Design and Implementation of a Charitable Regulation Regime

Brian Galle

Georgetown University Law Center

Charitable organizations are often billed as private alternatives to the governmental provision of public goods (Weisbrod 1975). Yet in developed countries many aspects of charities are subject to government dictate, or at least oversight. Government rules may circumscribe, or at least incentivize, the organizational structure of the firm, the rights and obligations of its stakeholders, what activities it will engage in or not, how much it will spend or save, even to whom it will communicate and the content of those communications. In a federated system such as the United States, a single multi-state firm might be answerable to literally dozens of separate regulators.

Why is regulation of charity so pervasive? Is regulation justifiable from a perspective of economic theory? How can it be squared with the fundamentally private—that is, non-governmental—nature of charitable firms? This Chapter explores five major questions in the design and implementation of regimes for regulation of charity. The analysis is centered in transaction-cost economics, building off the more general discussion in Steinberg & Galle, this volume.

I first consider the bedrock issue of what, if anything, justifies the extensive modern role government regulation plays in the private nonprofit sector. In many respects the question is not particularly different for charitable firms than it is for commercial operations. Like other organizations, charities may experience agency costs and asymmetric information, and can generate externalities for one another and the general public. All these require some collective-action solution, if not necessarily a governmental one. Unlike many commercial firms, however, charities in many developed countries are subsidized by the state, and these subsidies offer additional reasons for public oversight.

The second and third sections are closely related, and examine from different perspectives the extent to which regulation of charity need be provided by government, rather than by private auditors or other monitors. The second section reviews the alternative of "voluntary regulation" or "self-regulation," which is to say the agreed-upon use of hired third-party monitors to ensure compliance with contractual or other promises by each firm, or with collectively-agreed on standards of behavior. In the third section, I evaluate arguments about whether private parties should be granted the right to sue charitable organizations to enforce compliance. In both sections I conclude that, while active government monitoring is likely essential to any effective regime, there also are openings for important contributions from private oversight.

The fourth section considers a recurring tension in public supervision of charities, namely the problem with which we began: how can charities represent a diverse array of private views when closely supervised by a possibly unsympathetic government? Courts and scholars of charity law tend to favor minimalist, bright-line, and procedure-based rules for charity governance, on the theory that these approaches reduce the room for bureaucratic discretion. I argue, to the contrary, that other institutional design choices can strike a better balance between safeguarding public interests and minimizing damage to the charitable sector.

Lastly, in the fifth section, I examine what little is known about charitable compliance with law. The section provides an overview of compliance theory and evidence in the context of commercial firms, as well as the limited evidence available for charity.

Why Government Regulation?

For many commentators, the charitable sector represents a potential alternative to government. Taxation and public provision, of course, is a standard solution to the public goods problem. But private organizations may also be able to elicit individual contributions for collective goals if they can reach a threshold of available resources, such that the appeal of being able to build on those resources overcomes free-riding problems. Organizations also may encourage contributions for the good of the group by using what Olson called "selective incentives," or private goods paid back to the contributor. Naming rights and other, more modest, acknowledgments provide prominent contemporary examples.

Many societies, however, may not be satisfied with the limited quantity and variety of public goods available through this combination of government services and unsubsidized private production of public goods. As Weisbrod (xx) explained, government subsidies can help to overcome free riding, and to facilitate public goods production for individuals who do not share the preferences of the median voter. The story can be told in more nuanced form (see Galle & Steinberg, this volume, for a more complete account), but this basic point captures the essence of most subsidy theories. A similar view, grounded in political rather than economic theory, sees the charitable sector as a source of diverse voices and viewpoints that would otherwise go unheard in majoritarian political processes.

Public and private views of the charitable sector imply somewhat distinctive approaches to any possible government regulation. Perhaps counter-intuitively, however, there are strong arguments for government regulation under either approach. Further, in one hybrid view, government enforcement efforts may serve as a source of subsidy for otherwise supported private entities.

Some basic regulation is straightforward in a formal subsidy regime. Public funds are socially costly to raise. Further, because the goal of most subsidies regimes is to achieve goals that would not be chosen by the median voter and her elected officials, most charity support regimes rely on the judgment of donors, not government officials, to allocate funds. This opens the possibility for abuses of the support regime. Most obviously, donors and managers may be tempted to claim government support for the pursuit of their own private ends, rather than for the production of public goods. Government rules to prevent misallocation of subsidized funds is therefore a necessity for any well-functioning charity support regime.

A subsidy regime may also justify additional forms of regulation. In most cases these additional steps involve at least some degree of tradeoff between improved outcomes (by whatever measure) and a possible reduction in charitable autonomy. Reducing the autonomy of charity managers brings concomitant risks to the independent judgment that motivates most support regimes. Careful design of the enforcement regime, however, can help to minimize risks that officials will use their enforcement power disproportionately to constrain organizations whose views they do not share. I return to these design considerations later in the Chapter.

