Efficient Funding: Auditing in the Nonprofit Sector

Natalie Privett, Feryal Erhun
Department of Management Science and Engineering, College of Engineering, Stanford University, Stanford, California 94305 [nprivett@stanfordalumni.org, ferhun@stanford.edu]

Nonprofit organizations are a critical part of society as well as a growing sector of the economy. For funders there is an increasing and pressing need to ensure that society reaps the most social benefit for their money while also developing the nonprofit sector as a whole. By routinely scrutinizing nonprofit reports in an effort to deduce whether a nonprofit organization is efficient, funders may believe that they are, in fact, giving responsibly. However, we find that these nonprofit reports are unreliable, supporting a myriad of empirical research and revealing that report-based funding methods do not facilitate efficient allocation of funds. In response, we develop audit contracts that put funders in a position to enact change. Auditing, perhaps obviously, supports funders; however, we find that it also benefits the population of nonprofits. Moreover, auditing results in improved efficiency for the nonprofit sector overall. Indeed, our conclusions indicate that nonprofits may want to work with funders to increase the use of auditing, consequently increasing efficiency for the sector overall and impacting society as a whole.

Key words: public sector; nonprofit sector; audit contracts; adverse selection; principal–agent framework; resource allocation; operational transparency

History: Received: May 11, 2009; accepted: April 3, 2011. Published online in Articles in Advance September 2, 2011.

1. Introduction
At the interface of operations management and public sector application lies a wealth of untapped research possibilities. Nonprofit operations in particular have not been traditionally considered within the realm of operations management research; however, recent trends in nonprofit organizations call for more substantial investigation (Brest and Harvey 2008, Bradley et al. 2003). In light of these discussions, this paper seeks to examine the relationship between a funder and nonprofits as contractual in nature to address the following key issues:

1. Do report-based funding methods facilitate the efficient allocation of funds?
2. How can audit contracts be implemented in the sector? What effects do these contracts have on both the funder and nonprofits?
3. Can auditing be used to improve the performance of the nonprofit sector as a whole?

Nonprofit organizations are a significant and growing segment of the economy in addition to making critical societal contributions. Nonprofits provide services to those who otherwise would not receive them, in many cases partnering with the government. They provide the opportunity for individuals to volunteer and give back to society, which has become a prominent part of many cultures. Nonprofits also provide a large amount of public goods that no one directly pays for but from which everyone reaps benefits, such as advocating for cleaner air. Beyond these social contributions, the nonprofit sector’s economic impact should not be underestimated. In 2009, the Internal Revenue Service (IRS) registered approximately 1.5 million nonprofit organizations, which account for 10% of employment in the United States (Sherlock and Gravelle 2009).

There are many entities that sustain the nonprofit sector financially. Foundations, themselves nonprofit organizations, and governments both act as key grant-making institutions. In 2007, the Bill and Melinda Gates Foundation alone paid approximately two billion dollars in grants, funding nonprofits working in global health, poverty, and education; and the National Science Foundation, a United States government entity, distributed billions of dollars to scientific research and educational activities. Worldwide, foundations have continued to grow and expand, most dramatically in the United States where more than 72,000 grant-making foundations granted a record $42.9 billion in 2007 (Foundation Center 2008, Prewitt 2006). Because funders want assurance that their donations are used wisely, efficiency is a critical measure. In the nonprofit sector, efficiency is about getting “[more] mileage out of the money [nonprofits] spend” (Herzlinger 1996, p. 98). The most common cost-centered operational definition of efficiency measures the ratio of expenses directly forwarding the mission to total expenses, numbers that can be drawn from the organization’s publicly available IRS Form 990. This efficiency communicates a relative
financial condition, highlighting similarities in financial goals while controlling for differences over time and across nonprofit organizations (Chabotar 1989), thus enhancing understanding of these organizations and helping to create informed decisions about financial support.

Funding methods, particularly grants, are in essence contracts exchanging money for some societal return where nonprofit organizations act as "sellers" of certain societal activities or delivables (Brest and Harvey 2008). To allocate grants, many funders scrutinize financial statements and public reporting forms, such as the IRS 990, in efforts to deduce whether a nonprofit organization is operationally efficient (Frumkin and Kim 2001). In doing so, funders may believe that they are, in fact, funneled funds to the highest-return organizations. This research describes and examines this basic funding situation, which is based merely on the nonprofit's reported cost-centered operational efficiency and the funder's allocation. However, under these very basic terms, our analysis shows that these nonprofit efficiency reports are unreliable. This result of a lack of efficiency reliability and its causes are well documented (e.g., Quality 990 2000, Schwinn and Williams 2001, Frumkin and Keating 2003, Trussell 2003, Jones and Roberts 2006, Krishnan et al. 2006, Gordon 2007, Keating et al. 2008), confirming that our results accurately describe the basic situation where funders are unable to effectively distinguish between efficient and inefficient organizations. Regardless of the basis for the unreliability, this theoretical result leads to the conclusion that a simple contract based solely on funds allocated in response to an efficiency report is not effective. Thus, in response to our first research question, we show that report-based funding methods, in fact, do not facilitate efficient allocation of funds.

Although the current reports of efficiencies are unreliable, a 2008 study on the nonprofit marketplace confirms that efficiencies indeed are verifiable (William and Flora Hewlett Foundation and McKinsey & Company 2008). Thus, in response to the second question, we develop audit contracts that can be implemented within the funder–nonprofit relationship. In these contracts, the funder has the option to audit the nonprofits' efficiencies after awarding them funds. In case there is a discrepancy between the efficiency that the nonprofit reports and its true efficiency observed after auditing, the funder may ask the nonprofit to pay a penalty back. We find that using audit contracts, the funder can guarantee truthful efficiency reports from the nonprofits and thus attain operational transparency. The nonprofit population also prefers these audit contracts. As a result, we may see funders and nonprofits working together to increase auditing within the sector and improve their respective situations. This can be seen at work in both the nonprofit sector push to increase voluntary self-auditing and legislative efforts requiring audits, such as the California Nonprofit Integrity Act of 2004 (California Government Code 2007).

The implementation of audit contracts can also improve the performance of the nonprofit sector overall. More specifically, the funder's choice of penalty positions the sector anywhere between the current, inefficient situation and the transparent, efficient situation. This implies that through implementation of audit contracts with appropriately chosen penalties, benchmark performance can be achieved for the nonprofit sector as a whole. Thus, the use of auditing may put funders in a position to enact change in the nonprofit sector and increase overall sector efficiency.

The remainder of this paper is organized as follows: Section 2 reviews the related literature. Section 3 describes the general model framework. Sections 4 and 5 develop the contracts, where incentive and reliability issues are explored in theoretical detail. Section 6 investigates the effects of these different contracts through both theory and numerical examples. In §7, results are extended to the cases of constrained budget, uncertainty in production, and cost of auditing. Finally, §8 concludes by reviewing this work in light of the original research questions presented in the introduction. For expositional simplicity, all proofs are presented in an online addendum (available at http://msom.pubs.informs.org/eCompanion.html). Throughout the paper, the operator $E_x[\cdot]$ is used as the expectation over random variable $x$ and vectors are represented with boldface.

2. Literature Review

With the nonprofit sector expanding in size and influence, it may be surprising that there is relatively little research, especially analytical research, involving nonprofit operations.

2.1. Operations Management in the Nonprofit Sector

McCardle et al. (2009) use a utility-based donor model to analyze the behavior of nonprofit donations in the presence of publicized tiered fundraising structures. Their model can be used by organizations to make effective decisions regarding the implementation of tiered fundraising structures, which they show to generate larger donations. Similarly, Toyasaki and Wakolbinger (2010) study whether an aid agency should earmark its donations while raising funds for an emergency. The authors conclude that earmarking leads to a lower overall fundraising cost percentage and should be used if fundraising costs and goals are high and donors' willingness to donate to a special
2.2. Nonprofit Evaluation

There is a large, multidisciplinary body of literature on nonprofit evaluation, which provides many different techniques for measuring outputs (i.e., more easily measured, tangible returns) and outcomes (i.e., less clearly measured, less tangible returns). For example, Stufflebeam (2001) identifies, describes, and assesses 22 different nonprofit evaluation approaches; and Martin and Kettner (1996) put such performance measurement into perspective by delving into the nuances of its practice. Aimed toward the nonprofit practice audience, Newcomer (1997) gives an overview of both design and use of evaluation and performance measures. Herzlinger (1996) calls for increased accountability to restore public trust, particularly through financial evaluation. She details four questions to perform financial assessment, which are implemented in The Boston Foundation’s assessment of and call to action for the fiscal health of Massachusetts nonprofits (Keating et al. 2008). As a more holistic financial evaluation framework, social cost-benefit analysis values projects, programs, etc. as the net social benefit minus the net social cost where all impacts must be monetized; Boardman et al. (2006) and Levin (1983) both provide reviews. Our research draws on these methods in developing the funder’s perspective of nonprofit contribution. Beyond this, our modeling of the funder’s utility, particularly the portions derived from the nonprofits funded, extends the theory and understanding of how these methods can be incorporated into strategic decision making and theoretical modeling.

