

# 5

---

## Economic Theories of Nonprofit Organizations

---

RICHARD STEINBERG

Those unfamiliar with economics, nonprofit organizations, or both may wonder what one subject has to do with the other. Economics does not, of course, provide the reader with everything he or she ought to know but it does provide insight on virtually every problem relating to the role, behavior, management, and regulation of the nonprofit sector. In turn, economics has become a richer discipline for confronting the special challenges posed by analysis of nonprofit organizations.

Economics is the study of choices under scarcity. This goes far beyond the study of things bought and sold and far beyond the financial consequences of decision-making. When consumers decide whether to spend all their limited money on goods and services for themselves or donate some of that money to nonprofit organizations, they are choosing how to allocate a scarce resource. When consumers choose to spend some of their scarce time as volunteers rather than laborers or leisure-takers, they are making an economic decision. When socially conscious entrepreneurs ponder whether to form a new organization, support an existing organization, or lobby government to meet some societal need, they are choosing how to allocate their time and other scarce resources at their command. When financially strapped nonprofit organizations decide to charge their indigent clients a little something rather than eliminate programs, they are making an economic decision forced by scarcity.

Economics as a discipline does not rule out study of the sector, but many simplified models studied by economists divert attention from philanthropic and nonprofit issues. Economists assume that each individual pursues his or her self-interest as they see it. In practice, this is often simplified as the assumption that each cares only about his or her personal consumption of goods and services. However, self-interest encompasses helping others one cares about, and

individual perceptions of self-interest can be socially determined. The charitable behavior of donors and volunteers reminds economists that more complicated models are necessary, paying back the analyst with insights that go far beyond understanding charity. In like fashion, economic models typically assume that for-profit firms maximize profits because that is what the owners want to do. Framed in this way, departures from profit maximization appear as "market failures" due to "agency problems" cured by providing the proper financial incentives. Nonprofit organizations cannot be analyzed from the same starting point. These organizations may indeed maximize their "profits" (or financial surplus or endowment) under some circumstances. However, this is a result that must be shown and its significance must be interpreted anew. Thus, for example, Slivinski (2002) shows that a mixture of financial and non-financial incentives is the best way to motivate nonprofit employees in specified circumstances.

This chapter provides an overview on the economics of nonprofit organizations. Many later chapters provide far more detail on the decisions to give or volunteer (Havens, O'Herlihy, and Schervish; Vesterlund; and Leete), the economic relations among the sectors (Brown and Slivinski; Smith and Grønberg; Galaskiewicz and Colman; Tuckman and Chang), cross-sectoral comparisons of organizational behavior in specific industries (Kendall, Knapp, and Forder; Schlesinger and Gray) and selected public policy issues (Brody; Simon, Dale, and Chisolm). Here, I survey the big picture and fill some holes left by the collection of other chapters.

In the next section, I provide definitions and distinctions used for the economic approach. Later I discuss an older set of theories, known as the "three-failures theory," regarding the distinctive role of nonprofit organizations in the broader economy. Then, I argue for development of a more comprehensive and integrated theory, lay out the parameters of such

a theory, and discuss progress to date in carrying out these ideas. I conclude by illustrating how economic theory provides insight for those designing appropriate public policy for the sector.

### DEFINITIONS AND DISTINCTIONS

Many definitions abound in discussion of "the sector," but in this chapter I use Hansmann's (1980) idea: A nonprofit organization is one precluded from distributing, in financial form, its surplus resources to those in control of the organization. By this definition, nonprofit organizations can earn and retain financial surplus ("profits") provided they do not pay dividend checks or their equivalent to the board of directors or top managers. Instead, the surplus is either retained (as endowment, reserves, or temporarily restricted funds), reinvested (in organizational expansion or the provision of charitable services), or given to other nonprofit organizations (as grants). Some nonprofit organizations derive all their resources from commercial operations, and in this sense are just as much "for profits" as any for-profit firm. The distinction is that they must retain or reinvest their profits.

Hansmann (1980) called this prohibition on profit distribution the "nondistribution constraint," and made it central to his theories of nonprofit behavior. As we will see, the nondistribution constraint is an essential part of other economic theories. The constraint provides a clear distinction that affects how the organization obtains resources, how it is controlled, how it behaves in the marketplace, how it is perceived by donors and clients, and how its employees are motivated. Hansmann also defined the companion "fair compensation constraint" that applies nondistribution to executive compensation.

This definition of nonprofit organization excludes consumer cooperatives and worker-owned firms, but includes mutual savings banks. All three distribute their profits to consumers or workers; however, members of the banks' boards of directors do not receive distributions and those receiving distributions have no rights of control. Some organizations are legally incorporated as nonprofits but secretly distribute their profits to those in control, the so-called for-profits-in-disguise. Nonprofits-in-disguise, the opposite phenomenon, are rare but do exist. For example, medical insurance providers in Puerto Rico are required to incorporate under the statutes governing for-profit corporations, but one has nonetheless written nondistribution into its articles of incorporation and bylaws.

The owners (governing board) of a nonprofit organization do not enjoy all the usual rights of ownership. Ben-Ner and Jones (1995) detail three components of property rights: the right to control the use of an asset, the right to retain any financial surpluses generated from that use, and the right to sell the first two rights to a new owner. Nonprofit owners have "attenuated property rights," meaning they are allowed to control the organizational assets and transfer that control, but not allowed to profit financially from using their other

rights. If the organization is sold or converted to a for-profit firm, the owners obtain fair market value for the organization's assets, then donate this value to another nonprofit (typically a "conversion foundation" that makes charitable grants). If the owners derive any benefits from control of the organization, they must be in nonfinancial form. These benefits may help only the owners (for example, the organization can have an opulent headquarters, hold board meetings at vacation resorts, and employ attractive but not especially productive staff) or may be enjoyed jointly by the owners (who like using their position of control to make the world a better place) and members of society at large.

Despite sharing a nondistribution constraint, nonprofits differ from one another in a variety of economically meaningful ways. First, some nonprofits deliver services whereas others (such as united fundraising organizations, foundations, and donor-advised funds) make grants and program-related loans to other nonprofits. Second, some nonprofits rely mostly on donations (gifts, grants, and volunteers), others on membership dues, and others on commercial activity (sales to the public or contractual provision of service to the government). Third, nonprofits differ in the way their governing boards are selected. Fourth, nonprofits differ in the services they provide. These services are enumerated in various industrial classification codes (like the North American Industry Classification System) or in taxonomies specific to the nonprofit sector (like the National Taxonomy of Exempt Organizations or the International Classification of Nonprofit Organizations).

Hansmann (1980) emphasized some of these distinctions in his four-way classification of nonprofit firms, summarized in table 5.1. The financing dimension of this classification simply asks whether most of the organizational resources come from donations or other sources. The governance dimension distinguishes mutual nonprofits (where the power to elect the board is in the hands of donors and customers) from entrepreneurial nonprofits (where boards are self-perpetuating or appointed). Hansmann proposes these distinctions as ideal types (real-world nonprofits may straddle the boundaries) and asks whether these types behave differently. Subsequent literature has focused on differences between donative and commercial nonprofits, but work to date has

TABLE 5.1. 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 Assoc. Consumers Union  Country clubs	National Geographic Society Educational Testing Service Hospitals Nursing homes

Source: Hansmann 1980, as adapted in Hansmann 1987a.

not analyzed the impact of the governance distinction. This is an unfortunate gap in the literature as political dynamics no doubt explain the role, objectives, performance, and life cycle of nonprofit organizations.

### THREE FAILURES: AN EARLY SUCCESS

Before Weisbrod's pathbreaking work in 1975, economists typically viewed nonprofit organizations in isolation (for example, Tullock 1966; Newhouse 1970; Feldstein 1971; Pauly and Redisch 1973). Suppose, they would start, a nonprofit organization was supplying a particular good or service. They would then ask how that nonprofit's behavior would differ from that of a for-profit firm. For example, Newhouse characterized the behavior of a nonprofit hospital that cared about both the quality and quantity of health services it delivered. This was a useful starting point, but it was never clear why nonprofits were supplying the service in question, and not some other kind of organization. Why did organizations with quantity/quality objectives emerge in the hospital industry, but not in automobile manufacturing or accounting services? Given that nonprofit hospitals did arise, why is the hospital industry also populated by for-profit and government hospitals? Would the nonprofit presence in the hospital industry end if special tax advantages were removed, if Medicare shifted to a prospective payment system, if the population aged, or if technology improved?

Weisbrod (1975) began the process of searching for a distinctive set of roles that nonprofit organizations could play in a mixed economy. He catalogued the known virtues of for-profit firms and governmental action, finding a role for nonprofits when these two other sectors are expected to fail (due to "market failure" and "government failure"). In brief, markets fail to provide adequate quantities of collective goods, governments provide these goods in accord with the wishes of the electorate, and those who want higher levels of service than government provides support nonprofit organizations. Hansmann (1980) added an additional shortcoming of markets to the mix: contract failure. In cases where the quantity or quality of service cannot be verified, markets take advantage of informational asymmetries. Organizations that cannot distribute profits, including private nonprofits and government agencies, provide a more trustworthy alternative. Salamon (1987) turned the process around, cataloguing the virtues of nonprofit organizations and finding a role for the other two sectors when nonprofits are expected to fail ("voluntary failure"). In conjunction, these three approaches began the stream of literature that has become known as "three-failures theory." Next, I discuss each of these failures in more detail, showing how each sector responds to failures by others. Then I summarize some empirical evidence and conclude with the limitations of the three-failures approach.

Figure 5.1 illustrates how the pieces of three-failures theory fit together. I cover these topics row by row, first discussing market success and three sources of market failure (the first two rows). Then, I discuss how government responds to

each of these failures and why, in turn, government fails to completely address the problems (rows three and four). Nonprofits respond to the three sources of market and government failure (row five), but their response is not completely adequate either (row six). In addition, nonprofits generate their own failures (row seven), and markets and governments respond to these (the feedback loop from the bottom to the top and third rows).

### Market Failure

Market failure is the best understood of the three failures, and concerns inefficiencies resulting from for-profit provision of goods and services. The term *inefficiency* as used here and elsewhere by economists has a very broad definition. Markets can be inefficient because they waste resources by using the wrong production processes (productive inefficiency), but they can also be inefficient because they waste resources by producing the wrong mixture of goods and services (allocative inefficiency). Allocative efficiency requires an output mixture that properly balances the relative benefits to consumers against the relative costs of production.<sup>1</sup> For-profit firms are, at least in theory, productively efficient but often produce the wrong mixture of outputs in three ways: some worthwhile goods are underprovided, access to some goods is overrestricted, and the quantity or quality of some delivered goods is different from what the consumer or client was promised.

Adam Smith, in exploring the virtues of for-profit markets through his famous "invisible hand," laid the groundwork for the theory of market failure. In more modern terms, this idea circulates as the first fundamental theorem of welfare economics, which asserts that when all goods are traded in perfectly competitive markets, equilibrium outcomes are efficient. As used in the theorem, a "good" is anything that consumers value either positively (goods) or negatively (bads), including services and tangible objects; "perfectly competitive markets" are those in which no individual buyer or seller believes they can affect market prices; "equilibrium" refers to a set of prices such that the amount of each good that consumers would like to buy exactly equals the amount producers would like to sell; and "efficient" is used in the broad economic sense. The first fundamental theorem does not say that free markets are efficient in all circumstances, only that they would be efficient in an imaginary world quite different from the real world. Subsequent work by Pigou (1932) and others established that equilibrium is generally inefficient if some goods are not traded at all, or are traded in imperfectly competitive markets. This failure limits the role of markets and defines a potential role for governments and nonprofit organizations.

