Some Implications of the Agency-Cost Theory of the Nonprofit Firm

Benjamin Leff
Washington College of Law, bleff@wcl.american.edu

Follow this and additional works at: https://digitalcommons.wcl.american.edu/facsch_bk_contributions

Part of the Business Organizations Law Commons

Recommended Citation
https://digitalcommons.wcl.american.edu/facsch_bk_contributions/163

This Book Chapter is brought to you for free and open access by the Scholarship & Research at Digital Commons @ American University Washington College of Law. It has been accepted for inclusion in Contributions to Books by an authorized administrator of Digital Commons @ American University Washington College of Law. For more information, please contact kclay@wcl.american.edu.
The Cambridge Handbook of Social Enterprise Law

Edited by

BENJAMIN MEANS
University of South Carolina School of Law

JOSEPH W. YOCKEY
University of Iowa College of Law
Some Implications of the Agency-Cost Theory of the Nonprofit Firm

Benjamin Moses Leff

I INTRODUCTION

Perhaps the most canonical work in the study of nonprofit organizations in the United States is a series of articles written in the early 1980s by Henry Hansmann. Hansmann argued that nonprofits arise as a response to a problem in contracting between firms and what he called patrons. When patrons want goods or services whose quality is difficult to assess, they face a potential “contract failure.” One possible solution to this contract failure is to impose a limitation on the profits that can go into the pockets of the providers of goods (the so-called nondistribution constraint), under the belief that, once the incentive to increase personal profits is removed, managers of the firm will generally use excess value to provide high-quality or more goods.

According to Hansmann, the nondistribution constraint is at the heart of what differentiates a nonprofit from other firms. The nonprofit form arises in situations in which it is easier or less expensive to monitor and enforce the nondistribution constraint than to monitor or enforce the quality of goods provided. The costs incurred in monitoring and enforcing the commitments of the firm are sometimes called “agency costs,” and so this theory of the purpose of the nondistribution constraint has been called the “agency-cost theory” of the nonprofit form. The agency-cost theory forms the basis of a significant branch of the scholarly analysis of the legal regulation of nonprofits.

Hansmann differentiated between so-called commercial nonprofits (which receive the majority of their revenue from fees from their consumers) and donative nonprofits (which receive the majority of their revenue from donors who are not consumers). Donative nonprofits tend to supply what might be called “social goods,” goods that provide a benefit to society at large or to a charitable class. Another way of thinking about donative nonprofits is that they supply a service to people who want to act altruistically, or even that they are in the business of providing a good called “altruism.” Hansmann argued that contract failure in donative nonprofits almost always results in the choice of the nonprofit form. But, of course, for-profit firms commonly and increasingly “sell altruism” to their customers whenever they


2 See M. Todd Henderson & Anup Malani, Corporate Philanthropy and the Market for Altruism, 109 COLUM. L. REV. 571, 585 (2009) (using the term “altruism” to signify an inextricable mix of “pure altruism,” which is the utility derived from helping others, and “warm glow,” which is the utility derived from feeling like or being perceived as helping others).
benefit persons unrelated to their patrons in order to satisfy their patrons' sense of justice or mercy. This is sometimes called "corporate social responsibility," and it is often recognized as good business practice, since consumers seem to have a taste for socially responsible goods, perhaps at least partially because of altruism. But it raises a question: how do "regular" for-profit businesses overcome contract failure to provide altruism to their customers?

Social enterprises are essentially firms that take this commitment to the general good more seriously than "regular" businesses. To persuade stakeholders that they are different, social enterprises need to make binding commitments that they will pursue social goals. If they want to be for-profit social enterprises, then they need to make such commitments without reliance on the nondistribution constraint.3

An application of the agency-cost theory to the social enterprise space yields some potentially useful observations. First, it is impossible to account for agency costs without understanding who is the principal in an agent-principal relationship. Literature on social enterprises often emphasizes the need for a firm to make credible commitments about social benefits to diverse stakeholders. But the agency-cost story plays out very differently if the relevant principal is an investor who wants her investment to serve a double or triple bottom line, or a customer who purchases goods relying on an enterprise's commitment to the social good (to take just two examples). The mechanisms that enable firms to make credible commitments to customers are often not the same as the mechanisms that enable firms to make credible commitments to shareholders.

Second, whatever stakeholder you are, it is easier to quantify, monitor, and enforce a social goal if that social goal is a "negative" restriction than if it is a "positive" benefit. For example, if firms want to provide tuna without killing dolphins, and what they need to do to accomplish that goal is use specific dolphin-safe nets, then the cost involved in quantifying, monitoring, and enforcing their commitment to use dolphin-safe nets is relatively low. On the other hand, if a firm's purpose is to increase the overall health and sustainability of the ocean's ecosystem, the process of quantifying, monitoring, and enforcing commitment to that goal is relatively difficult. Agency-cost theory tells us that the nonprofit form may be the most efficient means for a firm to deliver social goods when quantifying, monitoring, and enforcing progress toward a social goal is difficult or expensive. It is predictable that for-profit social enterprises will have the hardest time when trying to make credible commitments to multiple different stakeholders simultaneously, and when they are trying to commit to something other than simple negative restrictions.

This chapter proceeds in three parts. First, it discusses the theoretical framework on which it rests, introducing the reader to Hansmann's theory of contract failure and focusing on the elements of that theory that are most applicable to the social enterprise context. Second, it discusses some implications of the observation that mechanisms to quantify, monitor, and enforce a credible commitment to the social good are differentially successful depending on which stakeholder an enterprise seeks to make such commitments to. Third, it distinguishes between negative and positive commitments and uses the agency-cost theory to describe why that distinction is so important.

