VALUE CREATION BY FIRMS

by

P. MORAN*
and
S. GHOSHAL**

97/19/SM

* Phd Candidate at INSEAD, Boulevard de Constance, Fontainebleau 77305 Cedex, France.

** Robert P. Bauman Professor of Strategic Leadership at London Business School, Sussex Place, Regent's Park, London NW1 4SA, England.

A working paper in the INSEAD Working Paper Series is intended as a means whereby a faculty researcher's thoughts and findings may be communicated to interested readers. The paper should be considered preliminary in nature and may require revision.

Printed at INSEAD, Fontainebleau, France.

VALUE CREATION BY FIRMS

Peter Moran

INSEAD

Boulevard de Constance

77305 Fontainebleau, France

Tel: 33-1-60.72.40.00
Fax: 33.1-60.72.42.42

Sumantra Ghoshal

London Business School

Sussex Place, Regent's Park

London NW1 4SA, U.K.

Tel: 44-171-262-5050
Fax: 44-171-724-7875

Preliminary draft. Please do not quote or cite without written permission from the authors.

Comments and suggestions are welcome and may be addressed to either author.

VALUE CREATION BY FIRMS

Abstract

This paper outlines a theory for how organizations, in general, and business firms, in particular, create economic value for themselves, for their members and for society. Drawing on the ideas of Schumpeter (1934), we develop a framework to describe value creation as consisting of the processes of (a) resource combination and (b) exchange, and we use the framework to show how firms can create value from a society's stock of resources beyond what markets alone can. The theory also offers an explanation of why neither a market nor a firm, by itself, can achieve adaptive efficiency (North, 1990), and why institutional pluralism - in the form of a variety of firms of many different forms and sizes, coexisting in a state of creative tension in a broader institutional matrix - is necessary for the healthy progress of societies.

VALUE CREATION BY FIRMS

As Herbert Simon (1991) pointed out, it would be hard not to notice the "ubiquity of organizations" in modern societies, which he aptly referred to as "organizational economies." Another feature of the socio-economic landscape noted by Simon is perhaps less obvious than the ubiquity of organizations but is just as impossible to dismiss once attention is drawn to it. It is the coincidence in the number and size of firms with the magnitude of market transactions. If, as Simon colorfully put it, firms appear as solid green areas and market transactions show up as red lines, to a mythical Martian visitor (able to view Earth with a telescope that reveals patterns of economic activity) observing "central Africa, or the more rural portions of China or India, the green areas would be much smaller, and there would be large spaces inhabited by the little black dots we know as families and villages. But the red lines would be fainter and sparser in this case, too, because the black dots would be close to self-sufficiency, and only partially immersed in markets" (p. 28).

Were our visitor also able to measure economic activity, it could not help but notice the positive relationship between this coincidence of market-based and organization-based exchange and the economic prosperity (however measured) of the regions where this coincidence is most pronounced. A report sent home by the Martian would likely characterize the Earth's most prosperous and wealthiest regions and nations as comprising the combination of both our most advanced markets together with our healthiest and most productive firms. Few would be surprised or take issue with the inference, from this strong positive association between the wealth of nations and the relative proportion of successful firms operating in those nations, that society and firms can coexist and prosper alongside one another.

Although certainly remarkable, there is little that would be controversial in a characterization of our organizational economy as comprising both intensely competitive markets and healthy firms, coexisting in a constant state of vigorous but creative tension with one another; an evolving state of continuous interaction between and among firms, on the one hand, creating

and realizing new value, and markets, on the other, relentlessly forcing these same firms to surrender, over time, most of this value to others and, as a result, to never letup their relentless search for new ways to create and realize value, as part of their continuing struggle to remain viable and healthy. Indeed, this is the essence of the process of creative destruction that Schumpeter theorised and wrote about over a half century ago.

Despite the self-evident importance of wealth creation to social and economic progress and its obvious dependence on the conduct of organizations, in general, and of business organizations, in particular, there is little in the way of theory development, since the work of Joseph Schumpeter (1934), that relates the role of firms to the creation of wealth. As a result, the interplay that characterizes the state of tension in which individuals, organizations and markets interact and evolve to shape each others' behavior is poorly understood. In particular, the process through which this tension results in both creative and coercive forces that direct the attention and efforts of individuals in deciding what resources to deploy and how to deploy them is largely unexplored. This paper aims to fill this void, in that it (i) provides the outline of a theory that seeks to explain the institutional role of organizations, in general, and business firms, in particular, in creating value for their members and for society at large and, (ii) facilitates the exploration and understanding of how both firms and markets contribute collectively to a process of economic development that is both purposive and dynamically efficient (Ghoshal and Moran, 1996).

To develop a theory of value creation by firms, it is useful to first characterize the value creation process. In the next two sections of the paper we develop a framework that builds on Schumpeter's arguments to show (a) that new resource combinations are the source of new potential value to be created (section II) and (b) that exchange accounts for the actual realization of this potential value while simultaneously setting up the stage for the next round of resource combinations and, thus, for the next iteration of the process (section III). In other words, resource combination and exchange lie at the heart of the value creation process and in sections II and III we both describe how this process functions and also identify the conditions that facilitate and impede each of these two elements of the process.

This framework establishes the theoretical infrastructure for the analysis of the roles firms can play in this value creation process and of how both firms and markets collectively influence the process of economic development. As our analysis of the requirements for effective resource combination and exchange reveals, markets alone are able to create only a very small fraction of the value that can be created out of the stock of resources available in society. Because of their very different institutional nature and context, firms, operating in a state of creative tension with markets, substantially enhance the fraction of the total potential value that can be obtained out of society's resource endowments. In section IV, we describe this process of value creation by firms and, in section V, we integrate the firm's role with that of markets to explain why both firms and markets are needed to ensure that economies develop and progress in a way that achieves what Douglass North (1990) has described as "adaptive efficiency."[1]

RESOURCE COMBINATION AS THE SOURCE OF NEW POTENTIAL VALUE

Schumpeter defined the concept of "economic development" as "the carrying out of new combinations" (1934: 66) - that is, "to produce other things, or the same things by a different method [or to] combine these materials and forces differently" (p. 65). Such combinations represent "simply the different employment of the economic system's existing supplies of productive means" (p. 68). By implication, each time a new resource combination is carried out, some resources are withdrawn from means that are already productive, perhaps even <u>more</u> productive, initially, than the new combination toward which the resources are diverted. Consequently, even when the new combination represents the source of new <u>potential</u> value, while that <u>potential</u> for new value is being created (or destroyed),[2] an innovative resource deployment can also (indeed, is likely to, at least temporarily) actually <u>reduce</u> the level of value that is <u>realized</u>, at least for some period of time.[3]

Following Schumpeter, we use the term "new combinations" to mean those deployments of resources that constitute the "new combinations" which, according to Schumpeter, constitutes the source of economic development. Each time resources are deployed in making new

combinations, a new source of "potential value" is created and added to the economic system. As did Schumpeter, we also acknowledge the value derived from the process of "handing on" that occurs as others join the bandwagon in reproducing these new combinations. However, whereas Schumpeter explicitly excluded this process of routinization from his concept of economic development (1934: 66), we merely distinguish it from the process of potential value creation and refer to it as part of the process of realizing the potential value that is created by new combinations. Even though "new combinations" lead the charge of economic development, by blazing new, previously uncharted value realization trails, economic development, for our purposes, is observed only upon the "realization" of this potential value; without such realization, there is no economic progress. In other words, new potential value can be created only through new combinations. But, by themselves, new combinations can add no wealth to society. For wealth to increase we need some potential value to be exploited; that is, value must be generated, appropriated and eventually "handed-on.". We refer to this process of generating and appropriating wealth as "value realization." For an economic system to be in dynamic balance, both processes of "value creation" and "value realization" are needed. As we have defined it, no economic progress is possible without some "value realization." However, even without any "value creation," at least some development (i.e., wealth generation) is always possible in the short-term, i.e., until all resources find there way to some condition of Pareto optimality. Consequently, a system that is focused exclusively on "value realization" is likely to achieve greater "allocative efficiency" (North, 1990) at any given point in time than one that allows for value creation. Such a system, however, is also likely to suffer from a lower level of "adaptive efficiency."

Like Schumpeter, we also explicitly acknowledge that many resource deployments, including combinations, represent often mutually exclusive alternatives to other deployments of the same resources. When the decision to combine resources in new ways is taken, other possible deployments, including other new as well as old combinations are often foregone.[4] The nature of the potential value creation on which such economic development rests is influenced by the forces which determine which of the many possible resource combinations are made, in favor of alternative resource deployments, and those combinations/deployments

which are not made, because of the perceived value of alternative deployments. Obviously, any combination of resources that is expected to enhance potential value should be made. Yet at the same time, it should also be obvious that nowhere near all potential value creating combinations that are possible, for a given set of resources and a given set of preferences at any given time, are ever actually made. Moreover, many combinations that are made may actually destroy more potential value than they create.

In order to explore the value creation process in some detail we will first set up a general model of all resource deployments, of which new value creating combinations compose a subset. We define three conditions as necessary conditions that must be satisfied before any voluntary resource deployment (including "new combinations") is likely to be executed. Before any specific resource deployment is likely to occur voluntarily: (i) there must first exist some opportunity or means to make the deployment - i.e., the resource(s) (i.e., knowledge, good or service) to be deployed and the opportunity to deploy it (them) must exist and be available for such deployment; (ii) the party (or parties) with the opportunity or means to execute the prospective resource deployment must be motivated to make the deployment and (iii) this same party (or parties) must also perceive the opportunity and expect or otherwise hope for some value to be realized from the deployment. It is important to emphasize that although many resource deployments occur by accident (i.e., without meeting any of these three conditions) and, perhaps, many more are executed solely on the basis of some faulty perception of an opportunity for value realization (i.e., failure to meet either the first or the last condition), all three conditions must be met before any purposive action can lead to the realization of any potential value. Satisfaction of all three conditions does not ensure the realization of value, however.

For the moment, little need be said of the first of these conditions (i.e., the existence and availability of some stock of resources that would increase in value through deployment). Surely, the importance of the possession of or access to certain resources or to the rights to deploy them in certain ways is a major constraint to the execution of a very large set of potential value enhancing resource deployments. As we argue in the following section,

exchange or the transfer of resources and resource rights among individuals and groups is the primary mechanism by which this constraint is overcome but, as we suggest below, access to resources or resource rights is not the only limitation to the execution of beneficial deployments. It would be incorrect to assume, for instance, that absent any need for exchange or transfer, all beneficial value realizing resource deployments are sufficiently motivated to occur without additional help. Even many potential value realizing deployments of those resources that are all controlled by a single individual are often not made because one or both of the remaining conditions necessary for value realization are not satisfied.

That motivation is necessary for any purposive action to occur voluntarily is self evident. However, even the presence of some opportunity to deploy resources in a way that will create additional potential value or lead to its realization (i.e., satisfaction of the first condition for resource deployment) together with the recognition of the opportunity (i.e., satisfaction of the second condition) are not enough to ensure sufficient motivation for that value creating or realizing deployment to be executed. To be fully motivated, an opportunity must be associated with satisfactory means to appropriate some value from the deployment. That is, some party must not only perceive some potential value that could be derived from a particular resource deployment or combination, it must also have reason to believe that it, or some confederate party with whom the resource deployer identifies, will appropriate some of the value to be realized from such deployment. Appropriability requires the rights to deploy the resource(s) in a certain way and the rights and the means to realize value from such deployment. These rights include the rights: (i) to deploy the resource in a certain way - e.g., to consume it, combine it with or exchange it for other resource(s) or otherwise use it; (ii) to appropriate value from the deployment; and (iii) to limit access to or use of the resource by others.

While satisfaction of all three of these criteria (i.e., the opportunity for deployment and its motivation and preception) may not in itself be sufficient for all types of resource deployments, if satisfied adequately they go a long way toward overcoming most if not all

common hurdles to most potential value creating and value realizing resource deployments. The nature and extent to which knowledge is acquired, shared and used largely determines what resources are involved in the development process and how they are deployed. Given satisfaction of the first criterion (i.e., the existence of a resource and an opportunity to deploy it productively, relative to other opportunities for deployment), the satisfaction of the remaining two criteria is likely to be impeded or, worse, distorted by uncertainty. Uncertainty surrounds the nature of or awareness of the existence of many beneficial deployment opportunities and of all their relevant associated effects, including the returns from such deployments.