A straightforward and relatively value-neutral way to improve the productivity of charitable dollars is to minimize agency costs within each firm. Galle & Steinberg, this volume, provide a more detailed account of the many sources of agency slack that charity is prone to. Crucially, the quality of charitable outputs is very difficult for supporters to monitor, creating openings for opportunistic behavior by managers. Charitable beneficiaries will typically suffer the burden of agency costs of these kinds, both through diversion of resources from their most productive charitable use as well as through lower support from donors who doubt the utility of their donated time or dollars. In rare cases, it might be argued that agency costs are welfare-improving, such as in the instance when managers are more public-spirited than donors. But even in these instances, there is a potential loss in donor support when rules allow for greater agency slack (see Galle unpublished for evidence in the private foundation context). Thus, if government subsidies are motivated by concern for the welfare of beneficiaries, regulation to reduce agency costs offers important benefits.

Externalities provide a third well-grounded basis for regulation. Just as standard economic theory recommends government intervention to support the production of positive externalities, so too it counsels regulation of negative externalities. In the charitable context, asking regulators to identify "negative" externalities raises some potential tension with a regime that aims for government neutrality about charitable goals and methods. In the United States, for example, the Supreme Court has struggled with whether racially discriminatory organizations should be allowed to qualify for government subsidies.

There may be some forms of externality that all or almost all parties would agree are socially undesirable, however. Charities depend on trust and goodwill from the general public. The public may form views about the trustworthiness of the charitable sector based in part on overall impressions or news reports, rather than solely on the performance of each individual firm (for evidence, see []). Newsworthy misdeeds, then, create reputational spillovers for other firms, suggesting that each individual firm under-invests in efforts to mitigate misuse of donor or other customer funds.

Similarly, a strong anti-fraud or other consumer-protection regime for fundraising efforts likely mitigates serious potential externalities. Fundraising protections closely resemble modern regimes of securities regulation, in which robust markets depend on mechanisms to prevent the development of a "market for lemons." (Akerlof 1972; for application to securities regulation, see []). That is, given that investors or donors cannot observe easily the quality of what they are buying, a market in which fraud is pervasive would likely unravel. Investors would not offer market value for an uncertain proposition, and honest entrepreneurs would not sell at the discounted price investors would pay. That would leave a disproportionate share of dishonest entrepreneurs, further depressing the prices entrepreneurs would pay, and so on. Since honest firms cannot credibly distinguish themselves, dishonest firms impose negative reputational externalities on the entire market.

Although there is some scholarly dispute on this point in the law and finance literature, many of these same arguments for regulation could be made even in the absence of any express government subsidy. Rules to minimize agency costs and reduce the danger of a market for lemons greatly increase the expected value of all of the affected enterprises. This is a point of consensus. More controversial is the question whether these rules must be supplied by government, or whether they could instead be arranged between private parties by contract. Even in the case of externalities, it is argued, firms could contract with a third-party private regulator to ensure mutual compliance with rules of mutual benefit. This question is considered in more depth in the [next] section.

Government Regulation or/and "Self-Regulation"

Governments are not the only source of regulation for charitable activity. Nonprofits around the world are increasingly participating in various forms of "voluntary regulation" or "self-regulation." (For overviews of the trend see Bowman 2010, Gugerty 2010). In these regimes, individual firms agree to be subject to a set of rules or guidelines established by a third party, often itself organized by collective agreement among some set of the member firms. A similar structure is the "certification" approach, pioneered by commercial firms whose product is subject to consumer tastes for difficult-to-monitor inputs, such as "conflict-free" diamonds and "no sweat" clothing. Certification regimes use an outside monitor, such as an auditing firm, to inspect the certification-seeking firm and verify that it is meeting the standards required for certification. Malani & Posner (2008) propose that certification processes can replace U.S. government enforcement of most rules for charitable enterprises, and certification has also emerged as the central concept behind "social enterprise" or hybrid charitable/for-profit organizations.

Credibility poses the central challenge for any system of voluntary regulation. Again, charitable outputs are difficult to monitor for quality. Firm outsiders also typically cannot observe how insiders prioritize across available goals and projects, nor can they observe whether managers are working hard or "shirking." Firms that seek outside support, then, must find ways to convince supporters that insiders will not stray too far from the supporters' preferences (Hansmann 1980). Reputation alone will not credibly constrain managers, because casual supporters cannot monitor closely enough to impose meaningful reputational sanctions, and because managers may have (unobservable) short time horizons that lead them to sacrifice the firm's future reputation for present gains.

