

2

Economic Theories of Nonprofit Organization

HENRY HANSMANN

Serious work on the economics of the nonprofit sector began only in the early 1970s. This timing probably reflects, in part, the recent growth in the size and scope of the nonprofit sector. Until the 1950s, the sector was largely composed of traditional charities that received a substantial portion of their income from philanthropic contributions. Consequently, economic theorizing about the nonprofit sector, to the extent that it was undertaken at all, focused primarily on philanthropic behavior (for example, Dickinson 1962).

By the late 1960s, however, the character of the nonprofit sector had begun to change noticeably, its structure and performance assuming obvious importance for public policy. This change was most conspicuous in health care, particularly in the hospital industry. The implementation of Medicare and Medicaid in 1965 completed a process of evolution through which nonprofit hospitals were freed from dependence on charitable contributions and came to be potentially profitable institutions deriving virtually all their revenue from patient billings. Large publicly held business corporations owning chains of for-profit hospitals emerged for the first time. Simultaneously, hospital cost inflation appeared as a serious policy problem. The hospital industry in general, and the role and behavior of nonprofit hospitals in particular, thus became the subject of serious economic inquiry. It is not surprising, then, that the first efforts to develop economic models of nonprofit institutions focused almost exclusively on hospitals (for example, Newhouse 1970; Feldstein 1971; Lee 1971; Pauly & Redisch 1973).

A number of individuals provided helpful comments on an earlier draft, including Avner Ben-Ner, Eugene Fama, Estelle James, Michael Jensen, Richard Nelson, Susan Rose-Ackerman, John G. Simon, Richard Steinberg, Burton Weisbrod, and Dennis Young; I am grateful to them all. If the work of these or other authors is nevertheless mis- or underrepresented here, the responsibility is mine. I am also grateful to Walter Powell for his thoughtful editing.

Change was conspicuous in other parts of the nonprofit sector as well, however. Higher education, for example, underwent enormous expansion in the 1950s and 1960s and then fell into serious financial difficulty in the early 1970s. The live performing arts exhibited the paradox of constant fiscal crisis in the midst of rapid growth.¹ The day-care and nursing home industries, which had scarcely existed before World War II, became enormous. These industries—and many others like them—were all characterized by a mix of nonprofit, for-profit, and governmental firms, thus raising questions as to the relative functions and behavior of these three types of organization. Moreover, because all these industries received large and growing public subsidies, an understanding of their underlying economics was of obvious relevance for purposes of policy. The resulting prominence of such industries has led, over the past fifteen years, to the development of a substantial body of work concerning the economics of the nonprofit sector in general.

The economic theories of nonprofit organization appearing in the literature can conveniently, if somewhat artificially, be divided into two types: theories of the *role* of nonprofit institutions and theories of their *behavior*. Theories of the first type address such questions as these: Why do nonprofit organizations exist in our economy? What economic functions do they perform? Why, in particular, are nonprofit firms to be found in some industries and not in others? Why, among those industries in which nonprofit firms are found, does their market share—vis-à-vis both for-profit firms and governmental firms—vary so radically from one industry to another?

Theories of the second type address such questions as these: What objectives are pursued by nonprofit organiza-

1. The fiscal problems of the performing arts were documented by Baumol and Bowen (1965, 1966) in work that helped bring particular attention to the economics of that industry.

tions? What are the motivations of managers and entrepreneurs in the nonprofit sector? How do nonprofit organizations differ in these respects from for-profit and governmental organizations? How does the productive efficiency of nonprofit organizations differ from that of for-profit and governmental organizations? In what ways are such differences attributable to the special characteristics of the nonprofit form?

Ultimately, of course, questions of role and questions of behavior cannot be separated. To understand why it is that nonprofit firms arise in one industry and not in another, one must understand something about the firms' characteristic behavior. Nevertheless, economic theories of nonprofit institutions have tended to focus primarily on only one or the other of these two broad areas of concern, and thus the division will be employed here as a means of organizing the literature.

In this survey I shall focus primarily on firms organized as "true" nonprofits—that is, firms that are formally organized as either nonprofit corporations or charitable trusts. These organizations are all characterized by the fact that they are subject, by the laws of the state in which they were formed, to a constraint—which I shall call the "nondistribution constraint"—that prohibits the distribution of residual earnings to individuals who exercise control over the firm, such as officers, directors, or members (Hansmann 1980, 1981d). Note that nonprofits are *not* prohibited from earning profits; rather, they must simply devote any surplus to financing future services or distribute it to noncontrolling persons. Theories of the nonprofit firm are, then, essentially theories of the way in which the presence of a nondistribution constraint affects a firm's role or behavior. I shall not deal here, except for purposes of comparison, with cooperatives (producer or consumer), which are discussed in chapter 24, or with mutual companies, such as mutual insurance companies or banks; such organizations are empowered to distribute net earnings to their members and thus are not formally subject to a nondistribution constraint. Also, except for purposes of comparison, I shall not discuss public enterprise, but shall rather confine myself to private nonprofits.

TYPES OF NONPROFIT ORGANIZATIONS

The organizations that populate the nonprofit sector are structurally rather diverse. For ease of reference, I shall adopt here a classification scheme offered elsewhere (Hansmann 1980) under which firms are distinguished according to (1) their source of income and (2) the way in which they are controlled.

Nonprofits that receive a substantial portion of their income in the form of donations will be referred to here as "donative" nonprofits; firms whose income derives primarily or exclusively from sales of goods or services will be called "commercial" nonprofits. The Red Cross is an example of the former; most nonprofit hospitals and nursing homes today would be in the latter category. The term *patrons* will be used to denote those individuals who are the ultimate

TABLE 2.1 A FOUR-WAY CATEGORIZATION OF NONPROFIT FIRMS

	Mutual	Entrepreneurial
Donative	Common Cause National Audubon Society Political Clubs	CARE March of Dimes Art Museums
Commercial	American Automobile Association Consumers Union ^a Country clubs	National Geographic Society ^b Educational Testing Service Hospitals Nursing homes

Source: Adapted from Hausmann 1980.

a. Publisher of *Consumer Reports*

b. Publisher of *National Geographic*

source of the organization's income. Thus, in a donative nonprofit the patrons are the donors, whereas in a commercial nonprofit they are the firm's customers. In the case of nonprofits that have both donors and customers, the term comprises both.

Firms in which ultimate control (the power to elect the board of directors) is in the hands of the organization's patrons will be referred to as "mutual" nonprofits. Other nonprofits—including, in particular, those in which the board of directors is self-perpetuating—will be called "entrepreneurial" nonprofits.

The intersections of these two two-way classifications yield four types of nonprofits: donative mutual, donative entrepreneurial, commercial mutual, and commercial entrepreneurial. Table 2.1 gives some examples of each type.

The boundaries between the four categories are, of course, blurred. Many private universities, for example, depend heavily on both tuition and donations for their income and thus are to some extent both donative and commercial. Also, university boards of trustees commonly comprise some individuals who are elected by the alumni (who are past customers and present donors) and some who are self-perpetuating, with the result that the universities cannot be categorized as clearly mutual or clearly entrepreneurial. The four categories are, then, simply polar or ideal types, offered for the sake of clarifying discussion.

THE ROLE OF NONPROFIT ORGANIZATIONS

Several theories have been advanced to date to explain the economic role of nonprofit organizations. These theories are sometimes competing and sometimes complementary.

The Public Goods Theory

The first general economic theory of the role of nonprofit enterprise was offered by Weisbrod (1974, 1977), who suggested that nonprofits serve as private producers of public

goods (in economists' sense of that term).² Governmental entities, Weisbrod argued, will tend to provide public goods only at the level that satisfies the median voter; consequently, there will be some residual unsatisfied demand for public goods among those individuals whose taste for such goods is greater than the median.³ Nonprofit organizations arise to meet this residual demand by providing public goods in amounts supplemental to those provided by government.⁴

Weisbrod's theory captures an important phenomenon. Many nonprofit firms provide services that have the character of public goods, at least for a limited segment of the public. This is conspicuously true, for example, of those donative nonprofits (such as the American Heart Association, the National Cancer Society, and the March of Dimes) that collect private donations to finance medical research. As originally presented, however, the public goods theory left two questions open. First, the services provided by many nonprofits do not seem to be public goods but rather appear to be private ones. This is true especially of commercial nonprofits, whose share of the nonprofit sector has increased impressively in recent years. For example, the appendectomy performed in a nonprofit hospital, the child care provided by a nonprofit day-care center, the education provided by a nonprofit preparatory school, the nursing care provided by a nonprofit nursing home, and the entertainment provided by a nonprofit symphony orchestra are all difficult to characterize as public goods in the usual sense. Second, Weisbrod's theory stops short of explaining why nonprofit, rather than for-profit, firms arise to fill an unsatisfied demand for public goods. What is it about nonprofit firms that permits them to

2. A public good, in the economists' sense, is a good that has two special attributes: first, it costs no more to provide the good to many persons than it does to provide it to one, because one person's enjoyment of the good does not interfere with the ability of others to enjoy it at the same time; second, once the good has been provided to one person there is no easy way to prevent others from consuming it as well. Air pollution control, defense against nuclear attack, and radio broadcasts are common examples of public goods.