2.3. Capacity and Resource Allocation

Allocation is widely studied in both operations management and economics literature. More recently, researchers have tackled information problems in capacity and resource allocation, specifically the inefficiencies created because of incentive misalignments (e.g., Harris et al. 1982; Cachon and Lariviere 1999a, b; Rajan and Reichelstein 2004; Karabuk and Wu 2005), which are most in line with the goals of this paper. Cachon and Lariviere (1999a, b) study inter-firm capacity allocation problems where a single supplier is allocating capacity among retailers with private demand information. Karabuk and Wu (2005), Rajan and Reichelstein (2004), and Harris et al. (1982), on the other hand, examine intra-firm capacity and resource allocation problems. Karabuk and Wu (2005) study incentive alignment for a single decision maker allocating capacity to managers with private demand information using bonus payments and participation charges. Rajan and Reichelstein (2004) and Harris et al. (1982) model a common resource being allocated among divisions where the output of each division is an increasing function of allocation along with productivity and managerial effort, both of which are the divisional managers’ private information. Transfer pricing mechanisms are found optimal. Our funding model is certainly in line with these streams of literature; it presents an allocation of funds among horizontal agents with private cost information in the form of an efficiency ratio. In addition, our decision maker has outside opportunities and we seek to use auditing mechanisms to resolve asymmetric information. Because of the nature of the utilities in our model, our objective function incorporates the funder’s as well as part of the nonprofits’ objectives making the solution more socially optimal. Therefore, we are examining a problem with both shared and conflicting objectives between the principal and agent(s).

2.4. Incentives

In seeking to align incentives and resolve problems due to asymmetric information, many researchers rely on optimal contracts (i.e., contracts that include contingencies for each possible value of asymmetric information). Such contracts may be arbitrarily complex, making them difficult to write, implement, and administer in practice. Recently, we observed a trend in operations management literature where simple (in terms of the number of contingencies included)
yet efficient contracts are studied as an alternative to complex optimal contracts (e.g., Lariviire and Porteus 2001, Kayı¸s et al. 2011, Kalkancı et al. 2011). In a similar spirit, we also do not work with complex optimal contracts, but develop simple, implementable audit contracts that can be applied given the current state of the nonprofit sector.

2.5. Auditing
A wealth of economic literature exists around auditing, particularly in relieving incentive issues; we focus on a few representative pieces here. Baron and Besanko (1984) explore the use of auditing in the context of regulation of firms, specifically in cost reporting, where the regulator can order consumer refunds in the case of overstated costs. The authors explore auditing as a means to structure firm incentives for truthful cost reporting under asymmetric information. Using a related model, Laffont and Tirole (1986) analyze incentive contracts, both regulatory and procurement, with the addition of a moral hazard component, which prevents the regulator from rewarding high costs. Our scenario, however, is not regulatory in nature. Though we study auditing in the spirit of this literature, we do not discuss the details of performing such audits. Tools, methods, and examples of performing audits in the nonprofit sector can be found in Bridge et al. (2009) and Carman (2009). Our funder is seeking her own objectives, which, in fact, are both conflicting and shared with the nonprofits. These shared objectives between the funder and nonprofits complicate the situation and produce unique results. We also do not merely look at optimal contracts that achieve first-best performance as we examine the funder’s perspective, the nonprofit perspective, and the sector efficiency perspective. Instead, we devise alternative contracts to improve efficiency, as well as the situation for all parties and the sector.

3. Model Definition
The relationship between nonprofits and the funders who support them is not trivial. The funder supplies resources to nonprofits in the form of grant allocations with the expectation that these nonprofits produce returns in the form of social output or outcome. An example of output might be the number of shelters distributed following a natural disaster by WorldVision, a nonprofit humanitarian organization. An example of outcome might be increased appreciation and education regarding the environment by the Sierra Club, a nonprofit environmental organization. However, as discussed in §2, there is a growing body of literature on nonprofit evaluation, which provides many different techniques for measuring program outputs and outcomes. Considering this literature and our stylized model framework, we treat output and outcome equivalently. Accordingly, in the remainder of the paper, we use the term “output” to include both output and outcome.

Although both the funder and the nonprofits aim to maximize this social output, they each have other, possibly conflicting, components of their respective objectives. More specifically, the funder looks to maximize the value for her dollar whereas nonprofits look to maximize the dollars they receive (i.e., their allocation). Each nonprofit has an unobservable efficiency type, which describes how well the nonprofit converts allocations into outputs/outcomes furthering their mission. In addition, nonprofits can also exercise effort to further their mission; this effort is also unobservable. These unobservables are in tension with the funder’s objectives of giving responsibly and fulfilling a duty to the sector. Note that efficiency is characterized as a type as opposed to a decision variable for the nonprofit because although nonprofits can increase their efficiencies, these changes “don’t just happen” (Neuhoff and Searle 2008, p. 33). Instead these changes reflect a decision that entails a process much longer than the funding process; thus, they are outside the scope of this paper.

The scenario that we analyze consists of one funder and \( N \geq 1 \) nonprofits. All parties are expected utility maximizers. The funder does not know the true efficiency types of the nonprofits with certainty, but knows that each true efficiency type, \( \theta_i \in [0,1] \), is distributed with probability density function \( f(\theta_i) \) and cumulative distribution function \( F_i(\theta_i) \), \( i = 1, \ldots, N \). Note that although the support of the efficiency distribution is \( [0,1] \), an efficiency type of one does not necessarily mean that the nonprofit is channeling all of its funds to mission critical activities. Some funds may be used for activities such as fundraising and overhead, however, any upper bound can be normalized to one to accommodate other limits.

The efficiency definition adopted here describes only part of how effectively the nonprofits accomplish their mission; as such, it is a simplification of a complex concept. Nonprofit organizational effectiveness is multidimensional and, therefore, cannot be encapsulated in a single measure (Herman and Renz 1999). Yet this ratio is the most common cost-centered operational definition of efficiency in the literature and is also commonly employed by auditors, accreditors, media, and charity oversight analysts “to compare the operations of organizations with similar missions, with the goal of determining which organizations have the leanest operations” (Frumkin and Kim 2001, p. 270); for example, see Chabotar (1989), Herzlinger (1996), Hager and Flack (2004), Krishnan et al. (2006), and Gordon et al. (2007). For many analysts, this definition translates into accountability...
by providing insight into management and administration spending (Hager and Flack 2004), production costs, and translation of “the supply-chain into financial values through recording and analyzing the costs associated with products/activities undertaken” (Bagnoli and Megali 2011, p. 150).

The funding sequence of events for a given contract is as follows: (i) The funder announces some allocation scheme, possibly based on types, $A_i(\theta_i)$, for each nonprofit $i$. (ii) Nonprofit $i$ announces an efficiency, $\hat{\theta}_i$, perhaps through an application process. These efficiency types are also typically available through IRS data submitted by the nonprofits, such as the IRS 990 Form. This announcement may or may not be true. (iii) The funder allocates resources within the contractual framework to maximize her utility considering the nonprofits’ announcements and any constraint on her budget. (iv) Each nonprofit chooses an effort to maximize his utility based on this allocation.

Nonprofit $i$’s utility, $u_i$, and production (output), $y_i$, are determined by

$$u_i = \delta_i A_i - e_i^2 \over 2 + y_i \quad \text{(nonprofit utility)}, \quad (1)$$

$$y_i = 2\sqrt{e_i \theta_i A_i} \quad \text{(nonprofit output production)}, \quad (2)$$

where $\delta_i \geq 0$ is a multiplier describing how much the nonprofit values allocation as compared to output, $A_i$ is the allocation made from the funder to the nonprofit $i$, $e_i$ is nonprofit $i$’s effort, and $\theta_i$ is the true efficiency type of nonprofit $i$. The nonprofit’s utility in Equation (1) is increasing in the allocation he receives from the funder and in his output and is decreasing in the effort he exerts. For their survival, nonprofits depend on funds; therefore, the nonprofit’s utility is increasing in the allocation he receives. The nonprofit gains utility in output because output supports the nonprofit’s mission. Nonprofits may value allocation more than, less than, or the same as output; both output and allocation contribute to the nonprofit’s utility. This comparative value is described by the multiplier $\delta_i$. Although the nonprofit would like to support his mission, he prefers to do this with the least effort. This may be so that the effort, such as volunteer time, can be spread over other programs. The production function in Equation (2) indicates that the nonprofit must have positive effort, efficiency, and allocation to produce output. Furthermore, the nonprofit production function is concave in effort, $e_i$, efficiency type, $\theta_i$, and allocation, $A_i$, signifying diminishing returns and boundedness of production, similar to related models in the literature (e.g., Cachon 2003). We initially assume that output can be measured unambiguously; we relax this assumption and study production uncertainty in §7.2.