Third-party monitoring does not necessarily resolve the credibility problem. In any instance where the firm selects its own monitor, there is the possibility the firm will choose a lax auditor. Firms may also buy off auditors to obtain favorable ratings. As a result, the same credibility problem that confronts the individual firms also confronts the monitors themselves. Indeed, in some commercial certification markets there is now a set of meta-certification firms, whose task it is to monitor the monitors.

As Gugerty & Prakash (2010) show, third-party monitoring organizations face difficult and conflicting incentives. On the one hand, the value of obtaining certification or membership depends on the strength of the signal sent by the monitor's approval. These signals usually depend on both careful auditing of the member firms as well as serious and credible "swords" or penalties for failure to meet certification or membership requirements. At the same time, each monitor faces what economists might term a "participation constraint": the higher the possible costs of employing the monitor, the fewer firms will do so. The price of entry must be set low enough that the monitor can attract enough clients to cover its fixed costs, and, in all likelihood, to gain enough members to acquire public visibility and respect. This implies that the signaling benefits of participation in any voluntary third-party monitoring regime are limited by the participation constraint of the marginal member.

Some clubs may be able to overcome the participation constraint through the use of selective incentives, but these are themselves costly. Suppose that there are gains from expanding the club's membership, such as increased public visibility and trust, but that marginal gains from adding new members diminish rapidly after some threshold of public acceptance is reached. Clubs may pay new members to join until they exceed this critical threshold. At that point, additional members are likely to want to join in order to share in the network externalities, such as reputation, provided by other members, and quitting is costly because it requires the exiting firm to relinquish these same benefits. Once this stage is reached, the participation constraint of marginal firms is less binding, allowing for greater enforcement. Reaching this point, however, likely requires very costly investments in enticing firms to join and not to abstain from violating the rules of the club. These investments may well be wasted if the club fails to reach the critical threshold.

Voluntary regulation regimes may also struggle to impose credible penalties because of potential spillovers across members. Enforcement actions against one member of the compliance club may raise the salience of the potential for misdeeds by other members (Gugerty & Prakash 2010). For example, consumers may form impressions of groups of similar firms, rather than of individual firms. Thus, paradoxically, if the monitor brings any enforcement action for noncompliance, it risks diminishing the signaling value of membership. Outside observers, if they realize this dynamic, will not find the signal of voluntary membership credible. Monitors could perhaps pledge secret enforcement, but that, too, raises credibility problems, as well as the practical obstacle that enforcement targets might threaten to go public.

Lastly, the credibility problem seems especially acute in the case of voluntary regulation of externalities that extend beyond the member organizations. A voluntary regulation club has obvious incentives for the members to monitor one another for mutual injuries. But why would the club monitor itself for spillovers that burden mostly non-members? In the commercial context, the answer seems to be mostly that consumers have a distaste for certain forms of harmful conduct. In other instances, industries conclude that "self-regulation" may help to ward off costlier government oversight. Both these scenarios suggest that the monitor has especially strong reason to avoid public disclosure of non-compliance among member firms, at least if there is any significant reputational spillover that results from disclosure.

These various limitations on voluntary regulation establish a strong case for at least some minimal level of government enforcement, especially for regulation of externalities. Mandatory assignment to the government monitor greatly reduces the participation constraint problem, although there remains the threat that excessive compliance costs could discourage formation of charitable firms. Government regulation does not require payments to achieve a critical threshold for network effects, since it simply compels all firms to participate. Further, since the government monitor does not directly internalize the cost of reputational spillovers, enforcement activity by the government is more credible. In addition, there is evidence that individuals comply with legal rules because they believe in the legitimacy of law (e.g., Tyler 1990), while no such phenomenon has been documented for voluntary regulation regimes.

It might be argued that the differences between government and voluntary enforcement are overstated in the account I just offered. Conceivably, it might be possible to design a private enforcement regime in which individual firms do not have a choice of their monitor, although I am unaware of any such scheme existing in the real world. Firms could instead be randomly assigned to one of a group of third-party monitors. Participation in the random-assignment system would likely have to be mandatory rather than voluntary, otherwise the credibility problem and incentives simply shift one level up, but this could be overcome if participation were legally required or made a condition for certain benefits (see Galle 2017b for more full-fledged development of a similar proposal in the "social enterprise" context). At the same time, in a federated system such as the United States, firms can in effect shop for their own monitor by choosing which set of state laws, and which state Attorney General's office, they will operate under. But U.S. law limits this forum-shopping to some extent by requiring firms to be bound by the law of their physical location (Fremont-Smith 2004), rather than an unbounded choice of any of the available fifty state systems.