Uncertainty often exists to varying degrees in terms of the uncertain underlying value of the deployment itself or of the uncertain likelihood of the deployer benefiting from the deployment. Information costs are incurred in defining, maintaining and exercising the rights to the resource(s), as well as in perceiving of the opportunity. The greater these costs, the less a given deployment is likely to be executed and if excessive, these costs can even deter deployment opportunities with huge potential returns. This is especially true of opportunities like "new combinations" for which the likelihood of value realization is highly uncertain because the realization is dependent upon future uncertain deployments. Note, that <u>these costs affect the *deployment* of resources</u> (whether or not any exchange is involved) <u>and are incurred independent of any *transaction* taking place</u>. In fact, it may be true that some additional transaction (and its associated cost) must be incurred to lower the cost of a particular deployment. Hence, they have the potential to influence the exchange of the resources in ways that transaction costs alone may not.

We are now ready to consider that subset of all resource deployments that create new potential value - i.e., new combinations (see Figure 1). Like any other resource deployment, all potential value creating new combinations must satisfy the three conditions necessary for any deployment. That is, before any potential value creating new combination is likely to be made, the opportunity for such combinations must exist and the opportunity to realize value from the combination must be motivated and perceived. However, unlike other resource

deployments which lead to the immediate realization of value, the opportunity to realize value from new combinations comes not from the new combination itself but from the expectation of subsequent deployments (i.e., uses) of the newly combined resource(s). In fact, most potential value creating "new combinations" are made with some degree of sacrifice (viz., opportunity cost) to realized value.

- Figure 1 about here -

To more easily conceptualize how the resource deployment process that is represented by the model depicted in Figure 1 might differentially favor certain types of resource deployments over others (viz., value creating vs. value realizing), consider the n-dimensional matrix of resources that would be created if a vector of all (n) resources that existed at any one time in a system were crossed with itself n-times. The matrix that would result represents all potential combinations (i.e., from all possible combinations and permutations), including all new combinations as well as all recombinations of resources that could be possible in a world of no constraints on the availability or use of any resource in this set. The first box, on the left-hand side of Figure 1 - labelled "resource stocks" (all potential deployments) - includes all the elements of this matrix (i.e., the crossing of the vector of all resources with itself n-times). As such, it contains the set of all potential resource combinations that are implied by any given stock of resources in a system and serves as the first of three independent variables in our model. The box on the right-hand side of Figure 1 - i.e., labelled "resource flows (deployments executed)" - is the dependent variable. It is the set of all resource deployments, including combinations, and that occur as a result of purposive action. Now imagine the two mutually exclusive and exhaustive subsets of resource combinations that would be created (i.e., hypothetically) if each and every resource combination were somehow classifiable as value creating or value destroying. In a hypothetical world of hyper-rationality and no constraints on resource accessibility or deployability, the dependent variable would be identical to the first subset of all potential value creating combinations. That is, given purposive action with neither resource constraints nor rationality limits, <u>all possible resource combinations</u> that would create system-wide value could be expected to occur without fail.

Of course, in the real world, where people are "intendedly rational, but only limitedly so" (Simon, 1957, [1945]: xxiv), the set of executed combinations will differ substantially from the hypothetical ideal. First, only a very tiny fraction of the subset of potentially value adding combinations will actually wind up executed because most combination opportunities will either not be fully motivated or even perceived. That is, of those value adding combinations that are possible and that would be beneficial, given the capabilities and tastes of the parties involved, many will not occur either because the parties in control of the requisite resources are not in a position to benefit themselves from making the combination or because they do not see the opportunity or its value to them. Furthermore, not only will the set of executed combinations be much smaller than the value creating combinations that are possible, it will also include many combinations from the second - value destroying - subset of combinations. This is largely the result of distorted perceptions of what constitutes a value creating combination (see footnote # 2). Note also that, given bounded rationality, this difference from the hypothetical ideal would still be substantial, even if each and every individual in the system had unrestricted access to all resources and, therefore, had no need for exchange. We will return to this point later when we discuss the dual role of exchange.

Each time resources are newly combined or combined in new ways, the combination itself represents a new resource in the system. This new resource then becomes available for recombination with other resources in the system. As the accumulating stock of resources grows, the number of potential combinations that are possible quickly overwhelms any individual's or group's ability to consider them. Selection processes, that determine which combinations are considered and how they are screened, then take on added importance. As was already suggested, in the hypothetical world of no resource constraints or rationality limits, all value creating combinations could be expected to occur without fail. As constraints begin to limit the number of combinations that can be considered, an ability to discriminate higher potential value yielding combinations from lesser ones becomes necessary. Unfortunately, the criterion of efficiency does not enhance such discriminating ability. In fact, as we will further elaborate in section IV, the forces of efficiency, as well as the forces of salience and satisficing, are biased <u>against</u> such discriminating ability. Whereas

the number of deployment opportunities available for consideration can be expected to grow exponentially with the number of new combinations made, the number of deployments executed is likely to grow at a much slower pace. Unless the constraints that limit deployments fall at a rate that is equal to or faster than the rate at which deployment opportunities are expanded (which is highly unlikely), the already much larger number of opportunities foregone can be expected to grow faster than the number of deployment opportunities that can even be entertained, let alone be executed. As the opportunity for value realizing deployments grows, so too will the opportunity cost for many new potential value creating new combinations. Limited attention is likely to favor consideration of a generally increasing proportion of value realizing deployments relative to value creating new combinations, as well as consideration of those deployments and combinations involving resources that were themselves most recently deployed or created from past combinations and to disfavor deployments where appropriability is weak or unlikely (Teece, 1986).

Obviously, the resource deployment process does not take place in a closed system, influenced only by the variables shown in the model but is also affected by the context that surrounds the resource deployment decision. Context affects the physical, legal, and cognitive access to certain resources themselves, and to the rights to deploy them in certain ways. The role of context in shaping the resource deployment process is crucial for understanding how organizations, in general, and firms, more specifically, can achieve value creating resource combinations in a way that markets alone cannot (Ghoshal, Moran, and Almeida Costa, 1995) and, therefore, will feature prominently in our analysis in Section IV on the value creating role of firms.

THE DUAL ROLE OF EXCHANGE

More than fifty years ago Friedrich Hayek aptly observed, "practically every individual has some advantages over all others in that he possesses unique information of which beneficial use might be made, but of which use can be made only if the decisions depending on it are

left to him or are made with his active cooperation" (1945: 522). Indeed, Hayek saw the problem of ensuring the effective use of this "local" knowledge - i.e., knowledge that is held by all individuals but held only in fragmented, incomplete and often contradictory and continually changing bits - as "the economic problem of society;" which he succinctly summarized as, "not merely a problem of how to allocate "given" resources - if "given" is taken to mean given to a single mind .... It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, it is a problem of the utilization of knowledge not given to anyone in its totality" (Hayek, 1945: 519 - 520).

The solution that Hayek advocated was decentralization. "We need decentralization because only thus can we ensure that the knowledge of the particular circumstances of time and place will be promptly used" (1945: 524). Unfortunately, even the complete decentralization of all resource deployment decisions (i.e., decentralized access to resources and the rights to decide how they are deployed) does not ensure the full use of all knowledge that could be used beneficially. As we argued in the previous section, individuals must also be likely to appropriate the benefits from the deployment and must perceive some opportunity to benefit. Of the set of beneficial deployment opportunities available to each individual, only a fraction meet all three criteria. Moreover, there is no compelling reason to believe that the best opportunities for realizing value system-wide (or even for the individual) reside in this small subset of resource deployments that actors are likely to focus on.

Although Hayek stressed the "marvel" with which a system of prices economically communicates the essence of the information that is necessary for actors to adjust their resource deployment decisions, it is the <u>exchange</u> of those resources (which prices facilitate) that undergrids the contribution of the system toward economic development. That prices reflect real value at all, is only because they are continually transmitted and kept up-to-date through exchange. But whether or not prices are available (or even exist at all), exchange promotes the use of local knowledge and, thereby, is a major determinant in both the realization and the creation of potential value in most modern economies.

Thus far, we have argued, as did Schumpeter long ago, that new resource combinations are the source of new potential value in any economic system. This additional potential value comes at some sacrifice (at least in terms of opportunity cost) of realized value foregone and the new value can only be realized to the extent that the new combination itself is subsequently deployed in the future. Moreover, such realized value shows up as wealth enhancing economic value added to the economic system, only to the extent that it becomes part of the fabric of the economic system itself, as reflected in the system's enhanced social productivity. Exchange is the primary mechanism through which this potential value becomes realized. As such, exchange validates the value of the resources exchanged and, thereby, promotes, and sends a signal of the realization of some of the potential value that was previously created by new combinations. But exchange does more than just promote the realization of value and validate the process. It also assigns a value to and reorders the set of resources that are available for new combinations. Hence, exchange also influences the way in which resources are deployed and the path taken in creating value.

In this section we explore the dual role that exchange plays in economic development and argue that the forces that drive exchange necessarily (in the presence of transaction costs) exert a meaningful bias, which can as easily set a course that is inefficient as one that is efficient over time; thus setting the stage for the subsequent argument that no single institution, be it market or firm, is as likely to be capable of charting the most efficient course over time as an institutional matrix comprising many varied institutions.

Exchange is the Catalyst for Value Realization

The potential value that is created through new resource combinations cannot yet be considered as realized value added to the economic system. It exists only as intrinsic value. While such combinations may be (indeed, are likely to be) valued by the party or parties responsible for the new linkage, this potential value lies dormant or unrealized and does not become economic value unless and until it becomes realized through subsequent deployments that routinely make use of such new resource combinations. These deployments, aimed at realizing some of this newly created potential value, generally cannot

be executed without some form of prior exchange to acquire the enhanced resource to deploy and, thus, make the deployment possible.

It is useful at this point to distinguish among several types of value. We distinguish "economic value" from more intrinsic forms of value. Individuals may value anything at all - an object, good or service; knowledge, another person or group; or even an idea, thought or emotion - and many individuals may even assign similar values to the same things. However, even though intrinsic value is real and can often be inferred from behavior by the apparent weight or priority a party assigns to certain things relative to others, intrinsic value has no economic value unless and until such value is reflected in exchange (i.e., the transfer or sharing of this object, knowledge, idea, etc. to others). Exchange is the fundamental mechanism through which economic value is realized in any economic system.[5] Each time ideas, goods or services are exchanged, either by themselves or in combination with others, some of the potential economic value that was created (destroyed) by prior resource deployments is realized and added to the system.

This important role of exchange, to facilitate the continual reallocation of resources to their more productive uses and, thereby increase their social productivity, receives a great deal of attention in the field of economics. By encouraging actors to use resources productively, exchange contributes to economic development. The most productive use of resources at one point in time is not necessarily the most productive long-run use of those same resources, however. Another role of exchange, one receiving far less attention, helps to provide a continuing supply of new productive uses for these resources. By reordering the stock of resources that are available for redeployment at any one time, exchange makes it difficult (if not impossible) to ensure the allocation of all resources to their best current use and, thereby, makes the discovery of new (potentially more productive) uses more likely. In this role, exchange influences the nature and extent of the potential value that is created in the first place and, thereby, characterizes the ultimate paths of value realization for the economic system as a whole.

Exchange Sets the Stage for Value Creation

Exchange is not only motivated by past resource combinations and other deployments (e.g., through its effect on prices) but it also motivates new deployments of all kinds. As already suggested, exchange is often necessary and common when one wishes to combine or otherwise deploy resources with restricted rights (e.g., when one does not have the resources or the right to combine/deploy them as desired or can not appropriate the value realized by such combination/deployment). Exchange can also influence those deployments for which exchange itself is unnecessary (e.g., because the resources for deployment are available and the deployment is motivated) by enhancing one's perception of the deployment or its value. Hence, besides being the principal enabler of value realization, exchange can also be an important architect of value creation. It is so to the extent it gives rise to new unanticipated resource combinations that were previously not possible (without the resources), not perceived or not motivated (rights or means to appropriate value were missing).