The funder’s expected utility, $U_f$, is

$$U_f = \sum_{i=1}^{N} E_{\theta_i}[c_i y_i] + \alpha \left( B - \sum_{i=1}^{N} E_{\theta_i}[A_i] \right), \quad (3)$$

where $c_i > 0$ is the funder’s monetary valuation of each unit of nonprofit $i$’s output, $\alpha$ is the return on an outside opportunity, and $B$ is the funder’s budget. The first term of the funder’s utility, $\sum_{i=1}^{N} E_{\theta_i}[c_i y_i]$, is the total expected value from the nonprofits’ output. The funder would like to maximize each nonprofit $i$’s output, $y_i$, and each output is worth some amount $c_i$. The $c_i$ term is individualized to each nonprofit and may incorporate the extent to which the work of nonprofit $i$ supports the mission of the funder. The funder only observes the output, $y_i$, and efficiency report, $\hat{\theta}_i$, of each nonprofit, but does not observe the nonprofit’s effort choice, $e_i$, nor true efficiency, $\theta_i$. The second term of the funder’s utility, $\alpha(B - \sum_{i=1}^{N} E_{\theta_i}[A_i])$, represents an outside opportunity. Although the funds allocated may increase the nonprofits’ output, these funds are also valued by the funder considering that they could be used for other opportunities. In fact, the funder may decide that the outside opportunity is worth more than the (social) return of investing that money into a particular nonprofit. For example, de Véricourt and Lobo (2009) investigate using this money in a for-profit venture or investing it in order to supplement the income stream. Note that the expectation of $A_i$, over $\theta_i$, is because the funder’s allocation scheme may depend on nonprofit types as outlined in the sequence of events.

Unlike a for-profit corporation that is presumed to distill the objectives of their shareholders, managers, employees, and clients into a single measure of shareholder value, “the nonprofit has no single primary interest group that is invariably and clearly defined, homogenous with respect to interests, and whose goals are easily expressible and transferable into the organization for assessment of alternative courses of action” (Speckbacher 2003, p. 268). Because of this varied abundance of stakeholders and purposes, quantifying and modeling the objectives of a nonprofit organization can be ambiguous and contentious. Our formulations are thus some first attempts to model the funding dynamics between funders and nonprofits and were developed through an understanding of the field, which we gained through personal communication (Foundation Center 2009, Philanthropy and Civil Society Research Workshop 2010, Meredith 2009), field reports, and academic and trade publications. Furthermore, our modeling assumptions, such as funder utility increasing in output and effort (Equation (3)) and efficiency modeled as asymmetric information, are consistent with relevant economics literature (e.g., Easley and
O’Hara 1983). Similar to related models in the literature (e.g., Cachon 2003), we express effort cost as a convex function (Equation (1)) because it is progressively more difficult to increase the effort, which is limited by capacity.

Based on our sequence of events, the nonprofit problem is solved first followed by that of the funder by backward induction. The nonprofit $i$ maximizes its utility based on the given allocation:

$$\max_{e_i \geq 0} u_i = \max_{e_i \geq 0} \left\{ \delta_i A_i - \frac{e_i^2}{2} + y_i \right\}, \text{ where } y_i = 2\sqrt{\epsilon_i\theta_i A_i}. $$

Substituting the production function, we obtain $u_i = \delta_i A_i - \frac{e_i^2}{2} + 2\sqrt{\epsilon_i\theta_i A_i}$. The effort, $e_i^*$, that maximizes the nonprofit $i$’s utility can easily be found as $e_i^* = (A_i\theta_i)^{1/3}$. Through substitution of $e_i^*$, reformulations of the nonprofit $i$’s utility and output functions as well as the funder’s utility function are obtained:

$$u_i^* = \delta_i A_i + \frac{1}{3}(A_i\theta_i)^{2/3} \quad \text{(nonprofit utility)}, \quad (4)$$
$$y_i^* = 2(A_i\theta_i)^{2/3} \quad \text{(nonprofit output)}, \quad (5)$$
$$U_f = \sum_{i=1}^N [2\epsilon_i(A_i\theta_i)^{2/3} - \alpha A_i] + aB \quad \text{(funder utility)}. \quad (6)$$

In this paper, we do not work with complex contracts in the traditional sense of principal–agent models. The idea here is to understand the situation and develop contracts that can be applicable given the current operational policies, practices, and trends of the nonprofit sector. This principal–agent model is different from others in that it incorporates both shared and conflicting portions of the respective objectives of the principal and the agents as well as unobservable characteristics. The utility and objective functions here in the nonprofit setting differ from those more familiar functions of the for-profit world. Our nonprofit objective function incorporates several perspectives: allocation from a managerial financial sustainability perspective, effort from an employee perspective, and production from the client and funder perspective. This heterogeneity of interests and objectives, partly shared with the principal or funder, is indeed a new avenue of exploration for both nonprofit research and “supplier–manufacturer” type game theoretic models. However, it is these differences that lead to unique results.

In the model analysis and remainder of the paper, we first assume an unconstrained budget in §§4–6. That is, we assume $\sum_{i=1}^N A_i < B$ so that we omit the budget terms in the above formulations. In §7.1, we explore and extend our results to the case of budget constraint, i.e., $\sum_{i=1}^N A_i = B$. In the following section, we begin investigating the funder–nonprofit relationship in order to answer our originally posed research questions. We start by describing the basic funding situation with the report-based contract as well as establish a benchmark, first-best contract for comparison purposes. These will motivate us to study audit contracts in §5.

### 4. Report-Based Contract

The funder would like to make the most of her money and, therefore, would like to allocate funds based on the efficiency types of nonprofits, i.e., $A_i(\theta_i)$. Given Equations (4) and (6) and using the revelation principle (Laffont and Martimort 2002) as is common in such models, the funder’s problem can be formulated as follows:

$$\max_{\lambda_i(\theta_i)} \sum_{i=1}^N E_i \left[ 2\epsilon_i(A_i(\theta_i)^{2/3} - \alpha A_i(\theta_i) \right] \quad (7)$$

subject to

$$u_i^*(\theta_i) \geq 0 \quad \forall i; \forall \theta_i \in [0, 1] \quad \text{(IR)},$$
$$u_i^*(\theta_i | \theta_i) \geq u_i^*(\hat{\theta}_i | \theta_i) \quad \forall i; \forall \hat{\theta}_i, \theta_i \in [0, 1] \quad \text{(IC)},$$
$$A_i(\theta_i) \geq 0 \quad \forall i; \forall \theta_i \in [0, 1] \quad \text{(nonnegativity)}, \quad (9)$$

where $\theta_i$ is nonprofit $i$’s true efficiency type, $\hat{\theta}_i$ is nonprofit $i$’s announcement of efficiency type, and $u_i^*(\hat{\theta}_i | \theta_i)$ is nonprofit $i$’s utility when his true efficiency type is $\theta_i$ but his announced efficiency type is $\hat{\theta}_i$; $u_i^*(\hat{\theta}_i | \theta_i) = \delta_i A_i(\hat{\theta}_i) + 3/2(A_i(\hat{\theta}_i)\theta_i)^{2/3}$. The individual rationality (IR) constraints ensure nonnegative utility for each nonprofit within the contract. The incentive compatibility (IC) constraints constrain each nonprofit to truthfully report their efficiency type by use of a utility-based incentive. Consequently, the funder’s problem is expressed as an adverse selection problem where the funder cannot observe the amount of effort, $e_i$, the nonprofit $i$ will exert nor the nonprofit $i$’s efficiency type, $\theta_i$. Note that in this formulation the funder allocates funds in response to the nonprofits’ reported efficiencies. Thus, the resulting report-based contract describes the situation where the nonprofits’ efficiency types are not verified:

**Theorem 1.** The report-based contract takes the form $A_i^R(\theta_i) = A_i^R = (64/27)\phi_i^3 \left[ E_i(\theta_i^{2/3}) \right]^3$, where $\phi_i = c_i/\alpha$ and $i = 1, \ldots, N$.

This solution indicates an intuitive relation: In a scenario where the true efficiency type is unknown, the funder would want to allocate more funds to the nonprofit that is more likely to be a higher efficiency type. The result also indicates that the funder is unable to differentiate among the nonprofits’ efficiency types as she bases her allocation solely on her expectation.
of each nonprofit’s type, the only reliable information she possesses. The nonprofits do not have an incentive to truthfully report their type, and, therefore, their reports are unreliable. The report-based contract, which is based merely on reporting without verification, consequently offers the funder no observability of efficiency types and no operational transparency within this contracting context. Theorem 1 is intuitive in yet another way: Nonprofit \( i \) will receive more funding as the funder values his output more (i.e., as \( c_i \) increases), but will receive less funding as the outside opportunity becomes more lucrative (i.e., as \( \alpha \) increases). In other words, when the outside opportunity is more valuable compared to nonprofit output, the funder will choose to allocate less funds to the nonprofit in order to gain from the outside opportunity.