3 As other chapters in this volume make clear, many social enterprises are nonprofits; many social enterprise innovations involve some form of "asset lock" or nondistribution constraint or partial nondistribution constraint; and many advocates of social enterprises think the nondistribution constraint (or a modified version of it) is essential to the success of social enterprises. See, e.g., Muhammed Yunus, Building Social Businesses: The New Kind of Capitalism That Serves Humanity's Most Pressing Needs (2000) (arguing that social businesses must not have investors from outside the community served). But it is the branch of the social enterprise movement that sees the future in for-profit enterprises, including benefit corporations, that is the subject of this chapter.
II THE THEORY OF CONTRACT FAILURE

Hansmann argued that the nonprofit form exists to solve a problem of "contract failure." Contract failure arises when one or more of the preconditions for an efficiently functioning market do not exist. Hansmann argued that the nonprofit form is sometimes selected when a firm is trying to provide goods or services to its "patrons," but those patrons are unable to ascertain information that is essential for them to do business with the firm. Hansmann further argued that the "most important of these conditions is that consumers can, without undue cost or effort" (1) compare the prices of various providers, (2) make an agreement with the firm about the goods or services provided and the price, and (3) ensure that the firm complied with the agreement. These three conditions impact the ability of a patron to effectively measure, monitor, and enforce the quality of the goods or services provided by the firm.

When these conditions do not exist — that is, when measuring, monitoring, and enforcing the quality of social goods can only be accomplished with prohibitively high costs or effort — then contract failure is likely to arise. Because these conditions have generally been described as "transaction costs" or "agency costs," Hansmann's contract-failure theory has often been described as an "agency-cost theory" of the nonprofit firm. When agency costs are prohibitively high, patrons need some mechanism to prevent the firm from providing low-quality services, or in the extreme case, not bothering to provide any services at all in exchange for payment from the patrons.

Hansmann uses the term "nondistribution constraint" to describe the solution to contract failure that nonprofits provide. It is the set of rules that prevent nonprofits from distributing their profits. But the nondistribution constraint does not actually provide such a mechanism directly. All it does is remove the profit motive, which, in the absence of effective monitoring, is the strongest incentive for managers of the firm to "skimp on the services it promises, or even to neglect to perform them entirely." By removing the ability of nonprofit managers to divert the revenue of the firm into their pockets in the form of profits, the nondistribution constraint facilitates at least some trust. But contracting with a nonprofit firm comes with a cost, because it is widely accepted that the profit motive is generally effective in incentivizing firm managers to provide services in the most cost-effective manner.

4 Hansmann, Role, supra note 1, at 841.
5 See id. at 845.
6 For a more detailed discussion of the propriety of the terms "transaction costs" and "agency costs," see Benjamin Moses Leff, The Case Against For-Profit Charity, 42 SETON HALL L. REV. 819, 833-36 (2012). One seminal article on agency costs defines an agency relationship as "a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision-making authority to the agent." Michael C. Jensen & William H. Meckling, Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, 3 J. FIN. ECON. 305, 308 (1976). If an agency relationship requires only that some decision-making be delegated, then presumably the purchase of altruism generally arises in situations that meet the definition, in which case "agency costs" is an adequate term for the costs of measuring, monitoring, and enforcing the provision of altruism. Although one could imagine a situation in which the terms under which one purchased altruism were so completely specified that no agency relationship existed. In that case, one would more accurately call the costs incurred "transaction costs."
7 Hansmann, Role, supra note 1, at 847.
8 See, e.g., Michael Krashinsky, Transaction Costs and a Theory of the Nonprofit Organization, in THE ECONOMICS OF NONPROFIT INSTITUTIONS: STUDIES IN STRUCTURE AND POLICY 114, 116 (Susan Rose-Ackerman ed., 1986) ("Unfortunately, abandoning the profit motive is also costly. The same drive for profits that induces entrepreneurs to cut corners on quality also induces them to cut costs, adopt new technology, and respond rapidly to changes in demand."); see also Burton A. Weisbrod, The Nonprofit Economy 43 (1968) (describing the nondistribution constraint, which is "a wedge between performance and reward" as "akin to tying one arm of a fighter behind his back").
is distinctly a “second-best” solution to the problem of contract failure. So Hansmann’s theory can be described as follows: “the nonprofit form will be chosen whenever the cost of monitoring and enforcing a specific level of product quality exceeds the gains that are expected to accrue from providing the management of the [firm] with strong incentives to implement cost-saving efficiencies.”

Hansmann distinguished between “donative” and “commercial” nonprofits, but this distinction potentially masks a conceptually more important one. Hansmann defines donative nonprofits as those that receive the majority of their revenue from donation, while commercial nonprofits receive the majority of their revenue from selling goods or services. But it is also the case that donative nonprofits are more likely to provide services to third parties unrelated to their donors, and commercial nonprofits are more likely to provide services for the consumption of their customers. Hansmann’s example of a donative nonprofit that provides services to unrelated third parties is CARE, an organization that collects donations in the United States and uses them to reduce poverty in Africa. Organizations like CARE provide services to beneficiaries who do not pay for them, but they also can be understood to provide something to their donors, who do pay for the services provided to others. In order to analyze the market dynamics of such organizations, some think it useful to conceptualize these organizations as providing some good to their donors, and this good is generally called “altruism” or “warm glow.”

Looked at in this way, an organization like CARE is selling altruism to its donors, in the same way that an organization like Ford is selling cars to its customers, except that altruism differs from cars in a variety of ways that are relevant to our analysis.

First, altruism is different from “normal” goods because the “quality” of altruism often depends on the benefits provided to people who are far from, and not in contact with, the purchaser of the altruism. That makes it the quintessential case for Hansmann of a good whose quality is costly for a patron to assess. Second, altruism is arguably a “public good.” Public goods are those goods that are “nonrivalrous” and “nonexcludible.” A nonrivalrous good is one where the provision of it to some stakeholders does not diminish its availability for others. An alternative description of the theory is this one: “the nonprofit form will be chosen whenever it is impossible or prohibitively costly for the Patron to distinguish profits derived from efficiently providing a high-quality product and profits derived from shirking on quality.”

Hansmann presented the donative and commercial classification as “ideal types rather than mutually exclusive and exhaustive categories.” Hansmann, Role, supra note 1, at 841.