In any event, I describe the case for governmental regulation as perhaps establishing grounds only for "minimal" regulation because it conceivable that both government and voluntary regulation work best in tandem. Voluntary regulation is inherently constrained, but may nonetheless be effective within certain bounds. It can therefore serve as supplement to government standards for firms and industries with preferences for greater enforcement than the median voter would choose. That is, just as charity provides private supplements to other forms of public goods, voluntary charity enforcement can supplement public enforcement. Voluntary enforcement may also facilitate variations in enforcement rules and intensity across industries, if for political economy reasons governments find it difficult to set different rules for different sub-fields.

Choosing how intensive government regulation should be then depends on how varied is the need for regulation across the charitable sector and how effective voluntary regulation proves to be. Government regulators may want to deliberately set standards less stringent than the median sub-field would demand, to allow space for voluntary regulation to achieve optimal standards for a wider set of sub-fields. On the other hand, depending on how limited voluntary regulation proves to be, a government "floor" set too low might leave the combined regulation regime still below optimal for sub-fields with the most acute need for signals of trustworthiness.

Government rules need not be floors, however. A common hybrid approach is for government to set out default rules, but to allow individual firms and their stakeholders to contract around the default term. For example, in the United States, state law establishes duties of care and loyalty for firm managers, but most of these duties can be modified by the organization's governing documents. This approach allows flexibility in the case of firms with relatively small number of supporters who expect to be able to monitor the firm closely, and therefore for whom the credibility of the governance regime is less central. Again, however, government likely will want to make mandatory rules that guard against externalities.

A final rationale for government over private regulation is simply as a form of subsidy. Drafting effective contract terms is time-consuming and costly, and monitoring compliance with them may be even moreso. There can be large fixed costs to negotiating entry into a private regulation regime, which could make such these approaches impractical for small firms. Government regulation, then, helps to distribute the cost of a trustworthy charitable sector to all taxpayers, rather than leaving it to weigh on charitable beneficiaries or their supporters. Government subsidies may also be combined with private standard setting—for instance, government could pay to incentive firms to join standard-setting organizations, helping them to reach the critical threshold at which they are self-sustaining.

Private Enforcement: "Standing to Sue"

Up until this point I have treated regulation as a unified concept that includes both the making of governing rules for the firm as well as the authority to monitor for and enforce compliance with those rules. These tasks, however, need not be bundled together. Governments could allow firms to set their own rules, but assign government personnel to enforce compliance. Alternately, governments could authorize private actors to monitor compliance with government-made rules. Both these hybrids are common. State attorneys general in the United States may bring suit to enforce against an organization its own bespoke organizational documents. And in some instances private actors can commence judicial proceedings against a firm or its insiders to compel obedience to law. However, "standing" or authorization to sue is still a minority position, and the subject of ongoing controversy. This section summarizes the contours of the debate.

The modern case for donor standing traces to the claim in Hansmann (1981) that private suits help to reduce agency costs between the firm and its supporters. Hansmann argued that the nondistribution constraint, although it reduced agency costs, still left open manager other possible sources of agency slack. He therefore urged that state law should permit major donors, at least, to sue the firm or its managers. Other leading commentators have argued for a right or "standing" to sue on similar agency-cost grounds (Langbein 1995, Sitkoff 2004).

Nonetheless, relatively few U.S. firms are at this time subject to suits by donors. Before 2002 there were no more than four states that authorized such suits. The Uniform Trust Code, first promulgated late in 2000, permits settlors of charitable trusts to sue the trustees. About half the states have adopted the UTC, or at least that piece of it that grants charitable donors standing, but only a small fraction of all charities are organized as trusts. Several courts have rejected efforts to extend the UTC's position to charitable corporations. Similarly, although many large donors now negotiate "contracts" that purport to allow individual suits against the firm (Brody 2007, Dale et al. in press), these contracts have not been recognized by U.S. courts.

Support for donor standing is hardly unanimous. In the commercial firm context, many commentators tend to oppose donor standing because of the availability of putatively less costly instruments for constraining managers. Commentators tend to reject stakeholder suits for commercial firms because of the availability of control mechanisms, such as incentive pay and the market for corporate control (see Becht, Bolton, & Roell 2003 for a review), that nonprofits cannot access. Further, while donors could in theory use staged financing to discipline managers (Schizer 2008), that tactic carries a heavy tax penalty (see Galle 2016 for more extended discussion). Thus, the case for donor suits seems much stronger than arguments for shareholder litigation.