Samuelson (1954) analyzed "pure public goods," defined as goods or services that are both nonrival (consumption by one person does not diminish any other person's consumption of that good) and nonexcludable (keeping some individuals from consuming the good is costly or impossible once it has been produced). A defensive militia provides a pure

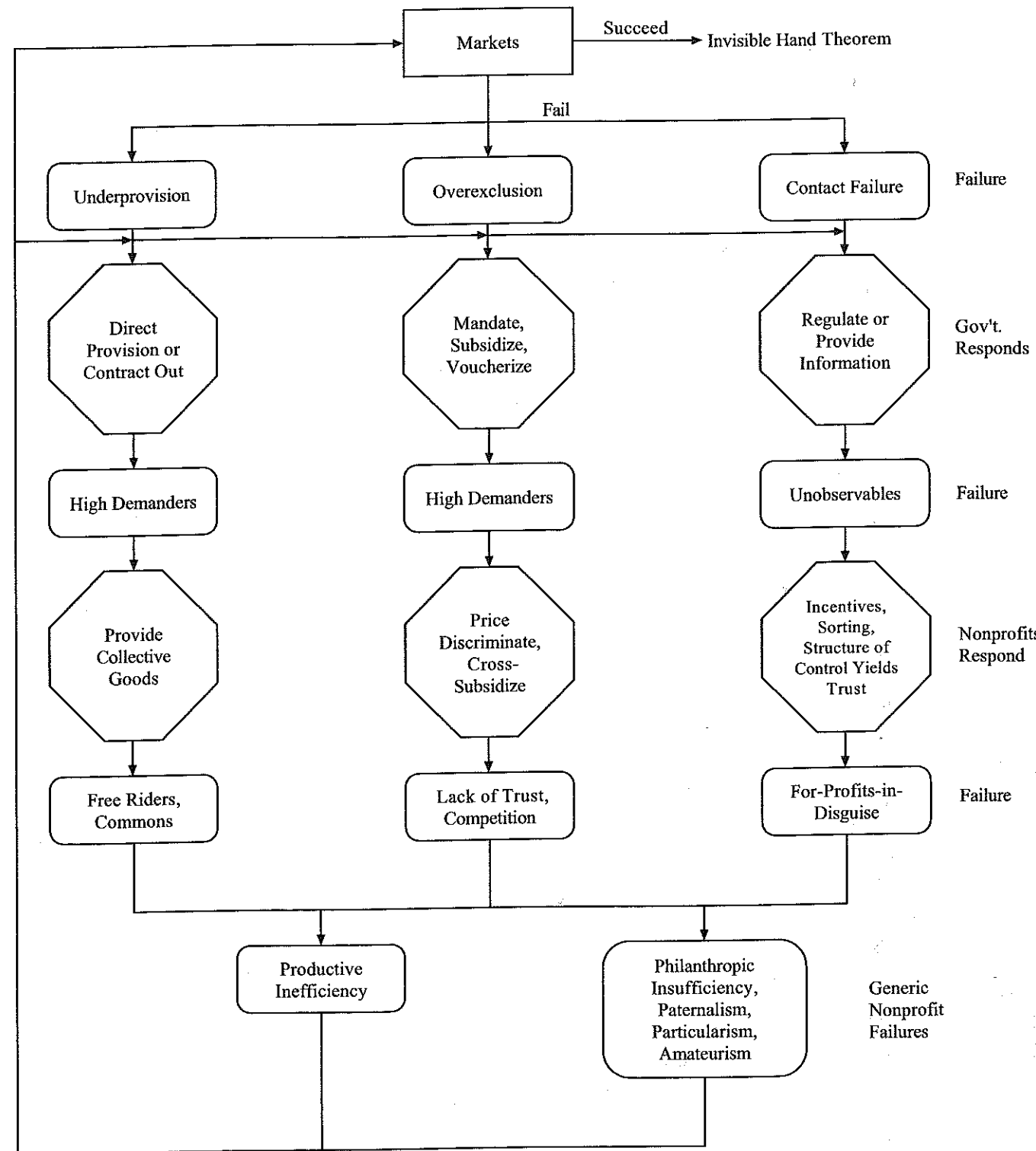


FIGURE 5.1. SCHEMATIC OF THREE-FAILURES THEORY

public good—birth of a new citizen does not diminish the quantity of protection enjoyed by other citizens (nonrivalry), and excluding any citizen from the protection provided by this army would be difficult (nonexcludability). For-profit firms do not produce pure public goods because consumers have the motive and opportunity to consume them without paying. This sort of market failure justifies governmental provision of national defense, hence the name “public goods.” However, some “public goods” are produced by private nonprofits, so it has become increasingly popular to call them collective, rather than public, goods.

Some collective goods, like the performing arts, are excludable though still nonrival. Keeping nonpaying customers from enjoying a performance is possible, but once the performance is staged, it does not diminish the enjoyment of existing consumers to let an additional person enjoy the show. Markets for excludable collective goods fail in a different way—although the market may provide the good, it limits consumption to paying customers. However, excluding borderline nonpurchasers (those who would enjoy the show but are not willing or able to pay the required entrance fee) is inefficient. Letting them attend would help them, hurt no one, and consume no additional scarce resources. Nonetheless, for-profit firms fear loss of revenue from paying customers if they let the borderline nonpurchasers in, resulting in an overexclusion market failure. Excludable collective goods may also suffer from the underprovision problem, although this is not our primary focus here.

A private good (for example, a surgical procedure) is the opposite of a pure public good—consumption is rival (no one but the patient is cured of appendicitis when the appendix is removed) and nonpayers can be kept from consuming it. Markets fail to provide the right mixture of private goods in certain cases of asymmetric information, where the seller of services knows more about the quantity or quality of delivered services than the buyer does. Hansmann (1980) extended the original insights of Nelson and Krashinsky (1973) as the theory of “contract failure.” Contract failure exists when: “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 quality or quantity 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” (Hansmann 1987a:29).

Contract failure arises when truthful information about the quality or quantity of delivered services cannot be purchased. Consider three cases where contract failure arises. First, suppose that instead of donating a sum of money to a nonprofit and asking it to feed the hungry in some foreign land, the prospective donor tries to purchase the same service from a for-profit. That customer is paying so that additional hungry people will be fed. The customer cannot easily observe whether they are fed, but could hire someone to find out how many people were being fed by the company in question. Still, perhaps some of these people are being fed

because of another customer’s purchase of food aid. The customer cannot learn whether his or her purchase *added* to the sum of people being fed, so that any explicit or implicit contract to buy food this way would fail.

The second example is a long-term care facility such as a nursing home. Weisbrod and Schlesinger (1986) noted that nursing homes, and services more generally, are bundles of easy-to-observe (type I) and hard-to-observe (type II) characteristics. Although for-profit firms can be trusted to provide the promised level of type I characteristics (room size, presence of medical staff), contract failure is likely for type II characteristics (whether residents are treated with due respect; whether sedatives are administered properly). Two other factors also contribute to contract failure here—the fact that the purchaser of services (often the adult children) is not the consumer of services (the resident) and the fact that if experience proves that contract failure is present in a particular facility, switching to another health provider is difficult for medical, social, and financial reasons.<sup>2</sup> These factors also affect other types of services such as day care for children and inpatient psychiatric care.

The third example comes when governments contract with private agencies to provide social services. Social services are complex and include many type II characteristics that matter to government contractors (e.g., Paulson 1988). For example, it is hard to tell whether foster home placements represent the best available match between caregivers and children. It is also difficult to figure out whether difficult-to-treat clients are steered toward other providers to cut costs or misclassified to reap higher contractual payments.

**Governmental Responses to Market Failure**

I detailed three sources of market failure above: underprovision of collective goods, overexclusion from excludable public goods, and contract failure. Each of these provides roles for the other two sectors,<sup>3</sup> and here I detail the governmental response. Governments solve the underprovision problem by either producing collective goods or paying a private-sector organization to produce them (contracting out). In effect, government payments complete the market, allowing trading for the collective of beneficiaries. For example, governments directly provide transportation infrastructure (highways and airports), collective recreational and conservation activities (parks), and reduction in the risk of theft or bodily harm (police). Periodically, government privatization efforts have led to the contracting-out of garbage collection, prison facilities, postwar reconstruction, and even public primary and secondary education.

Governments also address the market failure of overexclusion from excludable public goods in a variety of ways. First, when government produces the excludable public good, it sets the terms of exclusion. Many museums and zoos do not require payment of an admission fee, offer fee-free days periodically, or exempt favored groups (such as schoolchildren) from fee requirements. Second, government regulates for-profit providers, sometimes mandating



that nonpayers retain access to the collective good (for example, emergency phone service). Third, government gives selected groups special subsidies that enable them to compete for access to excludable public goods (such as housing vouchers for the indigent or work-study positions for eligible college attendees).

Governments address contract failure in a variety of ways. First, governments facilitate the enforcement of contracts, reducing the number of markets that fail in this way. Second, governments regulate the representations that firms make to the public through truth-in-advertising and fraud laws, labeling requirements, and the like. Third, governments limit entry into markets that suffer from asymmetric information problems through licensing and bonding requirements. Finally, government warns consumers of particular abuses and teaches consumers how to detect mistreatment. When government is the contractor, it deals with contract failure by monitoring for-profit providers more intensely (Ferris and Graddy 1991) and by negotiating longer, more detailed, and more complicated contracts (DeHoog 1984).

#### Government Failure

Subjecting government to the same formal scrutiny as markets, one uncovers a variety of sources of "government failure" (e.g., Wolf 1993). Many disciplines have contributed to the theory of government failure, which includes both efficiency and other considerations. I focus here on those government failures germane to three-failures theory.

Alternative models of political decision-making can be used to predict the levels of collective goods provided (or paid for) by government. Whatever the form of government, one result pervades—some citizens will be dissatisfied with the level, quality, or style of collective goods provided publicly. Tastes, tax burdens, and income differ, creating differing opinions on the ideal level of government spending. For private goods, differences in opinion are easily satisfied—consumers simply buy different amounts. For collective goods, what one person consumes is automatically consumed by everyone else, so that it is technically impossible to adapt to diversity of opinions in this way. Erik Lindahl developed a solution to this problem, today called Lindahl pricing in his honor, whereby each consumer is charged an individualized price for the collective good based on the intensity of their preference for it. Those who, all else equal, would like to buy a large amount of the collective good are charged a high price, those that would like to buy a little receive a low price. At the Lindahl prices, everyone would like to buy the same quantity and diversity in preferences is fully adapted to. The problem is that Lindahl prices are impractical and rarely used in the real world.<sup>4</sup>

Weisbrod (1975) focused on this sort of government failure in developing his theory of the role of the nonprofit sector. The word *failure* is used in a different sense here, to say that governments fail to provide collective goods at the level "high demanders" (those who would like to see the largest quantity or highest quality) would like. Whether or not gov-

ernment provision is efficient in the broad economic sense,<sup>5</sup> it will usually fail to meet the desires of some consumers. In particular, one theory suggests that public good provision in a democracy will be at the level the median-preference voter prefers.<sup>6</sup> Then, half the voters would like to see more provided, half less, and only the middle guy is completely satisfied with the government's choice.

Diversity of opinions leads to unsatisfied demand for collective goods by high demanders. However, the problem is reduced by the interplay of two factors—multi-level governments and migration. Some collective goods are consumed by everyone everywhere (a breathable atmosphere); others, like fire protection, are "local collective goods." Everyone in the United States consumes the same level of national defense, but the quality of fire protection, schooling, streets and roads, and public picnic areas is decided locally in accord with local preferences. To the extent that those who prefer differing levels of collective goods locate in different communities, voter dissatisfaction with governmental provision levels is reduced. Tiebout (1956) considered the logical extreme, with costless migration between an infinity of jurisdictions that provide local collective goods. In this limiting case ("Tiebout equilibrium"), dissatisfied voters move to a community of like-minded voters, and everyone is satisfied with the level of collective goods provided by their local government. Tiebout equilibrium describes some features of the real world, where families with children pick communities with good schools and childless couples pick communities with lower property taxes, but as a practical matter, some dissatisfaction with government always remains.

The market underprovision problem arises because consumers have both the incentive and opportunity to consume collective goods without paying. Government can solve this problem by coercing payments as current taxes (or future taxes if deficit spending is used), but this causes two types of failure absent from markets. First, individual payments are no longer voluntary as they are in markets. The ballot box provides a collective check on tax rules, but individual voters would rarely pay their full tax obligations if payment were not mandatory. Second, individual payments are not linked to either the amount of the collective good consumed or the individual consumer's willingness to pay for that good. Those with and without schoolchildren pay for the schools. The amount they pay is determined by the value of local property owned, income, or purchases of taxable goods rather than anything directly connected to their benefit from local public schools.

A second problem is that government cannot regulate abuses it cannot detect. Government is impotent precisely when contract failure is at its worst. This problem affects government regulation of for-profit sales of goods and services to the public and also affects sales to the government, as when government contracts with for-profits to provide social services.

Another limitation on government is self-imposed through constitutional restrictions on government action. When the government is prohibited from responding to majority demands for particular goods and services (most prominently

religious services in the United States) even the median-preference voter will be dissatisfied. Conversely, sometimes the government prohibits other sectors from responding to consumer demands for particular goods and services (James 1989). Then, the government role is too large compared with the unconstrained optimum.