Public goods are those goods that are “nonrivalrous” and “nonexcludible.” A nonrivalrous good is one where the provision of it to some stakeholders does not diminish its availability for others. An alternative description of the theory is this one: “the nonprofit form will be chosen whenever it is impossible or prohibitively costly for the Patron to distinguish profits derived from efficiently providing a high-quality product and profits derived from shirking on quality.”

Hansmann distinguished between “donative” and “commercial” nonprofits, but this distinction potentially masks a conceptually more important one. Hansmann defines donative nonprofits as those that receive the majority of their revenue from donation, while commercial nonprofits receive the majority of their revenue from selling goods or services. But it is also the case that donative nonprofits are more likely to provide services to third parties unrelated to their donors, and commercial nonprofits are more likely to provide services for the consumption of their customers. Hansmann’s example of a donative nonprofit that provides services to unrelated third parties is CARE, an organization that collects donations in the United States and uses them to reduce poverty in Africa. Organizations like CARE provide services to beneficiaries who do not pay for them, but they also can be understood to provide something to their donors, who do pay for the services provided to others. In order to analyze the market dynamics of such organizations, some think it useful to conceptualize these organizations as providing some good to their donors, and this good is generally called “altruism” or “warm glow.” Looked at in this way, an organization like CARE is selling altruism to its donors, in the same way that an organization like Ford is selling cars to its customers, except that altruism differs from cars in a variety of ways that are relevant to our analysis.

First, altruism is different from “normal” goods because the “quality” of altruism often depends on the benefits provided to people who are far from, and not in contact with, the purchaser of the altruism. That makes it the quintessential case for Hansmann of a good whose quality is costly for a patron to assess. Second, altruism is arguably a “public good.” Public goods are those goods that are “nonrivalrous” and “nonexcludible.” A nonrivalrous good is one where the provision of it to some stakeholders does not diminish its availability for others. An alternative description of the theory is this one: “the nonprofit form will be chosen whenever it is impossible or prohibitively costly for the Patron to distinguish profits derived from efficiently providing a high-quality product and profits derived from shirking on quality.”

Hansmann presented the donative and commercial classification as “ideal types rather than mutually exclusive and exhaustive categories.” Hansmann, Role, supra note 1, at 841.

See Hansmann, Role, supra note 1, at 846-48.

See Henderson & Malani, supra note 2, at 838 (using the term “altruism” to signify an inextricable mix of “pure altruism,” which is the utility derived from helping others, and “warm glow,” which is the utility derived from feeling like or being perceived as helping others), citing James Andreoni, Giving with Impure Altruism: Applications to Charities and Ricardian Equivalence, 97 Pol. Econ. 1447, 1448-49, 1455 (1989) (distinguishing between “altruism,” which is “demand for more of the public good,” and “warm glow,” which is the receipt of status, acclaim, or good feeling that arises from knowing that one has “done their bit”). The difference between altruism and warm glow has important implications beyond the scope of this chapter. The fact that warm glow can be produced by the false belief that a benefit has been provided, while altruism can only be produced when a benefit is actually provided, has very important implications for the amount of investment that a donor needs to make in agency costs, which is at the heart of the question of whether the nonprofit form is the most efficient means of delivering altruism. See infra note 21 (discussing the ways in which purchasers of “warm glow” may be relatively uncommitted to ensuring themselves that their purchases do any actual good).

This distance from beneficiary and patron/donor continues after the good has been provided. So, absent some potentially costly monitoring or enforcement mechanism, the patron may never learn if the beneficiary received any goods or services.

I say “arguably” because “warm glow” includes the feeling that one caused the benefit, and that may not be “nonrivalrous” or “nonexcludible.” Similarly, to the degree that warm glow also includes the utility derived from being perceived as acting altruistically, it is distinctly rivalrous and excludible.
Some Implications of the Agency-Cost Theory

other stakeholders. If my contribution to CARE saves the life of a poor child, I have received some benefit in the form of altruism. But once I have made the world slightly better by saving the child's life, I have equally benefited all other persons who are altruistic because, to the degree to which they desire this particular improvement to the world, all of them have simultaneously had their desires satisfied. The fact that a public good is "nonexcludible" means that it is impossible (or prohibitively expensive) to exclude "free-riders" from enjoying the good once it has been supplied. Because of the nonexcludable nature of altruism, patrons who pay for it subsidize anyone who wants the world to benefit but doesn't want to pay for it. These thrifty altruists are benefited by the provision of food for poor children when others pay for it, and that is the free-rider problem associated with altruism as a public good.

Finally, there is a third important characteristic of altruism that is an effect of its non-excludibility: the difficulty in measuring the marginal benefit of any payment for it. Because altruism is (largely) nonexcludable, a rational donor would want some assurance that her donation is being used to increase the quality or quantity of the charitable goods being supplied. If CARE is already saving the lives of 10,000 poor children using donations from other donors, my donation is wasted from an altruistic point of view unless I can be assured that it will increase the number of lives saved above 10,000. I gain no altruism unless my contribution benefits someone, and there is no benefit from my contribution unless the marginal amount of benefit provided by the firm is increased. The nondistribution constraint is some protection against this type of misuse of additional funds, since it purports to remove the ability of charity managers to divert additional funds into their own pockets, thus removing the strongest incentive not to use additional funds for additional benefit.

While Hansmann thought of donative nonprofits as the dominant provider of altruism (though he doesn't use that term), it is very common for other firms to provide altruism to their patrons, at least if we have a broad enough conception of what altruism is. M. Todd Henderson and Anup Malani have argued that regular business corporations provide their stakeholders with altruism when they benefit unrelated third parties more than is required by law. Corporate social responsibility, enhanced environmental or labor practices, and fair-trade or other global supply-chain-monitoring mechanisms all provide a benefit primarily to someone other than the firms' patrons. Under this broad definition of altruism, all of these
practices constitute the provision of altruism to such patrons. In that sense, commercial nonprofits and commercial for-profits are suppliers of altruism, even though they are primarily supported by fees paid by their customers.