3. Logrolling and other devices, of course, often lead to establishment of government programs that cater to supramedian demands. Consequently, the median voter model should probably not be taken too literally here. Nevertheless, extremely intense or idiosyncratic demands for public goods are unlikely to be fully satisfied by governmental programs.

4. Weisbrod's theory has recently been illustrated and refined, with an emphasis on welfare considerations, in a formal model developed by Weiss (1986). In that model, Weiss demonstrates that, while a Pareto superior allocation of resources might well result when high demanders of a public good supplement public production with privately financed production, this is not a necessary result; it is possible that, even where there is cooperation between the public and private providers of the public good, the welfare of the high demanders will be lower when they can undertake supplemental private production than when they cannot. The reason for this result is that the low demanders, foreseeing the incentive for the high demanders to supplement public production with their own private production, might vote to support a substantially lower level of public production than they would otherwise and free ride on the private production, which will, as a consequence, be larger (and costlier to the high demanders) than it would be otherwise.

serve as private suppliers of public goods when proprietary firms cannot or will not?

The Contract Failure Theory

The elements of a somewhat different theory of the role of nonprofits were set forth in an essay on day care by Nelson and Krashinsky (1973; Nelson 1977), who noted that the quality of service offered by a day-care center can be difficult for a parent to judge. Consequently, they suggested, parents might wish to patronize a service provider in which they can place more trust than they can in a proprietary firm, which they might reasonably fear could take advantage of them by providing services of inferior quality. The strong presence of nonprofit firms in the day-care industry, they argued, could perhaps be explained as a response to this demand. Similar notions had been hinted at in an earlier essay on health care by Arrow (1963), who suggested in passing that hospitals may be nonprofit in part as a response to the asymmetry in information between patients and providers of health care.

The theme advanced by Nelson and Krashinsky was fleshed out and generalized in an article by Hansmann (1980), where it is argued that nonprofits of all types typically arise in situations in which, owing either to the circumstances under which a service is purchased or consumed or to the nature of the service itself, consumers feel unable to evaluate accurately the quantity or quality of the service a firm produces for them. In such circumstances, a for-profit firm has both the incentive and the opportunity to take advantage of customers by providing less service to them than was promised and paid for. A nonprofit firm, in contrast, offers consumers the advantage that, owing to the nondistribution constraint, those who control the organization are constrained in their ability to benefit personally from providing low-quality services and thus have less incentive to take advantage of their customers than do the managers of a for-profit firm.⁵ Nonprofits arise (or, rather, have a comparative survival advantage over for-profit firms) where the value of such protection outweighs the inefficiencies that evidently accompany the nonprofit form, such as limited access to capital and poor incentives for cost minimization (see below). Because this theory suggests, in essence, that nonprofits arise where ordinary contractual mechanisms do not provide consumers with adequate means to police producers,

5. The emphasis in the text here is on the role of the nondistribution constraint as a direct bar to opportunistic conduct on the part of a nonprofit's managers. The nondistribution constraint might also, however, serve the same function through indirect means by screening for managers who place an unusually low value on pecuniary compensation and an unusually high value on having the organization they run produce large quantities of services or services that are of especially high quality. A simple model along these lines is offered by Hansmann (1980, Appendix). Data that lend some support to such a theory are presented, in the context of public interest law firms, by Weisbrod (1983). Young (1983; this volume, chap. 10) discusses screening for entrepreneurs at length, exploring a rich set of personal characteristics for which nonprofit firms might serve as a screen.

it has been termed the "contract failure" theory of the role of nonprofits (Hansmann 1980).

Donative Nonprofits

Although the contract failure theory has its roots in the work of authors (Arrow 1963; Nelson & Krashinsky 1973) who are primarily concerned with the role of commercial nonprofits, its most obvious application is in fact to donative nonprofits (Hansmann 1980; Thompson 1980; Fama & Jensen 1983a).⁶ A donor is, in an important sense, a purchaser of services, differing from the customers of commercial nonprofits (and of for-profit firms) only in that the services he or she is purchasing are either (1) delivery of goods to a third party (as in the case of charities for the relief of the poor or distressed) or (2) collective consumption goods produced in such aggregate magnitude that the increment purchased by a single individual cannot be easily discerned. In either case, the purchaser is in a poor position to determine whether the seller has actually performed the services promised; hence the purchaser has an incentive to patronize a nonprofit firm.

For example, individuals commonly contribute to CARE in order to provide food to malnourished individuals overseas. A for-profit firm could conceivably offer a similar arrangement, promising to provide a specified quantity of food to such people in return for a contribution of a given amount. The difficulty is that the purchaser (donor), who has no contact with the intended beneficiaries, has little or no ability to determine whether the firm performs the service at all, much less whether the firm performs it well. In such circumstances, a proprietary firm might well succumb to the temptation to provide less or worse service than was promised.

The situation is similar with public goods. If an individual contributes to, say, a listener-sponsored radio station, then, unlike the situation with CARE, she is at least among the recipients of the service and can tell whether it is being rendered adequately. What she cannot tell is whether her contribution of fifty dollars in fact purchased a marginal increment of corresponding value in the quantity or quality of service provided by the station or simply went into somebody's pocket. A for-profit firm that operated such a radio station would have an incentive to solicit payments far in excess of the amounts necessary to provide their programming. In situations such as these, the nonprofit organizational form, owing to the nondistribution constraint, offers the individual some additional assurance that her payment is in fact being used to provide the services she wishes to purchase.⁷

6. Fama and Jensen (1983a, 342) seek to distinguish their briefly sketched theory of donative nonprofits from that offered by Hansmann (1980). The difference, however, is difficult to discern.

7. The same arguments presumably apply to situations in which individuals donate their own labor or other goods or services in kind. If a volunteer were to donate his services to a for-profit hospital, for example, he might find it difficult to determine whether the result was in fact an equivalent increase in the services rendered by the hospital without a corresponding increase in price or whether, alternatively, the owners

As this example suggests, the contract failure theory is complementary to the public goods theory described above. Indeed, the public goods theory can be seen as a special case of the contract failure theory. For the reasons described by Weisbrod, there may be residual demand for public goods—such as noncommercial broadcasting—that is unsatisfied by government. Yet even if individuals are prepared to overcome their incentive to free ride and will donate toward financing of a public good, they will have an incentive to contribute to a nonprofit rather than a for-profit firm because of the monitoring problems just described.

We have been proceeding here on the implicit assumption that the donors to the nonprofit firm will be private persons. In many cases, however, the government is an important donor, and in some cases it is the only donor. Sometimes government donations are direct, as in the case of grants made by the National Endowment for the Arts to nonprofit performing arts companies or (now discontinued) Hill-Burton Act capital grants to nonprofit hospitals. In other instances, government donations are indirect, as in the case of tax exemption or reduced postal rates for nonprofits. Regardless of the way in which such donations are made, however, the government is often subject to the same problems of contract failure that face a private donor: it cannot easily determine directly whether its donation is being devoted in its entirety to the purposes for which it was made. Consequently, the government, like a private donor, has an incentive to confine its subsidies to nonprofit rather than for-profit firms, and it commonly does so. And this, in turn, creates further demand for the services of nonprofit firms.

Commercial Nonprofits

The contract failure theory can also help explain the role of commercial nonprofits. The types of services that commercial nonprofits commonly provide—such as day care, nursing care, and education—are often complex and difficult for the purchaser to evaluate. Further, the actual purchaser of the service is often not the individual to whom the service is directly rendered and thus is at a disadvantage in judging the quality of performance: parents buy day care for their children, and relatives or the state buy nursing care for the elderly. Finally, the services provided by commercial nonprofits are commonly provided on a continuing long-term basis, and the costs to the recipient of switching from one firm to another are often considerable. Consequently, purchasers are to some extent locked in to a particular firm once they have begun patronizing it, and thus the firm, if unconstrained, is in a position to behave opportunistically.⁸ For all

used him as a replacement for labor they would otherwise have paid for and thus simply increased their own profits. Consequently, individuals generally volunteer their services only to nonprofit organizations.

8. Ellman (1982) offers useful terminology for making distinctions between different forms of contract failure. On the one hand, there are problems of "quality monitoring," which involve situations in which the consumer can determine whether performance took place but has

these reasons, patrons might have an incentive to patronize a firm subject to a nondistribution constraint as additional protection against exploitation.⁹

Where commercial nonprofits are concerned, contract failure is presumably a less serious problem than with donative nonprofits. Consequently, it is not surprising that commercial nonprofits nearly always share their market with for-profit firms providing similar services. For example, roughly 20 percent of all private hospitals, 60 percent of all private day-care centers, and 80 percent of all private nursing homes are for-profit enterprises (Hansmann 1985a). If the contract failure theory explains the presence of commercial nonprofits in these industries, then the presence of both types of firms may reflect some division of the market: patrons who are reasonably confident of their ability to police the quality of the services they receive patronize the for-profit firms, whereas those who are less confident in this respect patronize the nonprofit firms, perhaps paying a premium for the service on account of the productive inefficiencies associated with the nonprofit form.