Recall from our discussion of Figure 1 how few of the potential value creating combinations (that are possible even without the need for exchange) would actually take place because of limitations in any single party's ability to perceive the value of resource combination opportunities. For example, when a single party, like a sovereign or king, has unlimited access to all resources and resource use rights (and no need for exchange), that party is motivated to execute, indeed, encourage the execution of all of the value creating/realizing combinations that are possible. All such potential value adding combinations do not occur, however, because the party controlling the resources (e.g., the sovereign) does not perceive all the opportunities to create value. If, however, resources and resource use rights were distributed to other parties and exchange were possible, more local knowledge would be exploited to expand the set of value adding resource combinations that would be attractive. Here, exchange greatly expands not only the set of potential value creating resource combining opportunities that are seen as attractive but also the number of these that are executed and exploited in subsequent value realizing exchanges.

Summarizing then, as the principal mechanism through which value is realized by and, thereby, added to any economic system, exchange is the primary conduit for building wealth. This is true whether the economic system we are concerned with is a small group, a nation or the global economy as a whole. Each time two or more parties enter into voluntary exchange, their constituent stock of resources is altered and the opportunity/need for new value realizing resource combinations/deployments is created. Often, wealth increases for the parties involved. When the costs incurred as externalities by others (i.e., those not directly involved in the exchange) are less than the net value added enjoyed by the transacting parties, wealth increases for the system itself. Over time, the cumulative process of exchange accounts for all economic development. While the rate at which development proceeds has typically been the primary concern of those interested in development, the path that development takes is also influenced by exchange. That is, the nature of the evolving pattern of exchange also influences the determination of which values, inherent in any economic system's universe of resources, are realized over time and which are inhibited from realization. This includes, not only which technological trajectories are favored over others but also the biases which favor the development of certain social and economic classes of people and geographic regions and not others.

Social welfare is maximized for any economic system when all potential exchanges that are mutually beneficial to all exchange parties are actually consummated and the resource combinations/deployments that are made possible by such exchanges are executed. Unfortunately, just as only a few of the resource combinations that could create value are ever conceived or executed, only a tiny fraction of all potential exchanges that would benefit all parties at the time of exchange ever do, in fact, occur. Because exchange is a prerequisite for most resource combinations, each missed opportunity also serves as a potential roadblock to all other exchanges that the unconsummated exchange would have otherwise made mutually beneficial to some parties, had it occurred. The economic system, then, at any given time falls far short of realizing the full potential value that exists in the resources that are available for exchange by all parties at that time. All of the value that is realized, however, comes about through exchange.

Forces Driving Exchange

Regardless of whether a given exchange is about to lead to the realization of existing potential value or the creation of new potential, the forces motivating exchange are no different than those motivating any other resource deployment decision. Like any other resource deployments (including new combinations and consumption), exchange requires that all three necessary conditions for deployment that we stipulated earlier be satisfied. That is, some potential opportunity for exchange must exist, and that opportunity (or some other) must also be perceived and be motivated. In addition, since exchange, by definition, requires more than one party, an additional and potentially demanding condition is added, viz., that the three conditions must be satisfied by all parties whose resources are to be included in the exchange. This additional condition is commonly referred to as the "double coincidence."

In a hypothetical world of no transaction costs each opportunity or need for exchange would independently provide all the motivation that is necessary for any exchange (of scarce goods) to take place, regardless of the type of deployment that is likely to follow. That is, absent transaction costs, exchange would follow the path of greatest value realization. However, in the real world of positive transaction costs, we are not so fortunate. Transaction costs are likely to manifest themselves by distorting an exchange opportunity's appropriability and/or one's perception of its appropriability or perception of the opportunity itself. The added constraint imposed by the "double coincidence," is likely to accrue potentially large additional transaction costs in locating suitable exchange partners and in assessing the availability and usefulness of the resources they have to exchange.

Like bounded rationality, indeed, in no small measure, because of it, transaction costs are an ever present part of the world we live in. Given the scarcity of human attention and the ubiquity of transaction costs, we are forced to focus that attention over a limited scope of activities in order to accomplish most of the things that we do. One consequence of attention focused on a few things, is a relative lack of attention on all other areas. Over centuries, institutions like markets and organizations have evolved to help us cope with the ubiquity of transaction costs. By collectively establishing (largely through our institutions) routine ways

of dealing with common activities, we have enabled ourselves to engage in a large number and variety of exchanges without the need to focus on every aspect of every single exchange. Institutions provide the focus for us, only in routine ways. While the benefits of our habits of routines are many, they are accompanied by a significant cost associated with changing those routines. The advantage of any single institution, whether it is a decentralized system of prices, a centralized authority system or some other set of complementary conventions and norms, is its ability to focus on certain activities, while ignoring others. The disadvantage is the cost in overcoming that focus to do other things.

Institutions Facilitate Exchange but Limit its Scope

To explore these limitations and the consequential biases of single institutions, consider the Venn diagram illustrated in Figure 2. The universe, represented by the large rectangle that contains the smaller circles in Figure 2, is the set of all possible resource deployments (including combinations) that exists at any one time in any economic system, given its distribution of resources (including technologies). Circle A represents all potential resource deployments that any party is motivated to execute, given the distribution of resources, resource rights, and individual preferences that exist in the system at that time. As such, it corresponds to the variable "opportunities motivated" in Figure 1 and includes all deployments that would be considered efficient at that instant in time (i.e., no single deployment outside of Circle A would improve the welfare of one more actor without also decreasing the welfare of another), as well as many deployments that are not efficient (i.e., even though all deployments inside Circle A would improve the welfare of the party or parties responsible, some deployments in Circle A would also decrease the welfare of at least one other actor).

- Figure 2 about here -

It is important to emphasize that many resource deployments falling outside of Circle A may hold great promise for future wealth generation, despite the fact that they may not represent the highest valued use for those particular resources at that time (i.e., these deployments are

inefficient), as the most efficient uses for these resources are specified by the current distribution of resources, resource rights and individual preferences. As Schumpeter pointed out, "A system - any system, economic or other - that at *every* given point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at *no* given point of time, because the latter's failure to do so may be a condition for the level or speed of long-run performance" (1942: 83, emphasis in original). Moreover, any change in the current distribution of resources, rights or preferences can easily change the composition of Circle A, i.e., motivating some deployments that currently reside outside and demotivating others that reside inside this set (Circle A). Indeed, a premise of our argument is that at least <u>some</u> of these less efficient deployments (i.e., residing outside of Circle A) would lead to greater long-term economic development. That is, total long-run social welfare will not be optimized if <u>only</u> those resource deployments represented by Circle A are executed and all deployments outside of circle A are systematically avoided. However, since we have defined Circle A to contain the total set of all motivated deployments, the only way any deployments outside of this area are likely to be executed is if resources, rights or individual perceptions or preferences get redistributed to motivate the opportunity (i.e., bring it into Circle A) or if the opportunity is misperceived as beneficial (to the deployer(s)) when it is not. Hence, to be efficient over the long-run, the system must be capable of responding to promising new deployment opportunities by somehow getting them motivated (i.e., inside of Circle A); that is, the system must be capable of shifting Circle A across the matrix of deployment opportunities to motivate the most promising among them.

Circle B represents all perceptions (including misperceptions) of beneficial (i.e., value creating or realizing) resource deployments that exist in the system at that time, given the existing distribution of resources, rights to deploy resources and to realize value from such deployments, and individual preferences. It corresponds to the variable "opportunities perceived" in Figure 1 and represents the total sum of all system-wide opportunities perceived at that time.

The intersection of circles A and B ($A \cap B$) represents all resource deployments for which each of the three necessary conditions of deployment are satisfied by at least one (but not necessarily the same) actor. In the model shown in Figure 1, $A \cap B$ corresponds to the total value that would be created/realized if all deployments that are motivated somewhere and perceived by someone in the system were to occur. The extent to which these deployments are likely to be executed depends upon the degree to which there is a one-to-one correspondence between the actor(s) who control the resource(s) required for each deployment opportunity, those who would benefit from such deployment and those who perceive the opportunity. Of course, because so many resources are scarce, most of the deployments that constitute the set $A \cap B$ are unlikely to satisfy this condition without at least some exchange.

Circle C, the largest circle in the diagram, separates out, conceptually, all these resource deployments that require some transfer of resource rights (inside Circle C) from all those that do not (outside Circle C). All deployments included in Circle C require exchange (or some other form of transfer) to link up the resources or their control with those who would benefit from and those who perceive the opportunity. Hence, that part of the intersection of circles A and B that lies <u>outside</u> of Circle C represents the resource deployments that individual actors are most likely to carry out, unencumbered by any need for exchange. Only the potential value from this tiny (relative to the rest of the part of $A \cap B$ that is also entirely within C) subset of deployments ($(A \cap B) - (A \cap B) \cap C$) can be expected to be created and realized without any exchange (or transfer) of resources, rights or perceptions of their usefulness.[6] All other deployment opportunities falling within $A \cap B$ (i.e., $(A \cap B) \cap C$) are constrained by the need for exchange.

For exchange to occur, the process that is depicted at the individual level in Figure 1 must occur for all prospective parties to the exchange. In other words, a "double coincidence" (or, more generally, a "multiple coincidence") of opportunity perception and opportunity motivation must take place for both (all) parties whose resource needs to be exchanged before the perceived resource deployments can occur. This "double coincidence" need not

occur simultaneously, however. Often the perception and motivation of one actor is all that is needed to induce the necessary perceptions and motivations in others.

Returning to our Venn diagram in Figure 2, those exchanges requiring resource deployments for which the double coincidence is satisfied is represented by the area in the smallest circle in the diagram, Circle D, which straddles the intersection of circles A and B and lies entirely within Circles B and C.[7] Circle D, then, represents the portion of all possible resource deployments that are most likely to occur through market exchange.

The Limited Scope of Market Exchange

Now that Figure 2 is specified, we can use it to more easily appreciate the nature and magnitude of the remaining problem left unaddressed by a system of only independent market exchange. The universe of all possible resource deployments (represented by the entire area within the large rectangle) represents all deployments that could conceivably be executed if all local knowledge were fully exploited - i.e., given the particular distribution of resources (including technologies), rights and individual tastes at a particular point in time. The subset of those resource deployments that are facilitated by markets and a system of prices is represented, in contrast, by the smallest circle "D" in the diagram. This is the set of resource deployments that satisfies all three of the necessary conditions in the deployment process plus the additional restrictive condition of the "double coincidence" which must exist before market exchanges can occur.[8]

Consider how limited the scope of market based exchange is, in terms of the proportion of potential value adding resource deployments that are supported by such exchange and the reasons for such limitations. As shown in Figure 2, the set of resource deployments that are most likely to be supported by market exchange (i.e., Circle D) excludes all but a tiny fraction of the resource deployments that could be made and that might otherwise add great value to the system. The nature of these deployments that are most likely to be excluded from the support of market based exchange can be classified into two general categories.

First, there are the potentially large portion of those value adding deployments for which all three conditions for deployment are met except for the "double coincidence" - i.e., the area within $A \cap B$ that is also outside of Circle D. This exclusion is likely to persist because the actors with the resources, ideas or rights needed to execute or benefit from a specific deployment are not sufficiently linked for each to satisfy all three conditions necessary for deployment and the necessary links are too costly to establish. We refer to this failure to meet the "double coincidence" as missing markets. Markets may be missing (or more precisely, incomplete) because even though the opportunity or need for deployment is recognized, the market conditions or conventions necessary for exchange are incomplete or missing entirely. These can include exchanges for which pricing is difficult, money is inappropriate, rights are unclear, inadequately specified or are not adequately protected or enforced by law, and so on. Although these deployments are desirable and would benefit the prospective parties (given their current distribution of resources rights and preferences), they are unlikely to occur because transaction cost barriers are just too high to overcome. The nature of these "market failures" stems in part from the conservative standards that must be applied by most markets in establishing exchange viability and not from the distribution of resources and rights per se. This is why money and credit, which do not seek to redistribute resource rights, can do much to overcome this type of constraint.

The second class of deployments that are systematically discouraged by markets are of a kind that are unavoidable in the presence of transaction costs; that is, when transaction costs are high a set of transactions is always discouraged by any single institution, whether a market or a firm. In the presence of transaction costs, any institution that induces behavior through a system of incentives (e.g., by allocating resources, assigning rights and restricting access) encourages the pursuit of some opportunities and necessarily discourages the pursuit of others. All deployments that are not motivated under the current regime or constitutional allocation of resources, rights and preferences that exists in the system (and, therefore, are located outside of Circle A) are discouraged. Although some of these deployments may be critical for future economic development, they are unlikely to come about from market exchange, given the current distribution of resources, rights and individual perceptions. A

similar set of unmotivated yet potentially value adding deployments exists for all institutions and each set is likely to be unique for its particular institution.

