Information, Competition, and Markets

By BARRY J. NALEBUFF AND JOSEPH E. STIGLITZ*

One of the dominant characteristics of modern capitalist economies is the important role played by competition: not the peculiar static form of pure price competition embodied in the Arrow-Debreu model, but rather a dynamic competition, more akin to the kind of competition represented by sports contests and other races (including patent races).

In recent years, there have been several attempts to explain why firms often base the pay of their workers and managers on relative performance. (See, for example, Edward Lazear and Sherwin Rosen, 1981.) Such compensation schemes become desirable when three conditions are satisfied: (a) The input (effort) of workers (managers) must not be directly observable, at least without cost. Thus firms must either expend resources to monitor inputs or devise reward structures in which compensation is a function of variables (such as output or profits) which are themselves functions of inputs but are less costly to observe. (b) The relationship between input and output must be stochastic, so that by observing output, one cannot perfectly infer what the input was. (c) Finally, the stochastic disturbances which affect the relationship between input and output of different firms must be correlated. By looking at the performance of one worker relative to that of others, one can make better inferences about his effort than one can make without using this information.

Not only can competition provide a basis of comparison, which enables the design of reward structures that can simultaneously provide a high level of incentives with relatively low level risk; but compensation schemes based on relative performance have the further advantage of automatically adjusting incentives to changes in the economic environment. (We refer to this as "built-in flexibility.") In a first best world, with perfect information concerning the nature of the technology (but where it is still costly to monitor individuals' activities), the compensation scheme would vary from time to time as the environment changed. Such changes in the compensation scheme are costly to implement and the information required to do so is seldom available. When a task is easier, the individual's rewards for performing the task should be reduced. If pay is based on *relative* performance, although all individuals perform better (when they exert the same level of effort), their compensation is automatically adjusted. Thus, teachers frequently grade on the curve and a significant fraction of the pay of successful salesmen often consists of bonuses based on relative performance.

I. Competition and Compensation

To see more clearly exactly how competition can provide the basis of the design of a better compensation scheme, we consider a simple example. Assume the government wishes to develop a bomber. Neither the government nor the potential developers know how much it should cost to build the bomber. With a fixed-fee contract (where the amount received by the developer is independent of his costs), the contractor will require a large risk premium to compensate for the large risk he must bear. The government can reduce the risk by sharing in the costs; but to the extent that it does this, it also reduces the contractor's incentives to save on costs. There is an alternative contractual arrangement that may be superior. Assume there are a number of potential contractors of the same ability, so that the costs faced by one firm (at any particular level of effort) will be identical to those faced by another firm. Assume the government lets out two contracts; it promises to pay firm A a fixed amount *plus*

*Harvard University and Princeton University, respectively. We are indebted to Steve Salop, Sherwin Rosen, Ed Lazear, Felix Fitzroy, Joe Farrell, and Oliver Hart for helpful conversations. Financial support of the National Science Foundation is gratefully acknowledged.

whatever it costs firm B to produce the bomber; and conversely for firm B. Since their costs are perfectly correlated, this scheme eliminates all risk. At the same time, the scheme has perfect incentives: if firm A can, by exerting extra effort, reduce costs it gets to keep the savings. Having two separate firms provides us with information which simply would not be available otherwise. This information allows the implementation of a compensation scheme with lower risk and better incentives. There is a cost: the government has had to pay for duplicate research expenditure, but this may still be less expensive than the alternative.[1]

A. *General Theorems on the Optimality of Contests and Relative Performance Schemes*

A natural question to raise at this point is, are there circumstances in which an appropriately designed contest or relative performance scheme can attain a first best allocation, that is, the same level of effort in every state as would obtain if information about the state were freely available and effort were costless to monitor, without the worker needing to bear any risk.

We consider a general structure in which the individual's output Q_i (assumed to be observable) is a function of a common random variable θ, his effort μ_i, and an idiosyncratic random variable ε_i; for simplicity, we assume a linear relationship of the form:[2]

$$(1) \quad Q_i = \mu_i\theta + \varepsilon_i;\ E\varepsilon_i = E\varepsilon_i\varepsilon_j = E\varepsilon_j\theta = 0.$$

The worker observes θ before he decides on effort; but the manager-owner of the firm can neither observe θ, μ_i, nor ε_i; he can only observe the output of each of his workers. We now show that if the distribution of ε_i is compact (as a normalization, we assume it ranges from -1 to $+1$), then with only two workers, it is possible to design an incentive structure which attains the first best. (This result is stronger than that of the earlier example, since the technology of the two individuals is not perfectly correlated.)