#### Nonprofit Responses to Market and Government Failure

Markets fail and governments are only partly successful in addressing this failure. The twin failures in each of the three areas discussed above provide the basis for three theories of the role of nonprofit organizations in a multisector economy. Weisbrod (1975) put all this together for collective goods, Hansmann (1981a) for excludable collective goods, and Hansmann (1980) for contract failure. In this section I discuss each in turn, then the failures of nonprofit organizations that lead to roles for the other two sectors.

Weisbrod (1975) observed that nonprofits in almost every nonprofit industry, especially donative nonprofits, provide collective goods. Consider first those nonprofits that help the needy. At first glance, this seems to be a private good—soup is consumed by one individual, and nobody else can be nourished unless more soup is made. Soup is indeed a private good, but "helping the needy" can be enjoyed by everyone who cares about this group of people. Other altruists enjoy the fact that anyone has helped these people. If, in addition, helping the needy reduces crime rates, selfish individuals would also consume this collective good.

Consider next medical research. If a nonprofit organization finds a medicinal cure for cancer, the medicine itself would be a private good—pills cannot be collectively consumed in any useful manner. However, the knowledge that produced that cure is a collective good—the knowledge is not used up by anyone's pill consumption. A patent system makes that collective good excludable, and so we do see for-profit medical research firms. However, for-profits practice overexclusion that public health systems are only partly successful in countering.

The performing arts provide another example of an excludable collective good. Most of the costs of an opera or concert presentation are bound up in rehearsals, sets, costumes, and hall rental. If the hall is not full, an additional consumer can enjoy the show at no cost to existing consumers. Zoos and museums are similar in this respect.

Education produces both collective and private benefits. A good education enhances lifetime earnings—a private benefit. Education also improves the quality of democratic decisions and provides a common language and set of understandings that helps business and social interactions. Health care is another industry that produces mostly private benefits, but the treatment and prevention of contagious diseases provides a collective benefit for those who have not yet suffered from them. Advocacy is simultaneously about collective benefits for the advocating group and collective costs for their opponents.

Many forms of nonprofits, particularly those governed by members or in the religion and education industries, nurture

repeated interactions among stakeholders. This is hardly unique to nonprofits, but Ben-Ner and Gui (2003) argue that nonprofit organizations create better personal relationships among stakeholders than for-profits. Thus, nonprofits often serve as creators of the collective good "social capital," a network of relations that facilitates joint action. Religious and other nonprofits complement these relational contributors to social capital by nurturing moral codes and behaviors.

After observing that nonprofits commonly provide collective goods, that government provision of these goods is limited, and that Lindahl prices are not available in the real world, Weisbrod (1975) suggested that high demanders who are dissatisfied with government turn to the nonprofit sector to meet their desires for higher levels of service provision. Governments meet majority demands, and nonprofits meet those demands that do not yet or will never obtain majority support. Sometimes, the service in question is innovative, and the majority is reluctant to support it due to its newness. Then, nonprofits pioneer the idea, and government takes over funding or provision after the idea is proven (as in the Head Start program of early childhood education for disadvantaged groups). Sometimes, the disagreement is only over the size or quality of the collective good, and then nonprofit funding will persist. Sometimes the disagreement is fierce on both sides of the issue, such as on public funding of family planning clinics that discuss the abortion option. Then advocacy groups will be supported separately by high demanders for public funding (pro-choice) and low demanders (pro-life). Sometimes the disagreement is over ideological or cultural attributes (as in the arts [Steinberg 1990b] or primary and secondary education), where there is no clear way to arrange citizens as high and low demanders but there is a majority opinion and a set of private alternatives representing the various minority-support approaches (religious schools, Montessori academies, military schools, and so on).

Migration adds a few twists to the Weisbrod story. James's (1989) discussion points out that the nonprofit option is preferred when assembling dispersed communities of interest is easier than assembling stratified communities satisfied with local government provision. In the basic Weisbrod story, high demanders can supplement public spending on collective goods but low demanders are stuck. Wolpert (1977) noted that low demanders can migrate to a community that has lower taxes and expenditures on the collective good. This means that over time, the average preference for the collective good will increase. At some point, the difference between the preferences of the median voter and the preferences of the high demanders may shrink sufficiently that the nonprofit organization shifts from donative finance to government provision-of-service contracts.

High demanders pay for the supplement to public provision of collective goods with their donations. Unlike taxes, these payments are voluntary, and so in this sense nonprofits provide a superior alternative to governments.<sup>7</sup> However, donations suffer from the same disconnection between consumption and enjoyment of the collective good and amount of payments that taxes do. Perhaps, because the payments

are voluntary, the disconnection is less severe, but notions of obligation, patronage, fairness, and competition for prestige may structure voluntary payments on other bases.

Why do high demanders choose the nonprofit recipient rather than donating to the other sectors? Weisbrod did not explicitly model this decision, but his insights are formally developed in later work by Hansmann (1981a) and Bilodeau and Slivinski (1998). Donations to for-profits are rare because of contract failure.<sup>8</sup> Donations to government are rare for a similar reason—donors fear that governments react to their donation by cutting their own tax-financed support for the collective good in question. However, this same fear seems to affect donations to nonprofit organizations. In response to one's donation, other donors might reduce their contributions or government might reduce its grants and contracts with the recipient organization. Bilodeau and Slivinski (1998) point out that the nondistribution constraint keeps previous donors (notably the founding entrepreneur) from withdrawing their contributions in response to contributions by others, but the problem remains for those who donate contemporaneously. No doubt the repeated structure of interaction that fosters social capital helps, but this would not seem to explain, say, one-time donations to a nonmembership organization. Vesterlund's chapter in this volume explores these questions in greater depth, but puzzles remain whenever we try to make our intuitions on these matters explicit.

How do nonprofits deal with the overexclusion problem? First, we need to understand the market failure a bit better. If for-profit firms knew the maximum amount each consumer was willing to pay and had the power to charge a different price for each attendee ("perfect price discrimination") then overexclusion would not occur. Those willing to pay a penny would be charged a penny, those willing to pay a million dollars would be charged a million dollars, and everyone who enjoyed the excludable collective good in the slightest amount would voluntarily pay the entry fee. However, for-profits do not know enough about consumer willingness-to-pay to practice this strategy. Consumers would not reveal this information to a for-profit firm because it would be used against them. Profit maximizers would collect more than is necessary to provide the collective good in order to maximize their distribution of profits. Ben-Ner (1986) argues that the same does not apply to nonprofits, because of both the nondistribution constraint and the typical structure of nonprofit governance. Consumers might reveal their willingness-to-pay directly, enabling the nonprofit to establish more effective sliding-scale fee structures or financial aid to college students, or they might reveal this information implicitly through the donations they make on top of their purchase. Thus, Hansmann (1981a) refers to donations for excludable collective goods as "voluntary price discrimination."

Hansmann (1981a) analyzes how nonprofits respond to the combined problem of underprovision and overexclusion affecting high-cultural organizations in the performing arts such as opera companies, symphony orchestras, and dance

troupes. When, as often happens, most of the costs of production do not depend on the number of performances and consumer demand is moderately high, we have a good that is socially beneficial but won't be provided at all by markets. No single price will attract enough consumers to cover the costs of production, and for-profits cannot successfully practice price discrimination for two reasons. First, as I noted above, Ben-Ner found that consumers would not trust for-profits with the necessary information. Second, tickets can be resold. Nobody would buy tickets directly if they were charged more than others for the same quality seat. Instead, they would ask someone charged a lower price to buy extra tickets for them. Voluntary price discrimination is needed, and that can be provided only by a nonprofit.

Another way that nonprofits deal with the overexclusion problem is to use cross-subsidization (James 1983, 1998; Weisbrod 1998). Here, rather than charge different consumers different prices for the same product, the nonprofit charges higher prices for some products to generate financial surplus that can be used to lower the price, and so reduce exclusion, for other services. For example, zoos and aquariums use profits from gift shops and other concessions to finance an average admission fee that is far below the profit-maximizing level (Cain and Meritt 1998). Cross-subsidization is done for many reasons, of which one is to reduce exclusion from collective goods, and is discussed further in the chapters in this volume by Brown and Slivinski and by Tuckman and Chang.

Nonprofits help to solve contract failure in five ways. First, the nondistribution of profits reduces (or eliminates, depending upon the details of enforcement) the financial benefits from delivering less than the promised quality or quantity of services. Second, the nondistribution constraint affects the rewards of founding and controlling a nonprofit rather than another kind of organization. A process of "entrepreneurial sorting" takes place, and those residing in the nonprofit sector will have different personal objectives regarding what they want to accomplish in their role. Hansmann (1980) wrote about both these arguments, speculating that the sorting would enhance the trustworthiness of nonprofit firms. As we will see in later sections, this is not necessarily the case, although sorting is likely to be important.

Third, nonprofits are often managed by "demand-side stakeholders" (Ben-Ner 1986), those who care about the organizational output quantity or quality and not just their financial returns. Donors, members, and clients are demand-side stakeholders, who presumably want the organization to offer higher quality than others charging the same price and lower prices than others producing the same quality. This is in contrast with for-profit organizations, controlled by "supply-side stakeholders" (stock and debt holders) who want high prices (given the quality) and low quality (given the price). Nonprofit organizations and consumer cooperatives (which are owned by the consumers of the organization's output and receive profits as member dividends) are the two kinds of "patron-controlled organizations" in Ben-Ner's terminology. Both are more trustworthy than for-profits. As he

points out, it is good to have your child in a nonprofit day-care center but even better if the center-owner's children are also customers.

Fourth, nonprofit organizations are immune from financially based takeover bids as they do not have shares of stock that can be traded for profit. Thus, the dedication of the founding entrepreneurs to trustworthy behavior is not endangered by organizational transformation. This argument and its limitations have not, to my knowledge, been much explored in the existing literature.

Finally, the existence of some trustworthy nonprofits can have spillover benefits on the trustworthiness of competitors. Hirth (1999) develops a formal model of contract failure in which two types of nonprofit firms (trustworthy and opportunistic), one type of for-profit firm (opportunistic), and two types of consumers (informed and uninformed) are present. Informed consumers can detect contract failure when it occurs, whereas uninformed cannot. Opportunistic nonprofits are "for-profits-in-disguise," organizations that claim to be nonprofits and have received approval for all the tax and other benefits that accompany this status but secretly distribute their profits. The relative proportion of trustworthy and opportunistic nonprofits depends upon enforcement of the nondistribution constraint. Hirth shows that, depending upon the proportions of each type and enforcement, a market consisting of only nonprofits is often trustworthy whereas a market consisting only of for-profits is often not trustworthy. His most interesting result stems from the sorting, not of entrepreneurs, but of consumers. When for-profits and nonprofits compete with each other in the same market, they may both be trustworthy. Uninformed consumers, knowing their inability to detect contract failure, will patronize nonprofit organizations exclusively. This means that the for-profit's pool of customers contains a higher-than-average share of informed consumers, and so it will no longer pay to try to cheat them. The presence of a nonprofit organization creates a spillover benefit, making competing for-profits equally trustworthy.

#### Voluntary Failure

The third failure in the three-failures theory is by nonprofit organizations. Again, nonprofit organizations fail for many reasons, only some of which concern economic efficiency or the issues discussed here. Salamon (1987) was the first to organize four of these ideas as his theory of "voluntary failure," although we will use this label to include additional arguments. His four sources of failure included philanthropic insufficiency, philanthropic particularism, philanthropic paternalism (or, as I prefer, "parentalism"), and philanthropic amateurism.

Philanthropic insufficiency suggests reasons why nonprofit organizations have difficulty addressing the underprovision of collective goods, particularly in recessions, when the need is greatest. Voluntary action faces the "free-riding problem," discussed in more detail in Vesterlund (this volume). This includes several interrelated issues. First, as dis-

cussed above, donors might fear that rather than adding to total provision of some collective good, their donations would enable governments or other donors to withdraw their own contributions. Second, potential donors enjoy the collective good whether or not they contribute. Third, donors may not consider the external benefits they confer on others when they contribute, weighing only their own enjoyment of the collective good in their tabulation of the costs and benefits of their donation. My summary of this literature is that philanthropic insufficiency is a problem, but it is not as severe a problem as the simplest economic theories would suggest. In any case, nonprofit organizations can take many specific actions to reduce the importance of this problem.