It is important and, perhaps, ironic to note a fundamental difference that exists between these two classes of resource deployments that lie outside the reach of market exchange. The first class is impeded by transaction costs that interfere with our perceptions of what is efficient - e.g., our often frustrated ability to find the best exchange parties or conditions to put our resources to their known (by someone) best use. These costs reduce the extent to which Circle B overlaps with Circle A and to which Circle D resides wholly within this intersection. The second class requires a change in the distribution of rights in order to be efficient. The transaction costs associated with this class are often inordinately more difficult to overcome than those associated with the first class. Often they require a trade-off of a certain decrease in the welfare of some (and are, therefore, inefficient) for a less certain potential to increase the welfare of others. Exchanges that do occur to support deployments in this class (i.e., outside of Circle A in Figure 2), are usually made either by mistake or else with no (rational) expectation of return. Structural changes that reduce transaction costs for deployments in the first class often increase transaction costs for the second class. Hence, reductions in some transaction costs may only serve to make these deployments even less likely and, thereby, lock the system in further to its current state of incentives.

For either class of these deployments (i.e. outside of Circle D) to be systematically exploited (particularly as some transaction costs decline), some other institutional support (besides markets alone) is required. It is this support that organizations, generally, and firms, more specifically, provide and it is through such support that these institutions help create value for society beyond what markets alone can create. A theory of value creation, however, requires a more precise understanding of how firms can and do engender such exchange. It is an explication of this process that we now turn to in the next section.

THE VALUE CREATING ROLE OF FIRMS

In the previous section we suggested that the scope for market exchange is not only limited but that it is also biased in terms of the nature of resource deployments that are likely to be supported by such exchange. We identified two sources of bias, one that is specific to the way most markets (at least those of the more advanced economies) have evolved to relax the constraining need to satisfy the "double coincidence;" and the other, the set of unmotivated deployment opportunities that is common to all institutions and whose elements is unique to each, as it evolves to become more efficient.

Regarding the first, (i.e., the conventions and norms that have evolved to help us satisfy the "double coincidence" for many exchanges), market institutions, and the conventions supporting them (e.g., the price system, money and credit, norms of reciprocity, laws that establish and enforce the security of rights) have evolved over centuries, like many other long lived institutions, to create a coherent institutional logic which is designed to improve the predictability and security of many exchanges. In the markets of the most developed economies, those conventions have evolved that support and reinforce a market logic which enables actors to enter into and exit from a variety of exchange relations at relatively little cost and, thereby, preserve their independence from all other actors. The very advantage of independence (i.e., of individual actors), which makes it easier (and therefore efficient) for these market's participants to adapt autonomously to changing conditions without the need to consult others, necessarily restricts the form of viability which must exist around each exchange transaction. Consequently, market exchanges, in general, must satisfy the condition of "reciprocal viability" (Coleman, 1990), that is, the resources received in an exchange must be of value to each party so that the exchange meets the criteria of the "double coincidence." Because of this, market exchanges are limited to only those parties who can mutually locate the resources and the parties required for exchange. Hence, resources that are already commonly exchanged are much more accessible than those that are not.

The second cause of limited market exchange is common to all institutions, as they strive to be more efficient in the presence of high transaction costs. That is, each institution favors the

conduct of a unique set of economic activities over all other activities and the set of favored activities comprises those that are more efficient, as defined by that institution. Both of these limitations are easy to see in terms of the Venn diagram presented in Figure 2. The first (i.e., failure to reach missing markets), stems from the necessarily conservative confinement of deployment opportunities to only those that can be supported by exchanges that satisfy the stringent condition of the "double coincidence;" thereby, excluding a large number of beneficial deployment opportunities that are also perceived and motivated (i.e., impeding access to opportunities that are contained in $A \cap B$ but not in D). The second (i.e., failure to adapt institutional incentives to new opportunities), implies some degree of lock-in to the current set of opportunities that are motivated (i.e., impeding the adaptive periodic reconstruction of Circle A).

Firms broaden the scope of exchange in ways that systematically address both of these market limitations. In doing so, firms contribute (by adding value) to the economic system they operate in. They do so, not by importing the institutional logic of markets (i.e., the means markets have evolved to cope with the "double coincidence" and to enhance their allocative efficiency) as the dominant logic to guide this activity but by evolving unique institutional logics of their own. That is, each firm creates its own unique institutional logic for overcoming the market's stringent demands for viability and for circumventing (at least for awhile) the severely constraining forces that exist in the market and in other institutions (including all other firms) and which strongly encourage the deployment of certain resources in certain (i.e., efficient, for this institution) ways. Hence, at any instant in time, of all the vital value adding activities engaged in by a firm, some will be executed more efficiently in the firm than they could elsewhere (i.e., in the market or in other institutions) and others will be conducted less efficiently than they would elsewhere. But <u>both</u> of these sets belong in the firm. For the first set, it is the firm's ability to be more efficient that is the source of the value added; for the second set, it is the firm's ability to discriminate among the external forces motivating these activities and to hold off those that push for greater efficiency, at the expense of some desirable activity.

In effect, the "organizational advantage" (Ghoshal and Moran, 1996) stems from a firm's ability to pursue resource deployment strategies that are difficult (i.e., costly) or impossible to pursue in markets, or any other single institution, alone. Such strategies include those which require resources that (i) are difficult (i.e., costly) to acquire or accumulate through market exchange (e.g., because prices or even markets are "missing," Ghoshal and Moran, 1996); or are difficult to coordinate the use of among independent actors, subject to the stringent demands of "reciprocal viability;" as well as those that (ii) cannot be created, accumulated or deployed in ways that viably satisfy the market's stringent demands for efficiency yet appear promising to those with requisite local knowledge. The organization's advantage in overcoming the market's first constraint (i.e., "reciprocal viability") dramatically broadens the scope of resources that are exchanged and considered for deployment within firms relative to markets. It's advantage in overcoming the second constraint that is present in any institution (i.e., institutional efficiency) enables each firm to influence the path of value creation and, thereby, provide society and each of its members with a vehicle for influencing the economic system of values and its value development (i.e., creation and realization) process, a vehicle that is unavailable in markets alone, where all exchanges are influenced and biased by the same (i.e., market) institutional context.

Broadening the Scope of Exchange: A Bridge to Missing Markets

The institutional conventions and norms that have evolved in society's most developed markets as a consequence of the need to satisfy the "double coincidence," have led to a standard of viability that suggests that every market based exchange relationship must be "reciprocally viable" (Coleman, 1990). In other words, each actor must have a positive account balance in each exchange relation that it is a part of. "Reciprocally viable" relations end when one party finds the relation to be no longer beneficial. In firms and other types of social institutions, the possibility exists for resources to be transferred or exchanged under other, less restrictive, forms of viability. It is an organization's internal institutional context that permits the organization to ensure the viability of its members under less (or more) restrictive conditions. This relaxed demand for viability, in turn, enables the organization to

broaden, beyond that which is achievable outside of the organization's institutional context, the scope of resources that are considered for exchange, as well as those that are ultimately exchanged and deployed.[9]

Two such less restrictive conditions of viability have been referred to by Coleman as "independent viability" and "global viability." "Independent viability" requires only that each actor have an overall positive account balance with the organization as a whole and not with each other actor with which it exchanges. "Global viability" is even less restrictive, in that individual actors themselves do not all require a positive account balance. Only the system of relations as a whole must have a positive balance for it to be globally viable. Because the organization itself is an "implicit third party" to every exchange relation, members are able to enter into and maintain relations that may be beneficial to the organization itself even if they are not directly beneficial to them (Coleman, 1993). The benefit comes indirectly to these members, through their relationship with the organization itself. In all these relations, where viability is not reciprocal and is, therefore, less restrictive, viability depends on one or more additional relations. For this reason and for simplicity's sake, we refer to all such (i.e., non "reciprocally viable") relations as relations that are characterized by "interdependent viability."

Interdependent viability dramatically expands the circle of exchange that takes place among members inside their organizations. By permitting individuals and groups to enter into voluntary exchanges that benefit the organization but benefit themselves only indirectly, organizations open up and make accessible to their members a much broader range of resource deployments (including exchanges) than would be possible were exchange required to satisfy the stringent condition of the "double coincidence" (see Figure 2). Because members can enter into exchanges whose value can be appropriated by the organization, their circle of appropriability is broadened significantly to represent the entire appropriability regime of the organization itself.

The potential impact of this broadening of viability for a single 2-party relation is shown in Figure 3 (where subscripts 1 and 2 refer to parties 1 and 2, respectively). Each party's

potential for exchange is represented by the set of resource deployment opportunities for which it is likely to appropriate some value (i.e., Circle A) and its individual perception of its resource deployment opportunities (i.e., Circle B). The stringent demand for reciprocal viability that is necessary for these same two parties to engage in market exchanges yields such exchange only to the extent of the very small area represented by the intersection of all four circles (i.e., $((A_1 \cap B_1) \cap (A_2 \cap B_2))$ in the center of Figure 3a.

- Figure 3 about here -

Once the existence of an organization shifts the potential for some exchange to occur outside of markets, interdependent viability becomes a possible (but not necessarily efficient or even effective) replacement for reciprocal viability (i.e., satisfaction of the "double coincidence") as a sufficient condition for exchange. The introduction of the viability of the organization itself as the "implicit third party" is represented by the addition of Circle A_F in Figure 3. The interdependent viability that is created by this introduction immediately permits both parties to expand their exchanges with the other to include all exchanges (and other resource deployments) that they perceive as beneficial, not only to themselves but also to the organization itself (e.g., $(B_1 \cap B_2 \cap A_F)$ in Figure 3b). It is straightforward to show that, once an organization's members are individually motivated to execute resource deployments (including exchanges) on the basis of the potential appropriability not just for them but for the organization as well, a complementary incentive to share local knowledge (a form of resource deployment) - to enhance one's perception of opportunities - with others in the organization is created. For example, in the 2-person relation depicted in Figure 3, resource deployment opportunities that are perceived by <u>either</u> party are more likely to be executed because the perceiving party is motivated to inform and convince the other party of the opportunity (e.g., $((B_1 \cap A_F) \cup (B_2 \cap A_F))$ in Figure 3c).

It is also easy to see this convergence of individual opportunity perception with organizational appropriability extending further to include resource deployments (including exchanges) that are executed based on the set of perceived exchange opportunities of <u>other</u> members of the organization, including those that lie outside of the set of perceptions of

either of the two focal parties (e.g., $((B_1 \cap A_F) \cup (B_2 \cap A_F) \cup (B_3 \cap A_F) \ldots)$ in Figure 3d). Over time, as members increasingly refocus their attention on the set of opportunities that are appropriable to the organization (but not necessarily to themselves), their circle of exchange will continue to expand to include resource deployment opportunities that are beneficial to the firm but which previously (i.e., until such refocusing of attention) fell outside of the scope of opportunities perceived by <u>any</u> organizational members (e.g., the shaded area of Circle A_F $((B_1 \cap A_F) \cup (B_2 \cap A_F) \cup (B_3 \cap A_F) \ldots)$ converges to all of A_F in Figure 3d).

Both the institutional and the system-wide impact of other institutions on the viability of exchange can more easily be appreciated with the help of Figure 2. The institutional context of any institution, be it a particular firm or the market the firm operates in, defines two unique sets of resource deployment opportunities that can be characterized as "opportunities motivated" and "opportunities perceived." As shown in Figure 2 and discussed earlier, these two sets are likely to overlap for some opportunities and not for others (e.g., $A \cap B$).

To the extent the firm is structured as a unique incentive system (Holmstrom and Milgrom, 1994), its matrix of resource deployment opportunities that are motivated and perceived (i.e., Circles A and B, respectively) are likely to differ significantly from that of other institutions, including all other firms and markets. Even those prospective exchanges for which the double coincidence is satisfied is likely to differ from institution to institution. Consequently at least a portion of the scope of resource deployments supported by the firm's institutional exchange structure is likely to reside outside of the Circle D that <u>would</u> exist for the institutional matrix (i.e., comprising all other firms and markets) if the focal firm did not exist in a unique institutional form. Obviously, the more firms there are, each similarly characterized by its own institutional context, the broader the scope of exchange for the system as a whole. This holds true even if all exchanges within all firms are also subject to the norms of reciprocal viability, as well as when they are less constrained by some form of interdependent viability.