I previously defined social enterprises as “firms that take [their] commitment to the general good more seriously than ‘regular’ businesses.” But another way of thinking of a social enterprise is that it is a business that is operated to supply a significant amount of altruism to some stakeholders along with its supply of a good or service. Just as regular for-profit businesses can provide altruism to a variety of stakeholders – including investors, labor, management, customers, neighbors, supply-chain participants, and others – social enterprises can likewise provide altruism to this broad variety of stakeholders. Each stakeholder can benefit from the receipt of some altruism when the firm increases the well-being of some unrelated others. The cost of this altruism can similarly be paid for by a variety of stakeholders. Investors can receive lower returns on their investments of capital than they would in a firm that did not also supply them with altruism; labor or management might be willing to work for lower wages than they would elsewhere; customers may pay more for goods or services that come with some altruism on the side; etc. In a social enterprise, customers or investors “buy altruism” by subsidizing the wages of workers, for example, so they might be assured that workers are receiving a living wage. On the other hand, no altruism is produced if stakeholders subsidize other stakeholders whom they do not intend to subsidize. So if customers pay a higher price for goods with the intention of subsidizing workers’ wages, but instead subsidize investors’ return, then altruism is not produced and social value diminishes.

If it is common for regular businesses, not just social enterprises, to provide altruism to their stakeholders, then they must somehow solve the agency-cost problems associated with such provision. After all, the same free-rider problems associated with the sale of altruism exist here as in other contexts, but there is no nondistribution constraint to ameliorate the effect. For-profit businesses use a variety of mechanisms to make credible commitments to their stakeholders that they are actually behaving in the socially beneficial ways that they claim. For example, for-profit firms sometimes simply make claims about their socially beneficial practices, relying on their individual brand to persuade consumers. At other times, they may rely on a legal regime that defines a socially beneficial practice and at least nominally enforces compliance with the regime once a firm has claimed that it is conforming to the practice, as when a firm claims that a good it sells is “organic.” When branding and legal regimes are insufficient, a firm may seek “certification” from a third party who attests to the firm’s compliance with a socially beneficial practice.

9 Not all definitions of social enterprises require that the firm differentiate itself from “regular” businesses. For example, some definitions require only that the firm seeks to “do well” (financially) while doing “good” (socially). Joseph W. Yockey, Does Social Enterprise Law Matter?, 66 ALA. L. REV. 767, 769 (2014). Other definitions require that the primary goal of a social enterprise is social or charitable, with financial goals being only secondary. For the purposes of this chapter, I prefer the definitions supplied in the text.
20 See Henderson & Malani, supra note 2, at 574, 589.
21 Stakeholders may not really care all that much if the incremental cost of their share of the altruism produced is not salient to them. For example, customers may be happy to pay more for Starbucks coffee for a wide variety of reasons (it has an ungodly amount of sugar in it, it’s on every corner, it is engaged in a war on Christmas) such that the fact that they purport to sell “fair-trade” coffee is an insignificant additional benefit. For many consumers, the question of whether “fair trade” actually benefits the farmers is only of tangential benefit, even if it makes the coffee seem a little better to them. In this case, it is unnecessary for Starbucks to make credible commitments to its consumers about the nature or degree of social benefit that buying “fair-trade” coffee produces.
Some Implications of the Agency-Cost Theory

Social enterprises may use the same commitment devices as regular businesses, but they may also seek commitment devices that distinguish them from regular businesses to make stronger claims of social benefit to their stakeholders. Many of the chapters in this volume explore the special commitment devices designed to enable social enterprises to make their distinguishing promises, a prime example being distinctive business forms like the cooperative, the British Community Interest Company, the American (Public) Benefit Corporation, and others. It is well beyond the scope of this chapter to explore the variety of mechanisms that firms, whether social enterprises or regular businesses, use to make credible commitments to stakeholders about the production of altruism. But the “agency-cost” theory of the nonprofit firm provides a theoretical basis from which to make a few basic observations about the relative usefulness of various social enterprise business forms.

III FIRST OBSERVATION: IDENTIFICATION OF THE "PRINCIPAL" MATTERS

In the nonprofit context, it is often said that an agency-cost analysis is complicated by the fact that the “principal” in the agency-cost analysis is not any identifiable stakeholder, but instead the charitable mission of the organization. Since a charitable mission cannot guard against shirking by an agent, there is no straightforward way to create a mechanism that enables principals to monitor agents to ensure the maximization of social benefit. Instead, some proxy must be used. It is tempting to suggest that beneficiaries of a social mission may serve as proxies for the social mission, guarding against misuse of the firm’s resources or drift from the mission. But in many cases, beneficiaries are either too diffuse or too powerless to serve as adequate guardians. In the case of some social missions, like preserving the environment or protecting animal welfare, it may be impossible to directly empower beneficiaries. In the context of donative nonprofits, therefore, a combination of relatively disinterested directors and donors is made to play the role of guardian. The nondistribution constraint is perceived as the primary tool for making sure that donors are primarily altruistic and that motives of directors and employees are sufficiently cabined to prevent misdirection of the firm’s efforts.

The removal of the nondistribution constraint complicates the use of proxies for social missions. In a social enterprise, there may be no donors, only investors, who by definition have “mixed” motives. Directors and managers may also have financial interests that at times come into conflict with their pursuit of social ends. The goal of credible commitment devices in the social enterprise context, then, is to make sure that those who seek altruism can assure themselves that whatever they paid for their altruism has not been diverted to other stakeholders whose interests are not altruistically sought. Those devices will differ depending on which stakeholders serve as the proxy for the social mission of the firm.