The report-based contract succeeds in modeling the basic funding situation where there is no reliability in efficiency reports and no transparency in nonprofit operations. Bradley et al. (2003) find nonprofits misreporting their efficiencies in situations where potential funders are investigating this measure to make decisions about giving and fund allocation. Misreports of efficiency may be due to managerial motivation, executive compensation incentives, use of efficiency in funding decisions, charity ratings, media attention, or reputational pressures (Frumkin and Keating 2003, Trussel 2003, Jones and Roberts 2006, Krishnan et al. 2006, Keating et al. 2008). A myriad of other literature related to unreliable IRS data and expense shifting also supports our theoretical result of a lack of observability and reliability of efficiency types in the current funding situation. The IRS Form 990 is the chief source of the financial data from which the efficiency measure is derived. However, the 990 is typically not verified. In fact, Schwinn and Williams (2001) assert that the IRS only examined 1.3% of 990s in 1999. Furthermore, the IRS Form 990 allows nonprofits to record certain administrative and fundraising costs as offsets to revenue rather than as expenses. As a result, nonprofits may portray their operations as more efficient in the Form 990 (Frumkin and Keating 2003). Jones and Roberts (2006, p. 161) report that charities do indeed employ expense shifting to “offset 16% to 28% of potential changes in program ratio [efficiency].” Herzlinger (1996) claims that this nonexistence of accountability perpetuates ineffective and inefficient organizations as well as other problems. Froelich and Knoepfle (1996, p. 49) summarize that overall, “the nonprofit situation [is] characterized by extensive reporting, but very weak monitoring.” This underlines the lack of reliability and, consequently, the need to make the relationship between the funder and nonprofits more reliable through stronger monitoring. More than a decade later, the situation has not changed considerably. Even today, by and large, “funders have influence, but not control [over nonprofits]” (Meredith 2009). Therefore, with our next set of contracts detailed in §5, we will consider monitoring by funders as a way to increase accountability in the nonprofit sector.

Before moving into the next section, we analyze a special case of the report-based contract, which is a benchmark model where efficiency types, \( \theta_i \), are observable. Although this perfect observation of efficiency types is unrealistic, it is for this reason the funder’s first-best solution as it eliminates the ineffectiveness due to the information asymmetry and allows for complete transparency.

\[ A_{FB}^{\theta} (\theta_i) = \frac{\phi_i}{\alpha} \theta_i^3, \]

where \( \phi_i = c_i / \alpha \).

We note that the funder allocates to all nonprofits with a nonzero efficiency \( \theta > 0 \). Each nonprofit’s allocation is based on his efficiency type; if the nonprofit is of a very low efficiency type, his funding is negligible. This situation of allocating money to every nonprofit may be unrealistic as allocating a few dollars would certainly not be worth the administrative costs. However, the funder may prescreen nonprofit applications, assigning \( c_i = 0 \) to nonprofits with too low of an efficiency report. Also the funder may want to have a minimum funding level where she only funds nonprofits whose allocations surpass this level, thus excluding some nonprofits from consideration.

5. Audit-Based Contracts

Whereas the first-best contract supposes the funder can observe all nonprofit efficiency types (complete transparency), the report-based contract results in the funder observing none of the nonprofit efficiency types (no transparency). Now we turn our attention to investigating contracts that will enable the funder to position herself anywhere between and including both the first-best and report-based contracts in terms of both observability of nonprofit efficiency types as well as contract performance. In this setting, the funder allocates her funds to the nonprofits based on their announced efficiency type. However, we know from the analysis of the report-based contract in §4 that the funder will not be able to differentiate; i.e., a nonprofit’s announcement of efficiency type is not reliable. In this auditing scenario, the funder can verify the nonprofit’s efficiency type by auditing after allocating funds and then impose a penalty upon nonprofits who misreport. Note that dealing with ex post auditing is different from dealing with ex ante observable efficiency types because
the funder must provide an incentive (here a penalty) to the nonprofits to observe their efficiency within the auditing framework.

The funder’s objective does not change from that formulated in Equation (7) in §4. In particular, an auditing cost is not included in the funder’s objective function because there are several situations where this would not be appropriate. For example, nonprofits may choose to self-audit using a credible third-party auditor. Also, funders may not have specific auditing costs or may have auditing costs and budgets separate from the contracts themselves. However, in §7.3 we extend this analysis to include auditing costs. The constraint set also continues to include the same IR and nonnegativity constraints found in Equations (8) and (10) of our original model. However, the IC constraint in Equation (9) is updated for the audit contracts. It now incorporates an imposed penalty, \( P_i \geq 0 \), for misreports, and is as follows for all \( i \) and for all \( \theta_i, \theta_i \in [0, 1] \):

\[
u^*_i(\theta_i | \theta_i) \geq u^*_i(\hat{\theta}_i | \theta_i) - PA_i(\hat{\theta}_i) \quad \text{(audit IC).} \tag{12}
\]

The left-hand side of the audit IC constraint is still the nonprofit utility when reporting truthfully. However, the right-hand side is the nonprofit utility from misreporting minus a penalty imposed on the nonprofit’s allocation. We saw in §4 that without this penalty modification, the nonprofit has no incentive to report truthfully. Consequently, the goal of this penalization for misreporting is to create an incentive for the nonprofit to report truthfully.

The funder has two levers she can use to guarantee nonprofit incentive compatibility: she can adjust allocation, \( A_i(\theta_i) \), or penalty, \( P_i \). Accordingly, she will offer contracts including both allocation and penalty, which take the form \((P_i, A_i(\theta_i))\). The best situation for the funder is to attain first-best allocations with incentive compatibility for all efficiency types, which can be achieved by the optimal penalty detailed in Equation (13). The corresponding optimal contract, which utilizes this optimal penalty and first-best allocations to achieve first-best performance and observability, is also outlined in Theorem 2.

**THEOREM 2.** The optimal audit-based contract takes the form \((P^*_i, A^*_i(\theta_i))\), where

\[
P^*_i = \delta_i + \frac{3}{4\delta_i \sqrt{3} \phi_i + 3} \tag{13}
\]

enforces incentive compatibility for all possible efficiency types of nonprofit \( i \) using first-best allocations, \( A^*_i(\theta_i) \), given by Equation (11).

The optimal audit-based contract enforces incentive compatibility by setting a high penalty \((P^*_i > \delta_i)\) for all types and decreasing the right-hand side of Equation (12). As the penalty \( P^*_i \) is greater than \( \delta_i \), each nonprofit will need to repay more than their valuation of what was originally granted to them in the event that they misreport. This arises because the nonprofit not only values the allocation but also the output he would get with that allocation; hence, he should be penalized accordingly. However, this penalty can also be thought of as, at least in part, a loss of future funds due to a loss of reputation. The optimal penalty increases in \( \delta_i \). As the nonprofit values allocation more, the funder penalizes more severely to enforce incentive compatibility. The optimal penalty intuitively decreases in \( \phi_i = c_i / \alpha \). As the funder values nonprofit output more, she penalizes less severely—again reflecting her increased regard for even the less-efficient nonprofit output. We summarize these observations in Corollary 2.

**COROLLARY 2.** The following are true for the incentive compatibility enforcing penalty, \( P^*_i \):

(i) \( P^*_i > \delta_i \);

(ii) \( P^*_i \) is increasing in \( \delta_i \); and

(iii) \( P^*_i \) is decreasing in \( \phi_i = c_i / \alpha \).

Although the optimal contract’s penalty, \( P^*_i \), enforces incentive compatibility for all types and achieves benchmark performance, it may have undesirable, if not unrealistic, characteristics. Under this contract, the funder is required to audit all types of nonprofits, which may create implementation problems. Furthermore, the funder imposes a penalty greater than \( \delta_i \) and thus requires the nonprofits to repay more than the value granted to them. Instead, the funder may want to specify a less severe penalty, \( P_i \leq P^*_i \), to implement and as a result may choose not to audit low types. Factors such as funder preference, policies, or auditing resources may dictate this penalty choice. Another factor might be auditing costs, which are further explored in §7.3. Utilizing a suboptimal penalty, \( P_i \leq P^*_i \), requires the funder to exercise her other lever of allocation to enforce incentive compatibility. Because the lower efficiency types have the greatest incentive to misreport, the funder must now compensate these types to prevent such misreports and guarantee incentive compatibility.