At the limit, however, This bridge that firms provide to missing markets is necessarily limited in the scope of resource deployment opportunities it can support. The exploitation of

much of $A \cap B$ can require many institutions. Any systematic exploitation of opportunities, beyond the area represented by $A \cap B$, is constrained, not by the transaction costs that mask one's ability to act efficiently, but by those transaction costs that impede the adaptation of incentives to more promising opportunities that might be motivated, given a different distribution of resources, rights or individual perceptions or preferences. The presence of more institutions (e.g., more firms) significantly expands the scope of exchange beyond Circle D, to much of $A \cap B$ and beyond to many other opportunities in Circle B. However, even this expanded scope still must exclude all deployment opportunities that remain outside of Circle B. Moreover, those deployments within Circle B but outside of $A \cap B$, regardless how promising they may be, are not motivated because they are inefficient, given the current distribution of resources, rights and preferences. For the system to systematically adjust itself to respond to these opportunities, some mechanism is required to ensure that Circle A can be adjusted to reflect these opportunities. Until appropriate adjustments are made, all deployments that reside outside of Circle A, regardless of their long term value realization potential, must pay the price of inefficiency. The remainder of the resource combinations that require exchange and for which exchange is possible are unlikely to occur without additional appropriate institutional support; support that can both respect and respond to market forces and, at the same time, is able to change them. Providing such support in order to clear the path for these adjustments is another important role of firms.

Challenging and Changing the Course of Market Efficiency

A world with no transaction costs would need neither organizations nor laws (Coase, 1992); that is, institutions matter only when exchange is costly (North, 1990).[10] When exchange is costless, markets alone are sufficient to ensure that resources are allocated to their most productive current uses <u>and</u> that the allocation process adapts smoothly to the continuous evolutionary cycle of new resource uses displacing the old. Indeed, absent transaction costs, no strong case can be made for the primacy of organizations or markets. Either would do, providing transaction costs were not created in the process.

In the world that we know, however, transaction costs cannot be avoided completely. They represent a large, pervasive and growing part of the economy in America and probably every other developed economy as well (Wallis and North, 1986). In our world of high transaction costs, institutions matter and they matter a great deal. As "the rules of the game in a society or, more formally, . . . the humanly devised constraints that shape human interaction, . . . [institutions] structure incentives in human exchange, whether political, social or economic. Institutional change shapes the ways societies evolve through time and hence is the key to understanding historical change" (North, 1990: 3).

In structuring incentives, all institutions, whether market or organizational, and the conventions and norms that have evolved to support them, largely determine what economic activity is efficient and what is inefficient, given the institutional structure these institutions have helped to put in place. As long as transaction costs are high, we can say with reasonable certainty that two institutions that differ significantly in the behaviors they influence, must classify many otherwise identical activities differently in terms of efficiency; some activities, efficient in one institution, are not in another and others, that are less efficient in the former institution, are more efficient in the latter.

We suggested in the previous section that, if an economic system is to be efficient in the long-run, it must be adaptive to motivate new value creating/realizing resource deployment opportunities as they emerge. In terms of the Venn diagram in Figure 2, this implies that adaptive efficiency requires a degree of responsiveness in Circle A that enables it to shift smoothly, as necessary, across the universe of all deployment opportunities to reflect the need to adjust institutional supports to enhance the efficiency of the most promising deployment opportunities over those that are less promising. Transaction costs not only impede this responsiveness, they distort it, as well. Many of the same incentives which encourage actors to enhance the system's allocative efficiency also lock the system into the status quo of rules, norms and conventions, which makes the achievement of adaptive efficiency more difficult. It follows then, that any single institution, be it market or firm, is

seriously limited in its ability to support the scope of exchange that is necessary to promote all value adding resource deployment opportunities in a systematic and unbiased way.

Of course, Figure 2 presents only a static slice of the resource deployment opportunities that exist at a single instant in time. For the argument to have any weight, these limitations and biases must hold over time and remain relatively unresponsive to new deployment opportunities that arise. As long as we can be reasonably assured that, Circles B and D can easily shift (i.e., adjust in response to deployments executed) to reflect the perceptions of new promising resource deployment opportunities as they arise, and that Circle A will adjust to reflect such awareness, the bias of a single institution to narrowly focus on only a few deployment opportunities at a time presents no particular problem for efficient deployments over the long run. Indeed, in the hypothetical world of no transaction costs, the biased focus that is provided by any institution would represent a viable means for systematically allocating and reallocating scarce resources to their most productive uses over time; focusing the use of scarce resources to maximize their output in the short term and appropriately adjusting that focus over time to reflect the greatest opportunities as they emerge.

Although a world without transaction costs is difficult to imagine, with Figure 2 we can easily conceptualize how resource rights and individual perceptions might adapt to new value adding opportunities to enhance adaptive efficiency. Absent transaction costs, both the set of appropriable deployments (Circle A), as well as people's perceptions of these deployment opportunities (Circle B) would shift effortlessly (i.e., costlessly) across the universe of deployment opportunities to enable actors to successfully negotiate with each other until all of their local knowledge is fully exploited. That is, without transaction costs, the "double coincidence" would not pose much of a constraint. First, information can be expected to be exchanged costlessly until all resource deployment opportunities represented by ($A \cap B$) would occur without fail. Next, institutional barriers posed by the distribution of resources/rights at any given time would easily be negotiated away to facilitate the expectation of newly recognized opportunities that would have remained inefficient only because of these institutional barriers (i.e., Circle A would shift more into those parts of

Circle B that are true opportunities). At the same time, this vigorous shifting of resource rights and its consequent opening of new areas of exchange would open up more new opportunities to individuals' perceptions (i.e., each causing a shift in Circle B) only to be followed closely by a concomitant shift of Circle A. Over time, more and more of the universe of opportunities would be exposed to the perceptions and local knowledge of many and resource rights would be costlessly renegotiated and realigned to exploit these opportunities.

In the real world of pervasive and high transaction costs, however, this adjustment process is not likely to occur spontaneously or even automatically, if it occurs at all. In such a world, issues of resource rights and value appropriability merit careful attention.[11] Unfortunately, the efficient allocation of resources does not ensure their deployment in ways that will lead to the greatest future realization of value or economic progress. Rather, as North (1990) observed, in striving to make current productivity of resources as efficient as possible, the forces of allocative efficiency favor and, thereby, reinforce the system's current distribution of rights. He argued "We are far from understanding how to achieve adaptively efficient economies because allocative efficiency and adaptive efficiency may not always be consistent. Allocatively efficient rules would make today's firms and decisions secure - but frequently at the expense of the creative destruction process that Schumpeter had in mind. Moreover, the very nature of the political process encourages the growth of constraints that favor today's influential bargaining groups" (1990: 81 - 82) .

Transaction costs permit the same forces that give rise to efficient behavior (in a single institution) today to persist beyond their usefulness and to thwart (at least to some extent) the institutional changes that will be necessary to support efficient behavior tomorrow. In terms of Figure 2, when transaction costs are high, Circle A encounters significant friction as it attempts to shift responsively to motivate promising new value creating/realizing deployment opportunities.

For economic development to be efficient in the long run, then, some countervailing force is needed to bring about the necessary institutional change. Organizations, in general, and

firms, in particular, play an important role in providing such a countervailing force; one that is necessary for both markets and firms to achieve adaptive efficiency.

With the help of Figure 4 we can now begin to pull the strands of our argument together to illustrate how these biases of market exchange are overcome by the introduction of other institutional forms of organization to the system. The two largest circles on the left, in Figure 4 (Circles A_M and B_M) correspond to Circles A and B in Figure 2. That is, they represent, respectively, all systematic deployment opportunities that are motivated by and perceived in the system before the influence of any other institution (i.e., a firm) is introduced. Circles A_F and B_F represent the influence of the institutional context that comes with the introduction of a focal firm. Circle A_F comprises the unique set of deployment opportunities that are motivated by the introduction of a firm and its unique institutional context. It differs from Circle A_M in that opportunities to deploy resources available to the firm are influenced by the firm's own system of incentives and constitution of rights, as well as those of the market. Hence, some opportunities motivated within the firm may not be motivated externally and, conversely, other activities that may be motivated to occur outside the firm are not so motivated to take place within.

- Figure 4 about here -

This difference in resource deployment opportunities that are motivated inside (i.e., Circle A_F) and outside (i.e., Circle A_M) of the focal firm gives rise to different perceptions of deployment opportunities inside the firm from those that exist outside the firm. Once a firm is added to the institutional matrix, however, the set of opportunities perceived in the system is expanded to include those that exist in the firm (i.e., $B_{M_{t+1}} = B_{M_t} \cup B_{F_t}$). From Figure 4 we can see that the set of resource deployments that are likely to be supported by exchange inside the firm (i.e. $A_F \cap B_F$) differ from those that are likely to occur outside the firm in one of two ways. Either, these deployments are motivated to occur outside the firm (i.e., are inside Circle A_M) or they are not (i.e., are outside Circle A_M). To the extent these opportunities fall within $A_M \cap B_M$ but remain outside of D_M they expand the scope of market exchange beyond those that satisfy the stringent conditions of reciprocal viability. Of those

opportunities within $A_F \cap B_F$ that fall outside of $A_M \cap B_M$, they occur in relative defiance to the forces of market efficiency. If the execution of these deployments leads to the realization of value which then brings about a change in market incentives or resource rights, that deployment and the firm responsible for motivating its execution will have challenged and changed the course of adaptive efficiency.

To summarize our argument thus far, the potential for value creation in most economic systems vastly exceeds the system's capacity for realizing this value. Consequently, only a tiny fraction of the possible value that resides in the constituent resources of an economic system at any time is ever created or realized. The reason for this, Hayek (1945) taught us, lies in our failure to exploit local knowledge - knowledge which collectively contains the information required to deploy resources in value creating ways but which only resides in fragmented and costly to acquire bits. This failure has been attributed to transaction costs - the costs incurred in coping with faulty perceptive abilities and sticky resource rights, in the face of uncertainty and limits to individuals' attention and cognitive abilities (Coase, 1991; 1992; North and Wallis, 1994). Exchange serves as a primary if not the only vehicle or mechanism through which any of the value that is inherent in the allocation of resources to their highest and best use is ever realized in any economic system. However, because nearly all exchanges are subject to the influence of transaction costs and institutions greatly affect the nature and magnitude of costs that surround particular exchanges, the path followed by any developing economy is unavoidably biased by the institutions put in place to make that development more efficient.

MARKETS AND FIRMS

Most economic theories put forth to explain the role of firms frame the problem of coordination as a choice among competing organizational forms (e.g., Williamson, 1975). The choice is usually one of markets or firms or some hybrid institution (usually defined as more or less "market like" or "firm like") and the unit of analysis is typically a particular

activity like a specific resource deployment (e.g., the transaction). Since organizations differ in their coordination attributes, one is likely to be more efficient than all others in mediating the focal activity. Hence, efficiency, in general, and market efficiency, in particular, is the common criterion for selection (i.e., the institutional arrangement that provides for the most productive use of resources). By implication then, market efficiency is presumed to be the standard by which solutions to Hayek's problem of making fuller use of fragmented, incomplete knowledge referred to earlier, are to be found.

Although certainly part of the solution, as we have argued in this paper, the goal of market efficiency (or, more specifically, the forces - incentives - encouraging such an outcome) cannot be a complete solution. Indeed, without other institutional support, it is as much part of the problem as it is a necessary part of any solution. As Coase observed, "If the costs of making an exchange are greater than the gains which that exchange would bring, that exchange would not take place and the greater production that would flow from specialization would not be realized. In this way transaction costs affect not only contractual arrangements, but also what goods and services are produced" (Coase, 1992: 716). Because institutions cannot help but influence the relative costs of making some exchanges over others, institutions too inevitably affect not just the way economic activity is organized but also what resources are exchanged and deployed in an economy.