For example, if the goal of a commitment device is to assure investors that their investment will be used for a social end, then the investor is the proxy for the social mission of the firm. Presumably, in that case, mechanisms to empower investors to identify, monitor, and enforce the social mission will be sought. Agency costs for such mechanisms may be low, especially if there are a small number of investors who are both highly invested in the firm and have the ability to monitor its activities. However, since investors will also be seeking a financial return (in the absence of the nondistribution constraint), there is no obvious way to ensure that investors as a class will prioritize the social mission of the firm over their financial interests.
When there are more than just a few investors, agency costs go up. Since each investor might value the social and financial goals of the firm differently, an investor who really values the social mission cannot rely on other investors to monitor or enforce the enterprise's social mission. In fact, other investors might be working at cross purposes. This problem is exacerbated if there are multiple classes of investors with differing abilities to monitor the firm's behavior. The worst-case scenario would be a diffuse class of small investors who value the social mission highly and a small class of large investors who value it less. But agency-cost analysis suggests that even within a single investor the existence of the profit motive might make one more likely to permit diversions from the firm's social mission when they conflict with financial goals. The whole point of the agency-cost analysis is that it will be costly to identify, monitor, and enforce social commitments, and that therefore one might choose to not know enough to really understand if the firm is pursuing its social mission. If it plausibly is, and you're making money, that might be enough to justify the investment even if it means that insufficient resources are invested in ensuring that the firm pursues its social mission.

If the commitment device is meant to assure customers of a social enterprise that the firm is advancing social goals, then commitment devices that empower investors may be useless, or at least less useful. This might occur in contexts in which social enterprises are selling goods or services to customers bundled with some altruism. That is, the customer is purchasing something for herself (like a cup of coffee) while also seeking to benefit some distant third party (like a coffee grower or the environment). Customers need commitment devices that use proxies other than investors to guard the social mission of the firm, since investors are by definition imperfectly altruistic (and may even not be altruistic at all). Certifications may suffice in certain circumstances, if the certification system is sufficiently robust and the customers value their altruism sufficiently to invest in identifying, monitoring, and enforcing compliance with the certification. But customers, like investors, are receiving a mixed good in these cases. They may not even be aware of how they differentially value the quality of the good itself and the quality of the altruism that accompanies it. Even more problematic, they may be heterogeneous in how they differentially value each component of what they purchase. Some customers may value the altruism strongly, but if they wrongly assume that other customers feel the same way, they may misperceive the amount of individual effort required to police the claims made by the firm. Since customers can often perceive the quality of the good relatively easily (this coffee tastes good), but not the quality of the altruism (how much deforestation is acceptable among independent growers?), it is predictable that firms will end up catering to the least committed altruists, no matter what their intentions are.

Much of the social enterprise literature imagines an altruistic entrepreneur who may serve as proxy for the social mission of the firm. After all, why would such a person invest their energy in a social enterprise when they could presumably get even richer in a "regular" business? Obviously, entrepreneurs in social enterprises have mixed motives just like everyone else, and so mechanisms that use them as proxies for the social mission of the firm are unlikely to be able to make credible commitments to other stakeholders that the firm will diligently pursue its social mission.

To summarize: the mechanisms that enable firms to make credible commitments to customers are not the same as the mechanisms that enable firms to make credible commitments to investors. If the commitments are to benefit society at large, or, even more amorphously, the environment, then both types of commitments are proxies for the social good. Since, in the absence of the nondistribution constraint, all stakeholders have mixed
Some Implications of the Agency-Cost Theory

interests in the output of the firm, no proxy is likely to be an adequate guardian of that social
good.24

On the other hand, if all that is intended is to maximize the social welfare of a specific
group of stakeholders, like employees, agency-cost theory suggests that empowering those
stakeholders may well be an easy way to serve that social end. As several chapters in this
volume point out, social enterprise business forms in continental Europe are more likely to
take traditional forms that empower workers or consumers, like cooperatives. Using business
forms that democratically empower specific stakeholders (like workers) over other stake­
holders (like investors) should function to increase the chances that those favored stake­
holders’ interests (whether social or financial) are maximized, There is arguably no need for
a nondistribution constraint when the beneficiaries and the principals are the same stake­
holders. There are obviously still agency-cost problems, since stakeholders as a class must still
monitor their agents, and they might have additional trouble attracting investors if they expect
them to be at least partially altruistic. But those agency costs are likely to be less burdensome
than those that Hansmann argued regularly result in contract failure.

The point of applying agency-cost logic to social enterprises in this way is simply to be
aware that different stakeholders have different interests. The nondistribution constraint is the
time-tested method of binding a variety of stakeholders to a
generalized
commitment to
restrict certain financial incentives in the interest of the general good. It allows all stake­
holders to trust that financial incentives are not diminishing the common pursuit of altruism,
whatever exactly that means. When the nondistribution constraint is removed, as it is for for­
profit social enterprises, trust must be derived from some other commitment device.

IV OBSERVATION TWO: NEGATIVE RESTRICTIONS VERSUS POSITIVE BENEFITS

Paying attention to the agency-cost theory of the nonprofit firm leads to a second observation
about the agency costs associated with different types of socially beneficial commitments.
Mechanisms that enable one to identify, monitor, and enforce compliance with a
negative restriction are easier to develop than those that enable one to identify, monitor, and enforce
a commitment to maximizing a positive social benefit. This distinction between negative
restrictions and positive benefits has been observed many times.25 But it is worth drawing out
the implications of this distinction since it is so central to the types of mechanisms that are
likely to be effective for enterprises seeking to make social commitments, and for legal reform
intended to facilitate such mechanisms.

A Agency Costs and Negative Restrictions

Brian Galle coins a nice term to capture the difference between negative restrictions and
positive benefit. He observes that regular businesses, especially in some sectors, inadvertently
produce what might be called “cold glow” along with the products or services they

24 As I discuss in much greater depth in an earlier work, the government is also a stakeholder when benefits, like
more lax regulation or tax incentives, are provided to social enterprise firms. When all stakeholders have mixed
motives (altruistic as well as financial), the government cannot use any of them as a proxy for the social benefit
of the firm, and its own agency-cost calculation counsels against providing benefits in such circumstances. See Leff,
supra note 6, at 854–65.

intentionally offer.\textsuperscript{26} Cold glow arises when goods or services are provided in a context in which stakeholders are concerned about possible negative externalities or bad behavior by the producer. For example, once it is known that tuna fishing often results in killing dolphins, the value of tuna is diminished by cold glow. In order to reduce cold glow, producers need to develop some mechanism to make credible commitments to their consumers that dolphins are not harmed during the production of their tuna. Once it is known that apparel is produced in sweatshops that employ children, sellers of apparel may seek to find commitment devices that reduce the cold glow arising from consumers’ knowledge of (and distaste for) child labor or other exploitative labor practices. Under the Henderson/Malani definition of “altruism,” reduction of cold glow is as much altruistic as affirmatively attempting to benefit third parties.\textsuperscript{27} Any attempt to benefit unrelated third parties beyond the requirements of the law constitutes an expression of altruism.\textsuperscript{28} Galle observes that regular businesses are relatively adept at making use of mechanisms to reduce cold glow,\textsuperscript{29} and the agency costs associated with such mechanisms appear to be modest.