Philanthropic particularism refers to the tendency of nonprofit organizations to focus on particular ethnic, religious, geographic, or ideologic groups, leading to duplication in some cases and gaps in coverage in others. To some extent, particularism is a natural consequence of fighting philanthropic insufficiency—it is easier to solve the free-rider problem in a community of similar individuals that repeatedly interact. Paternalism refers to the tendency of those who choose to work or volunteer for nonprofit organizations to treat problems as they perceive them, rather than as the clients perceive them. This is unlike government action, where clients have at least some small say through the ballot box. Amateurism refers to the tendency to rely less on credentialed workers, perhaps appropriate if client needs stem from moral problems rather than societal and technical factors. All these issues are discussed further in the chapters by Clemens and by Grønberg and Smith in this volume and by Douglas (1987).

How do nonprofit organizations fall short of curing the three market failures we have discussed? With respect to the underprovision problem, philanthropic insufficiency obviously limits the nonprofit ability to respond. In addition, when too many organizations representing too many causes compete for scarce donations, this causes problems. One organization's solicitation efforts can increase the costs of fundraising at other organizations (and so decrease the net funds available for collective good provision). This "commons externality" was first highlighted by Rose-Ackerman (1982) and is discussed further in Brown and Slivinski (this volume).

Nonprofit ability to solve the overexclusion problem depends upon whether consumers really trust the nonprofits enough to reveal their willingness to pay. Without such trust, nonprofits cannot charge high prices to high demanders, and so absent other sources of finance (such as government grants) they cannot subsidize prices below costs for the low demanders. In addition, there must be limited competition for high demanders. If, for example, a competing for-profit firm picked a price (or set of prices) only modestly above costs, the nonprofit could not set a higher price than its same-quality competitor. This would reduce the nonprofit's ability to subsidize prices for low demanders (Steinberg and Weisbrod 2005).

Nonprofits differ in their ability to combat contract fail-

ure for a variety of reasons. First, governments rarely devote substantial resources to monitoring and enforcing the non-distribution constraint. Second, when governments do enforce the constraint, detecting covert distributions is difficult. Nonetheless, it may be easier to detect distributions than shortfalls in output quantity or quality. Weisbrod (1988:22–23) concisely summarized the argument and its limitations: “We regulate what we can monitor easily, and we monitor what we can gauge usefully and inexpensively. If and when regulation of nonprofits per se is easier than direct regulation of outputs, production processes, or the distribution of output, the nonprofit form of institution is attractive.”

Thus, nonprofit markets will contain mixtures of genuine nonprofits, which do not distribute, and for-profits-in-disguise, which do.

In the basic contract-failure story, nonprofits do not cut corners and so have higher costs of production. They nonetheless break even because consumers are willing to pay more for the presumably higher-quality nonprofit outputs. However, if consumers believe that some “nonprofits” are actually for-profits-in-disguise, they will no longer be willing to pay such a large premium for products certified by the nonprofit label. Then, the bad drives out the good. Nonprofits that wish to provide the promised quantity or quality at the market price will have increasing trouble breaking even, and may compromise on quality or leave the market to the for-profits-in-disguise. Depending on the assumed details of entry, enforcement ability, and other subsidies granted to organizations that purport to be nonprofit, four outcomes are possible. First, as detailed above in the discussion of Hirth (1999), the honest nonprofits might force the opportunistic ones to behave well. Second, nonprofits and for-profits may occupy different niches, with nonprofits selling high quality at a high price, for-profits selling lower quality at a lower price, and no contract failure. This possibility is discussed in an alternative model in Hirth (1999). Third, honest nonprofits and for-profits-in-disguise may coexist over time, with average trustworthiness higher than it would be without some honest nonprofits and lower than it would be without some for-profits-in-disguise. Finally, for-profits-in-disguise can eliminate honest nonprofits, and so the nonprofit label becomes useless in solving contract failure (Steinberg 1993b).

Third, patron control does not work well when contract failure occurs for a private good (Ben-Ner 1986, 1987). We do not see nonprofit auto-repair shops because the manager can shortchange other consumers to improve the quality of the repair services he or she consumes. Patron control works best when important elements of product quality are jointly consumed by the manager/consumer and all other consumers.

Fourth, although the nondistribution constraint removes one incentive to shortchange donors and consumers, it does not remove other incentives. The organizational mission might differ from what the donor most desires. For example, donors might want to support increased student aid, whereas

university administrators want to support increased faculty research. Restricting their gifts in a formal legal sense is relatively easy for donors, but difficult in a more relevant economic sense because budgets can be reallocated to account for actual and likely gifts. Knowing that donors will support financial aid, the university can budget less of its other discretionary resources to the task, so that in effect the donations support additional research rather than additional aid.

At least some collective good is being provided in the preceding example. This does not have to be the case. Nonprofit managers may use donations and profits from sales to give themselves perks that do not help accomplish the organizational mission but are nonetheless legal forms of distribution. For example, nonprofits may locate in prestigious high-rent districts where they build magnificent headquarters. Nonprofit executives may travel first-class to conferences in exotic locations. Alternatively, incompetence, inattention, and indolence can flourish, protected from market competition by the donations, subsidies, and higher consumer willingness-to-pay that accompany the nonprofit label. In all these ways, well-meaning nonprofits can have the same erosive effects on trustworthiness as for-profits-in-disguise. Thus, whether nonprofits are trustworthy depends as well on the dedication of the board to the organizational mission and its vigilance in monitoring the managerial chain of authority (see Ostrower and Stone, this volume, for further discussion).

Finally, the nonprofit label cannot signal trustworthiness or elicit the financial support necessary to act trustworthily if consumers and donors are unaware of it. Surveys find that consumers do not always know whether the organizations they deal with are nonprofit or for-profit (Permut 1981; Mauser 1993). However, consumers may be aware of organizational characteristics associated with nonprofit status that serve as signals of trustworthiness (such as the religious affiliation of a day-care center). The mechanism Hirth (1999) suggests can still work if only some uninformed consumers are unaware of whether they are dealing with nonprofit or for-profit providers. Finally, regulators are certainly aware of sector, and often choose more stringent regulations for for-profit providers (Hansmann 2003). Further discussion of all these challenges to contract failure theory can be found in the chapters by Brody and by Schlesinger and Gray in this volume, Ortmann and Schlesinger (2003), and Hansmann (2003).

Productive and allocative inefficiencies provide additional forms of voluntary failure. The idea, which I have previously labeled the “property rights approach” (Steinberg 1987), is that the attenuated ownership structure of nonprofit organizations reduces owner incentives to care about things that the for-profit market does well. Productive inefficiency arises because the owners do not get to keep a share of financial residuals and so do not labor to keep costs down. As noted above, nonprofit managers may choose higher-cost perk-laden means of production, although here the emphasis is on social costs that occur whether or not consumers are misinformed.<sup>9</sup> The owners would not benefit financially

from lower costs, and there are no hostile takeover bids to force them to pay attention. Whether and to what extent this form of inefficiency occurs is quite controversial. The reader should see the chapters by Brown and Slivinski and by Schlesinger and Gray in this volume for further discussion.

Second, attenuated ownership raises the cost of capital, all else equal. Nonprofit organizations cannot sell meaningful shares of stock to raise capital, and so must rely more heavily on debt. Hansmann (1981b) argues that this raises the cost of capital, although the exemption of U.S. nonprofits from the corporate income tax serves as a crude corrective. Higher capital costs lead to an inefficient mix of inputs and inadequate or slow response to increases in demand for outputs. This creates both productive and allocative inefficiency. Again, this form of voluntary failure is controversial, and the reader should see Brown and Slivinski (this volume) and Bilodeau and Steinberg (forthcoming) for further details.

Third, attenuated ownership means that owners who follow changing consumer tastes and demands and innovate accordingly are not rewarded for this attention. Sometimes, nonprofit organizations are proud of their failures here, as in higher education. Nonprofit universities are proud to provide, paternalistically, what they think their students need rather than what they shortsightedly want. However, the pressures of competition with new for-profit universities and less-paternalistic universities are eroding this difference. The definition of economic efficiency does not allow for paternalism, so whether one regards these differences as good or bad, they show up as a form of inefficiency.

#### Closing the Circle: Reacting to Voluntary Failure

Three-failures theory does not presume that any sector is “first” and the other sectors react to its failures. Rather, the approach arranges the three sectors around a circle, with each reacting to the failures of its two neighbors. Weisbrod’s (1975) exposition, followed here, has nonprofits responding to failures by the other two sectors, which allows the argument to proceed linearly but perhaps falls short in developing our intuition about the whole circle. Salamon (1987) recognized this shortcoming, and began his exposition with nonprofits as the first sector, whose failures are addressed by government and for-profits. This brought the new insights regarding voluntary failure. Regardless, we need now to specify how the other two sectors respond to this voluntary failure.

Historically, the Salamon approach may be more accurate, although it is a bit hard to tell because the definitions of the respective sectors have been even fuzzier in the past. In modern times, it is often (but not universally) the nonprofit sector that is the first to respond to a natural disaster or the first to carry out a social innovation because of the natural inertia in government action. Salamon (1987) noted that governmental action requires, in order, public arousal, information gathering, passage of laws, and establishment of a

bureaucracy to carry out those laws. Some of these steps are carried out in advance for response to natural disasters, so it remains an open question whether one expects nonprofits to be the first responders. Regardless, voluntary failure limits that response and government supplies additional resources in a less particularistic and more credentialed way.

Salamon did not address how for-profits close the circle. Other literature makes clear that for-profits respond well to productive inefficiency wherever it occurs and to allocative inefficiency in markets for private and some excludable collective goods. Low-cost production, innovation, and attention to consumer demands are the hallmarks of the sector because the owners benefit and because these actions eliminate the takeover threat. When for-profits and nonprofits coexist, competition forces nonprofits to respond likewise or go out of business. This competition is limited because of a variety of “cushions”—subsidies, tax exemptions, and the like provided to nonprofits but not their for-profit competitors—that allow nonprofits to function distinctively (Steinberg 1991, 1993b).

#### Empirical Evidence

Evidence supports many propositions discussed above. Clearly, nonprofits provide collective goods, but Weisbrod’s model suggests that we can predict the relative roles of government and nonprofits in financing these goods. Specifically, he argues that the more heterogeneous a society is, the more dissatisfaction there will be with government provision levels and therefore nonprofit financing of collective goods will be larger. Note that he is not talking about the size of the nonprofit sector, which is largely paid for with government money as grants and contracts, but only about the donatively financed portion of nonprofit expenditures. Thus, the result of Salamon and Anheier (1998) that the nonprofit sector is smaller in more heterogeneous countries is not quite on point (Steinberg and Young 1998). In contrast, James (1993) finds that after controlling for government subsidies to private education, more heterogeneous countries rely more heavily on nonprofit primary and secondary education institutions. She tests various measures of heterogeneity. Religious heterogeneity has the largest effect, and measures of linguistic heterogeneity and income diversity have smaller and less-statistically significant effects. Feigenbaum (1980) used variation over time to explain state spending on income redistribution and total donations in the United States. She finds no statistically significant impacts of heterogeneity on government spending, but a very significant positive correlation between heterogeneity in age and donations. Finally, Chang and Tuckman (1996) find that nonprofits in racially diverse communities rely on donations for a greater share of their revenues.

Many other chapters in this volume present evidence on contract failure (especially Brown and Slivinski; Schlesinger and Gray; and Kendall, Knapp, and Forder), so I will present only an overview here. It is extremely hard to test contract failure theory because it concerns unobservables.



If the author of any study could reliably detect differences in the trustworthiness of organizations, presumably governments could too and they would directly regulate the behavior in question, rather than merely reducing the temptation to cheat. Faced with this dilemma, five strategies have been employed, with the greatest volume of studies testing contract failure in the health-care, day-care, and nursing-home industries. First, some studies look for differences in characteristics that will be observable by some consumers (the informed) but not all. This strategy is particularly effective where the unobservable characteristic is valued differently by different consumers, so that uniform government regulations would reduce desirable diversity in the marketplace. Second, some studies have used indirect tests based on the number of complaints filed with the government determined to be nonactionable because they concerned matters on which regulations had not been set. Third, some studies have used indirect tests based on how the respective sectoral market shares have changed when technological improvements have reduced the cost of monitoring or when state regulations have been changed. Fourth, studies have compared the experiences of different types of consumers—those who search extensively for a provider versus those who do not, or those who are deemed to be at special risk for exploitation versus those who are not. Finally, some studies have simply asked consumers or government contractors why they make the sectoral choices they do.