Although this theory is plausible as applied to most types of commercial nonprofits, it does not, interestingly, seem particularly persuasive when applied to hospitals, which constitute (in terms of GNP) the largest class of nonprofit institutions. There are two reasons for this. First, the hospital itself does not provide the patient care services that are the most sensitive and difficult to evaluate—namely, the services of the attending physicians. Rather, the physicians are usually independent contractors who deal separately with the patient. The hospital itself is largely confined to providing relatively simple services such as room and board, nursing care, and medicines. Second, the patient herself does not order the hospital services she receives; rather, they are ordered and monitored for her by a skilled and knowledgeable purchasing agent, namely, her physician. Consequently, it is not at all obvious that the nondistribution constraint offers the hospital patient any special protection that she would clearly be lacking without it.

Why, then, are hospitals nonprofit? It may be that, if we allow for a little historical lag, the contract failure theory in fact explains it. Until the end of the nineteenth century, hospitals were almost exclusively donative institutions serving the poor; the prosperous were treated in doctors' offices or in their own homes. The nonprofit form was therefore

efficient for the reasons of contract failure discussed above with respect to donative institutions in general. Then, however, a revolution in medical technology turned hospitals into places where people of all classes went for treatment of serious illness. Subsequently, the development of public hospitals took from the nonprofit hospitals much of the burden of caring for the poor. Finally, the spread of private, and more recently public, health insurance made it possible for the great majority of patients to pay their hospital bills without the aid of charity. The result is that today—which is to say, since the appearance of Medicare and Medicaid in 1965—most nonprofit hospitals have become more or less pure commercial nonprofits, receiving no appreciable portion of their income through donations and providing little or no charity care. The continuing predominance of nonprofit firms may simply be the consequence of institutional lag and of the various subsidies and exemptions that continue to be available to nonprofit but not to for-profit hospitals (Hansmann 1980, 866–68; Clark 1980). Indeed, since the late 1960s there has been substantial entry of large for-profit firms into the industry.

Contract Failure as an Agency Problem

In essence, the contract failure theory views the nonprofit firm as a response to agency problems. In situations like those just described, the purchaser (donor) is in the role of a principal who cannot easily monitor the performance of the agent (here, the firm) that has contracted to provide services to her. Consequently, there is a strong incentive to embed the relationship in a contractual framework, or "governance structure" (Williamson 1979), that mitigates the incentives of the agent to act contrary to the interests of the principal. The nonprofit corporate form, with its nondistribution constraint, serves this purpose.

It is worth noting, in this respect, that the relationship between the donors to a donative nonprofit firm and the managers of such a firm is analogous to the already much analyzed agency relationship between the shareholders in a publicly held business corporation and the managers of the corporation (see, for example, Jensen & Meckling 1976; Fama & Jensen 1983a, 1983b). The purchaser of a share of newly issued stock in a widely held business corporation, like a donor to CARE, is in no position to see for himself how the management is using the corporation's funds in general, much less what use is being made of his own marginal contribution to the corporation's assets. The shareholder is simply turning over funds to the corporation's management to be combined with other such funds and used however management chooses, subject only to the general constraints that (1) management will seek to obtain a reasonable rate of return for the shareholders on their contributed funds, and (2) management will take for itself no more than reasonable compensation for services rendered. These two constraints are precisely parallel to those that bind the management of a nonprofit firm, differing only in that, in the case of the nonprofit enterprise, the first constraint is replaced by one

difficulty judging the quality of the goods or services delivered. On the other hand, there are problems of "marginal impact monitoring," which involve situations in which the consumer can judge the quality of services performed by the firm but has difficulty determining whether the quantity or quality of services produced is higher than it would have been if he had not contributed. Commercial nonprofits presumably arise primarily in situations in which quality monitoring is a problem. Listener-supported broadcasting, in turn, presents a clear problem of marginal impact monitoring, and charities like CARE seem to involve both quality monitoring and marginal impact monitoring problems.

9. See Williamson (1979) for analysis of other contractual and organizational devices for mitigating opportunistic behavior in long-term complex transactions.

calling for management to devote the corporation's funds to the purposes specified in its charter. As in the case of the nonprofit firm, these two constraints are imposed upon the management of a business corporation by the terms of the corporation's charter and the legal framework in which that charter is embedded. Moreover, it is the second of these two constraints, which is effectively a nondistribution constraint, that has the more bite of the two; the very forgiving "business judgment rule" that the law applies to the decisions of corporate management makes the first constraint a largely nominal obligation.

In short, in the business corporation as in the nonprofit corporation, the only real contractual check on the behavior of the corporation's management is embodied in the nondistribution constraint imposed on management by the corporation's charter. The difference between the two types of corporations lies primarily in the class of individuals in whose favor the nondistribution constraint runs: the patrons (customers) or the investors of equity capital.

There are, to be sure, some important differences in the way the obligations of management are enforced in these two types of firms. The patrons of a nonprofit firm lack the mechanism of the derivative suit to enforce the nondistribution constraint against management. Rather, in most states only the state attorney general and/or the tax authorities have the right to bring suit in case of managerial malfeasance. Further, only in the case of mutual nonprofits do the patrons have any voting rights. And finally, patrons in both commercial and mutual nonprofits lack the advantage of a market for corporate control¹⁰ as a means of sanctioning management.

Easley and O'Hara (1983) have sought to capture the contract failure theory in a formal model, treating it as a principal/agent problem. In this model the manager of the firm (the agent) has sole knowledge of the firm's level of output, the firm's cost function, and the extent to which the manager's own effort exceeds some minimal observable level; a customer of the firm (the principal) knows none of these things but can only verify that the manager has expended the minimal level of effort. The authors interpret a for-profit firm as one that contracts with the customer only in terms of price and output, the firm (or rather its manager/owner) promising to produce a given level of output in return for a given price. In this model, such a for-profit firm will produce no output (since the customer cannot observe it); rather, the manager will simply pocket the whole pur-

chase price, expending no effort and using none of the purchase price to cover other costs of production.

A nonprofit firm, in turn, is interpreted as a contract that specifies (1) the amount of compensation to be received by the manager, (2) that the remainder of the purchase price is to be devoted to other costs of production, and (3) that the manager is to expend at least the minimal observable level of effort—all of which features of the contract are assumed to be verifiable by the customer. Easley and O'Hara show that this contract will result in a positive level of output in those cases in which the manager's minimal observable effort level (together with the other inputs acquired by the firm with that part of the purchase price that does not go to the manager as compensation) is sufficient to produce such a positive level of output. Thus, in this model, the nonprofit firm performs more efficiently than the for-profit firm, since the nonprofit produces a positive level of output in at least some cases, whereas the for-profit firm always produces zero output.

What is most interesting about this result is the nature of the assumptions necessary to establish it. In order for the nonprofit form to perform more efficiently than the for-profit form when output is unobservable, it is not sufficient simply to put a verifiable cap on the manager's compensation; the manager's level of effort and her use of the remainder of the purchase price must also be observable. In short, in this model the nonprofit firm involves policing inputs rather than outputs. If inputs were also completely unobservable, the nonprofit form would do no better than the for-profit form; both would always produce zero output.

This model probably captures the essential features of reality. In effect, the nonprofit corporate form is a device whereby the state (via the tax and corporation law authorities), on behalf of the customer, undertakes a certain minimal level of policing of inputs and of managerial compensation.

Note that an entrepreneur (or manager, in the model) will presumably submit herself and her firm willingly to such policing when she realizes that otherwise she will receive no patronage at all (and when the return permitted her by the nonprofit form is greater than her opportunity cost). It is in this sense that the nonprofit form is essentially a contract voluntarily entered into between a firm (more accurately, those in control of the firm) and its customers.

Empirical Tests

Weisbrod and Schlesinger (1986) have undertaken empirical work to test the contract failure theory as applied to commercial nonprofits, using data on Wisconsin nursing homes. These authors used consumer complaints to regulatory authorities as a proxy for quality of service. They found that nonprofit nursing homes are the subject of significantly fewer complaints than their proprietary counterparts, and they interpret this result as tentative support for the conclusion that nonprofit homes are less likely than proprietary homes to exploit the information asymmetry that exists between the homes and their consumers. These results must, however, be

10. The "market for corporate control" refers to the process whereby one business corporation can acquire effective control of another (the target corporation) by purchasing a majority of the target's stock, generally through a tender offer to the target's shareholders. Where the target corporation has been managed inefficiently in the past, such an acquisition opens the way for the acquiring corporation to replace the old management of the target with new, more effective managers, thus raising the value of the target's stock and producing a profit for the acquiring corporation. The mere threat of such a takeover, it has been suggested, may be an important incentive for the management of business corporations to perform with reasonable efficiency (Manne 1965).

interpreted with caution, since consumer complaints are a very indirect measure of quality of service, and the authors' regressions do not control for price or cost of service.¹¹

Indeed, it is not obvious that the contract failure theory implies that, in equilibrium, nonprofit firms will exhibit a higher quality/price ratio than their for-profit competitors. If, as suggested above, patrons sort themselves among the two types of firms according to their ability to police the quality of service they receive, one would in fact expect to find, *ceteris paribus*, a lower quality/price ratio in nonprofit firms (since patrons of nonprofits are paying a premium for the added protection they receive). Yet such an effect may be obscured in empirical data by the fact that nonprofit firms, but not for-profit firms, have the benefit of tax exemption and other explicit and implicit subsidies, and these will tend to create an offsetting reduction in the cost of service.