The audit-based contract with a specified penalty \( P_i \) allows a more liberal choice of penalty, \( 0 < P_i \leq P^*_i \), and uses the lever of allocation to enforce incentive compatibility. We define the efficiency type \( \theta^*_i(P_i) \) as a cutoff type for a given penalty \( P_i \) for which the incentive compatibility constraint in Equation (12) holds with equality when implementing first-best allocations; i.e., for \( \theta_i \geq \theta^*_i(P_i) \) the incentive compatibility constraint will hold and for \( \theta_i < \theta^*_i(P_i) \) it will not. The following theorem shows that a cutoff efficiency type, \( \theta^*_i(P_i) \), always exists and outlines the audit-based contract.
Theorem 3. The audit-based contract with a specified penalty \( P_i \) takes the form

\[
(P_i^A, A_i^A(\theta_i)) = \begin{cases} (P_i, A_i^{FB}(\theta_i(P_i))) & \text{for } \theta_i \leq \theta_i^*(P_i), \\ (P_i, A_i^{FB}(\theta_i)) & \text{for } \theta_i > \theta_i^*(P_i), \end{cases}
\]

where given a penalty \( 0 < P_i < \delta_i \), the cutoff efficiency type of nonprofit \( i \), \( \theta_i^*(P_i) \), is

\[
\theta_i^*(P_i) = \frac{1}{\eta_i}(\theta \eta_i^5 v_i^{1/3} - v_i^{1/2} + 8(\delta_i - P_i) \eta_i^3 \phi_i + \sqrt{\eta_i^6 (16(\delta_i - P_i)^2 \eta_i \phi_i - 27)})
\]

with \( \phi_i = c_i/\alpha_i \), \( \eta_i = 9 + 8\delta_i \phi_i \), \( v_i = 4(\delta_i - P_i) \eta_i^2 \phi_i + \sqrt{\eta_i^6 (16(\delta_i - P_i)^2 \eta_i \phi_i - 27)} \), and \( A_i^{FB}(\theta_i) \) given in Equation (11). For \( \delta_i \leq P_i \leq P_i^* \), \( \theta_i^*(P_i) \) does exist, but in a different closed-form expression than that of Equation (14).

Using her lever of allocation, the funder gives efficiency types below the cutoff \( \theta_i^*(P_i) \) a constant allocation to induce incentive compatibility whereas high types receive first-best allocation. Because of this constant allocation, the funder loses some utility and observability in over-allocating funds to nonprofits with low efficiency types. The extent of this loss is dictated by \( \theta_i^*(P_i) \) and ultimately the funder’s choice of penalty \( P_i \). The funder can recover utility by exploiting the lever of her penalty choice to position herself. Specifically, as she increases her penalty \( P_i \) (lowers \( \theta_i^*(P_i) \)), the funder increases the range of types she can observe as well as her utility. However, when the penalty is poorly chosen and insufficiently high, this contract can result in negative utility for the funder, which would indicate that she would not use it under the particular conditions. Also, credibility issues may arise when the funder chooses a penalty greater than one as this requires the nonprofit to repay more than the actual amount of their original allocation. Nonetheless, the audit-based contract with a specified penalty may still be a preferred option for funders as it does not over-penalize any nonprofits yet still rewards the higher efficiency nonprofits with first-best allocations. This preference and its associated conditions are further explored in Section 6. Following from Theorem 3, we can make two observations on the behavior of the cutoff type:

Corollary 3. The cutoff type, \( \theta_i^*(P_i) \), is

(i) decreasing in the penalty, \( P_i \); and

(ii) decreasing in \( \phi_i = c_i/\alpha_i \).

The first point is intuitive: As the funder imposes a more harsh penalty, this larger penalty enforces incentive compatibility for more nonprofit efficiency types. Thus, fewer types need to be cut off. Because \( \phi_i = c_i/\alpha_i \) measures the value the funder places on nonprofit \( i \)’s output compared to her outside opportunity, as the funder values nonprofit \( i \)’s output more, she cuts off fewer nonprofit efficiency types reflecting her increased regard for even the less-efficient nonprofit output.

In summary, the optimal audit-based contract employs a strict but optimal penalty, in fact penalizing all types to repay amounts beyond their original grant, which enables first-best allocations for all types. The audit-based contract with a specified penalty, on the other hand, allows the funder to specify a more lenient penalty. However, the funder pays for her leniency with a limited loss of observability and performance. The lenient penalty causes only efficiency types above a cutoff, \( \theta^*_i(P_i) \), to be incentive compatible. Thus, types below \( \theta^*_i(P_i) \) gain rent as they are allocated increasingly inefficient allocations the further they are below \( \theta^*_i(P_i) \). The choice of penalty does allow the funder to at least position herself in terms of observability and utility with a higher penalty moving her closer to the first-best situation by decreasing the cutoff \( \theta^*_i(P_i) \).

6. Effect of Contract Type on Performance

Using our theoretical results with numerical examples as illustration, we investigate the effect of each contract on the funder, the nonprofits, and the nonprofit sector overall. All results are in expectation, and thus, the results labeled as nonprofit are indeed the results for the population of nonprofits. We focus on the audit-based contract with a specified penalty (Theorem 3), because the optimal audit-based contract always achieves first-best. We use bold capital letters for vectors; i.e., \( P = (P_1, \ldots, P_N) \) and \( \Theta^*(P) = (\theta_1^*(P_1), \ldots, \theta_N^*(P_N)) \). We use superscripts \( FB, R, \) and \( A \) to denote the first-best contract, the report-based contract, and the audit-based contract, respectively.

6.1. Funder Expected Utilities

Examining the funder’s expected utility from Equation (7) reveals the following conclusions:

Proposition 1. Regarding contract effects on the funder’s expected utility, \( U_f \):

(i) \( U_f^{FB} \geq \max\{U_f^R, U_f^A\} \);

(ii) there exists a threshold \( \Theta \) such that \( U_f^A \geq U_f^R \) for \( \Theta^*(P) \in [\Theta, 1] \) and \( U_f^R \geq U_f^A \) for \( \Theta^*(P) \in (\Theta, 1] \); and

(iii) \( U_f^A \) is decreasing in the cutoff type, \( \Theta^*(P) \), and increasing in the penalty, \( P_i \).

These conclusions are illustrated in Figure 1, which shows the percentage of expected first-best (benchmark) utility captured by each contract and is plotted over \( c_i/\alpha_i \), which describes how the funder values the nonprofit output over the outside opportunity. In this figure, nonprofits are a symmetric population with \( \theta_i \sim U[0, 1], \delta_i = 1, c_i = c, \) and \( P_i = P \) for all \( i \). Note
that the results remain unchanged for heterogeneous nonprofits and varying type spaces.

In Figure 1(a) two regions can be identified for the funder. Although the first-best contract is always superior to the report-based contract and audit-based contract (Proposition 1(i)), the funder’s preference between the report-based contract and the audit-based contract depends on the value of \( c_i/\alpha \) and the choice of penalty (Proposition 1(ii)). When the penalty is too low (\( P = 0.5 \) in Figure 1), the funder is overly lenient on low efficiency types and she suffers from her leniency; independent of the value of \( c_i/\alpha \), the report-based contract performs better than the audit-based contract. However, for higher values of penalty, as long as \( c_i/\alpha \) is not too low, the funder prefers the audit-based contract to the report-based contract. For low values of \( c_i/\alpha \), the funder’s benefit from funding nonprofits is low compared to the return on the outside option; therefore, even small inefficiencies in funding (i.e., leniency toward the low efficiency types) make the funder relatively worse off. We also note that, consistent with Proposition 1(iii), the funder’s percentage of expected first-best captured and thus expected utility under the audit-based contract are increasing in the penalty, \( P \). Finally, we observe that for reasonably large values of \( c_i/\alpha \), the optimal penalty (Equation (13) in Theorem 2) is almost equal to \( \delta_i \). Though not illustrated here, Corollary 2(ii) analytically shows that the optimal penalty is increasing in \( \delta_i \); i.e., as the nonprofits value their allocation more, the funder must penalize more to maintain incentive compatibility.


Likewise, examining the nonprofit population’s expected utility derived from Equation (4),

\[
u_N = \sum_{i=1}^{N} \mathbb{E}\left[ \delta_i A_i + \frac{3}{2} (A_i \theta_i)^{2/3} \right]
\]

(nonprofit population utility), reveals the following conclusions:

**Proposition 2.** Regarding contract effects on the nonprofit population’s expected utility, \( u_N \):

(i) \( u_N^{A} \geq u_N^{R} \geq u_N^{R^*} \); and

(ii) \( u_N^{A} \) is increasing in the cutoff type, \( \Theta^*(\mathbf{P}) \), and decreasing in the penalty, \( \mathbf{P} \).

The audit-based contract gives a larger, constant allocation to the low efficiency types (i.e., types below the cutoff \( \theta^*(\mathbf{P}) \)) while also giving first-best allocations to high efficiency types (i.e., types above the cutoff \( \theta^*(\mathbf{P}) \)) as shown by Theorem 3. That is, compared to the first-best, the nonprofits are awarded higher allocations by the audit-based contract, which is why the percentage of first-best captured utilities greater than 100% can be observed with the audit-based contract in Figure 1(b). Consistent with Proposition 2(i), a very intuitive nonprofit preference for the audit-based contract is therefore clear. However, this preference is decreasing as the penalty increases (Proposition 2(ii)); that is, as the cutoff \( \theta^*(\mathbf{P}) \) decreases and more types are forced to report truthfully and are awarded first-best allocation instead of the larger, constant allocation.