On balance, it is my opinion that the evidence supports the predictions of contract failure theory, but others have looked at the same studies and come to a different conclusion. The importance of contract failure versus various sorts of voluntary failure has not, to date, been well assessed, so the case for systematically preferring nonprofit providers is, at best, incomplete. The importance of contract failure likely varies across nonprofit industries, and changes with shifts in technology and governmental regulation.

Hundreds of studies attempt to test for efficiency differences between nonprofit and for-profit organizations, and many of these are reviewed in other chapters in this volume (notably Schlesinger and Gray; Kendall, Knapp, and Forder). Most consider the competition between generic nonprofits and for-profits, but some distinguish types within each sector (secular versus religiously affiliated, chain versus independent). A majority of studies conclude that nonprofits are less efficient, but many studies find either no difference or a difference in the opposite direction. I remain skeptical of the conclusions drawn by these studies because of methodological difficulties noted in these chapters and elsewhere. My biggest worry is that organizations produce a multiplicity of outputs, some of which are excluded from available data and some of which are, by their nature, impossible to measure objectively. To the extent that nonprofits produce more unmeasured outputs, the costs of the measured outputs will be overstated and nonprofits will seem less efficient. For example, hospitals produce a collective benefit that is typically omitted from empirical studies—the assurance that in case of a sudden increase in demand due to

epidemics or disasters, they will have the capacity to treat everyone (Holtmann 1983). Are unused hospital beds evidence of productive inefficiency, or of the efficient production of capacity insurance, an output excluded from empirical studies?

A related strand of literature assumes that nonprofit inefficiency is limited by competition with for-profit providers. Thus, for example, many studies (summarized and evaluated in more detail in Steinberg 1987) compare the cost per claim processed by nonprofit and for-profit health insurance firms. The studies find that nonprofit costs are higher by an amount that varies with the size of the tax breaks given to nonprofits but not their for-profit competitors. The authors of these studies argue that the two sectors are doing the same thing, one at higher costs, so that higher costs represent productive inefficiency. However, claims processing can be done carefully, with every valid claim approved and every fraudulent claim disapproved, or less carefully to cut costs or increase revenues. Without measuring the quality of claims processing by the two sectors, we cannot tell whether nonprofits are inefficient.

A third strand of literature focuses on the reaction of both sectors to changes in demand. Steinwald and Neuhauer (1970) and Hansmann (1987b) find that the nonprofit market share is lower in markets that are rapidly expanding, which is consistent with either a failure to pay attention, lack of capital for expansion, or paternalistic preferences. Hansmann (1996) argues that nonprofits are less likely to exit markets when demand for services decreases; a later article provides modest supporting evidence (Hansmann, Kessler, and McClellan 2003).

#### Shortcomings of the Three-Failures Theory

Three-failures theory, at least in my exposition, is incomplete.<sup>10</sup> The various pieces explain why consumers would want to buy from and donors donate to nonprofits, but do not explain why nonprofits are there for them to use. What is needed is a theory of the supply of this organizational form to complement the theories of demand. Unless we know why and when nonprofit organizations will be created, it is hard to assess whether they can play the roles we have discussed. Predicting the objectives and behavior of individual organizations is also hard. How will they respond to changes in public policy, competition, the economy, or technology? Understanding the coexistence of providers from each sector in the same service industry is also difficult. If nonprofits are more trustworthy, why do they not drive their competitors out of business? If nonprofits are less efficient, do they survive only because of subsidies? If they have counterbalancing trust advantages and efficiency disadvantage, is the nonprofit market share entirely arbitrary or do economic theories have more to say on the matter?

The second problem with three-failures theory is its excessive focus on efficiency, in the broad economic sense. Efficiency is certainly important, but it leaves out much and the other sorts of roles that nonprofit organizations can play

in a mixed-sector economy are ignored by this literature. Efficiency concerns the size of the economic pie—whether the most-valued mixture of outputs is produced. It says nothing about how the pie is shared by consumers, about distributional justice. The fair distribution of income is a much more controversial matter, but one where arguably for-profits fail and governmental redistribution is limited. Nonprofit missions talk of helping the indigent, of providing affordable housing, of assuring that nobody is denied medical care because of insufficient income, and the other sectors leave room for nonprofits to play this role. The literature has made only the barest of starts in understanding this nonprofit role (e.g., Clotfelter 1992; Steinberg and Weisbrod 1998), but it rests well within, if not exclusively within, the economists' toolbox.

Efficiency is defined with respect to preexisting consumer preferences. These preferences determine the value placed on various goods and services for use in determining the value-maximizing mix. Yet, the stated mission of many nonprofits is to change those preferences—to make people want to enjoy a habit of lifetime learning, worship God, preserve the environment, stop child abuse, or respect the decisions made by gun owners. Advertising, social marketing, and advocacy play a role in market efficiency, but I believe persuasion involves more than just informing those with preexisting preferences. Governments and for-profits also play roles in seeking to change preferences. Much more research should be conducted on the distinctive roles best played by each sector. Perhaps economists are not the best ones to conduct this sort of research, but the work that economists do should not blind us to the importance of these other roles.

Three-failures theory also leaves out other roles for the sector. For example, Mason (1996) talks of the instrumental, expressive, and affiliative roles of the sector. Economists study the instrumental role—the use of nonprofits to obtain stated objectives such as finding job placements or feeding and housing the indigent. However, work in the nonprofit sector is not just about doing. It is about making statements and being with others. Philanthropic amateurism and particularism may seem like failures when viewed through the instrumental lens, but they are valued roles for the sector when viewed through the expressive and affiliative lenses. Nonprofits play many additional roles, omitted here because they are best discussed by sociologists and political scientists (see the chapters in this volume by Grønberg and Smith and by Clemens). Nonetheless, these roles should always be kept in mind.

#### TOWARD A COMPLETE THEORY

Three-failures theory is incomplete, omitting the supply side and focusing on the efficiency roles of the sector to the relative exclusion of other roles. In this section, I address some of these holes. However, just filling the holes is not enough—the supply side interacts with the demand side so that, for example, whether nonprofits deal with the under-

provision problem depends on which theory of supply we adopt. We need an integrated approach, where all the various pieces are mutually consistent and jointly establish the respective roles of each sector. I begin this section by laying out the framework of a complete theory, then discuss the various supplements and partially integrated theories offered to date.

The theory of long-run supply by for-profit firms is well established. Every firm wants to maximize its profits. If demand increases, more firms enter the market, but each of these new entrants also wants to maximize profits. There is a logical separability between the number of firms and their respective objectives. This is sensible since for-profit firms that did not want to maximize their profits would not survive as such due to competition and the market for takeover bids. The situation is more complicated in the nonprofit sector, because no overwhelming force automatically makes new entrants have the same objectives as existing ones. Thus, I argue in my 1993a paper, a satisfactory model of nonprofit supply should simultaneously explain the decision to enter and the objectives of those who do enter. This aspect of theory integration leads to a clearer understanding, as the following example illustrates. A partial theory based on Hirth (1999) would conclude that under specified circumstances, nonprofit organizations would remain trustworthy despite the presence of a few for-profits-in-disguise. Another partial theory would say that if a trustworthy nonprofit organization receives a government subsidy, it would expand and the market would function better. Putting these two theories together without any explanation for where organizational objectives come from would lead to the conclusion that government subsidies to nonprofits improve the efficiency of the market. However, Hirth actually has an integrated theory with respect to this example, where an increase in government subsidies would attract more for-profits-in-disguise into the market. In turn, competition with for-profits-in-disguise makes it more difficult for legitimate nonprofits to remain trustworthy and survive. Subsidies can make contract failure worse. The theory that combines objectives and entry gets a different, and more accurate, conclusion than the theory that separates the two.

Some progress is made if, like the first models of nonprofits, one simply asserts some specific organizational objective and explores the consequences of having that objective. Much more progress is made if the theory includes those factors that determine which organizational objectives are likely to emerge and thrive. Then we can better understand why nonprofit organizations play different roles in different countries, at different points in time, or following a change in public policy. Nonprofits first respond to the changing conditions, using their preexisting objectives, and then their objectives evolve, provoking further response.

Figure 5.2 provides a schematic diagram to aid our thinking about complete and integrated theories.<sup>11</sup> The supply decision is made by those who are tempted to found a new nonprofit organization or those who might want to maintain or transform existing nonprofits. For convenience, I refer to

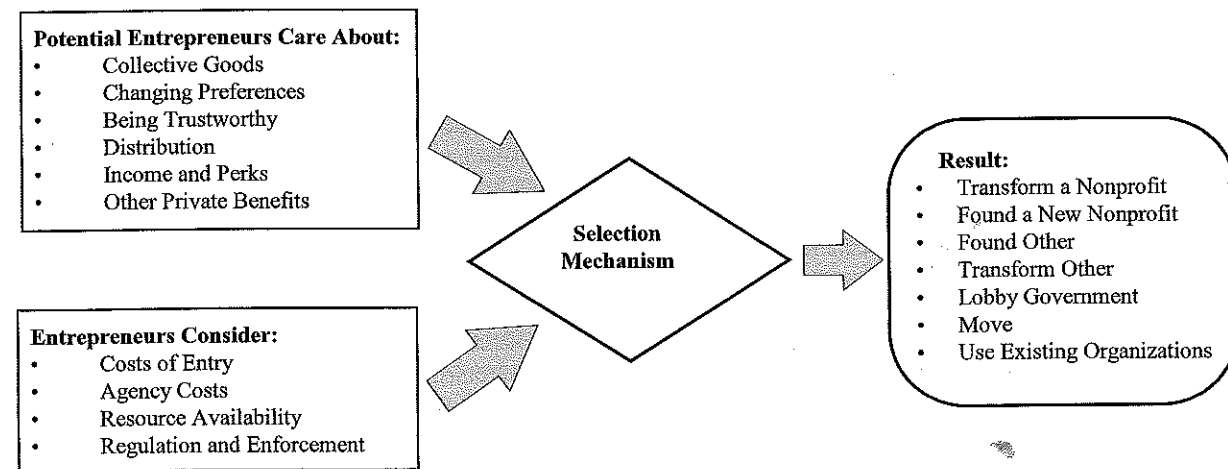


FIGURE 5.2. SCHEMATIC OF A COMPLETE THEORY

both kinds of actors as "entrepreneurs." Defined this way, entrepreneurs are a logical construct encompassing founders and agents of change, rather than specified persons.

The top left part of the figure tells us what potential entrepreneurs care about. Potential entrepreneurs consider how best to obtain their objective—by founding a new nonprofit, transforming the behavior of existing nonprofits, preserving the mission against pressures to change, founding or transforming a different kind of organization (for-profit, consumer cooperative, labor-managed firm, and so on), or working to change the decisions of government agencies. Alternatively, they take a nonentrepreneurial role—supporting existing organizations as donors, volunteers, and customers or moving to a location where existing organizations meet their needs. They decide by considering their own objectives and the various factors that make those objectives easier to achieve if they pick the nonprofit form. The factors that determine whether objectives can be accomplished through nonprofit entrepreneurship (founding or transforming) appear on the bottom left, and the possible outcomes of entrepreneurial decision-making appear on the right of figure 5.2.

Once each potential entrepreneur has pondered what to do if granted control of the organization, a selection mechanism or internal political process determines who is given that control and nonprofit behavior flows from this choice. The same process of decision-making leads to the emergence, or nonemergence, of competing organizations and simultaneously determines the objectives of competitors.

This approach is clearly inspired and influenced by a paper and later book by Young (1981, 1983). He argues that the objectives of nonprofit organizations are determined by the objectives of their entrepreneurs, who establish the organizational culture and/or write the articles of incorporation and the bylaws. He lists a variety of pure entrepreneurial

types, garnered from the management literature, and speculates on the factors that led some types to prefer to work in the different sectors and in different industries. This began the "entrepreneurial sorting" approach, which, I think, has yet to bear its fullest fruit. It is also inspired and influenced by Ben-Ner and Van Hoomissen (1991). Both will be discussed in context below.

#### Alternative Objectives

In the first part of figure 5.2, I list six categories of entrepreneurial objectives. Entrepreneurs might care about collective goods, either because, as a consumer, they want to guarantee that someone makes them available (as in Bilodeau and Slivinski 1996, 1997, 1998) or because they get a special warm glow from playing a personal role in providing them to the community (as in Eckel and Steinberg 1993; Steinberg and Eckel 1994; or Roomkin and Weisbrod 1999).