With the framework we have provided, we can now address the question that is commonly posed when discussing any theory of the firm. That is, if a firm can make a market economy more efficient, "Why can't a large firm (or a centrally planned economy) always operate at least as efficiently as a chaotic, unorganized market?" (Milgrom and Roberts, 1988). According to the argument we have made, the answer is clear. No single institution by itself is capable, in the presence of transaction costs, of bringing about adaptive efficiency. This is because a single institution seeks to adapt its constitution of resource rights to make them as efficient as possible, given the constitution of rights it must operate in. As long as there exists significant impediments to adapting the constitution to enable the pursuit of newly emerging opportunities, the process of a single institution adapting to forces of allocative

efficiency only serves to reinforce the broader constitution of rights and, thereby, impedes the process of adaptive efficiency. Institutional pluralism (i.e., a rich variety of institutional forms and sizes), on the other hand, helps to overcome this institutional straightjacket. With the help of Figure 4, we can see how markets and firms, operating in a state of creative tension, provide the necessary checks and balances to bring about adaptive efficiency.

One large firm, as a single institution, is no better at fully using knowledge than is one large market (i.e., as a single institution comprising the maximum - in terms of its own efficiency criteria - number of independent actors). As we have argued already, absent transaction costs, any institution would suffice, as long as transaction costs are not created in the process. But in the presence of transaction costs, neither institution, by itself, would be adaptively efficient. This is because there would be no areas of respite from which resources can be marshalled to redirect the course of efficiency (i.e., to guide the shifting of Circle A_M in Figure 4). Either institution (i.e., market or firm), left untempered by the countervailing force of other institutions, subjects us to an institutional straightjacket, one an iron cage of bureaucracy; the other, a treadmill of ever tightening competition; neither one offering sufficient freedom to perceive, experiment with and evaluate new ways to create and to realize value.

In terms of Figure 4, if markets were to exist in the presence of transaction costs and without the tension that comes from competing institutions, the set of resource deployment opportunities that are most likely to be supported by exchange are those represented by Circle D_M. Without some alternative institutional context to focus our motivation and perceptions on other deployment opportunities, many value creating combinations and value realizing deployments would go unexploited. This is so because there would be little to guide the adaptation of Circle A_M to accommodate new opportunities. It is the existence of other institutions, in general, and of firms, in particular, that provides this guidance.

As we argued in the last section, firms provide an institutional alternative that substantially broadens the scope for exchange beyond this set (i.e., Circle D_M) to support resource deployments that would otherwise not occur in markets alone. Firms do this by creating an

institutional context that specifies different sets of opportunities that are motivated and perceived and which are unique to each and every firm (Circles A_F and B_F represent these sets for a typical focal firm). Because the set of motivated and perceived opportunities differs from institution to institution (as shown by the different locations of $A_M \cap B_M$ and $A_F \cap B_F$), many more deployments than are likely to occur in a single institution can be expected in the presence of many institutions. In the process, the set of deployment opportunities motivated in the system as a whole (i.e., Circle A_M) is continually adapting, in a relatively efficient manner, to opportunities as they are perceived.

Institutional pluralism contributes to the process of adaptive efficiency in at least three ways. First, the scope of exchange is broadened to include more opportunities that are not exploited elsewhere. These include, (i) those opportunities that are already motivated and perceived in the market or in other institutions but fail to meet these institution's demands for viability (i.e., unexploited parts of the area indicated by the number "1" in Figure 4); (ii) Those opportunities that are already perceived elsewhere but not motivated anywhere else (i.e., the area indicated by the number "3"); and those opportunities that would not even be perceived without the existence of the focal firm (i.e., the area indicated by the number "4"). Second, some resources that are currently deployed efficiently elsewhere are made available for deployment within the firm under a different set of motivating conditions (the areas indicated by the numbers "2" - for markets and "1" for other firms). By replacing these motivating forces that encourage certain deployments (e.g., by muting external incentives, restricting access to resources or their use, Holmstrom and Milgrom, 1991) with forces that motivate alternative patterns of deployment, firms make it easier for value creating new combinations to be discovered.

Finally, each unique institutional context that is created by the existence of a firm provides a potentially attractive cite for actors who perceive deployment opportunities, that may be unknown to current members of the firm (i.e. the area indicated by the number "5") to join the firm and pursue his or her convictions at a lower cost than is available anywhere else (North, 1990) and, in the process, expand the degree to which Circles A_F and B_F overlap

(i.e. $A_F \cap B_F$). It is easy to see from this that in the course of many firms of different shapes and sizes engaging in this process, each broadening the scope of exchange in ways that allow it to focus on some fragmented bits of the local knowledge that Hayek talked about, more of the local knowledge that resides in the system is used and used in a way that also promotes adaptive efficiency.

Adaptive Efficiency Within the Institutional Matrix

Conventional economic theory conceptualizes the ideal structure of an economy as a market comprising as many decentralized independent actors as possible; each acting autonomously to best exploit its local knowledge. Firms are needed only to the extent they can satisfy market demands for efficiency better than markets themselves can. In other words, they are seen as a means of containing the damage of market failures. In contrast, we have shown in this paper why the value creation and realization potential of atomistic autonomous agents who can only exchange among themselves without the support of multiple institutions, is severely limited. When the complementary characteristics that make for efficient markets (i.e. large numbers of independent participants) are confronted with (i) strong uncertainty, which masks the true value potential in most new resource combinations; (ii) pervasive transaction costs, which distort perceptions of the nature of available resources and their accessibility; and (iii) the pressures of competition, which favor resource deployments that are as efficient as possible, the outcome is a relatively conservative (i.e., more efficient and less creative) set of resource deployment decisions.

Further, we cannot tell much about the efficiency of such a system. As Coleman (1993: 85 - 86) argued, "The concept of efficiency of an economic system is defined only within a particular distribution of resources or as I will call it, a particular constitutional allocation of rights and resources. If in a given system, with a given constitutional allocation, all externalities are internalized and transaction costs are reduced to zero, the system has achieved efficiency. But if rights are allocated differently, to persons with different interests, then the "efficient" outcomes of the system may differ . . . ." North too has emphasized the constraints of an institutional straightjacket that would be created by a market that was

somehow designed to be as efficient as possible - i.e., comprising a Pareto optimum distribution of actors, each organized to efficiently overcome one market failure or another: "If the basic institutional framework makes income redistribution (piracy) the preferred (most profitable) economic opportunity, we can expect a very different development of knowledge and skills than a productivity increasing (. . . [e.g., a] twentieth-century chemical manufacturer) economic opportunity would entail. . . . The incentives that are built into the institutional framework play the decisive role in shaping the kinds of skills and knowledge that pay off" (1990: 78).

As long as transaction costs continue to pervade economic activity, we cannot claim with any certainty that a market that is characterized by lower total transaction costs (and, therefore, more efficient) are superior to others To the extent that the former (more productive in the short run) system prices and constrains the deployment of resources in ways that make any transition to the latter system more difficult, it impedes long term development and the realization of more value. To overcome this limitation, some degrees of freedom are necessary to change the rules of the game. Factors like uncertainty and bounded rationality - sources of most transaction costs - provide the degrees of freedom required for firms to initiate and succeed in making progressive change.

They do so by replacing, muting of otherwise modifying market incentives and, thereby, redefining the motivation and conditions of efficiency for the economic activities that take place within their organizations. Seen from the perspective of the market logic (Ghoshal and Moran, 1996), such modifications of the market's high powered incentives are seen as a disadvantage that must be overcome, if firms are to be efficient. From the perspective we are advancing, the firm's ability to mute market incentives is not a disadvantage but an opportunity for firms to exploit. More generally, firms provide their members with an opportunity to establish a beachhead, from which forces can be marshalled to enable the firm and its member to challenge and to change the forces of market efficiency. In a narrow equilibrium sense, <u>given the current distribution of resources and allocation of rights</u>, this collective disregard of market incentives is irrational or, at least inefficient. However, from a

broader perspective - one that also recognizes no initial distribution of rights and resources as more "appropriate" than another - it is precisely this ability of firms to shield themselves and their members, albeit temporarily, from the strong market pressures to be efficient, that creates the forces of adaptive efficiency within the economy.

In essence, then, all participants embedded in a given institutional context (market or organization) are pressured to abide by a "defacto constitutional allocation" of rights and the pattern of resource distribution and deployment that emerges from it. All changes in this allocation, even those that are likely to enhance resource productivity and, therefore, create potential value, must force the realignment of the entire constellation of interdependent rights that compose this "constitutional allocation." To force such a dramatic and integrated change, the benefits must not only be positive and perceived by all necessary parties, the benefits must also be large enough to overcome all the transaction costs that support the (inefficient) continuation of the current constitution of rights. Hence, for adaptive efficiency to prevail, it must dominate not only the productivity of current period allocative efficiency but also overcome the baggage (i.e., transaction costs) of inefficient perceptual and objective distortions that support the rights represented in the status quo. By muting the incentives of the broader institutional matrix (North, 1990), firms are able to protect their members, somewhat from this tyranny of market (organizational) power.

This perspective on the value creating role of firms within an institutional matrix and the logic that guides their member's collective action stands in sharp contrast to that held by many scholars. For example, Williamson characterizes firms as just another type of "contractual instrument, a continuation of market relations, <u>by other means</u>" (Williamson, 1991: 162, emphasis in original). The institutional perspective that we have put forth, on the other hand, demonstrates how the <u>same means</u> (i.e., the same forces amplified and attenuated by different institutions) transforms the relations themselves and, thereby, the nature of what is exchanged among the parties in a particular set of relationships.

Moreover, this institutional view of firms echoes the seminal works of Barnard (1938) and Selznick (1957) and that of many other organizational theorists. It also positions

organizations in the broader socio-economic framework as institutions that exist not to continue market relations but whose raison d'être is to confront and challenge the very logic that drives the market forces of adaptation. In the process of pursuing their different logics in a constant state of tension, markets ensure a reasonably efficient use of scarce resources and organizations provide individuals (and society) with a powerful collective force for pursuing their convictions. Such a force influences the autonomous adaptation (to an emerging set of market forces) of all actors in the system. As a result, even the autonomous adaptation of markets is, in this way, driven (albeit indirectly but probably appropriately, given uncertainty) by a form of purposive adaptation (Ghoshal and Moran, 1996), the source of which lies in individual organizations.

ENDNOTES

[1]We adopt North's distinction between allocative and adaptive efficiency as follows, "In allocative efficiency, the standard neoclassical Pareto conditions obtain. Adaptive efficiency, on the other hand, is concerned with the kinds of rules that shape the way an economy evolves through time" (North, 1990: 80).

[2]In a hypothetical world of perfect rationality and zero transaction costs, value destroying resource combinations would be limited to new, irreversible (trial and error) like combinations for which their value creation/destruction potential is not known until the resources are so deployed. All value creating/destroying effects associated with subsequent transactions will be factored into any decision to repeat the combinations. No net value destroying combinations would rationally be repeated. Of course, in the real world of uncertainty, ambiguity, pervasive and high transaction costs and only intended rationality many value destroying resource combinations are likely to be attempted and repeated, even by the same parties (either by accident or for strategic purposes).

[3]In many respects this is just what happens in any process of investment, where short-term productive assets are diverted (sacrificed?) for the prospect of longer-term gains. The primary distinction we wish to emphasize between the resource combinations which constitute innovation in the Schumpeterian sense (1934; 1947), which we employ, and more conventional investment decisions is at the time that commitments are necessary to divert resources away from already productive assets, the rational for doing so is always <u>questionable</u> and can only be justified on some belief that is often based on faith or even nothing more than hope that the established logic dictating the most rational deployment of the resources will somehow be overturned, making room for a new logic which will then make the new resource combination the most rational.

[4]Even non depletable resources like information, whose deployment requires the application of other resources that are depletable or scarce (like attention), necessitate some sacrifice of

alternative deployments from which the scarce resource is diverted. Putting non depletable resources in the hands (and minds) of many actors greatly reduces this constraint.

[5]Regardless of the intrinsic value of ideas, goods or services, it is only through their exchange that economic value is in fact realized. For instance, individual learning is surely a major source of individual value and is also a prerequisite for probably most institutional value. By itself, however, individual learning does not contribute to institutional value unless and until it leads to exchange. Productivity gains from such learning, for instance, do not show up in added wealth (either to the source or to the system) until the enhanced productivity affects the output that ultimately contributes to wealth as a result of exchange. Organizational learning, on the other hand, for it to exist as a phenomenon at all, must be shared by individuals. Hence, it represents a form of institutional value.