An important conclusion drawn here is that high efficiency type nonprofits clearly prefer the audit-based contract over the report-based contract because they receive efficient allocations. In fact, these higher efficiency type nonprofits may want to work with funders on cultural changes that will increase auditing and thus increase utility for themselves. This can be seen at work in the nonprofit sector push to increase voluntary self-auditing and legislative efforts. Considering the government as a prominent funder, legislation aiming to elevate the standards of accountability and auditing are certainly evidence of funders seeking to create audit-based contract situations. An example of such legislation is the California Nonprofit
Integrity Act of 2004, which requires larger nonprofits to conduct objective, regular audits (California Government Code 2007).

6.3. Sector Expected Efficiency

Turning attention to the nonprofit sector as a whole, sector expected efficiency is defined as $\Pi = E[\sum A_i \theta_i / \sum A_i]$. This is the expected percentage of allocations going toward the production of output for the nonprofit sector as a whole, which is consistent with the definition of efficiency discussed in §1.

PROPOSITION 3. Regarding sector expected efficiency, $\Pi$, assuming a common distribution over nonprofit types, $\theta_i$, and $0 \leq P \leq P^*$, the following holds:

(i) $\Pi_{FB}^A \geq \Pi^A \geq \Pi_{FB}^R$;

(ii) as $\Theta^*(P) \rightarrow 0$, $\Pi^A \rightarrow \Pi_{FB}^A$ and as $\Theta^*(P) \rightarrow 1$, $\Pi^A \rightarrow \Pi^R$; and

(iii) $\Pi^A$ is decreasing in the cutoff type, $\Theta^*(P)$, and increasing in the penalty, $P$.

Figure 2 plots the expected sector efficiency against the cutoff $\theta^*(P)$. The figure illustrates that through the choice of penalty ($0 \leq P \leq P^*$) the funder can position the efficiency of the entire sector as stated in Proposition 3(ii). In fact, the sector can be positioned anywhere between and including both the first-best and report-based contract scenarios as made explicit by the equality possibilities in Proposition 3(i). Thus, the audit-based contract with penalties up to the optimal penalty, $P^*$, has significant potential to improve the efficiency of the sector overall. In this figure, nonprofits are a symmetric population with $\theta_i \sim U[0, 1], \delta_i = 1, c_i = c$, and $P_i = P$ for all $i$. Note that the results remain unchanged for heterogeneous nonprofits and varying type spaces.

In review, the funder prefers the audit-based contract under reasonable parameters and well-chosen penalties and can, in fact, achieve first-best performance. The nonprofit population strictly prefers the audit-based contract. In particular, high efficiency type nonprofits are clearly better off by the audit-based contract because they receive efficient allocations. They may want to work with funders on nonprofit sector cultural change to increase auditing, thus increasing their own utility. Through expected sector efficiency analysis, it was observed that the audit-based contract also has significant potential to improve the efficiency of the nonprofit sector overall. Because funders, nonprofits, and the sector may favor auditing under appropriate conditions, auditing and efficiency increases can and may become an industry effort.

7. Extensions

In this section, we study three extensions of our original model. We first extend the results of this paper to the constrained budget setting in §7.1. We then analyze uncertainty in measuring output in §7.2. Finally, in §7.3, we study a setting with audit costs.

7.1. The Budget Constrained Case

When funders face limited budgets, strategic allocation is even more critical. As stated by Brest and Harvey (2008, p. xiii), “whether you are giving away $100,000 or $1 billion a year, your funds are not unlimited, and a good strategy can multiply their impact many times over.” In this section, we extend our model to incorporate a budget constraint for the funder. Now the funder must allocate her resources to maximize her utility within the contractual framework considering both the nonprofits’ efficiency announcements and the constraint on her budget. The funder’s objective function from Equation (7) thus becomes

$$\max U_f = \max_{A_i(\theta)} \left\{ \sum_{i=1}^{N} E_{\theta} [2c_i(A_i(\theta)\theta_i)^{2/3} - \alpha A_i(\theta)] + \alpha B \right\},$$

with the constraints detailed in Equations (8)–(10) of our original model plus the additional budget constraint of

$$B - \sum_{i=1}^{N} A_i(\theta_i) \geq 0 \ \forall \theta_i \in [0, 1] \ \text{ (budget)}.$$

In the rest of this section, we will extend the results of this paper for the setting with a budget constraint. Note that for this extension we define a “constrained budget” as $\sum_{i=1}^{N} A_i = B$; otherwise the system effectively operates with an unlimited budget. We use superscripts $FBC, RC$, and $AC$ to denote the first-best contract, the report-based contract, and the audit-based contract, respectively, under a constrained budget.
7.1.1. Report-Based Contract. The problem formulation of the report-based contract follows the formulation of the problem outlined above and previously in §4 with the funder’s objective function found in Equation (15) and constraints detailed in Equations (8)–(10) and (16).

**Proposition 4.** The report-based contract under a constrained budget takes the form $A_i^{RC}(\Theta) = A_i^{RC} = (B/\Gamma^{RC})c_i^2E[\theta_i^{2/3}]$, where $\Gamma^{RC} = \sum_{i=1}^N c_i^2E[\theta_i^{2/3}]$.

This solution results in the same conclusions drawn from Theorem 1 where the funder is unable to differentiate among the nonprofits’ efficiency types and bases her allocation solely on her expectation of each nonprofit’s type. Thus, the report-based contract still offers the funder no observability of efficiency types and no operational transparency within this contracting context.

For this budget-constrained case, we again analyze the first-best, benchmark model where efficiency types, $\theta_i$, are observable. This analysis results in Corollary 4:

**Corollary 4.** The benchmark first-best contract under a constrained budget takes the form

$$A_i^{FBC}(\Theta) = \frac{B}{\Gamma^{FBC}}c_i^2\phi_i^2, \quad (17)$$

where $\Gamma^{FBC} = \sum_{i=1}^N c_i^2\phi_i^2$ and $i = 1, \ldots, N$.

We note that the funder still allocates to all nonprofits with a nonzero efficiency ($\theta_i > 0$) based on their efficiency types. Notice also that under a constrained budget, the allocation is dependent upon the efficiency types of all nonprofits. Therefore, each nonprofit introduces an externality to the others by his efficiency type. For instance, in the case of two nonprofits sharing the funder’s tight budget, the first-best allocation to one nonprofit decreases as the other nonprofit’s type increases.

7.1.2. Audit-Based Contracts. The auditing scenario remains consistent with §5 with the updated funder objective function (Equation (15)) and constraints (Equations (8), (10), (12), and (16)). In exploring audit contracts under a budget constraint, the discussion runs parallel to §5. As such, an optimal audit-based contract analogous to Theorem 2 as is follows:

**Proposition 5.** Under a constrained budget, the optimal audit-based contract still takes the form $(P_i^{C*}, A_i^{FBC}(\theta_i))$ where there exists a $P_i^{C*}$ that enforces incentive compatibility for all possible efficiency types of nonprofit $i$ under a constrained budget using first-best allocations, $A_i^{FBC}(\theta_i)$, given by Equation (17). For $P_i^{C*}$, the following are true:

(i) $P_i^{C*} > \delta_i$;
(ii) $P_i^{C*}$ is increasing in $\delta_i$;
(iii) as the budget, $B$, decreases, the penalty, $P_i^{C*}$, increases; and
(iv) as the number of nonprofits, $N$, goes to infinity, so does the constrained penalty, $P_i^{C*}$.

The optimal contract displayed in Theorem 2 itself does not change apart from its components as detailed in Corollary 4 and Proposition 5. Thus, the interpretation and conclusions in §5 still continue to hold. Just as in Corollary 2(i), Proposition 5(ii) indicates that all nonprofits must be over-penalized in order to enforce incentive compatibility for all types, implying that each nonprofit will need to repay more than their valuation of what was originally granted to them in the event that they misreport. Furthermore, this penalty should increase in $\delta_i$ (Proposition 5(iii)). The penalty intuitively decreases in $B$ (Proposition 5(iii)). When the funder’s budget is tight, she needs to manage her money more effectively, which dictates a higher penalty. Such a scenario can occur, for example, when the number of nonprofits is large (Proposition 5(iv)). Thus, budget certainly has a profound effect on funding scenarios and decisions. The sheer number of nonprofits applying for grants, let alone monetary awards, can constrain a funder’s budget.