Some definitions of social enterprises are broad enough to include these cold glow-reducing firms.\textsuperscript{30} Others may exclude “regular” cold glow reduction, but include firms which have a significant purpose of advancing a social good through sustainable or responsible practices that are, for all intents and purposes, the same as “ordinary” cold glow reduction. The key distinguishing factor for cold glow reduction is that firms are making commitments to abide by a specified code of conduct rather than to maximizing social benefits. So they may avoid harming dolphins by using dolphin-safe nets, but they are not claiming that they are maximizing the welfare of dolphins or the ocean’s ecosystem. As long as the firm uses dolphin-safe nets, it has successfully pursued its social mission. Likewise, a firm can avoid harming children simply by refusing to employ them (and refusing to use sourcing policies that are known to lead to increased child labor, like permitting workers to produce goods in their own homes). They do not have to justify the social benefit of using dolphin-safe nets or of not employing children to their stakeholders because the harms of doing otherwise are believed to be well understood.

In the context of impact investing, this distinction between negative “screens” and positive benefits is widely recognized. Negative screens are investment rules that identify harmful industries in which socially responsible funds refuse to invest. For example, funds may choose not to invest in the gun industry, the tobacco industry, or the fossil fuels industry. These investment rules are relatively easy to identify and communicate to stakeholders, and those who choose to invest in a no-tobacco company fund have presumably made their own


\textsuperscript{27} It may well be that Henderson/Malani’s definition of all positive externalities as a separate “altruistic” good is less coherent than treating the ethical production of a good (however defined) as a component of its quality. The distinction is potentially important because Henderson/Malani view altruism as a public good, whereas the product-specific quality of ethical production may well not share the characteristics of a public good.

\textsuperscript{28} As defined earlier, supra note 18, legal requirements form this baseline against which altruism is measured. It’s useful to have some baseline, although on reflection the baseline probably falls apart. If dolphin-safe nets are required by international law, but many fishing operations flout the law, then buying certified dolphin-safe tuna is an expression of altruism as much as if there was no such international law.

\textsuperscript{29} See Galle, supra note 26, at 2027 (“Firms have learned how to contract around many of these problems. As ample literature now demonstrates, firms can hire outside auditors or other third-party verification systems to oversee the firm’s compliance with cold-glow reducing goals. If the goals are readily defined, quantified, and measured, then both contract drafting and enforcement can be fairly cheap.”).

\textsuperscript{30} Yockey, supra note 19, at 769 (“The prototypical social enterprise is a for-profit firm that relies on market-based strategies and techniques to advance a specific social mission.”).
decision that tobacco companies do not operate in the public interest. Likewise, it is relatively
easy to communicate a policy of not hiring child labor, and the supposed benefits of such
a policy are already well understood by consumers who may favor a company that has such
a policy.

One would predict that the agency costs associated with negative restrictions would be
relatively low. Once the socially harmful practice and its solution are identified, all that
remains is convincing stakeholders that the rule has been followed. Obviously, this task can
be daunting, and firms could no doubt be doing much more to assure their stakeholders that
they are following these rules. But it is not inconceivable that ordinary contractual mechan­
isms or special social enterprise commitment devices would succeed, given how modest the
firm's commitment is.

It is also very important to recognize that customers can use ordinary market mechanisms
to assess the efficiency by which a firm provides goods subject to a negative restriction.
Customers can compare the price of the good produced subject to the restriction and the
price of comparable goods produced without any restriction and choose the higher-cost good
if they value the restriction more than the difference in cost. Consumers make these
determinations every day without worrying about the quantity of social good produced by
their purchases. In that case, so long as the firm has earned the consumer's trust that it is
abiding by the restriction, there are no additional agency costs. The situation is even better if
more than one firm is bound by the same restriction. If both Starbucks and Peet’s serve fair­
trade coffee, a consumer can compare prices of coffee produced subject to the same restric­
tion from each firm. In that case, we would predict competition between the two firms to
produce efficiencies in the production of fair-trade coffee, and a consumer need not worry
about directly monitoring agency costs.

It is possible to imagine a world in which some consumers or investors believe that ordinary
business practices – or the particular business practices in common use in contemporary
global capitalism – are themselves harmful. The problem is not the use of particular tuna­
fishing technology that harms dolphins, but the whole apparatus of contemporary business
practice, in which multinational corporations pursue their financial ends on a global scale.
For these consumers or investors, contemporary capitalism itself functions like a cold-glow
industry. They are looking for a “negative screen” for all the ordinary business practices
that cause harm to the social good or the environment. They need a generalized cold-glow­
reducing mechanism. It is possible that these consumers and investors are looking to social
enterprise to provide an all-purpose mechanism to reduce the cold glow that arises from
global capitalism.

Because of the importance of market pricing mechanisms in comparing similarly restricted
enterprises providing similar goods, generalized cold glow reduction is likely to produce
much higher agency costs than the kind of specific mechanisms found in the regular business
context. That is, if the goal is to find a mechanism to reduce the general cold glow produced
by global capitalism, then agency costs might be much higher. That is because cold-glow­
reducing mechanisms demand consensus on the precise contours of the restrictions so firms
that abide by the restrictions can compete against each other on the same terms.

The consumer needs to already accept the precise nature of the evil to be avoided, and
what would constitute an acceptable avoidance of such evil. If no two firms in the same
industry have exactly the same set of restrictions, then it is impossible for consumers or
investors to use market mechanisms to price the benefit of abiding by the restrictions, and so
agency costs are likely to be unconstrained. Even worse, if the restrictions are somehow
aspirational, rather than strict, and firms must balance one goal against another, then there is no easy way to even determine how well a firm has abided by the restriction, and agency costs should be even more unconstrained.