In any event, the contract failure theory is a theory of consumer expectations, not of actual performance. Individuals who are uncertain of their ability to monitor quality might patronize nonprofit firms in preference to for-profit firms in the belief that the nonprofits are most trustworthy, yet be mistaken in that belief. An incongruity between performance and consumer expectations that persisted over the long run would, however, require some explaining.

Some efforts have been made to test the contract failure theory as applied to commercial nonprofits by determining, through surveys, whether patrons in fact believe that commercial nonprofits are more to be trusted than their for-profit competitors. The results are thin and ambiguous, though arguably somewhat supportive of the contract failure theory (Newton 1980; Permut 1981; Hansmann 1981a).

Subsidy Theories

In most industries in which they are common, nonprofit firms benefit from a variety of explicit and implicit subsidies, including exemption from federal, state, and local taxes, special postal rates, financing via tax-exempt bonds, and favorable treatment under the unemployment tax system. It is often suggested that such subsidies are in large part responsible for the proliferation of nonprofit firms (for example, Fama & Jensen, 1983a, 344), particularly in those industries in which nonprofits compete with for-profit firms.

Given the structure and administration of these subsidies, however, there is reason to doubt that they have had much effect in determining the industries in which nonprofits have and have not developed. In general, the scope of the subsidies seems to have adjusted over the years to include the new industries into which nonprofits have proliferated, rather than vice versa (Hansmann 1980). On the other hand, it seems reasonable to expect that the presence of these subsidies has had an impact on the overall extent of nonprofit development in those industries in which such firms appear.

11. The regressions do, however, control for the certification level of the home—that is, whether the home has been certified as a "skilled," "intermediate," or "personal and residential" care facility.

An empirical study using cross-sectional (state-by-state) data on four industries in which nonprofit firms and for-profit firms compete—hospitals, nursing homes, private primary and secondary education, and postsecondary vocational education—in fact provides tentative evidence that the availability of state property, sales, and income tax exemptions has a significant effect in enhancing the market share of nonprofit firms vis-à-vis their proprietary competitors (Hansmann 1985a).

The Consumer Control Theory

There are some types of nonprofits—in particular, some types of mutual nonprofits—that do not seem to have arisen in response to contract failure.

For example, it appears that exclusive social clubs, such as country clubs, constitute a distinct exception to the contract failure theory (Hansmann 1980, 1986). In such organizations, the patrons seem as capable of judging the quality of services as they would at, say, a resort hotel. The nonprofit form is evidently adopted here simply as a means of establishing patron control over the enterprise. Such control serves the purpose of preventing monopolistic exploitation of the patrons by the owners of the firm. The source of such monopoly power in social clubs is the personal characteristics of the members of the club. A substantial part of the appeal of belonging to an exclusive club lies in the opportunity to associate with the other members, who presumably have qualities or connections that make them unusually attractive companions. Consequently, if such a club were for-profit, its owner would have an incentive to charge a membership fee high enough not just to cover costs but also to capture some portion of the value to each member of associating with the other members. That is, so long as individuals who would make equally desirable clubmates were insufficiently numerous to populate a number of competing clubs, the owner of a proprietary club could charge a monopoly price to each member for the privilege of associating with the other members. Thus the members as a group have an incentive to exercise control over the club themselves to avoid such exploitation. Exclusive social clubs, under this view, therefore play an economic role that has more in common with that of consumer cooperatives—which, as discussed below, typically seem to be formed to cope with problems of simple monopoly—than with that of other types of nonprofits.

Ben-Ner (1986) takes a broader view of the role of patron control, arguing that *most* nonprofit organizations are formed primarily in order to provide consumers with direct control over the firm from which they purchase goods or services. He points, in particular, to three possible circumstances in which consumers might desire to have direct control over a firm rather than simply exercise control via the market. The first is contract failure (asymmetric information about quantity or quality of output), although Ben-Ner focuses on consumer control as a means of eliminating the information asymmetry rather than on the nondistribution constraint as a means of

curtailing incentives for the firm to exploit that asymmetry. The second circumstance is that in which the firm is a monopolist and, although product quality is easily observable, there is a broad range of potential quality levels for the product, only one of which can be chosen. The problem here is that market signals alone may lead the firm to choose a quality level that appeals to marginal rather than average consumer evaluations of quality; direct consumer control could mitigate this problem. The third circumstance is that in which the firm produces price-excludable collective consumption goods. In such a case, consumer control might lead to a superior form of price discrimination, and thus higher aggregate welfare, than would control by profit-seeking investors.

Ben-Ner gives few examples of industries in which these factors constitute important sources of the demand for nonprofit, as opposed to for-profit, enterprise. Moreover, the few examples he does offer—such as the performing arts—may be better explained by other theories (see Hansmann 1981b). Consequently, although the factors examined by Ben-Ner may possibly play an important role in some industries, it is not obvious that they have broad application.

In developing his theories, Ben-Ner does not distinguish between nonprofit organizations and consumer cooperatives, but rather suggests that his theory explains the appearance of both types of firms. As the following section suggests, however, these two organizational forms generally seem to occupy distinct economic niches, and thus we need a theory of role that distinguishes between them.

Nonprofits versus Other Forms of Limited-Profit Enterprise

Nonprofits are not the only common form of profit-constrained enterprise. Privately owned public utilities typically operate under a form of price regulation designed to permit them no more than a competitive return on invested capital. Limited dividend companies, which are restricted by contract or statute to a stated maximum rate of cash return on equity, are common in the construction and operation of publicly subsidized housing. Producer and consumer cooperatives are constrained by the cooperative corporation statutes to pay a return on capital shares that does not exceed a specified percentage rate. And finally, cost-plus contracts, which provide for no more (and no less) than a stated rate of return to the seller, are common in situations such as defense procurement.

One might be tempted to suppose that all such forms of limited-profit enterprise, being so similar in form, must play similar economic roles. In fact, this is not the case. To be sure, limited-dividend companies do seem to occupy a role similar to that of donative nonprofits: in particular, they seem to be used by the government as a means of ensuring that public subsidies are passed through to housing consumers rather than accruing entirely to developers. Regulated utilities, however, are a response to the potential for pricing abuses that accompany natural monopoly, a role that nonprofits seem rarely to play. (Indeed, the industries in which

nonprofits are commonly found are almost all characterized by a substantial number of competing suppliers.) Consumer cooperatives also generally seem to represent a response to monopoly. Like public utilities, and in contrast to nonprofits, they usually sell only simple standardized goods and hence do not typically seem to arise as a response to contract failure (Hansmann 1980; Heflebower 1980). There are exceptions, however. For example, mutual life insurance companies—which are formally structured as consumer cooperatives—originally arose in large part as a response to contract failure in the insurance market (Hansmann 1985b).

Cost-plus contracts, in turn, commonly serve as a device for shifting risk to the purchaser when *both* parties face ex ante cost uncertainty. Easley and O'Hara (1984) argue that a particular form of cost-plus contract—the cost-plus-variable-fee contract—may arise not exclusively as a risk-sharing device but also or instead as a response to situations of information asymmetry in which producers know more about cost of performance than consumers do. They are careful to distinguish this situation, however, from the type of information asymmetry concerning quality of performance that seems to give rise to nonprofits.

Nonprofit versus Governmental Enterprise

As the preceding discussion suggests, most work on the role of nonprofit enterprise has focused on the choice of the nonprofit versus the for-profit form of organization. In particular, this has been true of the work that has sought to explain the development of nonprofits as a response to contract failure, subsidies, or a need for consumer control to counter monopoly power. Relatively little work has been done to date comparing and contrasting the role of nonprofit and governmental enterprise. This is unfortunate because nonprofit firms typically operate in industries in which the organization of firms as governmental entities is a serious alternative. In fact, in the United States, governmental firms have a significant share of the market in many industries in which nonprofits are common, including hospital care, nursing care, primary and secondary education, and postsecondary and vocational education. Moreover, many of the activities that in this country are performed in substantial part by nonprofits are performed in most other developed countries almost exclusively by governmental firms; health care, higher education, and the performing arts are conspicuous examples.

An important explanation for this gap in existing theory undoubtedly lies in the fact that contemporary economic theory offers a much more coherent view of the role of for-profit enterprise than it does of the role of governmental enterprise, and thus the proprietary form of organization offers a much firmer basis for comparison than does governmental organization. Nevertheless, there has been some useful work illuminating various aspects of the relationship of nonprofit and governmental enterprise.

To begin with, Weisbrod's work on the public goods theory (1974, 1977), discussed above, suggests that non-

profits tend to serve a gap-filling role vis-à-vis governmental enterprise, meeting some of the supramedian or idiosyncratic demand for public goods that is left unmet by government provision. This theory leads to the prediction that the market share of nonprofit versus governmental firms will be larger in those jurisdictions in which demand is unusually heterogeneous. Lee and Weisbrod (1977) have sought to test this implication with respect to hospitals using cross-sectional (state-by-state) U.S. data. In particular, they regressed nonprofit hospitals as a fraction of total nonprofit and governmental hospitals against various proxies for heterogeneity of demand, including variance within the population in age, education, income, and religion. The results are mixed, but arguably mildly supportive of the theory. James, whose work is discussed in chapter 22, has also tried to test this implication of the public goods theory by exploring the relative shares of nonprofit and governmental provision of services in several foreign countries and seeking to correlate these relative shares with the apparent heterogeneity of the populations involved. She too finds some support for the theory.