The second category of objectives involves changing the preferences or consumption behavior of others. This includes a variety of missions—conversion of others to one's faith or ideology (James 1982, 1986, 1989; James and Rose-Ackerman 1986), cultivation of tastes for the arts (Throsby and Withers 1979; Hansmann 1981a), and various forms of social engineering, such as encouraging people to stop smoking, enjoy safe sex, use their seat belts, or exercise more. Strong believers would have sufficient motivation to incur the substantial costs required of entrepreneurs, and may prefer the nonprofit form if doing so signals a sincerity that facilitates the conversion of others.

Third, some entrepreneurs care about not abusing the trust consumers and donors place in them. Thus, some models of nonprofit behavior include trustworthy entrepreneurs among the set of potential founders and managers (such as Schiff and Weisbrod 1991). Fourth, potential entrepreneurs

may care about the distribution of income, either directly or about changes in that distribution that they can take credit for (Steinberg and Weisbrod 2005).

The fifth category is income plus perquisites. This objective is used in the property rights literature (e.g., Frech 1976) but also appears in Preston 1988; Gassler 1989; Schiff and Weisbrod 1991; Eckel and Steinberg 1993; Steinberg and Eckel 1994; and Glaeser and Shleifer 2001. By perquisites, I refer here to job attributes other than monetary compensation that are valued by the entrepreneur but not by others, including first-class travel, fancy offices, and on-the-job leisure. In a sense, the warm glow from providing public goods is also a perquisite. This is a job attribute the entrepreneur might value, but it also provides benefits to others so, for example, Steinberg and Eckel (1994) call this a "public-benefit perk."

The last category is a catchall, intended to remind us of all the benefits that economists typically leave out of their models. This includes desires for power, control, expression, affiliation, legitimation, and the like. Many entrepreneurial types listed in Young's (1981) original typology belong here, including "artists" (who value the creative act in reshuffling organizational blocks or creating a whole representing their vision), professionals (who value the pursuit and development of new ideas), searchers (out to prove themselves), independents (who want to avoid sharing authority and decision-making), conservers (loyalists who innovate only during crises), and power seekers (who either like to control the individuals that work for them or want a larger stage to wield power from).

My classification of entrepreneurial motives does not necessarily show the motives and behavior of the organization for three reasons. First, many entrepreneurs will have multiple motives. This idea is exploited in Eckel and Steinberg (1993) and Steinberg and Eckel (1994), where potential entrepreneurs care about varying mixtures of collective goods and private-benefit perks. Second, other factors (such as the dependence of resources on the nonprofit output mix, contractual stipulations, and public regulation) may force entrepreneurs with one objective to pursue alternative goals. For example, Preston's (1988) entrepreneurs care only about their own income, but donors reward those entrepreneurs that provide collective goods with higher incomes. Finally, the internal political dynamics of the organization may cause the organizational objectives to stray from those of the original entrepreneurs.

Many existing models of nonprofit organizations postulate an organizational objective consistent with mixtures of these entrepreneurial motivations. For example, Newhouse's (1970) nonprofits maximize a mixture of the quantity and quality of their output. He does not explicitly incorporate entrepreneurs. However, this organizational objective could stem from entrepreneurs that care to change the preference and consumption of others (refining their tastes so that they appreciate a "higher quality" of artistic expression). Alternatively, it could result from those that want to be trustworthy

(and so provide the promised high level of quality despite opportunities to do otherwise), from those that value provision of an excludable collective good (who deal with the overexclusion problem through quantity maximization), or from those that care about prestige (an element of the catchall category). Tullock (1966) and Niskanen (1971) postulated that some nonprofits care about maximizing the budget under their control. This organizational preference might stem from entrepreneurs that care about prestige or income (which are generally higher the larger the organization). Lee (1971), Pauly and Redisch (1973), and Feigenbaum (1987) have nonprofits that care about their use of preferred inputs, which can be doctors, high-tech equipment, or disabled employees. Finally, Malani, Philipson, and David (2003) provide a general model, where three of the hybrid motivations provided in other literature are special cases. Unfortunately, none of these models start with entrepreneurial preferences or otherwise incorporate a model of how and when organizational objectives would change.

#### Factors Hindering or Aiding Accomplishment of Objectives through the Nonprofit Form

The theory of market failure already tells us why entrepreneurs who want to provide collective goods or operate trustily would prefer the nonprofit form. True, a for-profit firm could donate all its profits to charity, but especially for publicly held firms (where shares of stock are openly traded), the threat of a takeover bid limits such behavior. Nonprofit organizations are different. Nonprofits that want to provide collective goods are immune from takeover and any initial investment by the entrepreneur is supplemented by the donations of others. This bit of foreshadowing suggests that we enumerate the factors that promote or hinder the entrepreneur's ability to accomplish his objectives through the nonprofit organizational form.

First, there are costs of entry (or of transforming the mission of an existing nonprofit). Ben-Ner and Van Hoomissen (1991) detail these costs as including: (a) identifying and assembling a collection of willing stakeholders; (b) determining whether collective demand is sufficient to cover costs; (c) organizing production decisions; (d) inducing stakeholders to truthfully reveal their preferences; and (e) establishing a governance mechanism to ensure stakeholder control against free-riding, internal incentive problems, and the like. A new organization is formed if the expected flow of net benefits exceeds the flow of net benefits from the next best alternative.<sup>12</sup> Thus, for collective goods, the entrepreneur compares the expected net benefits from current governmentally supported provision with those if a new organization is created and makes his choice accordingly. If government is not supporting the entrepreneurial cause at all, the entrepreneur compares the expected net benefits from creating a nonprofit with the net benefits from lobbying government to meet the need. In this calculation, the entrepreneur recalls Salamon's (1987) list of the transactions costs



of governmental action: public arousal, information gathering, passage of laws, and establishment of a bureaucracy to carry out those laws.

Most of the costs of founding a nonprofit are reduced if the founder is a member of a group of like-minded individuals who repeatedly interact with each other and so gain mutual trust. This happens through clubs, alumni groups, and most importantly religious congregations, so it is not surprising that many nonprofits are founded through a seed organization like a congregation.<sup>13</sup> In contrast, the costs are increased if the pool of potential stakeholders is diverse. Although all might be high-demanders, if they thoroughly disagree among themselves about what expenditure level on collective goods would be optimal, they might not agree to mutually form a new organization to meet their high demands. This suggests a revision of Weisbrod's theory—preferences for the collective good must be heterogeneous (otherwise government suffices) but lumpy, with a cluster of agreeable high-demanders willing and able to work together (Ben-Ner and Van Hoomissen 1991).

Agency costs result whenever the entrepreneur (the principal) requires the assistance of others (agents) whose objectives are not thoroughly aligned with his own. The principal uses costly mechanisms that realign the agent's incentives (such as profit-sharing plans), monitors the performance of agents, and accepts that some failure will remain after incentives and monitoring are carried out. Agency costs are likely to differ with the prospective choices that an entrepreneur makes. Profit-sharing plans are somewhat restricted in the nonprofit sector by the nondistribution constraint, but to the extent nonprofit workers are motivated by accomplishment of the nonprofit mission, they receive a distribution in kind that plays a similar role (Slivinski 2002; see also Brown and Slivinski, this volume, and Bilodeau and Steinberg, forthcoming).

One agency cost recognizes the finite duration of entrepreneurial control. Entrepreneurs can lose control during their lifetime if their for-profit firm is taken over or if their nonprofit board moves in a different direction, and certainly lose control after their deaths. Closely held corporations, with dynastic control, reduce this problem in the for-profit world, but for at least some purposes, establishment of a charitable trust or corporation provides a better solution.

The availability of resources also determines entrepreneurial choice. In the for-profit sector, entrepreneurs can obtain capital by issuing shares of stock or obtaining loans. Nonprofits can do only the latter, and because any loans are not backed by at-risk shareholder investments, the cost of debt is higher for nonprofits (Hansmann 1981b). However, nonprofits can use grants and donations, volunteer labor, and (depending upon local laws) tax-exempt bonds to obtain capital, and can accumulate retained earnings for investment more quickly if they are exempt from corporate income taxes. Resources from sales or from fundraising depend, in part, on competition from other organizations, so that entrepreneurs should also consider the likely evolution of competitors over the lifetime of their mission.

Finally, the way in which organizations in the various sectors are regulated, and the quality of enforcement of those regulations, determines where the entrepreneurial objective can be best accomplished. We have already seen how these factors affect the emergence of for-profits-in-disguise; the other ways they matter are numerous and self-evident.

The integrated approach I advocate assumes that the entrepreneur picks the organizational form that best accomplishes his objectives. That said, nonprofits arise in situations where it is very difficult to accomplish objectives using any kind of organization. The correspondence between organizational behavior and entrepreneurial objectives need not be very close, as entrepreneurs are hindered in this "least-worst" situation by transactions costs, agency costs, resource dependencies, and governmental regulations that make it hard to transform objectives into results.

### Fitting the Pieces Together

There are many ways to fit these various pieces together correctly and gain deep insight into the role of the nonprofit sector. The literature has only begun this process. Perhaps the best effort to date is found in a series of papers by Bilodeau and Slivinski (1996, 1997, 1998) and by Bilodeau (2000), employing variations on a common structure. In all these models, a group of individuals cares about a collective good. Comparing the benefits of founding a new nonprofit or for-profit to provide that good with those from supporting organizations founded by another member of the group, one of them chooses to be the entrepreneur. She invests an initial sum of money toward providing that good, adds any resources obtainable through donations by others, and then devotes the total amount to the collective good. If the entrepreneur picks the for-profit form, she is allowed to withdraw some of the initial investment after seeing how much other people donate, but if she picks the nonprofit form, any such withdrawal would violate the nondistribution constraint.

One puzzle for any theory of supply is why the founder would choose to permanently constrain her future option to receive profits. Dividends can always be donated, and it would seem that the entrepreneur would like to retain the right to keep some dividends if the venture proves more than profitable enough to support the collective good. Bilodeau and Slivinski provide an answer to this puzzle: other donors consider the organization's sector in making their decisions. If they knew that the entrepreneur could withdraw part or all of her initial investment in response to their donation, they would be reluctant to give. Thus, those entrepreneurs that need the donations of others to accomplish their goal would want to give up the right to receive dividends.<sup>14</sup> The various papers by Bilodeau and Slivinski extend this approach to explain how nonprofits compete with each other or form a united campaign, how they compete in commercial markets also populated by for-profits, and how enforcement of the fair compensation constraint affects performance.

Several papers tackle aspects of entrepreneurial sorting. In Bilodeau and Slivinski (1996), potential entrepreneurs

differ in entrepreneurial costs and the value they place on the collective good, but they focus on who emerges as a nonprofit founder without considering other entrepreneurial options across the sectors. Gassler (1989) and Schiff and Weisbrod (1991) assume that entrepreneurs are of two types (those valuing only profits, and those valuing other things). These types sort perfectly across the for-profit and nonprofit sectors respectively. Steinberg and Eckel (1994) assume that potential entrepreneurs value three things: income, private-benefit perks, and public-benefit perks, varying in the relative importance placed on the last two factors. They show both short-run effects of competition and tax policy (responses by preexisting nonprofits) and long-run effects (due to changes in the type of entrepreneur who locates in each sector). It is unclear at this time how much their results will generalize.

### Evidence

What evidence is available on the prevalence of various entrepreneurial and organizational objectives? One could simply ask those in control what they are trying to do, or collect and analyze organizational mission statements. I am not aware of any studies that systematically survey founders, managers, and board members or that categorize mission statements according to the set of objectives detailed above,<sup>15</sup> and doing so would be worthwhile. However, this approach is problematic. Mission statements specify multiple, often competing, objectives without detailing how one objective is weighed against another. Mission statements are intentionally vague, as too much specificity risks alienating selected groups of stakeholders. The stated mission of a nonprofit organization will sometimes differ from its real objectives.