We assume, for simplicity, that workers have additively separable utility functions of the form

$$(2) \qquad U(Y) - V(\mu),$$

where Y is income. With perfect information, a first best allocation requires $Y = \bar{Y}$ (individuals obtain perfect insurance) and

$$(3) \qquad \theta U(\bar{Y}) = V'(\mu).$$

(The marginal rate of substitution between leisure and goods equal the marginal rate of transformation.) The solution to (3) we refer to as the optimal level of effort, and denote by $\mu^*(\theta)$. Assume that after the individual has observed θ, but before he has allocated his effort (and, in particular, before the output has been produced), the individual is asked to announce a "goal." He is told if he comes within one unit of making his goal, he will have a given income; if he fails to come within a unit of meeting his goal, he receives minus infinity (or a suitably large punishment). The higher the goal he sets, the higher the income he will receive, provided he attains it. Finally, he is told that his pay will depend on the announcement of others as well; they observe, of course, exactly the same common random variable that he does. We now show that there exists a compensation structure of the form indicated which provides perfect incentives and eliminates all risk.

Consider the compensation scheme which pays the ith individual

$$(4) \qquad Y^i = \frac{\phi(\hat{\theta}^i)}{U'(\bar{Y})} - \frac{\phi(\hat{\theta}^{-i})}{U'(\bar{Y})} + \bar{Y} \quad \text{if} \quad Q_i \geq \mu^*(\hat{\theta}^i)\hat{\theta}^i - 1,$$
$$-\infty \text{ otherwise;}$$

[1] With only two firms, there is the further danger of collusion. As the number of firms increases, the likelihood of collusion may decrease, but the excessive waste from duplication increases. In other contexts, however, such as natural monopolies with average cost curves only slightly declining, the loss in efficiency from having several competitors may be slight.

[2] What is crucial about this specification is that the common random variable affects the marginal productivity of labor; as a result, in the first best allocation, effort will change from state to state. Since one of the issues with which we are concerned is the extent to which compensation schemes provide the appropriate incentives to change the level of effort in response to changes in circumstances, it is important that the model analyzed have this feature.

where $\hat{\theta}^i$ is the ith individual's announcement of θ, $\bar{\theta}^{-i}$ is the average of the announcements of other individuals. If each contestant tells the truth, the individual faces no risk, since $\phi(\hat{\theta}^i)=\phi(\bar{\theta}^{-i})$. Moreover, the individual will, if he announces θ, always make his target; he will choose his level of effort so that in the worst event ($\varepsilon=-1$), he just makes it. The return to announcing a higher value of θ is then

$$(5) \qquad U'(\bar{Y})\phi'/U'(\bar{Y})-V'd\mu/d\hat{\theta}.$$

To guarantee meeting the quota,

$$(6) \quad \mu^*(\hat{\theta})\hat{\theta}=\mu\theta,$$

$$(7) \quad d\mu/d\hat{\theta}=(\hat{\theta}/\theta)(d\mu^*/d\hat{\theta})+\mu^*(\hat{\theta})/\theta.$$

Hence if

$$(8) \qquad \phi'=V'[d\mu^*/d\theta+\mu^*(\theta)/\theta],$$

equation (5) will equal zero when $\hat{\theta}=\theta$. By integrating (8), we obtain a ϕ function which, when used in the compensation scheme (4), provides perfect incentives and eliminates all risks.

Though the use of targets in conjunction with a relative performance scheme can, under some idealized conditions, provide the basis of an extremely effective compensation system, it has its limitations: if, for instance, individuals obtain information about their idiosyncratic random variable at different times, or if the support of the idiosyncratic random variable differs for different individuals or is not known by the employer. (If individuals know ε before the announcement occurs or if the announcement must be made before θ is known, the announcement conveys no information additional to that which is conveyed by observing Q.)

B. *Further Results on Contests*

In our forthcoming article, we investigate the design of compensation schemes which do not employ announcements. We have shown how contests can be designed to provide the first best level of effort in every state of nature. In general, such contests do impose a greater risk on the individual than he would bear if monitoring effort were costless. There are two conditions under which contests may attain a first best outcome: (i) if the agents are risk neutral; or (ii) when there are a large number of contestants. Although generalized relative performance schemes include, as special cases, individualistic schemes (compensation schemes where pay depends only on the individual's own performance) and thus must be at least as good as such schemes, determining circumstances under which simple relative performance schemes (such as contests) do better than the simple kinds of individualistic schemes (such as piece rates) often found in practice is a far more difficult question. We show that contests will be preferred to (even nonlinear) individualistic schemes when the risk associated with the common environmental variable is large (relative to that of the idiosyncratic random variable). This is a theme to which we shall return later.