Further considerations bearing on the choice of nonprofit versus governmental organization are offered by Nelson and Krashinsky (1973) and by Hansmann (1980). For example, governmental firms have the advantage, through use of the taxing power, of more reliable access to capital and to operating revenues (especially in the case of public goods). Also, governmental organizations are usually linked by an organizational chain of command to the central executive of the government in order to provide the government with the requisite degree of information and control. This chain of command can serve as an additional mechanism for ensuring accountability in situations of contract failure. On the other hand, it also imposes a degree of bureaucratization that can make governmental organizations more costly and less flexible than their nonprofit counterparts. The private nonprofit form of organization has the corresponding advantage that it permits the development of a number of independent firms and thus promises greater competition and responsiveness to market forces. Moreover, a nonprofit firm can be more easily tailored to serve a narrow patronage, since it need not respond to the interests of the public at large. These and other factors that might affect the relative market shares of governmental and nonprofit organizations are explored in a cross-national context by James in chapter 22.

Further questions involve the relationship between governmental action and donative nonprofits. Governmental policy can affect the amount and direction of activity undertaken by donative nonprofits in various ways. Much attention has been given in recent years to exploring, both theoretically and empirically, the extent to which the charitable deduction incorporated in the personal income tax serves to encourage larger donations and the way in which these increased donations are distributed across different types of charities (Feldstein 1975; Clotfelter & Salamon 1982; Jencks, this volume, chap. 18). Less well explored, but of equal interest, are the ways in which direct government grants to nonprofits, or governmental provision of services in competition with those

provided by nonprofits, may affect, positively or negatively, the amount or types of activity undertaken by nonprofits. Similarly, it is of interest to inquiry why it is that governments in some cases provide services directly and in other cases provide the same or similar services by means of grants to private nonprofit organizations (see Rose-Ackerman 1981; James, this volume, chap. 22).

The Role of Donative Financing

Another set of interesting questions concerns the role of donative financing. The contract failure theory provides a potential explanation for the fact that donatively financed organizations are almost universally organized as nonprofits: by definition, donations involve payments that, though usually intended to be used for specific purposes, are not made with the expectation that they will be used simply to finance private goods for the donor. Consequently, the donor is very likely to experience difficulty in overseeing the use made of his donation and feel the need for the kind of protection afforded by the nonprofit form. In itself, however, the contract failure theory does not explain why it is that some services are donatively financed and others are not. To be sure, some services—such as redistribution to the poor or the provision of public goods—must by their very nature be donatively financed if they are to be provided privately at all. But in the case of some services that are commonly provided by donative nonprofits, it is not obvious that either redistribution to the poor or the production of public goods is involved. In such cases, closer consideration sometimes suggests that donative financing has arisen as a means of coping with special types of market imperfections that are peculiar to particular industries.

Price Discrimination

It is interesting to inquire, for example, why donative financing plays such a large role in the high-culture live performing arts. The services provided by such organizations, after all, are seldom rendered to the poor and are not easily characterized as public goods whose benefits spill over to individuals who do not pay the price of admission.

One likely explanation is that donative financing in the performing arts serves as a form of voluntary price discrimination, the need for which is dictated by the unusual cost and demand structure in that industry (Hansmann 1981b). In the high-culture live performing arts, fixed costs (primarily those of preparing a show prior to the first performance, including the cost of rehearsals, costumes, and stage sets) are a large proportion of the total costs of a production; once a production has been staged, the marginal cost of adding another performance to the run or of admitting another person to the audience for a performance that has not sold out is relatively small. This is, of course, in part a reflection of the fact that the potential audience for the high-culture performing arts is limited, even in the largest cities. It appears that, as a conse-

quence, for many productions there is no single ticket price that can cover total costs. If costs are to be met, some form of price discrimination must be employed so that high demanders pay more than low demanders for a given performance. Transferability of tickets, however, puts limits on the amount of price discrimination that can be accomplished through ticket pricing. Yet *voluntary* price discrimination has proven possible here: ticket purchasers with unusually high demand for performing arts productions can simply be asked to contribute some portion of the consumer surplus they would otherwise enjoy at the nominal ticket price—and, interestingly, a large proportion is in fact willing to do so.

The audiences for the popular performing arts such as movies and Broadway theater—in contrast to those for opera, symphonic music, and ballet—are large enough so that fixed costs can be spread widely, and thus fixed costs are low relative to marginal costs. Consequently, price discrimination is unnecessary for the viability of such productions, and they are usually produced by for-profit firms.

Although the performing arts seem to offer the best illustration of voluntary price discrimination, this may also be part of the function played by donative financing in other parts of the nonprofit sector, such as museums (which also experience fixed costs that are high relative to marginal costs), higher education, and health care.

Implicit Loans

The substantial role of donative financing in private education raises similar questions. In part it may serve to finance public goods or the provision of education to the poor. These explanations do not seem compelling, however, in the case of private primary or secondary schools or in the case of four-year private colleges, many of which emphasize teaching rather than research and have (at least until recently) served almost exclusively the relatively well-to-do. Further, such explanations do not entirely square with the fact that donations come largely from alumni of these colleges.

An alternative explanation, more consistent with such phenomena, is that donative financing in higher education serves at least in part as a system of voluntary repayments under an implicit loan system that has arisen to compensate for the absence of adequate loan markets for acquisition of human capital (Hansmann 1980). Many individuals for whom the present value of the long-run returns from higher education exceed the cost of that education are unable to finance it out of their own or their family's existing assets. If these individuals could take out a long-term loan against their future earnings, then this would be a worthwhile strategy for financing their education. Yet, since an individual cannot pledge human capital as security for such a loan (owing to laws against peonage, among other things), lenders will offer an inadequate supply of such loans. Private nonprofit schools provide a crude substitute for such loans. They supply education to many students at rates below cost, in return for an implicit commitment on the part of the students that they will

“repay” the school through donations during the course of their lives after graduation.

Option Demand

Weisbrod (1964; Weisbrod & Lee 1977) has argued that donations to nonprofits may in part reflect what he calls “option demand.” In particular, he suggests, this may help explain why hospitals are organized as donative nonprofits. “An individual's uncertainty with respect to demand for hospital services that may become critical to life means that he will be willing to pay a sum to secure the physical availability of those facilities in the future. An option demand may be said to exist for stand-by capacity, which is capacity in excess of the expected level of utilization” (Weisbrod & Lee 1977, 94). Of course, the mere fact that future demand is unpredictable need not in itself lead to market failure. Simply because one's own future demand, or even the entire market's future demand, for personal computers, four-bedroom apartments, or penicillin is uncertain does not mean that for-profit producers will supply them at an inefficient level. Some reason must be given to explain why for-profit firms, in the face of uncertain demand, will provide an inefficiently low level of capacity. One such reason has recently been offered by Holtmann (1983), though Holtmann makes no reference to Weisbrod's earlier option demand theory.

Holtmann develops a model in which demand is stochastic and in which a producing firm must choose its price and its maximum capacity level (which will subsequently represent a fixed cost for the firm) before the level of demand is revealed. The socially efficient behavior for the firm is to select a capacity level for which marginal expected (social) benefits equal marginal expected costs, and then set price equal to marginal cost. Such a policy will, however, produce negative returns for the firm regardless of the level of demand that subsequently materializes, and hence it will not be chosen by a for-profit firm. Without developing the point formally, Holtmann suggests that a donatively financed nonprofit firm will choose a lower price and larger capacity level than will a for-profit firm, and hence will come closer to the social optimum. Hence, Holtmann intimates, donative nonprofit firms might arise to meet the *ex ante* demand for capacity that for-profit firms will leave unsatisfied.

Other Motivations for Donating

There are, to be sure, many other reasons for donating besides those surveyed here. For example, donations to performing arts organizations may often be a form of conspicuous consumption (a type of signaling). Donations to one's alma mater may in part be inspired by a desire to maintain its institutional prominence in order to ensure that one's own degree will retain its status or quality. And donations to performing arts organizations, local hospitals, and one's alma mater may in part be, in effect, dues for membership in a club—the club of active supporters of the institution involved—which may be valuable for companionship or con-

tacts. I have focused here on voluntary price discrimination, implicit loans, and option demand—in addition to the familiar functions of redistribution and financing public goods—simply because these are functions served by donations that (1) are frequently overlooked, (2) come to light most clearly when nonprofits are examined with an economist's special facility for appreciating the functions and limits of markets, and (3) have been explicitly developed in the existing economics literature.

Why Not Free Ride?

We would also like to understand why, and under what conditions, individuals make contributions rather than succumb to the temptation to act as free riders in situations such as those just discussed. The contract failure theory, after all, suggests only why it is that, *given* that an individual wishes to make a donation, he is likely to direct that donation to a nonprofit rather than a for-profit firm; it does not explain why individuals are willing to make donations in the first place. Yet the question is obviously an important one: Americans *do* donate a substantial portion of their income to nonprofit organizations; moreover, they commonly make such donations in response to impersonal (for example, through-the-mail) appeals. At present, most of the wisdom we have on this subject focuses on aggregate phenomena and especially on the responsiveness of donations to changes in income and in price—in particular, tax incentives. (The available data and theories are surveyed in chapter 18.)

