demsetz, Harold, 1998. "Review: Oliver Hart, 'Firms, Contracts, and Financial Structure'," *Journal of Political Economy* 106: 446-452.

Dittmer, Timothy

Grossman, Sanford, and Oliver Hart, 1986. "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy* 94: 691-719.f

Holmstrom, Bengt, 1999. "The Firm as a Subeconomy," *Journal of Law, Economics & Organization* 15: 74-102.

Holmstrom Bengt, Paul Milgrom (1991) “Multitask Principal-Agent Analysis: incentives Contracts, asset ownership and job design”, *Journal of Law, Economics and Organisation*, 7, 24-52.

Holmstrom, Bengt, and Paul Milgrom, 1994. “The Firm as an Incentive System,” *American Economic Review*, 84: No. 4, 972-991.

Jensen, Michael C. and William H. Meckling, 1976, “Theory of the Firm  
Managerial Behavior, Agency Costs and Ownership Structure,” *Journal of Financial Economics* 3, No. 3: 305-60.

Klein, Benjamin, Robert C. Crawford and Armen A. Alchian, 1978 “Vertical Integration, Appropriable Rents, and the Competitive Contracting Process,” *Journal of Law and Economics*, October: 21, 297-326.

Klein, Benjamin, (2000). “Fisher-General Motors and the Nature of the Firm,” *Journal of Law and Economics*, (April) 43: 105-142.

Masten, Scott E. “Contractual Choice,” in *Encyclopedia of Law and Economics*, B. Boukaert & G. De Geest (eds.), Edward Elgar Publishing and the University of Ghent, 1999.

McManus, John. 1975. "The Cost of Alternative Economic Organization," *The Canadian Journal of Economics*: 8, 334-350.

Rock, Edward, and Michael Wachter. 2001. "Islands of Conscious power: law, Norms and Self Governing corporations," Forthcoming in: Symposium, Norms and corporate Law, University of Pennsylvania Law Review, Spring 2001.

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.