[6]It should be emphasized, however, that many other value adding (i.e., value creating and realizing) resource deployment opportunities that also do not <u>require</u> exchange are likely to exist; some efficient - i.e., of which all are contained in the set of motivated deployment opportunities (Circle A) that remains outside of Circle C ($A - (A \cap C)$) - and some not efficient - i.e., the entire area that remains outside of both Circles A and C, as well as the set of motivated deployment opportunities that are not efficient but are included in $A - (A \cap C)$. Among these, only some of the deployment opportunities in Circle B (i.e., those that are <u>misperceived</u> as beneficial by the party or parties with the means to execute the deployment) are likely to occur, albeit mistakenly and at a cost to the deployer(s). These deployments are unlikely to occur, however, without some additional transfer of resource rights. The reason for this apparent paradox is, despite the fact that no exchange is <u>required</u> for these beneficial resource deployments to occur, few would actually take place because of limitations in any single party's ability to perceive the value of these resource deployment opportunities. Even though exchange is actually unnecessary, without it there is little likelihood that sufficient linking will occur among those actors - (i) who, like the sovereign, in our earlier example, may have the requisite resources/rights but are unable to perceive of a deployment

opportunity and (ii) those who, like his or her subjects, may perceive of the deployment opportunity but are unable to access the resources or the benefit from their deployment - to allow them to benefit from each others' underutilzed resource. Paradoxically, then, given transaction costs, resource scarcity can actually promote many exchanges that lead to enhanced levels of value creation and value realization that otherwise would not be achieved if resources were more abundant to begin with!

[7]In a hypothetical regime of zero transaction costs, we can expect Circle D to exist almost entirely within $A \cap B$. To account fully for all resource deployments that are likely to occur, in the presence of transaction costs, however, we must make allowances for misperceptions of efficiency and for uncertainty or differences in expectations concerning the future value of a given resource. To reflect this uncertainty, Circle D would generally extend beyond the boundary of Circle A (but always remain within Circles B and C). Increasing transaction costs decrease the portion of A and B that intersect ($A \cap B$). It is possible, at least theoretically if transaction costs are high enough, that the extent to which Circle D would be expected to straddle the boundary between $A \cap B$ and B - ($A \cap B$) could be random (i.e., D could as easily reside completely outside of $A \cap B$ as inside - while still remaining entirely within B and C). As transaction costs go to zero, more of Circles A and B will overlap and Circle D will shift from extending significantly beyond ($A \cap B$) into the rest of B to fall almost completely within ($A \cap B$) when all transaction costs are zero.

[8]The need for a "double coincidence" is most commonly associated with barter trade but must, in principle, be satisfied in all exchanges. To be sure the evolution of institutional conventions, such as the innovation and use of money and credit, have greatly reduced, but have by no means eliminated this constraint. These conventions only make it <u>easier</u> to satisfy the demanding conditions implied by the "double coincidence; they do not make their satisfaction unnecessary. In exchanges of goods for money, both the money and the goods exchanged are perceived by the parties to the exchange as valuable and worthy of exchange. That is, the opportunity to productively deploy the resources exchanged is implicitly

perceived and motivated for both the party receiving the resource, on the one hand, and for the party accepting money in exchange, on the other hand; otherwise no exchange would occur. Exchanges involving money, although certainly large and growing in absolute terms, are not necessarily a large or growing percentage of all exchanges (e.g., of goods, services, ideas or knowledge or skills) that take place. The pervasiveness of non-monetary exchange is particularly acute inside organizations. We consider these exchanges and their implications in greater detail in the next section.

[9]We do not mean to imply that all or even many exchange relations in all organizations are characterized by a relaxed demand for reciprocal viability. Indeed, the viability of many intraorganizational relations may be even more (not less) restrictive than in markets. The point we seek to make here is that organizations have a unique ability to affect the viability of their internal relations in ways that the larger institutions in which organizations are embedded cannot, or can only with great difficulty.

[10]A world of zero transaction costs is hard to imagine. Such a world, as pictured by Coase (1988), must be a world that comprises only omniscient actors without limited rationality. Because bounded rationality necessarily impedes the exchange of information (Cohen and Levinthal, 1989; 1990), only if rationality is unbounded (and there are no other transaction costs) could all of eternity be experienced in a split-second, as Coase (1988) avers would be possible in such a world.

[11]For example, individuals in most economic systems do not have the right to acquire any resource they want or to use all resources that they might acquire in any way they desire. Nor do they always have the means to appropriate all the value that is to be realized from such use nor the rights to transfer the resources to others who might have all the requisite rights and means to realize and appropriate their resources' recognized potential value. A variety of laws and regulations (of which patent and copyright laws are common examples) usually exist to protect the rights of some and thereby necessarily restrict the rights of others. In the

absence of these conditions, parties have little incentive or motivation to engage in otherwise mutually beneficial exchanges and the potential value that would accrue to such exchanges will remain unrealized, regardless of the potential value that may be inherent in the resource deployment that such exchanges would make possible. The constantly evolving conditions of exchange, which continuously alter the potential value that is inherent in a resource (for a given state of time and sets of technologies and individual preferences) may cause some value to vanish (i.e., be rendered unrealizable), without ever being realized.

REFERENCES

Barnard, Chester 1938 The functions of the executive. Cambridge, MA: Harvard University Press.

Coase, R. H. 1988 The firm, the market, and the law. Chicago: University of Chicago Press.

Coase, R. H. 1991 "The nature of the firm (1937)." In Oliver E. Williamson and Sidney G. Winter (eds.), The Nature of the Firm: Origins, Evolution, and Development, 18-33. New York: Oxford University Press.

Coase, R. H. 1992 "The institutional structure of production." American Economic Review, 82(4), 713-719.

Cohen, Wesley M., and Daniel A. Levinthal 1989 "Innovation and learning: The two faces of R&D." The Economic Journal, (September), 569-596.

Cohen, Wesley M., and Daniel A. Levinthal 1990 "Absorptive capacity: A new perspective on learning and innovation." Administrative Science Quarterly, 35(1), 128-152.

Coleman, James S. 1990 Foundations of social theory. Cambridge, MA: Harvard University Press.

Coleman, James S. 1993 "Properties of rational organizations." In S. Lindenberg and Hein Schreuder (eds.), Interdisciplinary Perspectives on Organization Studies, 79-90. Oxford: Pergamon Press.

Ghoshal, Sumantra, and Peter Moran 1996 "Bad for practice: A critique of the transaction cost theory." Academy of Management Review, 21(1), 13 - 47.

Ghoshal, Sumantra, Peter Moran, and Luis Almeida Costa 1995 "The essence of the megacorporation: Shared context, not structural hierarchy." Journal of Institutional and Theoretical Economics, 151(4 - December), 748 - 759.

Hayek, Friedrich A. 1945 "The use of knowledge in society." American Economic Review, 35(4), 519 - 530.

Holmstrom, Bengt, and Paul Milgrom 1991 "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." Journal of Law, Economics, & Organization, 7(Special Issue), Supplement, 24-52.

Holmstrom, Bengt, and Paul Milgrom 1994 "The firm as an incentive system." American Economic Review, 84(4), 972-991.

Milgrom, Paul, and John Roberts 1988 "Economic theories of the firm: Past present and future." Canadian Journal of Economics, 21(3), 444 - 458.

North, Douglass C. 1990 Institutions, institutional change and economic performance. Cambridge, England: Cambridge University Press.

North, Douglass C., and John J. Wallis 1994 "Integrating institutional change and technical change in economic history: A transaction cost approach." Journal of Institutional and Theoretical Economics, 150(4), 609 - 624.

Schumpeter, Joseph A. 1934 The theory of economic development. Cambridge, MA: Harvard University Press.

Schumpeter, Joseph A. 1942 Capitalism, socialism and democracy. London: Unwin University Books.

Schumpeter, Joseph A. 1947 "The creative response in economic history." The Journal of Economic History, 7(2), 149 - 159.

Selznick, Philip 1957 Leadership in administration. Evanston, IL: Row Peterson.

Simon, Herbert A. 1957, [1945] Administrative behavior. New York: Free Press.

Simon, Herbert A. 1991 "Organizations and markets." Journal of Economic Perspectives, 5(2-Spring), 25-44.

Teece, David J. 1986 "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy." Research Policy, 15(6), 285-305.

Wallis, John J., and Douglass C. North 1986 "Measuring the transaction sector in the American economy, 1870 - 1970." In Stanley L. Engerman and Robert E Gallman (eds.), Long-term Factors in American Economic Growth, 95 - 161. Chicago: University of Chicago Press.

Williamson, Oliver E. 1975 Markets and hierarchies: Analysis and antitrust implications. New York: Free Press.

Williamson, Oliver E. 1991 "Economic institutions: Spontaneous and intentional governance." Journal of Law, Economics, and Organization, 7(Special Issue), 159 - 187.