Instead of surveying the stated objectives of organizations, researchers have analyzed nonprofit behaviors to detect the "revealed objectives" of the organization—those objectives which organizations act as if they are trying to achieve. For example, Steinberg (1986) examines whether organizations act as if they want to maximize their net resources available for service provision (contributions minus fundraising costs) or their total budget (contributions) when they conduct fundraising campaigns. Service maximizers spend until the last dollar of fundraising expenditure brings in one dollar of added donations (with every previous dollar generating net resources), whereas budget maximizers spend until the last dollar brings in no additional donations. Steinberg finds that "welfare" organizations act like service maximizers; "education," "arts," and "research" organizations spend less on fundraising than they would if they were service maximizers, and "health" organizations act like budget maximizers. Other studies using similar methods challenge his results and the question remains open.<sup>16</sup>

Steinberg's study illustrates both the merits and drawbacks of the revealed objectives approach. Mismatches between stated and revealed objectives can be due to inaccurate statements, resource dependencies or regulations that

impinge on behaviors, managerial errors or lack of knowledge, or the inability to control the actions of employees, subcontractors, and volunteers. For example, those in control of arts organizations may want to maximize the net returns from their fundraising campaign but lack up-front capital to do so. Still, behavior counts. Those in control of health organizations may have no intention of maximizing their budget, but the fact that they appear to act like budget maximizers is important.

Lowry (1997) generalizes this approach to include the effect of other sorts of nonprofit expenditure on total revenues and to broaden the class of objectives studied. He theorizes, like James (1983), that nonprofit managers have favored, neutral, and disfavored activities. Managers may care about revenues for their own sake (as in budget maximization), for their ability to support increased provision of favored activities, or both. He then analyzes a panel of citizen environmental groups and finds that they spend too little on fundraising (suggesting either that they are capital-constrained service maximizers or that they view fund-raising as a disfavored activity) and too much on collective goods (suggesting that the provision of collective goods is a favored activity) compared with the surplus-maximizing expenditure levels. He also finds that spending on selective incentives and information (loosely, member benefits) is excessive, suggesting that this too is a favored activity. A similar approach is taken by Vitaliano (2003), who compared religious nonprofit, secular nonprofit, and government nursing homes in New York State. Twenty-one percent of these homes extended the quality and quantity of care beyond their estimated profit-maximizing level, suggesting the organizations acted as if they had quality and quantity objectives. The same share of each kind of organization departed from profit maximization. The remainder of the organizations acted like "for-profits-in-disguise," although Vitaliano concedes this could be due to insufficient revenues to do anything else rather than duplicity.

Other papers infer the organizational objectives from the structure of managerial compensation. For example, Ballou and Weisbrod (2003) examine bonuses paid to hospital executives in secular nonprofit, religious nonprofit, and government hospitals. Controlling for many factors, they find that government hospitals are least likely to pay a bonus based on the quality of care, secular nonprofits are somewhat more likely, and religious nonprofits are much more likely to do so. They conclude that secular and religious nonprofit hospitals have different objectives, but that government hospitals face different constraints so it is not clear whether the behavioral difference is due to objectives or constraints. Related studies (Roomkin and Weisbrod 1999; Erus and Weisbrod 2003) come to similar conclusions. Finally, Ehrenberg et al. (2000) look at contractual bonuses given to university presidents when they meet academic, research, and performance goals.

A variety of other approaches for uncovering objectives have been tried. Eldenburg et al. (2004) examine the factors that determine the composition of hospital boards, board

turnover, and CEO turnover to uncover objectives with respect to providing uncompensated care, generating excess revenues, and administrative costs. They find that for-profit, secular nonprofit, teaching nonprofit, religious nonprofit, government, and district hospitals place different weights on these three objectives. That said, uncovering the exact form of objectives from evidence of high turnover related to each factor is hard. Deneffe and Masson (2002) look at the response of hospital prices for private patients to changes in the Medicare, Medicaid, and charity caseload. They find that hospitals consider both profits and output as objectives. Lastly, Kapur and Weisbrod (2000) examine whether government and private nonprofit nursing homes and facilities for the mentally handicapped differ in their use of waiting lists and consumer satisfaction levels. Finding significant differences, they note that one explanation is that government pursues a supplier-of-last-resort objective function.

### IMPLICATIONS FOR PUBLIC POLICY

Economic theories about the role of the nonprofit sector have much to contribute to the public policy debates that swirl around the sector. Rather than treat any policy comprehensively (a task left to other chapters in this volume), I end this chapter by illustrating the ways in which economic theory can be used here.

How should we design nonprofit corporation statutes, and how should we enforce the nondistribution constraint? The theory of contract failure suggests that our decisions here affect the trustworthiness of the sector, and this is one of the first published policy applications (Hansmann 1981c). From the trust perspective, the value of the nonprofit label applies equally well to purely commercial ventures as to traditional charities. Nondistribution policy also affects the ease with which nonprofits can obtain capital, and so interacts with the plethora of other statutes that exempt nonprofits from corporate taxation and allow them to benefit from tax-exempt bonds. The manner in which nondistribution is applied to takeovers and conversions affects whether entrepreneurs will choose the nonprofit form in order to restrict changes in mission. Finally, the way in which the fair compensation constraint is applied to board and executive compensation is critical. If the constraint is set too high, it does nothing to stop distributions. If the constraint is too low, entrepreneurs who value collective goods will prefer to found for-profit firms (Bilodeau 2000).

How should we treat nonprofit monopolies? Combinations in restraint of trade are restricted by antitrust legislation, with few allowances made for differences among the sectors. This is perhaps unfortunate, as there are real differences in the roles played by the respective sectors and the manner in which they will use their monopoly profits (Steinberg 1993a; Eckel and Steinberg 1993). Regardless of sector, monopoly power results in underprovision of the monopolized good, a market failure that antitrust law seeks to address. For-profit monopolies distribute their profits to owners, but nonprofit monopolies may use those profits to fix a

different market failure (underprovision of collective goods or overexclusion from excludable public goods) or to supply redistributive financial aid. They may also waste those profits on private-benefit perks. Depending on the balance of these factors, nonprofit monopolies can be efficiency-enhancing or efficiency-diminishing, and ideally public policy should take account of this. Differences in the structure of control also change the motivation for merger and the suspicion with which authorities should view this activity. Patron-controlled nonprofits are more likely to be motivated by cost savings, whereas for-profits are more likely to be motivated by increased revenues resulting from underprovision.

When governments contract out for provision of social services, should they employ competitive bidding? Should they seek both nonprofit and for-profit bidders? Steinberg (1997) takes a first stab at these questions, noting that contract failure can make nonprofits a better option, but their productive inefficiency works in the opposite direction. Competitive bidding might help solve the latter problem, but reduces the ability of nonprofits to be trustworthy and still offer a winning bid, resulting perhaps in more for-profits-in-disguise.

Should we exempt nonprofit organizations from income, sales, value-added, and property taxes? Exemption gives nonprofits more resources, but this is hardly the most efficient way to deal with market failures. Direct subsidies for providing collective goods can be offered to organizations in both sectors (as in the U.S. tax credit for historic preservation), or government can provide the collective good itself. Nonprofit exemption from the corporate income tax rewards organizations in proportion to their capital stock, exemption from sales or value-added tax rewards them in proportion to commercial activities, and exemption from property taxes rewards them in proportion to the value of their property. None of these are particularly good at fostering the distinctive roles we have noted for the nonprofit sector.<sup>17</sup>

### ACKNOWLEDGMENTS

I am grateful to Wolfgang Bielefeld, Marc Bilodeau, Kirsten Grønby, Estelle James, Patrick Rooney, and Richard Turner for helpful comments and Barry Harvey for his help with the diagrams.

### NOTES

1. More precisely, an economy is a list of the quantities of each good affecting each person. An economy is efficient (syn. "Pareto-efficient") if it is feasible, and any alternative economy that at least one person ranks higher and no person ranks lower is not feasible. In general, there are an infinite number of efficient economies, differing in their distribution of income. Efficiency is socially desirable in a limited sense, necessary but not sufficient. Some specific inefficient economies might be socially preferred to other efficient economies. However, if we accept the mild judgment that one economy is better than another if everyone (weakly) prefers that economy, then there is always at least one efficient economy that would be preferred to any specific inefficient

one. Atkinson and Stiglitz (1980: chapter 11) provide a fine introduction to the subtleties of this topic.

2. If there is no lock-in effect, then contract failure does not arise. For-profit firms solve the problem by offering money-back guarantees.

3. There are other causes of market failure (such as monopoly) that are not discussed here because nonprofits do not seem to play a role in coping with them. Ben-Ner and Gui (2003) provide an alternative but more comprehensive discussion of the demand for nonprofit organizations in situations of market failure.

4. To use Lindahl prices, the government must know each person's preferences. However, consumers would have an incentive to misrepresent their preferences to secure a lower price.

5. If, as suggested below, the median preference voter is decisive and the warm-glow motivation (Vesterlund, this volume) is inoperative, then governmental provision is efficient if and only if the distribution of most-preferred provision levels is symmetric about the mean.

6. This result, known as the "median voter theorem," was proved by Bowen (1943) for direct democracy and by Hotelling (1929) and later Downs (1957) for representative democracy. For the former, the proof involves showing that a referendum proposing the median preference voter's most-preferred level of provision will defeat any alternative by a majority of voters; for the latter, that vote-maximizing politicians would select as their platform the median voter's ideal. Both proofs make restrictive assumptions, but the median voter theorem is a useful starting point for analysis.

7. Sometimes, tax laws are so poorly enforced that payment of taxes is almost purely voluntary. Conversely, sometimes social pressures to make donations are so strong that payment is almost coerced. Nonetheless it is often reasonable to view the nonprofit alternative as less coercive than government.

8. Contract failure is less important for certain types of volunteering than for donations of money (e.g., Steinberg 1990a). Volunteers observe and help decide how their labor is used, allowing them to obtain the incremental output they want. Thus, parents donate their time to for-profit day-care centers and adult children donate their time to for-profit hospitals and nursing homes to provide recreational, educational, and social services that would otherwise be lacking.

9. Provision of perks is not, by itself, inefficient (Schlesinger 1985). Nonprofit managers are efficiently producing a mixture of mission-related outputs and perks; they have no desire to produce perks inefficiently. The problem is really an allocative inefficiency, where (a) managerial compensation takes the form of a mixture of money and perquisites that, owing to the nondistribution constraint, is higher cost than the optimal mixture and (b) when nonprofits compete with for-profits but enjoy otherwise lower costs because of tax subsidies, these subsidies can be applied to perk production. The higher social cost of nonprofit production is not reflected in prices due to the subsidy, and so the market produces a mixture that includes relatively too much high-social-cost nonprofit output and too little low-social-cost for-profit output. Nonetheless, because these inefficiencies show up as higher cost,

the literature has errantly labeled this a form of productive inefficiency, a tradition we repeat here for consistency with the literature.

10. The solution to this problem lies in adding more to the three-failures theory. As a result, some received conclusions will change, but the three failures will remain a part of the story. Whether to call the result an enhanced three-failures theory or another name is entirely a matter of taste and semantics.

11. In private conversation, Wolfgang Bielefeld suggested a further broadening of the modeling agenda, creating a "blended" theory that would embed economic theories within a world shaped by political, sociologic, historical, and cultural factors. Thus, both the preferences of entrepreneurs and the factors affecting their decisions would be determined by the social structure, networks, history, and the like. In turn, the behaviors of nonprofit organizations at any point in time helps to determine the future evolution of networks, government regulations, and even cultural norms. This approach would integrate the distributional, preference-shaping, affiliative, and expressive roles of nonprofit organizations into the analysis. This approach would also help us understand the evolution of the roles played by the various sectors, and so seems well worthy of further development.

12. More precisely, Ben-Ner and Van Hoomissen argue that a self-provision coalition is formed if the expected flow of net benefits exceeds the next best alternative. This coalition can take the form of founding a nonprofit or a consumer cooperative, with additional factors skipped here governing that choice.

13. James (1982, 1986, 1989) observes that most secular nonprofits are founded out of such groups, especially religious congregations, but explains this as an attempt to gain converts.

14. More precisely, they show that the decision to incorporate as a nonprofit solves a moral hazard problem between the entrepreneur and other donors by acting as a commitment device in a three-stage game of perfect information. Donations still suffer from the free-rider problem, but they do not also suffer from this problem of moral hazard.

15. Tuckman and Chang (this volume) analyze selected mission statements to see what they reveal about nonprofit commercial activities. This is a start, but a more systematic and broad-ranging effort would be interesting.

16. See Weisbrod and Dominguez (1986), Posnett and Sandler (1989), Khanna, Posnett, and Sandler (1995), Okten and Weisbrod (2000), Khanna and Sandler (2000), Hewitt and Brown (2000), and Tinkelman 2004. These studies challenge Steinberg's (1986) specific findings, but do not agree on which charitable industries act like budget or service maximizers.