Demand-Side versus Supply-Side Theories

The various theories of the role of nonprofit enterprise that have been surveyed here are all essentially demand-side theories. That is, they present reasons consumers might choose to patronize nonprofit firms in preference to for-profit firms in particular industries. To date, much less systematic work has been done on developing supply-side theories that help explain why there is a supply of nonprofit firms in particular industries, and whether the current distribution of nonprofit firms across industries can be explained at least in part on the basis of differing conditions of supply. This is not to say, however, that there has been no work at all in this area; chapter 22, for example, offers some important observations on supply. Moreover, the behavioral theories discussed below also offer some insight into these issues.

THE BEHAVIOR OF NONPROFIT ORGANIZATIONS

The theories of the role of nonprofit organizations just surveyed are all based on the assumption that nonprofit firms are—or at least appear to their patrons to be—bound by a nondistribution constraint. This constraint, however, is consistent with a variety of forms of behavior on the part of nonprofit firms. Therefore, commitment to one of these theories of the role of nonprofit organizations does not necessarily

involve commitment to a particular theory of the behavior of nonprofit firms, and vice versa. Moreover, many of the early efforts to model the behavior of nonprofit firms—especially hospitals—were developed without concern for the reasons such firms developed and survived. Consequently, the behavioral models of nonprofit organizations developed to date have been to some degree disconnected from models of the role of such firms.

Optimizing Models

Following the neoclassical tradition, most models of the behavior of nonprofit firms have been optimizing models, typically focusing on firms in a particular industry. Hospitals have been the most common subject.

Choosing the maximand has been a problem in these models. In contrast to the case of the for-profit firm, there is obviously no reason to believe a priori that profit maximization is a reasonable goal to impute to the nonprofit firm. Most commonly, nonprofit firms have instead been assumed to maximize the quality and/or quantity of the service they produce. The first of these goals might seem reasonable for a nonprofit firm run by professionals who derive strong satisfaction from doing craftsmanlike work, independent of the needs or desires of their clientele. Quantity maximization, in turn, might be imputed to managers who are empire builders or who are altruists of a type that seeks to serve as broad a segment of the public as possible. Models of nonprofit firms that pursue one or both of these goals have been developed by Newhouse (1970) and Feldstein (1971) for hospitals, James and Neuberger (1981) for universities, James (1983) for nonprofits in general, and Hansmann (1981b) for performing arts organizations. Lee (1971), in contrast, presents a model of a hospital that maximizes (or, more accurately, *satisfices*) not output but rather its use of certain inputs.

Models of nonprofits that seek to maximize their budgets have also been common. Presumably budget maximization might be chosen as a goal because it enhances the apparent importance of (or justifies a higher salary for) the firm's managers or, alternatively, because it provides the preferred trade-off between quality and quantity maximization. Examples of budget-maximizing models have been offered by Tullock (1966), who considers a purely donative nonprofit, and Niskanen (1971, chap. 9), who considers a purely commercial nonprofit. Hansmann's previously mentioned paper on the performing arts (1981b) also models the behavior of a (partly donative and partly commercial) nonprofit budget maximizer.

Each of these optimizing models is employed by its author to some degree to explore the welfare implications of the type of behavior the model postulates. For example, Newhouse (1970) emphasizes that the quality/quantity-maximizing firm in his model will usually exhibit productive inefficiency when contrasted with the performance of a for-profit firm operating in an environment free of market failure. Hansmann's performing arts model assumes that the firm is operating under conditions of contract failure and that it must

adhere to the nondistribution constraint; the model then explores the socially optimal objective function for the firm, *given* this constraint. It turns out that quantity, quality, or budget maximizing may or may not constitute efficient behavior for the firm, depending on the structure of consumers' preferences and the way in which donations respond to firm behavior. In Tullock's model (1966), the budget-maximizing donative nonprofit overspends considerably (from a social welfare point of view) on promotion: at the margin, it spends more than a dollar in promotional expenses in order to solicit an additional dollar in donations.

Pauly and Redisch (1973) offer a model of a hospital that is operated to maximize the financial returns to its affiliated doctors. This is not the same thing as profit maximization for the firm. Rather, since doctors do not receive payment directly from the hospital but instead bill patients separately, this theory implies that hospitals will bill patients only enough to cover costs and will procure inputs that enhance the physicians' productivity. Pauly and Redisch then develop an explicit model of the hospital as a Ward-Domar-type producer cooperative (with the physicians as the worker/owners) and thus predict for a hospital the same behavior that characterizes other models of this type—behavior that involves considerable inefficiency in the short run in the form of perverse supply response. Since the work behavior and compensation of hospital-based physicians do not follow the simple fixed-effort and equal-sharing rules assumed in this class of producer cooperative models, it is not clear that in fact we should expect perverse supply response to be an empirically important phenomenon in hospitals. Nevertheless, the general view of hospitals as serving indirectly the financial interests of doctors may capture an important aspect of reality.

Productive Inefficiency

Optimizing models of the types just surveyed implicitly assume that the firms involved minimize costs. Another line of behavioral theory has argued that, whatever objectives nonprofits may pursue with respect to quantity or quality of output, they are inherently subject to productive inefficiency (that is, failure to minimize costs) owing to the absence of ownership claims to residual earnings (Alchian & Demsetz 1972; Hansmann 1980). This argument is clearest when applied to entrepreneurial nonprofits, which constitute the great majority of financially significant nonprofits. Those who control such organizations—whether the managers or the board of directors who appoint the managers—are unable, by virtue of the nondistribution constraint, to appropriate for themselves the net earnings obtained by reducing costs, and thus they have little pecuniary incentive to operate the organization in a manner that minimizes costs.¹² Of

12. In mutual nonprofits, ultimate control is by definition in the hands of the patrons of the organization, and the patrons have an incentive to have the organization minimize costs. If the organization has many patrons, however, transaction costs and free-rider problems

course, it could be that the managers of some nonprofits derive substantial utility from having the firm produce large amounts of output and thus have a desire to minimize costs that is independent of the income they derive from the firm. And there is reason to believe that nonprofit organizations tend to attract more managers of this type than do for-profit firms (see Young 1983 and chapter 10 in this volume). Nevertheless, nonprofit managers in general might be expected to indulge themselves in various perquisites of office—including some forms of nonpecuniary income as well as a more relaxed attitude toward their duties—to a greater extent than do their counterparts in for-profit firms. Clarkson (1972) presents empirical results comparing the behavior of nonprofit and for-profit hospitals that provide some support for this view.

It is almost certainly true that nonprofit firms are productively inefficient in the sense that, in the absence of subsidies or a substantial degree of market failure of some type (such as contract failure) in the product market, they will generally produce any given good or service at higher cost than would a for-profit firm. If it were otherwise, we would expect to find nonprofit firms operating successfully in a much broader range of industries than is actually the case. As emphasized in the preceding discussion of the role of nonprofits, nonprofit firms seem to have survivorship properties that are superior to for-profit firms only where particular forms of market failure give them an efficiency advantage sufficient to compensate for their failure to minimize costs. Thus, in general we do not find nonprofit firms producing, wholesaling, or retailing standard industrial goods or agricultural commodities (such as machine screws or cucumbers) for which contract failure is not a significant problem.

Supply Response

Empirical work (Steinwald & Neuhauser 1970; Hansmann 1985a) indicates strongly that nonprofit firms tend to respond much more slowly to increases in demand than do their for-profit counterparts. For example, in those industries populated by both nonprofit and for-profit firms, such as nursing care, hospital care, and primary and secondary education, the ratio of nonprofit to for-profit firms is much lower in markets in which demand has been expanding rapidly than it is in markets in which demand has remained stable or declined.

One likely explanation for this phenomenon is that, in comparison to for-profit firms, nonprofit firms are constrained in their access to capital. Unlike for-profit firms, nonprofit firms cannot raise capital by issuing equity shares; rather, they must rely on debt, donations, and retained earnings for this purpose—sources that, even in combination, offer a less responsive supply of capital than does the equity market.

may prevent the patrons from exercising effective authority over the firm's management, thus leading to poor incentives for cost minimization in these firms as well.

An alternative explanation for nonprofits' relatively poor supply response points to problems of entrepreneurship. Owing to the nondistribution constraint, nonprofit entrepreneurs are unable to capture the full return that can be gained by establishing a new firm or expanding an old one in the face of increased demand. Consequently, their incentive to undertake such entry or expansion is limited relative to that of entrepreneurs in the for-profit sector.

At present we cannot say to what extent each, or either, of these explanations accounts for the relatively poor supply response exhibited by nonprofits. There is empirical evidence that nonprofit firms are sometimes capital constrained (Ginsburg 1970), but we do not know precisely how this translates into supply response. And entrepreneurship in the nonprofit sector presents an even more elusive problem. Young (whose work is surveyed in chapter 10) has undertaken case studies of nonprofit entrepreneurship that indicate, as one might expect, a substantial range of motivation and behavior. He describes a set of personality types into which nonprofit entrepreneurs can be divided and suggests that certain of these personality types are selected for disproportionately by particular types of nonprofit firms. Not surprisingly, some of these personality types are inconsistent with a strong emphasis on expansion of services.