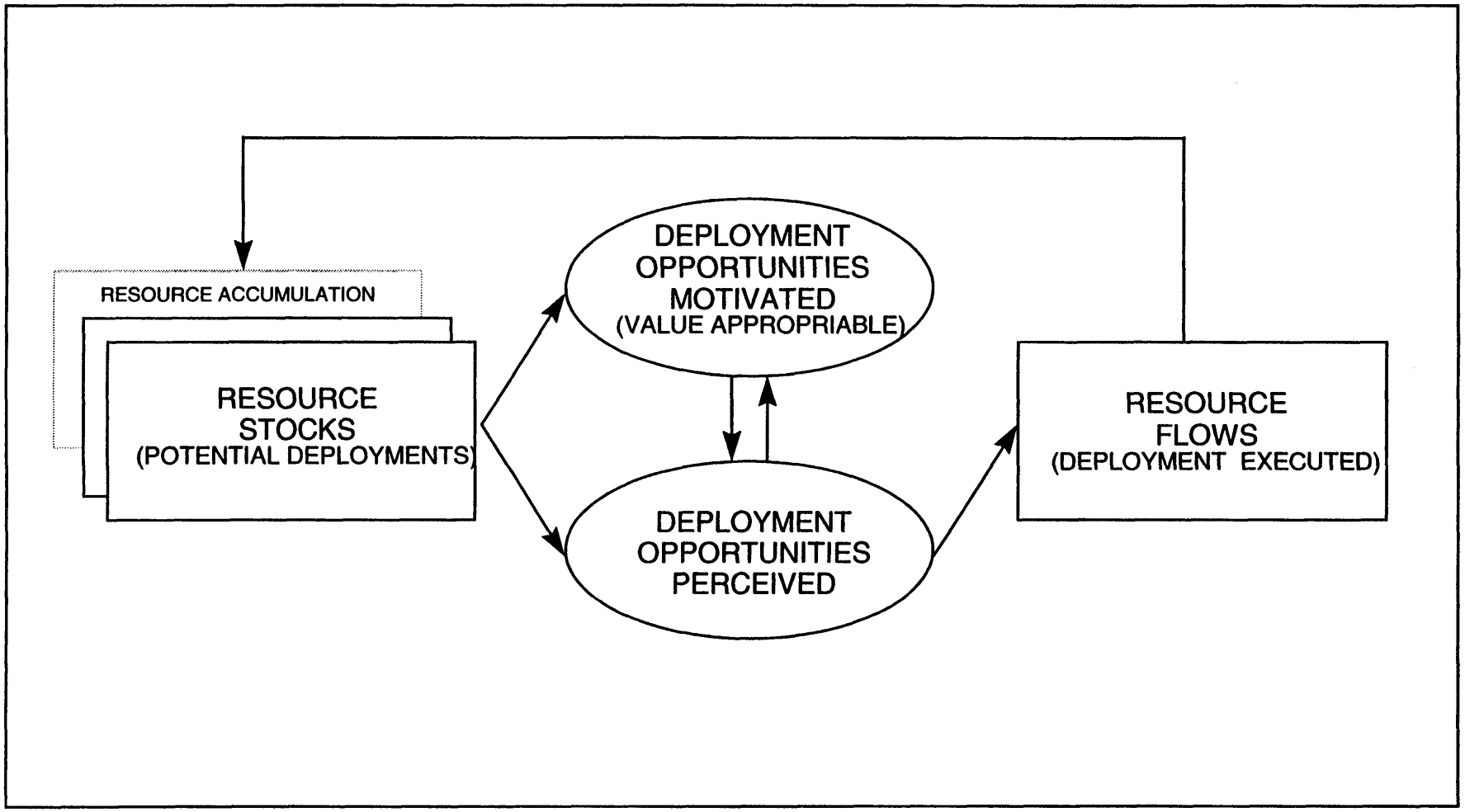

Figure 1
Resource Deployment Process

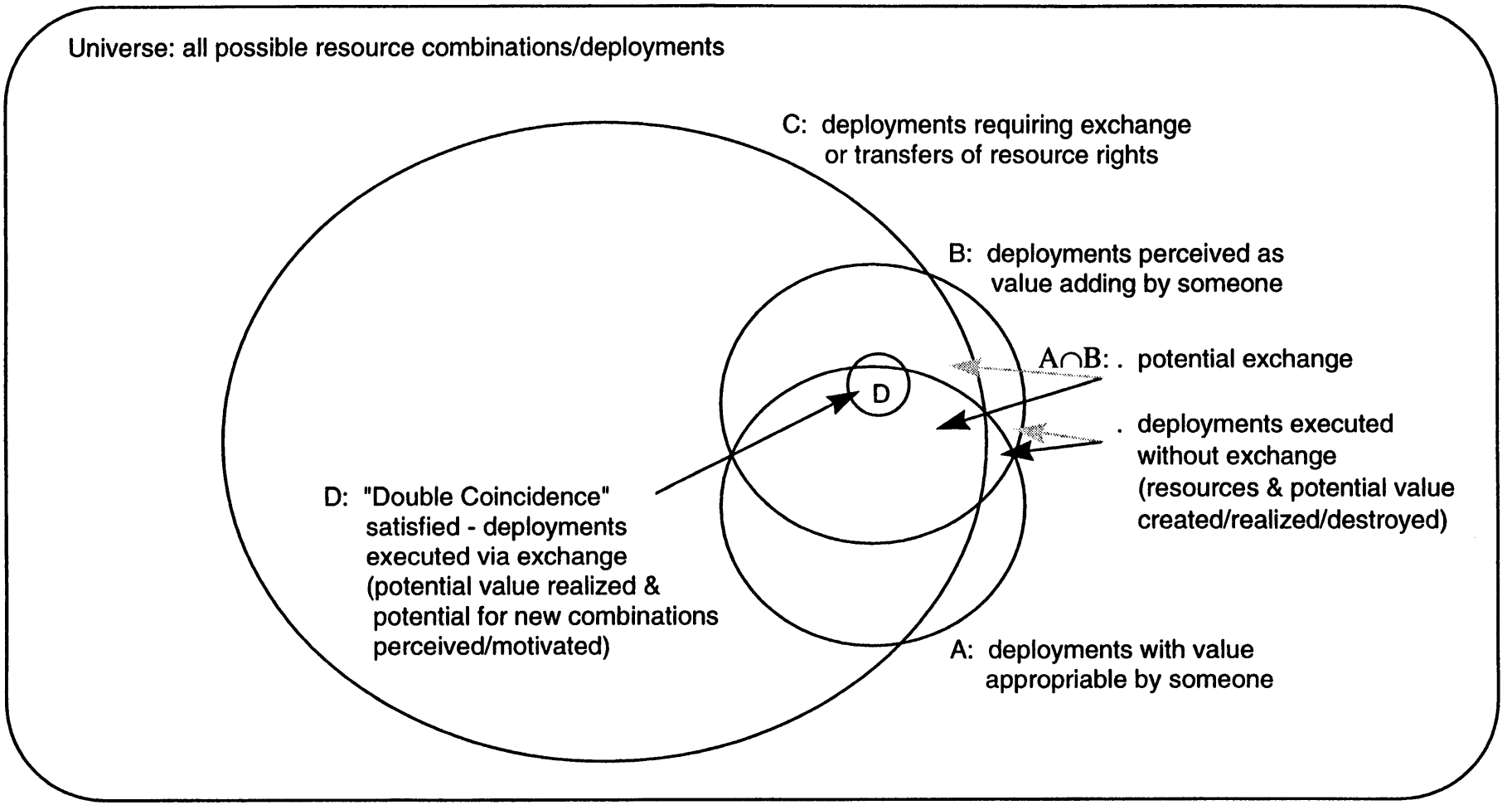

Figure 2
The Role of Exchange in Value Creation

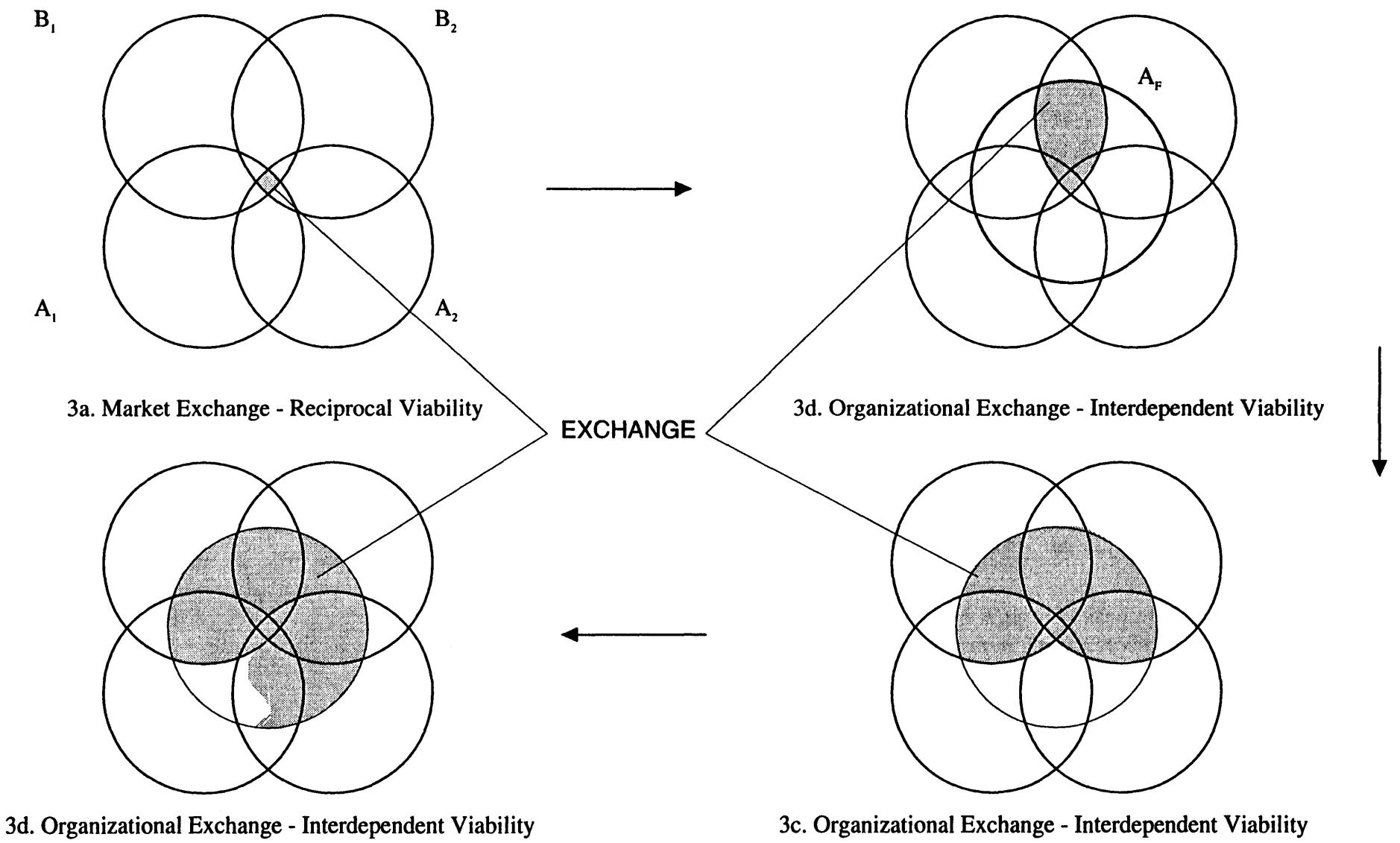

Figure 3
The Organizational Advantasge
A Widening Circle of Exchange

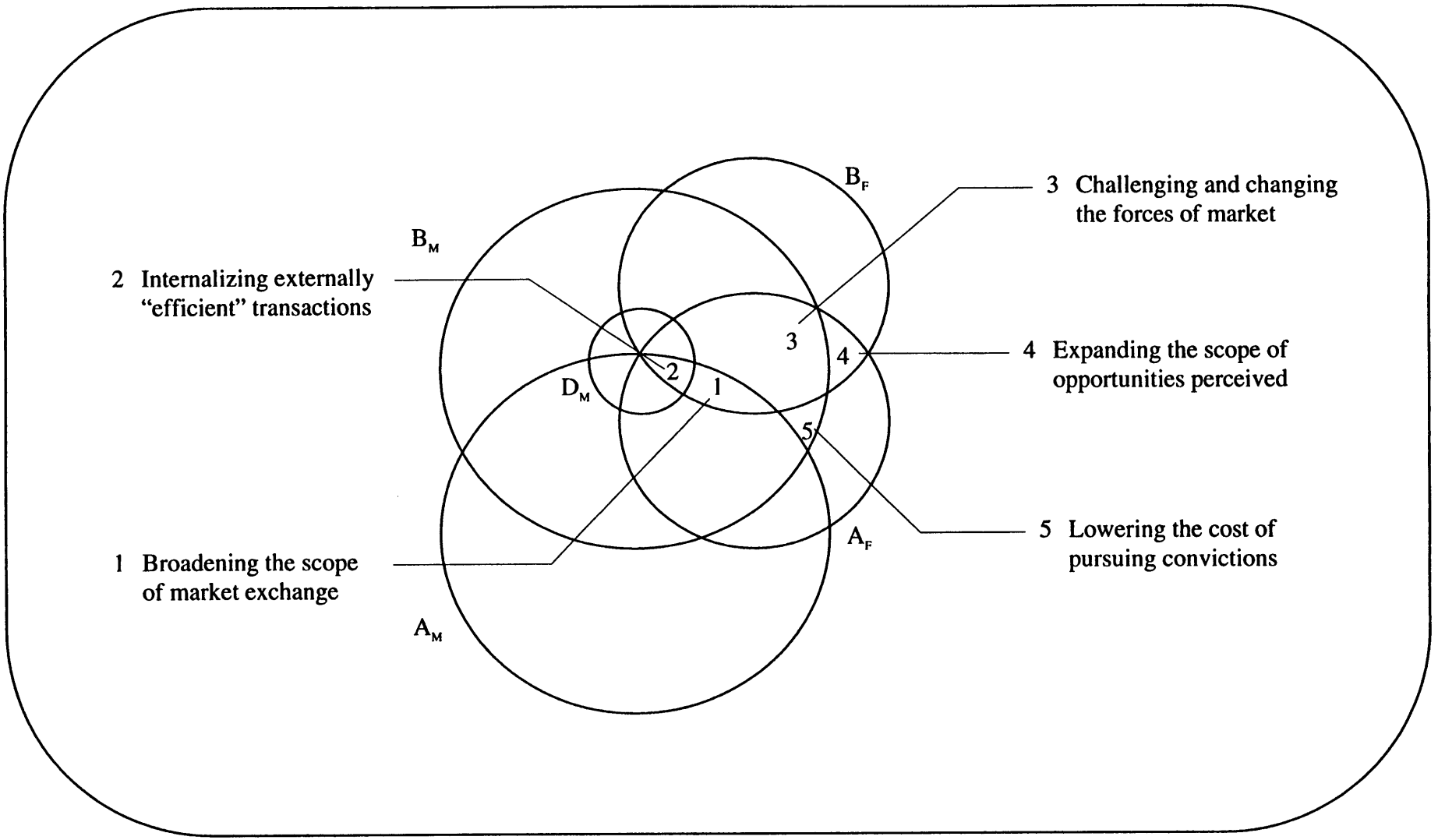

Figure 4
Markets and Firms:
Both are Necessary to Ensure Adaptive Efficiency
with Appropriate "Checks and Balances"