Imagine that “abusive labor practices” is the evil to avoid. If one imagines that a general characteristic of contemporary global capitalism is that labor is in an untenably weak bargaining position against capital, and so the ordinary functioning of the market is likely to produce exploitative labor relations, then what exactly is the restriction that would solve that problem? Presumably, one could require that firms pay their employees (and contractors, and the employees and contractors of all their sourcing companies) a living wage. But that demands consensus on what constitutes a living wage in each labor market in which the firm operates. One could require that all labor involved in the production of the good be produced in a jurisdiction that has robust labor laws to protect against exploitation of labor, but again, that demands consensus on what constitutes such regulations. And why demand that workers in countries that are less protective be denied the opportunity to earn a living? One could require a unionized workforce, or even that the firm be wholly worker owned, but patrons may disagree about which of these solutions is the most effective means of countering exploitative labor practices. If the problem seems overwhelming with regard to labor practices, how much more so is it with regard to environmental practices, especially if carbon production (or something equally ubiquitous and unavoidable) is the kind of general evil that one seeks to avoid? Suffice it to say that producing a general negative restriction for the evils of global capitalism would be difficult, and each component of that difficulty increases the agency costs associated with employing that mechanism as an effective general screen.

B Agency Costs and Positive Benefits

If the cost of enforcing negative restrictions seems relatively modest, at least for specific restrictions, the same cannot be said for the costs associated with positive benefits. The agency-cost story about positive social goals is completely different from the story about negative restrictions. When a social enterprise (or any firm) seeks to commit to its stakeholders that it will pursue a positive social good, it cannot rely on a simple restriction to make its case. Instead, it must somehow communicate to its stakeholders what it intends to do to benefit society and how those activities distinguish it from “ordinary” businesses. Many supporters of social enterprises imagine the social enterprise space as especially intended to support enterprises that pursue this kind of positive social benefit.

Any enterprise attempting to commit itself to a positive social benefit must solve several problems simultaneously. First, and perhaps most importantly, some metric must be identified to measure the social benefit provided. Attention to the measurement of social impact has been growing dramatically in the past several decades. This measurement takes place in all three sectors of the economy (government, nonprofit, and business). But, as anyone who has tried to develop an impact assessment system for a project or enterprise knows, there are numerous problems with creating an effective one. These problems include: identifying the right ultimate measure of impact, separating outputs (within the direct control of the...
Some Implications of the Agency-Cost Theory

When a firm tries to maximize social benefit, one of the biggest challenges is just deciding how to account for the social impact provided. Narrative reporting may capture rich detail about the ways in which the firm’s activities positively impact stakeholders or the environment, but it is so flexible and anecdotal that it cannot provide any real grounding for the reduction of agency costs. Rather, what is needed is some objective measurement system, which demands relative rigidity. But there are countless examples of measurement systems distorting the service being provided in perverse ways. So, even just choosing what to measure and how to measure it is a profound problem in many cases.

Once a measurement system is satisfactorily created, it is very difficult to effectively price the benefit unless there is a market that pursues the same specific social benefits using the same terms. If two firms are seeking to provide clean water to distressed communities, then one could imagine pricing clean water by volume or by number of people served. After all, it is easy to be convinced that clean water is essential to health, and so no ultimate measure of health benefits is required. But it gets harder when one tries to measure the benefits of clean water against the benefits of secure shelter, for example. The dream of creating a scale or measurement system that could assess very different social benefits against each other seems well beyond our capacity, at least with any precision. This gets even more difficult when environmental impacts are added to the mix.

In other words, imagine that one has successfully identified a social benefit and a way of adequately measuring the degree to which the firm’s actions advance that social benefit. If no other firms are providing a social benefit in a comparable way, using the same units of measurement, then it is very difficult for a consumer to know how to determine if the price of such a benefit is good. Without this market pricing mechanism, an ordinary consumer is unlikely to be able to demand a price that reflects an efficient provision of the good, and so agency costs are likely to be very high.

Even after all these measurement issues have been resolved in a satisfactory way, the enterprise arguably has to account for the marginal benefit provided, rather than just the overall benefit. As pointed out earlier, as long as there is no way to measure and report social benefit per dollar received, there is no way to adequately assess whether a firm is providing social benefits in an efficient manner. In that case, there is no check on agency costs.

All of this is not to say that there are no mechanisms for measuring positive social impact; there are. For several decades more and more enterprises (and governments) have been working diligently to devise ways of measuring social impact in a wide variety of contexts. However, the challenges of adequately measuring (as well as monitoring, reporting, and enforcing commitments to) social impact suggest that agency costs are likely to be highest in situations in which the firm is attempting not just to exclude certain known negative behaviors of other firms, but instead to affirmatively improve the world in some way. Firms

33 Controlled experiments are widely viewed as the most effective mechanism for determining the actual impact of any intervention, but it is widely understood that conducting such experiments is infeasible in many or most situations.  Id.

34 Some make a compelling case for assessing very different social benefits against each other in a very general way, even in the absence of any more granular method. See, e.g., WILLIAM MACASKILL, DOING GOOD BETTER: HOW EFFECTIVE ALTRUISM CAN HELP YOU MAKE A DIFFERENCE (2015); PETER SINGER, THE MOST GOOD YOU CAN DO: HOW EFFECTIVE ALTRUISM IS CHANGING IDEAS ABOUT LIVING ETHICALLY (2015).
with these types of goals are likely to find that some sort of nondistribution constraint is still the most efficient mechanism available to reduce agency costs.

C Current Mechanisms for Assessing Social Impact: IRIS, GIIRS, and B-Lab

If both the attempt to create generalized negative restrictions to screen out the harms of global capitalism and the attempt to create mechanisms to permit firms to make credible commitments to maximize positive social benefits are unlikely to reduce agency costs, where does that leave the social enterprise movement? Evaluating such progress is well beyond the scope of this chapter, but other scholars have done significant work critically evaluating the market mechanisms and legal reforms that constitute the current state of the social enterprise reform movement.