17. In the interests of space, I am leaving out many more economic arguments on both sides of the tax exemption question, as well as the noneconomic arguments. Personally, I am more supportive of exemption than my brief summary here might suggest. See Simon, Dale, and Chisolm (this volume), Steinberg (1991), and Steinberg and Bilodeau (1999) for a more balanced and comprehensive perspective.

## REFERENCES

- Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York: McGraw Hill.
- Ballou, Jeffrey, and Burton A. Weisbrod. 2003. "Managerial Rewards and the Behavior of For-Profit, Governmental, and Nonprofit Organizations: Evidence from the Hospital Industry." *Journal of Public Economics* 87:1895-1920.
- Ben-Ner, Avner. 1986. "Nonprofit Organizations: Why Do They Exist in Market Economies?" Pages 94-113 in *The Economics of Nonprofit Institutions*, edited by Susan Rose-Ackerman. New York: Oxford University Press.
- . 1987. "Birth, Death, and Bureaucratization in Nonprofit Organizations: An Economic Analysis." Mimeo, University of Minnesota.
- Ben-Ner, Avner, and Benedetto Gui. 2003. "The Theory of Nonprofit Organizations Revisited." Pages 3-26 in *The Study of Nonprofit Enterprise: Theories and Approaches*, edited by Helmut Anheier and Avner Ben-Ner. New York: Kluwer Academic/Plenum Publishers.
- Ben-Ner, Avner, and Derek Jones. 1995. "Employee Participation, Ownership, and Productivity: A Theoretical Framework." *Industrial Relations* 34:532-554.
- Ben-Ner, Avner, and Theresa Van Hoomissen. 1991. "Nonprofit Organizations in the Mixed Economy: A Demand and Supply Analysis." *Annals of Public and Cooperative Economics* 62:519-550.
- Bilodeau, Marc. 2000. "Profitable Nonprofit Firms." Department of Economics Working Paper, IUPUI, Indianapolis, Ind.
- Bilodeau, Marc, and Al Slivinski. 1996. "Volunteering Nonprofit Entrepreneurial Services." *Journal of Economic Behavior and Organization* 31:117-127.
- . 1997. "Rival Charities." *Journal of Public Economics* 66:449-467.
- . 1998. "Rational Nonprofit Entrepreneurship." *Journal of Economics and Management Strategy* 7:551-571.
- Bilodeau, Marc, and Richard Steinberg (1997). "Ransom of the Opera." Department of Economics Working Paper, IUPUI, Indianapolis, Ind.
- . Forthcoming. "Donative Nonprofit Organizations." In *Handbook on the Economics of Giving, Reciprocity, and Altruism*, edited by S.-C. Kolm and J. Mercier Ythier. New York: Elsevier/North Holland.
- Bowen, Howard R. 1943. "The Interpretation of Voting in the Allocation of Economic Resources." *Quarterly Journal of Economics* 58:27-48.
- Cain, Louis, and Dennis Meritt, Jr. 1998. "Zoos and Aquariums." Pages 217-232 in *To Profit or Not to Profit: The Commercial Transformation of the Nonprofit Sector*, edited by Burton A. Weisbrod. Cambridge, U.K.: Cambridge University Press.
- Chang, Cyril F., and Howard P. Tuckman. 1996. "The Goods Produced by Nonprofit Organizations." *Public Finance Quarterly* 24:25-43.
- Clotfelter, Charles T., ed. 1992. *Who Benefits from the Nonprofit Sector?* Chicago: University of Chicago Press.
- DeHoog, Ruth. 1984. *Contracting Out for Human Services: Economic, Political and Organization Perspectives*. Albany, N.Y.: State University of New York Press.
- Deneffe, Daniel, and Robert T. Masson. 2002. "What Do Not-For-Profit Hospitals Maximize?" *International Journal of Industrial Organization* 20:461-492.
- Douglas, James. 1987. "Political Theories of Nonprofit Organizations." Pages 43-54 in *The Nonprofit Sector: A Research Handbook*, edited by Walter W. Powell. New Haven, Conn.: Yale University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Eckel, Catherine C., and Richard Steinberg. 1993. "Competition, Performance, and Public Policy Towards Nonprofits." Pages 57-81 in *Nonprofit Organizations in a Market Economy*, edited by David Hammack and Dennis R. Young. San Francisco: Jossey-Bass.
- Ehrenberg, Ronald, John Cheslock, and Julia Epifantseva. 2000. "Paying Our Presidents: What Do Trustees Value?" Cambridge, Mass.: NBER Working Paper 7886.
- Eldenburg, Leslie, Benjamin E. Hermalin, Michael S. Weisbach, and Marta Wosinska. 2004. "Governance, Performance Objectives, and Organizational Form: Evidence from Hospitals." *Journal of Corporate Finance* 10:527-548.
- Erus, Burcay, and Burton A. Weisbrod. 2003. "Inferring Nonprofit Organization Objective Function from Compensation Structure: The 'Mixed' Hospital Industry." Pages 117-142 in *The Governance of Not-For-Profit Firms*, edited by Edward Glaeser. Chicago: University of Chicago Press.
- Feigenbaum, Susan. 1980. "The Case of Income Redistribution: A Theory of Government and Private Provision of Collective Goods." *Public Finance Quarterly* 8:3-22.
- . 1987. "Competition and Performance in the Nonprofit Sector: The Case of US Medical Research Charities." *Journal of Industrial Economics* 35:241-253.
- Feldstein, Martin. 1971. "Hospital Price Inflation: A Study of Nonprofit Price Dynamics." *American Economic Review* 61:853-72.
- Ferris, James, and Elizabeth Graddy. 1991. "Production Costs, Transaction Costs, and Local Government Contractor Choice." *Economic Inquiry* 29:541-554.
- Frech, H. E., III. 1976. "The Property Rights Theory of the Firm: Empirical Results from a Natural Experiment." *Journal of Political Economy* 84:143-152.
- Gassler, Robert Scott. 1989. "The Economics of the Nonprofit Motive: A Suggested Formulation of Objectives and Constraints for Firms and Nonprofit Enterprises." Vesalius College Working Paper. Vrije Universiteit Brussel, Brussels, Belgium.
- Glaeser, Edward L., and Andrei Schleifer (2001). "Not-For-Profit Entrepreneurs." *Journal of Public Economics* 81:99-116.
- Hansmann, Henry. 1980. "The Role of Nonprofit Enterprise." *Yale Law Journal* 89:835-901.
- . 1981a. "Nonprofit Enterprise in the Performing Arts." *Bell Journal of Economics* 12:341-361.
- . 1981b. "The Rationale for Exempting Nonprofit Corporations from the Corporate Income Tax." *Yale Law Journal* 91:45-100.
- . 1981c. "Reforming Nonprofit Corporation Law." *University of Pennsylvania Law Review* 129:497-623.
- . 1987a. "Economic Theories of Nonprofit Organization." Pages 27-42 in *The Nonprofit Sector: A Research Handbook*, edited by Walter W. Powell. New Haven, Conn.: Yale University Press.
- . 1987b. "The Effect of Tax Exemption and Other Factors on the Market Share of Nonprofit Versus For-Profit Firms." *National Tax Journal* 40:71-82.
- . 1996. "The Changing Roles of Public, Private, and Nonprofit Enterprise in Education, Health Care, and Other Human Services." Pages 245-275 in *Individual and Social Responsibility: Child Care, Education, Medical Care, and Long-Term Care in America*, edited by Victor Fuchs. Chicago: University of Chicago Press.
- . 2003. "The Role of Trust in Nonprofit Enterprise." Pages 115-124 in *The Study of the Nonprofit Enterprise: Theories and Approaches*, edited by Helmut Anheier and Avner Ben-Ner. New York: Kluwer/Plenum.
- Hansmann, Henry, Daniel Kessler, and Mark McClellan. 2003. "Ownership Form and Trapped Capital in the Hospital Industry." Pages 45-70 in *The Governance of Not-for-Profit Organizations*, edited by Edward L. Glaeser. Chicago: University of Chicago Press.
- Hewitt, Julie A., and Daniel K. Brown. 2000. "Agency Costs in Environmental Not-For-Profits." *Public Choice* 103:163-183.
- Hirth, Richard A. 1999. "Consumer Information and Competition Between Nonprofit and For-Profit Nursing Homes." *Journal of Health Economics* 18:219-240.
- Holtmann, Alphonse G. 1983. "A Theory of Non-Profit Firms." *Economica* 50:439-449.
- Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal* 39:41-57.
- James, Estelle. 1982. "The Nonprofit Sector in International Perspective: The Case of Sri Lanka." *Journal of Comparative Economics* 6:99-129.
- . 1983. "How Nonprofits Grow: A Model." *Journal of Policy Analysis and Management* 2:350-365.
- . 1986. "The Nonprofit Sector in Comparative Perspective." Pages 397-415 in *The Nonprofit Sector: A Research Handbook*, edited by Walter W. Powell. New Haven, Conn.: Yale University Press.
- . 1989. "Introduction." Pages 3-30 in *The Nonprofit Sector in International Perspective: Studies in Comparative Culture and Policy*, edited by Estelle James. New York: Oxford University Press.
- . 1993. "Why Do Different Countries Choose a Different Public-Private Mix of Educational Services?" *Journal of Human Resources* 28:571-592.
- . 1998. "Commercialism Among Nonprofits: Objectives, Opportunities, and Constraints." Pages 271-286 in *To Profit or Not to Profit: The Commercial Transformation of the Nonprofit Sector*, edited by Burton A. Weisbrod. Cambridge, U.K.: Cambridge University Press.
- James, Estelle, and Susan Rose-Ackerman. 1986. *The Nonprofit Enterprise in Market Economies*, a monograph in the series Fundamentals of Pure and Applied Economics and Encyclopedia of Economics, edited by J. Lesourne and H. Sonnenschein. London: Harwood Academic Publishers.
- Kapur, Kanika, and Burton A. Weisbrod. 2000. "The Roles of Government and Nonprofit Suppliers in Mixed Industries." *Public Finance Review* 28:275-308.
- Khanna, Jyoti, John Posnett, and Todd Sandler. 1995. "Charity Donations in the U.K.: New Evidence Based on Panel Data." *Journal of Public Economics* 56:257-272.
- Khanna, Jyoti, and Todd Sandler. 2000. "Partners in Giving: The Crowding-In Effects of UK Government Grants." *European Economic Review* 44:1543-1556.
- Lee, M. 1971. "A Conspicuous Consumption Theory of Hospital Behavior." *Southern Economics Journal* 38:48-58.
- Lowry, Robert C. 1997. "The Private Production of Public Goods: Organizational Maintenance, Managers' Objectives, and Collective Goals." *American Political Science Review* 91:308-323.
- Malani, Anup, Tomas Philipson, and Guy David. 2003. "Theories of Firm Behavior in the Non-Profit Sector: A Synthesis and Empirical Evaluation." In *The Governance of Not-for-Profit Organizations*, edited by Edward L. Glaeser. Chicago: University of Chicago Press.
- Mason, David E. 1996. *Leading and Managing the Expressive Dimension: Harnessing the Hidden Power Source of the Nonprofit Sector*. San Francisco: Jossey-Bass.
- Mauser, Elizabeth. 1993. "Is Organizational Form Important to Consumers and Managers?: An Application to the Day-Care Industry." Unpublished Ph.D. dissertation in Economics. Madison, Wisc.: University of Wisconsin.
- Nelson, Richard, and Michael Krashinsky. 1973. "Two Major Issues of Public Policy: Public Policy and the Organization of Supply." Pages 47-69 in *Public Subsidy for Day Care of Young Children*, edited by Richard Nelson and Dennis Young. Lexington, Mass.: D.C. Heath.
- Newhouse, Joseph. 1970. "Toward a Theory of Non-Profit Institutions: An Economic Model of a Hospital." *American Economic Review* 60:64-74.
- Niskanen, William A., Jr. 1971. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.
- Okten, Cagla, and Burton Weisbrod. 2000. "Determinants of Donations in Private Nonprofit Markets." *Journal of Public Economics* 75:255-272.
- Ortmann, Andreas, and Mark Schlesinger. 2003. "Trust, Repute, and the Role of Nonprofit Enterprise." Pages 77-114 in *The Study of the Nonprofit Enterprise: Theories and Approaches*, edited by Helmut Anheier and Avner Ben-Ner. New York: Kluwer/Plenum.