James in chapter 22 notes that, particularly in countries other than the United States, both the entrepreneurial initiative and the necessary capital for founding a nonprofit institution, such as a school, are commonly supplied by an existing organization that is already well established and well financed—such as a major religious sect. This observation underlines the importance of both factors, though it does not clearly indicate which is generally the more important bottleneck.

Income-Generating Behavior

To the extent that a nonprofit seeks to provide a service of a quantity or quality that cannot be supported by market demand, some form of subsidy must be found. One source of such a subsidy, evidently commonly used by nonprofits, is cross subsidization: one service is produced and sold by the nonprofit at a profit, which is then used to finance provision of another service that is more highly valued by the firm. The net returns earned on the subsidy-generating service may result from the fact that the nonprofit firm has some degree of market power in providing that service or from the fact that the nonprofit firm has lower costs than its competitors owing to tax exemption or some other form of governmental favor. James (1981, 1983) illustrates this form of behavior with a simple model of a multiproduct nonprofit firm that places different degrees of (either positive or negative) utility on the quantities of the various products it sells, and then determines price and output for the full set of products in a fashion that maximizes utility to the firm while meeting a breakeven constraint. Harris (1979) presents a model of a nonprofit hospital illustrating how cross subsidization can be employed to compensate for distortions and inequities in health insur-

ance coverage and offers empirical results suggesting that to some extent hospitals behave consistently with this model.

Another way to raise funds to pay for services whose production provides positive utility to the firm is to solicit donations. To be sure, as suggested earlier, we can view donations simply as a price that is paid by persons who wish to finance provision of services for third parties. In this sense, then, donations are not a subsidy, and efforts to increase donations are simply efforts to market the firm's goods—that is, a form of advertising. Nevertheless, donation-seeking behavior presents some interesting questions, especially from a welfare standpoint. In particular—depending on one's assumptions about donor information and behavior—nonprofits may have an incentive to expend inefficiently large amounts of funds on solicitation, as Tullock (1966) suggests. Since theoretical and empirical work on donation-seeking behavior is surveyed in chapter 7, however, the issue will not be addressed further here.

Patron Control

The discussion so far has proceeded largely as if all nonprofits were entrepreneurial nonprofits whose management is constrained in its behavior only by market forces and the nondistribution constraint. However, many nonprofits (namely, those we have called mutual nonprofits) are ultimately controlled, at least formally, by their patrons. Thus, it remains to ask whether, and how, patrons influence the behavior of mutual nonprofits through the exercise of their voting power—that is, through voice rather than exit, to use Hirschman's now familiar terminology (1970).

The only general theoretical treatment of this subject is offered by Ben-Ner (1986; this volume, chap. 24), who sees patron control as the principal *raison d'être* for nonprofit firms and thus devotes considerable attention to the possible behavior of customer-controlled firms. He focuses in particular on coalition formation among customer-members, arguing that high-demand customers can frequently be expected to dominate the firm and to set price and output parameters that maximize their own welfare while exploiting other customers to the extent permitted by competition.

A more narrowly focused treatment is offered in Hansmann (1986) of social clubs, colleges, hospital medical staffs, and other membership organizations in which the personal characteristics of one's fellow patrons (or employees) are an important factor in the utility derived from membership. A simple model is presented to illustrate the way in which member control interacts with competition to determine the size, fees, and membership characteristics of individual clubs. In that model, each individual is assumed to be characterized by a unidimensional variable denoting "status." Individuals join clubs in order to associate with other individuals, and the value of a given club's membership as companions is given by their average status: the higher the better. Assuming limited economies of scale (in terms of membership size) in the operation of clubs, and assuming that a given club must charge all its members the same fee,

free formation of clubs in this model results in roughly the pattern we see in reality: that is, a system of member-controlled clubs that are usually smaller than the size that minimizes average cost per member, and that are exclusive and stratified in the sense that the highest-status individuals will be in a single club of their own, the next highest will constitute the membership of a second club, and so on.

CONCLUSION: SOME POLICY APPLICATIONS

The theories concerning the role and behavior of nonprofit firms discussed above are of interest simply as a matter of positive social science. They are also of interest, however, from a policy perspective. Indeed, the most pressing current problems of policy concerning the nonprofit sector cannot be resolved intelligently without adopting one or another point of view concerning the role and behavior of nonprofit firms. This is not the place to consider policy problems in detail. But for purposes of illustration and as a means of providing some perspective on the theories that have been surveyed here, we shall look briefly at some examples.

As suggested earlier, the most dramatic development in the nonprofit sector in recent decades has been the rapid growth of commercial nonprofits. The appearance of large numbers of such firms, which derive their revenues largely from fees for service and commonly exist in competition with for-profit firms providing similar services, has brought with it some of the most difficult problems of policy that currently involve the nonprofit sector.

Tax Exemption

One important set of issues, for example, concerns tax exemption. At present, most nonprofit firms are exempt from taxation (including sales, property, and corporate income taxation) at the federal, state, and local levels. These exemptions were relatively unproblematic when they were first established many decades ago: most nonprofits were simple donative charities that provided either public goods or aid to the poor, thus offering a substantial rationale for public subsidy. In any event, the potential tax liability of the organizations involved was often quite small. Yet the scope of these exemptions has been extended to keep pace with the expansion of the nonprofit sector, so that today large numbers of commercial nonprofits are also exempt. And it is not obvious that the arguments for exempting traditional charities carry over to commercial nonprofits such as nursing homes or health maintenance organizations.

It is difficult to rationalize tax exemption for commercial nonprofits on the simple ground that the basic service they provide—nursing care for the elderly, day care, hospital care, or whatever—is for some reason worthy of subsidy in general, since that argument would seem to call for exempting not only the nonprofit firms in the industry but the for-profit firms as well. Of course, the exemption might be confined to nonprofit firms even under this rationale on the

theory that the nondistribution constraint ensures that the subsidy will actually be passed through to consumers (see the general discussion of government "donations" above). But, in industries like those in question, in which firms simply sell goods or services directly to consumers, it would seem that competition among competing firms would go far toward ensuring the same result for for-profit firms. Consequently, the exemption seems more easily justifiable if it can be established that nonprofit firms in the relevant industries offer a type of service that is different from that offered by their for-profit competitors and that would be undersupplied without subsidy.

A possible argument along these lines is that nonprofit firms provide services that have more of the character of public goods than do the services provided by for-profit firms in the same industry. Yet, as we have observed in discussing the public goods theory of nonprofits above, it is not at all obvious that this is the case for commercial nonprofits in most industries. An alternative possibility is that commercial nonprofits are in fact a response to contract failure and that they offer a higher degree of fiduciary responsibility toward their customers than do their for-profit competitors—a quality that is of special service to that subset of customers who do not trust their own ability to look out for their interests in the market. Yet even acceptance of the contract failure theory as applied to commercial nonprofits does not necessarily resolve the question of exemption. For we must ask why it is that customers who want the special protection of the nonprofit form cannot be left to seek it out and pay for it on their own. Do such customers constitute a class that is specially deserving of a subsidy? Or is it the case that such customers will myopically undervalue the special protection afforded by nonprofit firms and thus need a subsidy to encourage them to patronize such firms? Or is the subsidy provided by the exemption best justified as a way of compensating for problems of supply response among nonprofit firms that would otherwise develop too slowly to meet demand? (Note that the latter justification is persuasive only if the problem of supply response is primarily the result of lack of capital rather than lack of entrepreneurship.)

The object here is not to offer a resolution of these issues.¹³ Rather, it is simply to emphasize that, if one is to take a thoughtful position on whether to continue or revoke the exemption for commercial nonprofits in any given industry, one must necessarily think carefully about the role and behavior of the firms involved.

Outlawing For-Profit Firms

In recent years considerable attention has been devoted to abuses in the nursing home industry involving shoddy patient care and shady finances. These exposés have brought proposals from several prominent quarters for public policies designed to eliminate for-profit nursing homes (for example,

13. For further theoretical and empirical discussion, see Hansmann (1981c, 1985a).

by denying them licenses) on the theory that for-profit homes are the source of most of the abuses and that the industry would perform better if it were composed only of nonprofit firms (Etzioni 1976; New York Temporary State Commission . . . 1975). Indeed, such proposals have not been confined to the nursing home industry; public measures disadvantaging or outlawing for-profit as opposed to nonprofit firms have been enacted or proposed at various times as well for aspects of medical practice, legal practice, and higher education.¹⁴

To accept such proposals, it seems, one must accept strongly the contract failure theory of the role of nonprofits. Indeed, one must presumably believe not just that existing nonprofit nursing homes serve a fiduciary role toward their customers but that even those customers who currently patronize proprietary nursing homes need the protection of the nondistribution constraint and were misguided in choosing a for-profit rather than a nonprofit provider.

Further, to accept such proposals one must also believe that outlawing proprietary homes will not have a significant

14. For a more thorough discussion of such policies see Hansmann (1981d, 548–53) and Young (1983, 141–44).

effect on the character of nonprofit homes—for example, by forcing profit-motivated entrepreneurs to utilize the nonprofit form, thus creating a group of nominally nonprofit firms that actively seek to evade the nondistribution constraint.