On the other hand, there are a number of features of donor standing that raise potential concerns. Nonprofit managers are sometimes said to be attracted to their trade by a sense of mission or ideological commitment, and so may have little inclination to waste the scarce resources at their disposal (Roomkin & Weisbrod 1999, Rose-Ackerman 1996). Some commentators go so far as to suggest that being subject to suit would only lead managers to avoid desirable risk-taking or to compile wasteful paper trails, discouraging donations (Fischel & Bradley 1986 and Bainbridge 2002 tell this story in the for-profit context). Many courts have warned of the danger of time- and money-wasting nuisance suits (Fishman 1985). Reading between the lines of these decisions, one also senses a view that in some cases charitable assets, once parted from formal ownership by the donor, are better used (from a social point of view) if the donor retains only imperfect control. Indeed, severing ongoing control by a donor's family is a key condition of federal tax benefits for charitable bequests.

As with a number of other rules of nonprofit governance, then, standing to sue represents a balance between competing positions whose respective costs and benefits are not well studied or understood (Posner 2014). Granting the right to sue may reduce agency slack between donor and manager, and thereby probably encourage more gifts overall. But it may also contribute to needless administrative and litigation costs, and it is possible that managers would make better, or at least more public-regarding, decisions with less donor oversight. It is likely, as Atkinson (1988), Sitkoff (2004) and Brody (2007) observe, that these factors will balance out differently in different contexts.

One important dimension on which standing's impact likely varies is in the size of a firm's group of donors or other supporters. We might expect that free riding among a large class of potential litigants would make donor standing relatively ineffective as a means of controlling agency costs (Manne 1999, Sitkoff 2004). Donor standing could also be costlier with many donors. As the group of outsiders empowered to sue the firm increases, managers may tend to operate more defensively and conservatively (Bainbridge 2002), and donors may incur added costs of cross-monitoring donors with inconsistent goals (Triantis 2004). While some donors may value increased control over the firm, suits may also bring to light negative information that would decrease gifts from other donors (Fishman 1985).

At present, however, there are few data that speak to any of these factors. Researchers have found evidence of agency slack at least in very large nonprofit institutions, such as colleges and universities (see Galle & Walker 2014 for a survey of the field). Fisman & Hubbard (2005) and Desai & Yetman (2015) find cross-sectional correlations between state law and measures of firm agency costs, but it is not clear whether these results would be robust to methods that include state and firm fixed effects.

Galle (unpublished) examines the effect of changes in U.S. donor-standing law on donations and "overhead" expenses among private philanthropic organizations. I find that on average legal changes that expanded the rights of donors to sue were correlated with considerable increases in donations, as well as reduction in the share of firm resources spent on administrative costs. I interpret this to suggest that donor standing reduces agency costs, with concomitant increases in donor enthusiasm. It is not clear, however, that these shifts improve social welfare. For example, it is possible that donors do not value administrative costs even where those costs improve the quality of the firm's outputs. Further, I find no or contrary evidence of the impact of standing among the set of firms with a wide base of donor support, although this difference may be due to data limitations rather than a real difference in behavior.

Nonetheless, I argue that expanded donor standing deserves serious consideration because it represents one of the few effective mechanisms for nonprofit oversight. This is not to say that I advocate a one-size-fits-all approach. Again, evidence also suggests that standing may be more costly or less efficacious for large charities with many supporters. Lawmakers may wish to consider tailored standing rules that maximize standing's benefits while minimizing its expected costs, as advocated by Brody (2007), among others.

Another approach to the question of private standing is to consider its relationship to the alternative of government monitoring. There is a rich literature, in the United States especially, on the interrelationship between private and public enforcement generally. While there is of course much nuance, the central theme is again one of tradeoffs. Private suits may make it difficult for government regulators to strike what they view as the proper level of enforcement intensity, potentially resulting in over-deterrence, enforcement against those whom the government views as not serious wrongdoers, or socially wasteful "me too" suits (e.g., Pierce []). The other side of the coin is that government regulators may be "captured" or overly influenced by those who should be subject to regulation, so that private suits help to prevent corruption and rent-seeking. Private litigation also may expend the capacity of resource-constrained legislators.

How these rival considerations play out in a particular legal context depends on the interplay between the bureaucratic structure of the enforcement entity and the political economy of demand for public enforcement. Consider charitable enforcement at the sub-national level in the United States. Attorneys general hold primary enforcement responsibility for state nonprofit law. For the most part, the social benefit of AG charity supervision commands little attention except in the case of major scandal, and provides small benefits spread widely across the general population. At the same time, regulated firms are a relatively small and cohesive group that may well be willing to incur costs to minimize regulatory burdens (St. Clair 2016, for instance, finds evidence that firms manage reported revenues to avoid costly state reporting requirements). Standard public choice theory would predict minimal political enthusiasm for funding nonprofit oversight under these circumstances. Indeed, most AG offices have fewer than one attorney for every thousand charities in their portfolio (Galle 2017). Attorneys general are also elected officials, making them especially vulnerable to lobbying by targeted firms.