Analogous to Theorem 3, Proposition 6 below gives the cutoff efficiency type, $\theta_i^{C*}(P_i)$, for the budget-constrained case. Because of the complexity of the expressions, we limit our analysis to penalties less than $\delta_i$:

**Proposition 6.** Under a constrained budget, the audit-based contract with a specified penalty $P_i$ still takes the form

$$(P_i, A_i^{FBC}(\theta_i)) = \begin{cases} (P_i, A_i^{FBC}(\theta_i^*(P_i))) & \text{for } \theta_i < \theta_i^*(P_i), \\ (P_i, A_i^{FBC}(\theta_i)) & \text{for } \theta_i \geq \theta_i^*(P_i), \end{cases}$$

where given a penalty $0 < P \leq (\delta_1, \ldots, \delta_N)$, the cutoff efficiency type, $\Theta_i^{C*}(P)$, is as follows:

$$\Theta_i^{C*}(P) = \left\{ \theta_1, \ldots, \theta_N \right\} \delta_i \frac{B_c^3\phi_i^2}{\Lambda} + \frac{3B^{2/3}c_i^2\phi_i^{2/3}}{2\Lambda^{2/3}}$$

$$= \frac{B_c^3(\delta_i - P_i)}{c_i^2} + \frac{3c_i^2B^{2/3}\phi_i^{2/3}}{2(c_i^2 + Y_i^{1/3})},$$

where

$$\Lambda = \sum_{i=1}^N c_i^2 \max\{\theta_i, \theta_i^*(P_i)\}^2$$

and

$$Y_i = \sum_{j \neq i} c_j^2 \max\{\theta_j, \theta_j^*(P_i)\}^2.$$

The audit-based contracts displayed in Theorems 2 and 3 can be extended to the budget-constrained case by replacing their components with
$A_{i}^{FBC}(\max\{\Theta, \Theta^C(P)\})$ from Equation (17) and Propositions 5 and 6, respectively, where max is taken component-wise. Similar to its counterpart in the unconstrained case, the cutoff type in Proposition 6 is still decreasing in penalty, $P_i$. As with both the first-best and optimal audit-based contract allocations, the other nonprofits’ types introduce an externality to nonprofit $i$ brought to bear in $\hat{\Theta}^C_i(P_i)$.

7.1.3. Effect of Contract Types on Performance. Analogous to §6, we use numerical examples to illustrate the effect of each contract on the funder, the nonprofits, and the nonprofit sector overall. All results are in expectation, and thus, the results labeled as “nonprofit” are again the results for the population in expectation, and thus, the results labeled as “non-profit” are again the results for the population

of non-profits. We focus on the audit-based contract, because the optimal audit contract always achieves first-best. For the numerical examples, we assume that nonprofits are a symmetric population with $\theta_i \sim U[0, 1]$, $\delta_i = 1$, $c_i = 25$, and $P_i = P$ for all $i$; $\alpha = 1$. Note that the results remain unchanged for heterogeneous nonprofits and varying type spaces.

Figure 3 includes a range of budget values. Consequently, as the budget increases, the situation moves from the constrained budget case to an effectively unconstrained budget case. Therefore, for high budget values, Figure 3 is comparable to Figure 1 in §6. Figure 3 shows that all our observations for the funder and the nonprofits from §6 continue to hold when the budget is constrained. We further observe that a tight budget improves the performance of both the report-based and the audit-based contracts. Even the first-best does not have much room to distribute a tight budget effectively; hence, all contracts perform similarly. As the budget increases, we observe that the performances of the contracts start to diverge. In particular, the performance of the report-based contract significantly deteriorates. Hence, the value of the audit-based contract increases as the budget increases.

All in all, incorporating a budget constraint into the basic model is relatively straightforward, though more computationally demanding. Though we find that the audit-based contract is most beneficial in the unconstrained budget cases, the intuitions gained in this analysis can help even smaller funders to best understand their contractual options.

7.2. Uncertainty in Production

The assumption that nonprofit production, that is, nonprofit output or outcome, can be measured unambiguously may be unrealistic in some situations, especially given the discussion in §§2 and 3 concerning the challenges of outcome measurement and evaluation in the nonprofit sector. This assumption is now relaxed through the following updated production function:

$$y_i = 2\xi_i \sqrt{\epsilon_i} \theta_i A_i$$ (uncertain nonprofit output production), \hspace{1cm} (18)

where $\xi_i$ is a random variable with support $(0, \infty)$ and mean $\xi_i > 0$. This random variable $\xi_i$ can be interpreted as an unmeasurable portion of the production, more specifically random variation in the production process, random variation or error in the accuracy of output measurement, or a combination of these. Under this new model, the updated model equations are as follows:

$$u_i = \delta_i A_i - \frac{1}{2}(A_i \theta_i)^{2/3} \xi_i^{2/3} (\xi_i - 4\xi_i) \hspace{0.5cm} \text{(nonprofit utility)},$$

$$y_i = 2(A_i \theta_i)^{2/3} \xi_i^{1/3} \xi_i \hspace{0.5cm} \text{(nonprofit output)},$$

$$U_i = \sum_{i=1}^{N} E_{\theta_i, \xi_i} [2c_i \xi_i (A_i \theta_i)^{2/3} \xi_i^{1/3} - \alpha A_i] \hspace{0.5cm} \text{(funder utility)},$$

where each nonprofit $i$ has maximized his expected utility. The following propositions result from analysis similar to that which led to the theorems and corollaries of §§4 and 5:

**Proposition 7.** Under uncertain production where the uncertainty is defined as in Equation (18), the
7.3. Auditing Costs

As briefly mentioned in §3, an auditing cost is not included in the funder’s objective function formulation because there are several situations where this would not be appropriate. For example, nonprofits may choose to self-audit using a credible third-party auditor. Also, funders may not have specific auditing costs or may have auditing costs and budgets separate from the contracts themselves. In this extension section, however, we explore the impact of including the cost of auditing in the funder’s objective function.

It can first be noted that such an inclusion does not affect the report-based or first-best contracts from §4 because these assume no auditing. For the audit-based contracts from §5, the funder’s utility function can be updated to include auditing costs as follows:

\[
U_f = \sum_{i=1}^{N} \left[ c_i y_i - \alpha A_i - \gamma I_i^A \right] = \sum_{i=1}^{N} \left[ 2 c_i (A_i \theta_i)^{2/3} - \alpha A_i - \gamma I_i^A \right].
\]

(19)

In this formulation, the cost of auditing an individual nonprofit is \( \gamma \). Because the funder may not choose or need to audit all nonprofits, \( I_i^A \) is the indicator function of auditing nonprofit \( i \). Whereas this formulation does not affect the form of the first-best allocations, it does affect the funder’s choices of which nonprofits to fund and which nonprofits to audit.

Note that the optimal audit-based contract detailed in Theorem 2 is no longer optimal when there is an auditing cost, as discussed later in this section. However, if the funder is committed to auditing all nonprofits that she funds, a similar contract can be characterized as follows:

**Proposition 9.** With auditing costs \( \gamma > 0 \), the optimal audit-based contract achieves full transparency

\[
(P_i^*, A_i(\theta_i)) = \begin{cases} (\tilde{P}_i^*, A_i^{FB}(\theta_i)) & \text{for } \theta_i \geq \tilde{\theta}_i, \\ (\tilde{P}_i^*, 0) & \text{otherwise}, \end{cases}
\]

where \( \tilde{P}_i^* = \delta_i + 9/(8 \phi_i) \), \( \tilde{\theta}_i = \sqrt{2 c_i / \gamma} \), \( \phi_i = c_i / \alpha \), and \( A_i^{FB}(\theta_i) \) is given in Equation (11).

Thus, the funder will only offer the optimal contract to certain efficiency types, namely high efficiency types that satisfy \( \theta_i \geq \tilde{\theta}_i \), so that types below this receive no allocation. However, the higher penalty, \( \tilde{P}_i^* > P_i^* \), ensures that all types will still remain incentive compatible, even those with zero allocation. As such, this contract enables complete transparency, but potentially at a high cost, because all nonprofits are audited.

Alternatively, the funder may choose not to audit some of the nonprofits that she funds and use a
contract corresponding to the audit-based contract detailed in Theorem 3 where only the high efficiency type nonprofits must be audited. Consequently, this revised audit-based contract with funder specified penalty is as follows:

**Proposition 10.** With auditing costs of \( \gamma > 0 \), there exists a \( \theta_i^* \) such that the audit-based contract with a funder specified penalty, \( 0 < P_i < \bar{P}_i \), takes the form

\[
(P_i, \bar{A}_i(P_i)) = \begin{cases} (P_i, A_i^{FB}((\bar{\theta}_i))) & \text{for } \theta_i < \max\{\theta_i^*(P_i), \bar{\theta}_i\}, \\ (P_i, A_i^{FB}((\theta_i))) & \text{otherwise}, \end{cases}
\]

where \( \bar{\theta}_i \in [\theta_i^*(P_i), \bar{\theta}_i] \) if \( \theta_i^*(P_i) < \bar{\theta}_i \) and \( \bar{\theta}_i = \theta_i^*(P_i) \) otherwise, \( \theta_i^*(P_i) \) is defined by Theorem 3, \( \bar{\theta}_i \) by Proposition 9, and \( A_i^{FB}((\theta_i)) \) by Equation (11).

It still remains that the funder is dedicated to funding low efficiency type nonprofits in this contract. They do not present any additional costs of auditing, but instead only the cost of their inefficiency manifest as reduced output. The main differences between the settings with and without audit costs are due to the constant payment and the cutoff type. If the audit cost is low, then \( \theta_i^*(P_i) \) would most likely be higher than \( \bar{\theta}_i \), and the two settings are equivalent. However, if the audit cost is high, i.e., \( \theta_i^*(P_i) < \bar{\theta}_i \), then the funder needs to increase the cutoff type to \( \bar{\theta}_i \), and consequently the constant allocation to guarantee incentive compatibility for moderate efficiency types, i.e., the types she prefers to award first-best allocation but cannot because of the audit cost. Once again, if the audit cost is high, there may be unfavorable situations where the funder gives a high constant allocation under this contract, which would decrease contract performance and transparency.