And finally, before implementing such a proposal, one must consider the problem of supply response. If one goes no further than simply outlawing for-profit homes, then there will presumably be a long period of excess demand for the services of the remaining nonprofit firms. Thus, many of the elderly may simply go from having poor service to having no service. If capital constraints are the chief cause of poor supply response among nonprofits, this problem might be remedied by governmental provision of loan or grant capital to nonprofit firms. If, on the other hand, the supply response problem has its roots in the lack of incentives for nonprofit entrepreneurship, then capital subsidies in themselves might be unavailing, and the problem, if remediable at all, must be dealt with through more complex policies.

Thus here, as with tax exemption, intelligent policy must necessarily be based on a sophisticated understanding of the role and behavior of nonprofit firms. Recent work on the economics of nonprofit organizations holds the promise, at last, of yielding such an understanding.

REFERENCES

- Alchian, Armen, and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review* 62:777–95.
- Arrow, Kenneth. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53:941–73.
- Baumol, William, and William Bowen. 1965. "On the Performing Arts: The Anatomy of Their Economic Problems." *American Economic Review Papers and Proceedings* 55:495–502.
- . 1966. *Performing Arts—The Economic Dilemma*. Cambridge, Mass.: MIT Press.
- Ben-Ner, Avner. 1986. "Non-Profit Organizations: Why Do They Exist in Market Economies?" In *The Economics of Nonprofit Institutions: Studies in Structure and Policy*, ed. Susan Rose-Ackerman. Oxford: Oxford University Press.
- Clark, Robert. 1980. "Does the Nonprofit Form Fit the Hospital Industry?" *Harvard Law Review* 93:1416–89.
- Clarkson, Kenneth. 1972. "Some Implications of Property Rights in Hospital Management." *Journal of Law and Economics* 15:363–84.
- Clotfelter, Charles, and Lester Salamon. 1982. "The Impact of the 1981 Tax Act on Individual Charitable Giving." *National Tax Journal* 35:171–87.
- Dickinson, Frank. 1962. *Philanthropy and Public Policy*. New York: National Bureau of Economic Research.
- Easley, David, and Maureen O'Hara. 1983. "The Economic Role of the Nonprofit Firm." *Bell Journal of Economics* 14:531–38.
- . 1984. "An Information-Based Theory of the Firm." Unpublished manuscript; Cornell University Department of Economics.
- Ellman, Ira. 1982. "Another Theory of Nonprofit Corporations." *Michigan Law Review* 80:999–1050.
- Etzioni, Amitai. 1976. "Profit in Not-for-Profit Institutions." *Philanthropy Monthly* 9:22–34.
- Fama, Eugene, and Michael Jensen. 1983a. "Agency Problems and Residual Claims." *Journal of Law and Economics* 26:327–50.
- . 1983b. "Separation of Ownership and Control." *Journal of Law and Economics* 26:301–26.
- Feldstein, Martin. 1971. "Hospital Price Inflation: A Study of Nonprofit Price Dynamics." *American Economic Review* 61:853–72.
- . 1975. "The Income Tax and Charitable Contributions." *National Tax Journal* 28:81–97, 209–28.
- Ginsburg, Paul. 1970. *Capital in Non-Profit Hospitals*. Ph.D. diss., Harvard University.

- Hansmann, Henry. 1980. "The Role of Nonprofit Enterprise." *Yale Law Journal* 89:835-901.
- . 1981a. "Consumer Perception of Nonprofit Enterprise: Reply." *Yale Law Journal* 90:1633-38.
- . 1981b. "Nonprofit Enterprise in the Performing Arts." *Bell Journal of Economics* 12:341-61.
- . 1981c. "The Rationale for Exempting Nonprofit Organizations from Corporate Income Taxation." *Yale Law Journal* 91:54-100.
- . 1981d. "Reforming Nonprofit Corporation Law." *University of Pennsylvania Law Review* 129:497-623.
- . 1985a. "The Effect of Tax Exemption and Other Factors on Competition between Nonprofit and For-Profit Enterprise." Yale University, Program on Non-Profit Organizations Working Paper no. 65.
- . 1985b. "The Organization of Insurance Companies: Mutual versus Stock." *Journal of Law, Economics, and Organization* 1:125-53.
- . 1986. "Status Organizations." *Journal of Law, Economics, and Organization* 2:119-30.
- Harris, Jeffrey. 1979. "Pricing Rules for Hospitals." *Bell Journal of Economics* 10:224-43.
- Heflebower, Richard. 1980. *Cooperatives and Mutuals in the Market System*. Madison: University of Wisconsin Press.
- Hirschman, Albert. 1970. *Exit, Voice and Loyalty*. Cambridge, Mass.: Harvard University Press.
- Holtmann, A. G. 1983. "A Theory of Non-Profit Firms." *Economica* 50:439-49.
- James, Estelle. 1983. "How Nonprofits Grow: A Model." *Journal of Policy Analysis and Management* 2:350-65.
- James, Estelle, and Egon Neuberger. 1981. "The University Department as a Non-Profit Labor Cooperative." *Public Choice* 36:585-612.
- Jensen, Michael, and William Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3:305-60.
- Lee, A. James, and Burton Weisbrod. 1977. "Collective Goods and the Voluntary Sector: The Case of the Hospital Industry." In Burton Weisbrod, *The Voluntary Nonprofit Sector*. Lexington, Mass.: Lexington Books.
- Lee, M. L. 1971. "A Conspicuous Production Theory of Hospital Behavior." *Southern Economic Journal* 38:48-59.
- Manne, Henry. 1965. "Mergers and the Market for Corporate Control." *Journal of Political Economy* 73:110-20.
- Nelson, Richard. 1977. *The Moon and the Ghetto: An Essay on Public Policy Analysis*. New York: W. W. Norton & Co.
- Nelson, Richard, and Michael Krashinsky. 1973. "Two Major Issues of Public Policy: Public Policy and Organization of Supply." In *Public Subsidy for Day Care of Young Children*, edited by Richard Nelson and Dennis Young. Lexington, Mass.: D. C. Heath & Co.
- Newhouse, Joseph. 1970. "Toward a Theory of Non-Profit Institutions: An Economic Model of a Hospital." *American Economic Review* 60:64-74.
- Newton, Jamie. 1980. "Child Care Decision-Making Survey—Preliminary Report." Unpublished manuscript, Yale University, Program on Non-Profit Organizations.
- New York Temporary State Commission on Living Costs and the Economy. 1975. *Report on Nursing Homes and Health-Related Facilities in New York State*. N.p.
- Niskanen, William. 1971. *Bureaucracy and Representative Government*. Chicago: Aldine Publishing Co.
- Pauly, Mark P., and Michael R. Redisch. 1973. "The Not-for-Profit Hospital as a Physicians' Cooperative." *American Economic Review* 63:87-99.
- Permut, Steven. 1981. "Consumer Perceptions of Nonprofit Enterprise: A Comment on Hansmann." *Yale Law Journal* 90:1623-32.
- Rose-Ackerman, Susan. 1981. "Do Government Grants to Charity Reduce Private Donations?" In *Nonprofit Firms in a Three-Sector Economy*, edited by Michelle White. Washington, D.C.: Urban Institute.
- Steinwald, Bruce, and Duncan Neuhauser. 1970. "The Role of the Proprietary Hospital." *Law and Contemporary Problems* 35:817-38.
- Thompson, Earl. 1980. "Charity and Nonprofit Organizations." In *The Economics of Nonproprietary Organizations*, edited by Kenneth Clarkson and Donald Martin. Greenwich, Conn.: JAI Press.
- Tullock, Gordon. 1966. "Information without Profit." In *Papers on Non-Market Decision Making*, edited by Gordon Tullock. Charlottesville: Thomas Jefferson Center for Political Economy, University of Virginia.
- Weisbrod, Burton. 1964. "Collective-Consumption Services of Individual-Consumption Goods." *Quarterly Journal of Economics* 78:471-77.
- . 1974. "Toward a Theory of the Voluntary Non-Profit Sector in a Three-Sector Economy." In *Altruism, Morality, and Economic Theory*, edited by Edmund S. Phelps. New York: Russell Sage.
- . 1977. *The Voluntary Nonprofit Sector*. Lexington, Mass.: D. C. Heath & Co.
- . 1983. "Nonprofit and Proprietary Sector Behavior: Wage Differentials among Lawyers." *Journal of Labor Economics* 1:246-63.
- Weisbrod, Burton, and A. James Lee. 1977. "Collective Goods and the Voluntary Sector: The Case of the Hospital Industry." In Burton Weisbrod, *The Voluntary Nonprofit Sector*. Lexington, Mass.: Lexington Books.
- Weisbrod, Burton, and Mark Schlesinger. 1986. "Ownership Form and Behavior in Regulated Markets with Asymmetric Information." In *The Economics of Nonprofit Institutions: Studies in Structure and Policy*, ed. Susan Rose-Ackerman. Oxford: Oxford University Press.
- Weiss, Jeffrey. 1986. "Donations: Can They Reduce a Donor's Welfare?" In *The Economics of Nonprofit Institutions: Studies in Structure and Policy*, ed. Susan Rose-Ackerman. Oxford: Oxford University Press.
- Williamson, Oliver. 1979. "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics* 22:233-61.
- Young, Dennis. 1983. *If Not for Profit, for What?* Lexington, Mass.: D. C. Heath & Co.