In short, state nonprofit oversight seems to feature many of the factors that weigh in favor of private enforcement. Based on considerations similar to these, courts in some states have allowed suits by "beneficiaries" of charitable organizations, although without any great consistency in the definition of beneficiary. Typically, a private litigant must possess some unique interest distinct from that of the general public in order to obtain beneficiary standing. Sympathetic plaintiffs in instances where state attorneys general have been disinterested in enforcement, however, have sometimes found their way into court. While the political economy of federal enforcement is similar, and probably could similarly justify some role for private enforcement, the U.S. Supreme Court generally bars private actors from suing other private actors to enforce federal law, unless some statute specifically authorizes the suit; none does.

Why limit standing only to beneficiaries with distinctive interests? Given the likelihood of free riding among the general public, one might think that open-ended standing would be unlikely to result in any significant burden for most charities. Consider, however, that one goal of public support for charity is to encourage the production of relatively unpopular activity. Standing for all might invite vexatious litigation intended to curtail a controversial charity's activities or drain its resources. And, again, the wider the class of available litigants, the more opportunities for simply wasteful lawsuits, and perhaps the greater the need for supporters to cross-monitor those who might use lawsuits to influence the firm.

A compromise position might be to allow private suits upon approval from the relevant regulator, as in the proposals from some commentators to allow so-called "relator" actions (Chisolm 1995). As Galle (2003) observes, making private suit subject to administrative approval removes most of the downside risks of such suits. At the same time, of course, conditioning suits on the public enforcer's acquiescence greatly reduces the likelihood of private suits in instances where the enforcer is captured by the lawbreaker. Still, a conditional approach allows the AG to decide how best to balance the other competing factors. One can see something like this approach in some judicial decisions; courts are much more likely to permit beneficiary standing if the AG has taken no action.

Reconciling Government Discretion With Charitable Independence

Once again, charity law generally attempts to promote a diverse array of perspectives in the production of social policy. Many charities are popular across the political spectrum. Others, though, may pursue goals to which the average voter is indifferent or outright unfriendly. Nonprofit scholars argue that official discretion over charities should be highly constrained, in order to prevent officials from using their authority to shape charity to their own liking.

As I suggested earlier, these views are in some tension with many forms of government regulation. Even basic judgments, such as whether a given project actually serves enough of a useful public purpose to justify receiving the government's subsidy, quickly becomes entangled with officials' judgments about what to value. It may be challenging to design a system capable of making fine distinctions between different organizational goals without permitting space for government judgment about the desirability of those goals. Until the 1970's, for instance, the Internal Revenue Service routinely denied charitable status for environmental rights organizations. It was the Service's position that environmental advocacy served no useful social purpose.

Official discretion poses a challenge in other line-drawing scenarios. Political activity offers an especially difficult example. Information is a public good, especially information about the functioning of the polity, so that charities that inform the public sit very close to the core rationale for subsidizing charity. At the same time, information can also be used for propaganda or advocacy, either through falsehood or deliberate omission of contending facts and viewpoints. Further, many charities claim that engaged participation in the political process is an important feature of their mission. All these factors suggest reasons for caution in assigning line-drawing responsibility to officials who may have their own partisan commitments.

Current law in the United States reflects these tensions. In deciding what is charitable, officials cannot deny an organization charitable status because they dislike its goals, but instead can determine only whether the entity performs a function that the for-profit market does not. For example, rules for distinguishing permissible "educational" activity from prohibited lobbying at least purport to test only how the organization spreads its message, not the message's content. Although there is an exception to the definition of charity for violations of “public policy,” the Supreme Court has interpreted that exception to refer only to policies on which there is near-uniform social agreement. Thus far, the only disqualifying policies that the Court has identified are invidious racial discrimination and criminal enterprises; even charities that discriminate based on gender remain in the clear. The Court has suggested that the reason the public policy exception must remain narrow is because government officials cannot be trusted to decide which charitable endeavors are good policy.

This minimalist, highly proceduralized approach may be overly cautious. The harm of excluding the occasional borderline group does not obviously outweigh the benefits of curtailing wasteful subsidies. The social cost of asking regulators to withhold judgment may in some cases be more dramatic than simply waste. In the political context, opening the government treasury to support for political advocacy could transform the political landscape, offering in effect unlimited government matching grants for partisan activity, with predictably disproportionate gains for contributors with greater resources.