For example, Sarah Dadush has investigated the mechanisms being developed to enable impact investors to reduce agency costs involved in identifying social enterprise firms. She finds that the Impact Reporting and Investment Standards (IRIS) system, which purports to be a system to regularize the vocabulary used to assess social impact, is actually largely devoted to what I have been calling negative restrictions. Even so, these standards can be used by firms in multiple uncoordinated ways, reducing the ability of investors to easily find the kind of standardized information that could be predicted to significantly reduce agency costs. The Global Investing Rating System (GIIRS) has a more lofty goal in that it attempts to provide a score for each firm making use of IRIS and other measures. But it too generally avoids trying to quantify actual positive impact. For example, rated categories like governance policies and transparency presumably provide the preconditions for reporting social impact, but they do nothing to measure the social good produced. In that way, these policies function more like negative restrictions (against corporate secrecy) than actually certifying some social good produced. The system assumes that a particular process or policy will produce socially beneficial results and gives points based on whether the firm has adopted the policy. The same is true of employment policies like worker control. While some might be an altruistic end in themselves (like higher wages for workers), mostly they are proxies for actual social benefit. And, like other negative restrictions, they assume both the existence and the harm of firms that do not comply with them. Thus, efforts like IRIS and GIIRS function more like generalized negative screens than positive commitments to particular social benefits.

The same tension is present in debates about social enterprise business forms like (social) benefit corporations, as is discussed more fully in other chapters of this volume. Most benefit corporation statutes require that firms measure themselves against a comprehensive third-party standard that, by its very nature, cannot do more than provide a (large) set of negative restrictions, along with a set of positive policies intended to lead to socially beneficial outcomes. Jurisdictions that have adopted specific social benefit regimes as part of their

---


Dadush, supra note 35, at 406.

Although Dadush concludes that these mechanisms may be able to reduce some transaction costs, id. at 404-05.

Id. at 404.

benefit corporation statutes permit firms to attempt to quantify the social benefit they provide, but there is very little evidence of firms successfully doing so, except in very limited contexts. Most importantly for the purposes of this chapter, the difficulties involved in identifying appropriate metrics and committing to a method of reporting progress toward social goals suggest that agency costs are likely to be prohibitively high in these cases. Even more problematic is the veritable impossibility of using a market pricing mechanism to evaluate the cost of such progress (except in very limited circumstances) or to identify the marginal value of any customer’s or investor’s purchase of such benefit. Very little about the existing assessment technologies suggests promising solutions to these types of agency-cost issues. The promise is exciting, especially for consumers or investors who are chilled by the cold glow of contemporary global capitalism, but the progress to date is limited.

V Conclusion

Paying attention to the agency-cost theory of nonprofits gives rise to some observations that are relevant to social enterprises. In this chapter, I have argued that paying attention to agency costs suggests that commitment devices intended to permit one set of stakeholders to monitor whether a social mission has been pursued are not effective for other stakeholders. More importantly, the fact that all stakeholders in a for-profit social enterprise have mixed motives (financial and social) means that stakeholders cannot depend on the efforts to monitor and enforce social mission provided by other stakeholders, whether those stakeholders are of the same class (e.g., investor or consumer) or a different class.

Second, I have argued that there is a big difference from an agency-cost perspective between specific negative restrictions, a general negative restriction, and commitments for positive social benefits. The fact that for-profit firms are somewhat successful at making compelling commitments to their stakeholders about particular negative restrictions should not confuse social enterprise advocates into thinking that it is as easy to overcome agency-cost problems for either general negative restrictions or commitments for positive benefits. The existing social enterprise legal forms that attempt to do just that are unlikely to successfully cabin agency costs simply because of the challenge of what they are proposing. Of course, whether they succeed is an empirical question that can only be answered after some period of experimentation. The fact that theory urges caution should not be taken as a negative assessment of the movement. The promise of social enterprise is exciting. But there are reasons that unimpeded experimentation might be harmful as well.

Attempts to create new legal forms for for-profit social enterprise firms are an express attempt to gain a government imprimatur for the efforts of social enterprises. To the degree to which these new legal forms imply more enforcement of a firm’s social mission than they actually accomplish, there is a risk that stakeholders will in effect be deceived by the legitimacy bestowed by legislation. The fact that the legislation facilitates forms that have no real likelihood of reducing agency costs in the provision of social good means that it is likely that value will be diverted from its intended purposes. If stakeholders become frustrated with such diversions, it is not only the specific business or business form that will suffer, but arguably the whole social enterprise sector.

40 Which is why Charles Yablon, in his chapter in this volume, advocates permitting maximum experimentation among social enterprises. See Yablon, supra Chapter 22.
More than three decades ago, Henry Hansmann described the nonprofit form as a “collective contract between the organization and its patrons,” and he argued that “considerable economies can be realized” by placing transactions under this one umbrella contract. These economies are derived not just from the substance of the agreement—that it prevents managers from expropriating the excess value of the firm—but from its very nature as a single form contract enforced by the state. The parameters of that contract—what distinguishes impermissible profits from fair payment for services, for example—may be subject to dispute, but the system can only work if stakeholders trust that the nondistribution constraint is largely working as promised. The existence of multiple forms of the contract, each with its own strengths and weaknesses, has the potential to weaken the value of the unity of the nondistribution constraint. The nonprofit sector is huge and growing, and the nondistribution constraint appears to be adequately serving its purpose of facilitating “altruistic” transactions for the social good. To the degree to which new social enterprise business forms draw on the trust and legitimacy of the nonprofit sector, but then do away with the bedrock commitment device that underpins that sector, the cost of experimentation with new business forms may be high. It may negatively impact the nonprofit sector along with the emerging social enterprise sector, and that would be costly for all of us.

Hansmann, Role, supra note 1, at 853 (“Consequently, considerable economies can be realized by placing all such transactions under one collective contract between the organization and its patrons: the contract determined by the state’s nonprofit corporation law and policed by the state.”).