II. Markets and Competition

Markets provide reward structures which have some of the properties of contests and relative performance schemes. The exact nature of the compensation scheme provided by the market for the owner-manager of a firm depends critically on the nature of the production technology and the market equilibrium. Consider, for instance, two firms engaged in cost reducing *R&D*. If the production technology is constant returns to scale, with marginal cost c, the profits of the ith firm will depend on his costs and those of his rival: $\pi_i(c_1,c_2)$. The profit function will differ markedly depending on whether the duopoly equilibrium is best described as a Nash-Cournot quantity setting equilibrium or as an Edgeworth-Bertrand price setting equilibrium. This is an example of a relative performance compensation scheme. But while our earlier study considered the *design* of a compensation scheme, here we are concerned with *describing* the consequences of the compensation scheme which is always *implicit* in the market equilibrium.

A. *Market Reward Structures Have the Property of Built-In Flexibility*

If the cost functions facing different firms are correlated (there is a common environmental variable affecting all firms) when costs are low, price is low, and therefore profits will be less variable and rewards more commensurate with the difficulty of the task than in a noncompetitive environment. To see this, consider the following modification of the previous example: let the costs of production be

$$c_i = k - \theta\mu_i, \tag{9}$$

where μ_i is the (unobservable) level of managerial effort. There are a large number of firms, sufficiently large that they act as price takers. The "net" profit of the owner-manager $\hat{\pi}_i$ taking into account the utility cost of managerial effort, is[3]

$$\begin{aligned}\hat{\pi}_i &= [P - k + \mu_i\theta - V(\mu_i)]Q_i \\ &= \pi_i - V(\mu_i)Q_i,\end{aligned} \tag{10}$$

where Q_i is the output of the ith firm and P is the price of output. In competitive equilibrium

$$P = \underset{\{\mu_i\}}{Min}\{k - \mu_i\theta + V(\mu_i)\}; \tag{11}$$

and

$$\hat{\pi}_i = 0, \text{ for all producing firms.} \tag{12}$$

Thus, competition *forces* each of the competitors to expend the correct amount of effort at cost-reduction activities; and it does this in such a way as to eliminate all variability in net profits (they are identically zero).

An owner-manager monopolist would have the correct incentives for cost minimization, but would face considerable variability in his profits.

B. *The Consequences of the Separation of Ownership and Management*

But markedly different results obtain in the comparative performance of managers who do not own all the resources which they manage. Assume that there is a competitive supply of managers. The equilibrium in the competitive market remains as described above. Since there is no risk, the optimal contract entails the manager receiving 100 percent of the profits at the margin. (Effectively, the identity of ownership and control is an endogenous characteristic of this economy.) But consider the problem of the monopolist attempting to hire a manager for his enterprise. He wishes to choose a contract which maximizes his own expected utility, subject to the constraint of being able to hire the manager. For simplicity, we assume the manager is risk averse, but the owner is risk neutral.[4] Then the optimal contract will entail some risk sharing by the original owner. (The interests of owner and manager do not coincide; the separation of the two is again an endogenous feature of this economy.) If we limit ourselves to simple linear contracts, then the pay of the manager will be of the form $\alpha\pi_i + \beta$, assuming that profits of the firm are observable, but the input of the manager is not. (β is a fixed fee, which may be either negative or positive, but plays no essential role in the subsequent analysis.) Three consequences follow immediately from the fact that it is optimal for $0 < \alpha < 1$.

1) Managers will not expend the efficient amount of resources on cost reduction; they will set $\alpha\theta = V'(\mu)$.

2) Managers will not adjust their effort to changes in circumstances as much as they would in the competitive regime; when it is easier to reduce costs, they effectively enjoy some of the benefits of the increased ease in greater leisure (to a greater extent than this would be true in a competitive economy): $d\mu/d\theta = \alpha/V'' < 1/V''$. This phenomenon is sometimes referred to as managerial slack. The argument for why noncompetitive environments may experience managerial slack is perhaps even stronger than we have put it here. The owner of the firm may have knowledge about the "normal" state of nature; the contract thus may implicitly specify a normal level of effort, and a normal expected return. When the level of effort required to attain

[3]In this formulation, the effort expended is proportional to the level of output; there are neither increasing nor decreasing returns to scale in managerial technology. This assumption is made to avoid the difficulties which arise in comparing economies in which the number of managers (as opposed to aggregate managerial effort) is relevant.

[4]All that is really required is that the manager be more risk averse than the original owner.

this "normal level" is greater, the managers have an incentive to present evidence to that effect; while if the level of effort required is less than this normal level, the managers have no incentive to present that information. The natural asymmetries of information give rise to an asymmetry in response to unusually good and unusually bad states.

3) Managers still have to bear some risk.