In the presence of auditing costs, identifying the optimal penalty is not as straightforward as in the case of §5 where there are no such costs. However, one can use Propositions 9 and 10 to search over possible penalty values for the value that yields the highest funder utility. For example, when \( \delta_i = 1 \), \( c_i = 2 \), \( \alpha = 1 \), and \( \gamma = 1 \), the funder would prefer a penalty of 0.9750. Notice that this penalty is strictly less than one, indicating that granting allocations to all types and only auditing some (Proposition 10) is better than granting allocations to some types and auditing all (Proposition 9). Many different reasons may motivate a funder to select specific penalty values as discussed in §5, including such consideration of auditing costs.

As evident in Propositions 9 and 10, for both the optimal and audit contracts, the funder’s decision of whether to fund (offer the contract to) each nonprofit hinges on the cost of auditing, \( \gamma \), relative to the value presented to the funder by the nonprofit’s output, \( c_i \). The more highly valued the nonprofit output, the more likely the nonprofit is to be funded. As expected, the performance of the audit-based contract deteriorates as the audit cost increases relative to the value that the funder can generate. When the audit cost is too high, simply relying on reports of nonprofit efficiency, i.e., the report-based contract, may be a better approach for the funder.

8. Discussion and Conclusion

To conclude, our analysis has revealed that report-based funding methods do not facilitate efficient allocation of funds as they do not result in operational transparency in nonprofits nor resolve the asymmetric information issue. Audit-based contracts with sensibly chosen parameters, on the other hand, can achieve both performance and transparency comparable to the first-best. Furthermore, both the nonprofit population and the sector overall benefit from auditing. Therefore, it may not be surprising to see auditing on the rise within the sector, especially in the wake of the Sarbanes–Oxley Act. The California Nonprofit Integrity Act of 2004 requires independent audits for large nonprofits (California Government Code 2007). The Panel on the Nonprofit Sector (2007) recommends nonprofits to have independent audits. Furthermore, organizations like Independent Sector and BoardSource are recommending nonprofits to self-audit, self-regulate, and take proactive actions to sustain trust and confidence (BoardSource and Independent Sector 2006).

Thus, this research offers new understanding of how funders can use auditing information in their contractual relationship with nonprofits, that is, how funders can structure their grants and/or contracts efficiently and effectively around the information provided by audits, including third-party or independent audits.

Furthermore, our research sheds light on the implications of auditing and its effective use for nonprofits, funders, and the sector overall. Indeed, our conclusions regarding both funder and nonprofit preference for auditing not only uncover funders’ potential position as change agents in the nonprofit sector, but also indicate a potential for collaboration to change nonprofit culture from merely heavy reporting to efficient monitoring. Such a culture change will increase the use of auditing, transparency, and, thus, efficiency for the sector overall. Carman’s (2009, p. 387) recent empirical findings support what we find theoretically, suggesting that “funders might [...] be able to change the way they make funding decisions [...] by explicitly using evaluation and performance information to make funding decisions.”

We do not make any restrictions regarding the funders or nonprofits that we target in this work; however, our models and results would be most applicable and beneficial to large funding organizations. Such funders have both the power and means...
to initiate change both in the nonprofits they fund and the sector overall. Furthermore, as our results show, these types of funders have the most to gain by moving from their current funding methods to emerging strategies, such as auditing. Where funders engage with a potentially growing and changing pool of nonprofits, requiring them to constantly learn more about the organizations they fund, our model should prove most effective. Regarding the nonprofits funded, our model does not discriminate; it is suitable for both nascent and more established nonprofits. Our paper is thus complementary to the current research in operations management on the subjects of emergency relief and humanitarian logistics (Thomas and Kopczak 2005, Tomasini and Van Wassenhove 2009), where recent work focuses on efficiency and performance measurement (Beamon and Balci 2008, Van der Laan et al. 2009). For funding decisions in this area, auditing would be an additional step of monitoring with compelling benefits.

Our analysis has several limitations. First, it is limited to theoretical and numerical analysis. The nature of contracting makes it difficult to obtain data, especially on nonprofit true efficiencies in the midst of their reports. This underscores the incentives necessary in the contracts to obtain truthful data. As such, empirically establishing the impact of these contracts would be very difficult. However, data could still be useful in providing insights and enhancing numerical analysis. Also, we do make certain theoretical assumptions regarding production and utility functions. These assumptions, detailed in §3, are some first attempts to model funding dynamics in the nonprofit sector. Whereas preliminary analysis shows that the higher-level qualitative insights and conclusions, particularly the answers to the initial three research questions, will stay intact even under less restrictive assumptions, the functional form assumptions made here allow closed-form, clear results, comparisons, and conclusions. For instance, under the assumptions of this paper, optimal contract forms can be explicitly stated, which allows for insights into optimal penalties and contract comparisons. Such straightforward clarity would be compromised under less restrictive assumptions. Admittedly, though, some results are less robust and follow from specific assumptions made. For example, in §7.2, the multiplicative nature of the uncertainty allows for seamless extension of the contracts to a setting with uncertainty. One should expect to have different observations under different forms of uncertainty, and such an extension of our results would be valuable. Lastly, although we used the most common definition and measure of efficiency in the nonprofit sector, this is a recognized limitation. A different measure of operational efficiency, such as outcome per unit of funds, could be insightful. Although such outcome measures might be more difficult to quantify in some contexts, when properly implemented, they are indeed true operational measures. Auditing by no means is limited to the efficiency definition used in this paper as other operational efficiencies can also be monitored through auditing.

There are several potential extensions for exploration. One such extension is the use of long-term or multiyear contracts. Although not commonplace in nonprofit funding, the use of a long-term contract may enable the funder to more precisely develop beliefs about the nonprofits’ efficiency types as well as incentivize the nonprofits through future allocations. Issues such as commitment, renegotiation, and breach of contract will need to be considered when analyzing these contracts. The current economic downturn provides new opportunities to evaluate inefficiency in the nonprofit sector. However, proper consideration should be given to the additional stresses these contracts may impose on already financially burdened nonprofits. In this research, we characterize efficiency as a type as opposed to a decision variable for the nonprofit because, although nonprofits can increase their efficiencies, these changes happen only in the long term. Thus, another potential extension is to consider the efficiency as a decision variable. Multiyear contracts can be designed to encourage such long-term improvements on efficiency. Yet a third potential extension is the use of signaling. Perhaps the nonprofit proposal can be a signal of the nonprofit’s type, enabling the funder to have more precise beliefs about the efficiency type of the nonprofit. A signaling game is a bit of a diversion from the line of contractual analysis studied here, but it may prove productive. Finally, the evaluation of reputation impacts and efficiency announcements with a dynamic model, along with long-term contracts, would be fruitful in shedding light on sector dynamics.

The metaphor of a nonprofit as a “seller” of services to a funder (Brest and Harvey 2008) captures much about their relationship and lends naturally to viewing grant agreements as contracts: Funders have goals and contract with nonprofits to perform the activities necessary to, in part, achieve these goals. As such, our nonprofit problem and a more traditional operations management problem are not so distant. Corporations manufacture products and services, and nonprofits produce outputs and outcomes; operations management has a crucial role in both settings. There also exist differences between nonprofit organizations and for-profit corporations. Highlighted in this research is that unlike a for-profit corporation that is presumed to distill the objectives of their shareholders, managers, employees, and clients into a single measure of shareholder value, a nonprofit
organization has numerous and diverse stakeholders including funders, staff, volunteers, beneficiaries, and the public. Therefore, a proper appreciation for the difficulty that nonprofits face in providing socially important services combined with sound understanding of operations management techniques provide excellent opportunities for cross-sector learnings. Possibilities abound for research in areas involving and benefiting nonprofit organizations. This study of efficient funding is just one of these many prospects.

Electronic Companion
An electronic companion to this paper is available on the Manufacturing & Service Operations Management website (http://msom.pubs.informs.org/ecompanion.html).

Acknowledgments
This research was partially supported by National Science Foundation Grant NSF/CAREER-0547021 and the National Science Foundation Graduate Research Fellowship Program. The authors express gratitude to the anonymous reviewers whose suggestions significantly improved the content and the exposition of the paper. The authors also thank Denise Gammal, Enis Kahyasi, and seminar participants at the 2006 INFORMS Conference in San Francisco and the 2007 INFORMS Conference in Seattle for valuable comments and suggestions.

References
Foundation Center. 2009. E-mail communication with Natalie Privett, September 29.
Kahyasi, E., F. Erhun, E. L. Plambeck. 2011. Delegation vs. control of component procurement under asymmetric cost information and simple contracts. Working paper, Department of Management Science and Engineering and Graduate School of Business, Stanford University, Stanford, CA.