Furthermore, governmental review often improves, rather than threatens, charity. Organizational psychology suggests that groups perform better when they must engage with contending viewpoints. The need to justify charity to such officials may in itself lead to more internal deliberation within the charitable organization about the place of its goals within the larger society. For instance, an organization that excludes women should be made to justify its decision against the expressive harms that even its private decision may have on others. The very act of considering and offering an explanation that could be plausible to outsiders may begin to change the discriminatory attitudes of the organization's members.

More searching review can have other benefits as well. Excluding harmful and marginal charities is likely to boost the willingness of the public and government officials to support the sector. Approval by officials who are accountable to the public at large helps to ensure that charitable goals are balanced against the possibility of externalities or consumption of scarce resources that might be better devoted to other charities. Surviving a formal review can also help the organization to cement its case to a skeptical public.

On the other side of the ledger, wise institutional design can limit the risks of official discretion. Modern administrative law is built to constrain the individual impulses of administrators and to guide them into public-minded and scientific deliberation. Reporting and analysis obligations, making line enforcers answerable to other bureaucratic actors, and to judicial and legislative monitors, oblige the executive to justify its decisions based on public reasons, taking into account all reasonable competing viewpoints. To be sure, the effectiveness of judicial or other forms of review depend on the reviewers. Administrators may blatantly discriminate against some charitable enterprises, and courts at times may let them.

But stare decisis and norms of duty to follow it will help constrain judges who might be tempted to give a pass to harsh administrative treatment of organizations of which they disapprove. Agencies also must anticipate that judges may adhere to law rather than personal bias, and so it is in the agency's interest to structure its own internal deliberations in ways that give effect to contending viewpoints.

These bureaucratic checks may be hard to replicate in legal systems that rely mostly on courts to resolve disputes about charity. Silber (2001), for example, offers a historical account of New York courts' dismissive treatment of womens' rights organizations.

Lawmakers can also constrain the downside risks of biased judgments by their choice of enforcement mechanisms. For example, by making the penalties for non-compliance continuous, rather than discrete—having graduated penalties, instead of cliffs—the result is that in close cases, the consequences of a mistaken or biased judgment will be relatively modest (see Kaplow & Shavell 1996 for more discussion). That is, charitable missteps might be subject to small fines, rather than revocation of all the charity's favorable status. Lawmakers may also substitute ex ante screening tools for ex post penalties. Galle (2011) proposes that firms that present greater risks of violations—for instance, those that plan to engage in conduct that is close to violating rules against discrimination or partisan activity—can be required to incur more costs in justifying their eligibility for subsidies.

Even aside from these design options, any politicization of charity would not necessarily doom unpopular projects. Public choice theory predicts that an extremist charity that fails to win a subsidy is still quite likely to continue its mission precisely because of the extremity of its positions. Individuals with views that are distant outliers from the median social point of view know that they cannot free ride on the efforts of others; thus, each member is likely to contribute, regardless of subsidy. Each member will lobby more vigorously for political approval. Even if approval is not forthcoming, the government's refusal further marks the charity as an outsider, confirming to its members their inability to rely on other contributors and strengthening their own resolve to contribute.

Charitable Compliance with Law

Researchers have relatively little direct information about charitable compliance with legal rules. A long series of commentators have suggested that nonprofit compliance with law is in "crisis" (Eisenberg 1997:333) or otherwise minimal, and call regularly for fundamental rethinking of how law can be made a meaningful constraint on nonprofit behavior (Brakman Reiser 2005, Gary 1999, Gugerty & Prakash 2010, Manne 1999, Sugin 2015). These analyses typically draw on anecdotes derived from press accounts of misdeeds, along with certain stylized facts about the compliance and enforcement environment. A slightly wider lens, taking in theories of compliance among firms of all kinds, may offer a somewhat more optimistic assessment.

Nonprofit compliance skeptics draw largely on a model in which formal deterrence is a major driver of compliance. For example, Potoski & Prakash (2011), writing about the effectiveness of nonprofit-sector self-regulation, aver that "Effective ... programs must have monitoring and enforcement mechanisms to curb shirking" (see also Potoski & Prakash 2005 for evidence to that effect). As discussed earlier, the political economy of charity enforcement results in relatively modest enforcement effort. In the U.S., many aspects of nonprofit law, especially in the case of state law, offer no meaningful likelihood of sanction for noncompliance (Fremont-Smith 2004:2). Few potential litigants have legal authority, or "standing," to enforce state law against a noncompliant charity (Manne 1999, Galle unpublished). Due to resource constraints and political disinterest, enforcement by state or federal officials is typically modest given the size of the sector. It is natural, therefore, to predict similarly modest degree of compliance.

In the commercial firm context, commentators suggest that corporate compliance is a more complex process. Managers must acquire information about the expected costs of their alternatives, process that information in light of their existing values and norms, and then establish channels to convey their decision to line workers, who in turn must process the managers' signals, and so on (Coglianese & Kagan 2007, Langevoort 2002, Parker & Nielsen 2009).