C. *Imperfect Competition*

In the case of competitive economies, all the relevant information is embodied in the price; the owner of the firm does not have to base his manager's pay on the observed costs of his firm, say, relative to that of other firms. In the case of imperfectly competitive economies (for example, duopolies), the owner may wish to employ an incentive scheme which makes use of some of this detailed information, if it is available. Indeed, a slight modification of the first example given in this paper shows that it is possible with only two firms to design managerial incentive schemes in which the manager bears no risk yet has perfect incentives.

D. *Correlations Between Firms' Costs*

The success of the relative performance schemes analyzed in the previous section was based on the fact that all firms faced identical cost functions. Assume that there are two firms, each of which has a choice of two technologies (or any linear combination of the two). If a firm devotes λ of its resource to technology i, its cost function per unit output will be $F(\lambda, \theta_1, \theta_2)$ (where θ_i are random variables). Define λ^* as the mixture which minimizes the expected costs, $EF_\lambda(\lambda^*, \theta_1, \theta_2) = 0$. Assume that the manager's pay depends on the difference between his costs and that of his rival. Clearly, if one firm immitates his rival, then costs are identical, and there is no risk. But if $\lambda \neq \lambda^*$, the manager can move λ towards λ^*, and decrease mean costs, and thus increase his pay. Though this increases his risk, for small deviations he acts in a risk neutral manner; thus *the Nash equilibrium entails an efficient choice of techniques and no risk.*

There are three important qualifications to this result. First, assume firms can only choose technique 1 or technique 2. The mean return with technique 2 is higher. If firm A chooses technique 1, the B manager's risk by choosing technique 2 may be so much higher that he isn't sufficiently compensated by the increased mean. Thus, there may be an equilibrium in which both firms choose technique 1, and an equilibrium in which both firms choose technique 2. One of these may Pareto dominate the other.

Second, if one firm has a comparative advantage in technique 1, and the other firm in technique 2, then each firm will not choose the technique which minimized its expected costs, but rather will choose a technique which is somewhere between the cost-minimizing technique and the technique chosen by the other firm.

Third, if there is idiosyncratic risk, then even when the two firms imitate each other, risk is not eliminated; still, if the two firms face the *same* stochastic technology, the only equilibrium is that where costs are minimized.

E. *The Anarchy of the Market Place: Excessive Competition*

While the stories we have told here have pictured competition as reducing the risks faced by businessmen, businessmen often complain that unbridled competition forces them to bear an excessive amount of risk; they have, accordingly, often called upon the government to help regulate (stabilize) the market place. Some of these pleas are simply blatant attempts to cartelize the market, and to reap the monopoly rents which result. On the other hand, when the idiosyncratic risk is large, the variability in profits may be large.

In such situations, in our earlier studies of contests with risk-averse agents, relative performance schemes did not work well, since they imposed an excess amount of risk on the contestants. Similarly, in our earlier example of the development of the bomber, if the two researchers face different cost functions, then basing the compensation of one researcher on his performance relative to that of another imposes an additional source of risk. The market imposes similar risks on managers, even when the compensation scheme is not based directly on relative performance, but on profits; because profits will, to a large extent, reflect differences in the costs of the firm relative to its competitors (or differences in relative performance in some other dimension, such as quality or marketing). To attempt to alleviate these risks by making pay less dependent on performance is likely simultaneously to ameliorate incentives.

III. Conclusions

This paper has attempted to delineate a central role that competition plays: it allows the development of compensation schemes where pay is based on relative performance. Such compensation schemes have risk sharing, incentive, and built-in flexibility properties which make them superior to the best (individualistic) schemes which can be designed which do not make use of such information. The reward structures provided in competitive markets are, implicitly, related to relative performance. This provides an additional reason that competitive economies perform better than monopolies, a reason which is quite distinct from the loss in consumer's surplus arising from the monopolist's reduction in output. In particular, we have formulated a model in which monopolies are less efficient and less adaptable, and there is more managerial slack than in competitive economies. (In spite of the widespread belief that monopolies are less efficient than competitive firms—including the intertemporal inefficiencies arising from inadequate allocation of resources to *R&D*—in traditional neoclassical models, there is no managerial slack and monopolies are perfectly efficient.) We have indicated that there are limits to the extent to which the market may reduce risks: in some cases, competition may effectively increase it. An examination of the full consequences of our observations, including an investigation of the constrained optimality of the economy and the implications of our analysis for policy, must await another occasion. What should be apparent is that the perspectives into the functions of competition in market economies arising from the approach taken here stand in market contrast to those provided by the traditional competitive paradigm.

REFERENCES

Lazear, Edward P. and Rosen, Sherwin, "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, October 1981, *89*, 841–64.

Nalebuff, Barry J. and Stiglitz, Joseph E., "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, forthcoming.