While that much is common ground, commentators disagree considerably on the relative importance of different inputs. Sociologists and psychologists, for example, seem especially apt to discount the contribution of formal government sanctions, contending instead that compliance is a product mostly of the norms and preferences of corporate stakeholders and employees (Feldman 2007, Kagan, Gunningham, & Thornton 2011:40, Suchman & Edelman 1997:482–483, Vaughn 1998). Many lawyers, on the other hand, would echo the views of two respected scholars of administrative law, who remark that "There is little question that the ability to monitor compliance with legal requirements is a critical component of effective regulation" (Markell & Glicksman 2014).

In some respects, those who emphasize the role of formal deterrence are on solid empirical ground. A number of studies report that enforcement or anticipation of enforcement is a major driver of firm behavior, either directly or via experience-rated insurance systems (Ashby & Diacon 1996, Moore & Viscusi 1990, Toffel & Short 2008; see Vandenbergh 2003 for an overview). It is not necessarily the case that it is the government's sanction itself that triggers compliance, but rather some collateral consequence that flows from sanction, such as fear that enforcement will expose the violators to negative judgment by peers (Karpoff, Lott & Wehrly 2005, Vandenbergh 2003), cause conflict with important partners or other regulators (Coglianese & Kagan 2007), or damage the firm's reputation with customers or future investors (Anton et al. 1998, Arora & Cason 1996, Kagan, Gunningham, & Thornton 2011). All of these are consistent with the standard economic model of deterrence, which explicitly incorporates notions such as reputational harm (Rasmussen 1996). That is, whatever the exact mechanism, government enforcement actions (if not severe formal penalties) are needed to drive compliance (Ayres & Braithwaite 1992, Braithwaite 2002, Coglianese & Kagan 2007:*xix*).

On the other hand, survey respondents and interviewees sometimes state that fear of sanction or collateral consequence was less important to their compliance than other factors (Ellis & Simpson 1995, Paternoster & Simpson 1996, Simpson et al. 2013, Van Rooij 2015), or that both sets of motives were important (May 2004, Winter & May 2001; Etienne 2011:308 collects yet others). It is difficult to assess whether enforcement is always necessary to compliance, or instead whether governments select enforcement in those settings where it appears effective. A central identification problem is that in general we rely on government enforcement systems to collect compliance data, and thus it is difficult to observe outcomes under a counter-factual in which there is little or no enforcement activity (Parker & Nielsen 2009).

Researchers point to several factors that may contribute to compliance in the absence of formal deterrence. One is organizational culture. While the term "culture" has a variety of shades of meaning in the literature, in general it is meant to refer to a shared set of values or expectations within the organization (Parker & Nielsen 2009). Culture may bind, among other pathways, through norms or via a tournament-like mechanism in which employees identify and attempt to emulate the set of behaviors that are rewarded within the organization (see Langevoort 2002 for a detailed review). Culture can have many sources, including laws, public expectations, and intentional or unintentional examples set by top management (Smith, Simpson & Huang 2007, Suchman & Edelman 1997). For example, managers may comply with law because they believe in its legitimacy (Braithwaite 2003, Tyler 1990).

While some firm culture may be a reaction to existing law (Kagan, Gunningham, & Thornton 2011), it need not be. For example, compliant culture might develop in a regulatory vacuum precisely in order to reduce political demand for greater enforcement.

Relatedly, professional experts can contribute to corporate compliance, especially professionals holding an official compliance role (Beckenstein & Gabel 1983, Gramling et al. 2004, Suchman & Edelman 1997). Even where law's bite is uncertain, professionals may press for compliance in order to elevate their own importance in the firm (Suchman & Edelman 1997:500). Professionals may also have peer norms to comply (Rock 1997). And outside professionals, such as auditors and attorneys, have financial incentives to maintain a reputation for probity even in contexts where the odds that noncompliance would be detected is low (Jackall 1988:108–111).

Galle (2017) provides some evidence that these informal compliance mechanisms may play important roles in the behavior of charitable firms. I find that changes in legal rules governing the spending of U.S. philanthropic organizations result in marked changes in actual spending for those firms (and only those firms) that had been subject to a binding legal constraint. I interpret this as evidence of compliance with the pre-amendment rule. This measure of compliance is correlated with firm accounting expenditures, as well as with measures of the firm's compliance "culture" or internal controls.

References

Burton A. **Weisbrod**, Toward a Theory of the Voluntary Non-Profit Sector in a Three-Sector Economy, in Altruism, Morality, and Economic Theory 171, 175-83 (Edmund S. Phelps ed., 1975)