CMPO Working Paper Series No. 03/060

Not-for-profit Provision of Public Services

Patrick Francois

CentER, Tilburg University, The Netherlands; Dept of Economics, University of British Columbia, Vancouver, Canada; and CMPO, University of Bristol

September 2002

Abstract

Not-for-profit firms are greatly over-represented in the childcare, medical care, education and care for the aged sectors where service providing workers, as well as purchasers, seem to care about the level or quality of service being provided. Since all individuals who care about service levels receive non-excludable benefits, these services have a public good element. Such care can be used to motivate employees to perform tasks beyond their strict job description. But such care only motivates effort if workers believe it affects the final level of provision. Since nonprofit status ensures management is not directly concerned with profit, or not answerable to owners with such concerns, it ensures workers' efforts 'matter' by committing the firm to not expropriating extra worker effort to lower costs or raise profits. Nonprofit firms can thus motivate their work force in a way that for-profit firms cannot match.

JEL Classification: H11, H42, D64

Acknowledgements

This paper was given at a CMPO session of the Royal Economic Society Conference 2002 (University of Warwick) on 'The Private Delivery of Public Services'. It is forthcoming in the Economic Journal (2003). The author has benefited from the comments of Paul Grout, Ian Preston and seminar participants at CMPO, Bristol and Tilburg University.

Address for Correspondence
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver
British Columbia
Canada V6T 1Z1
francois@interchange.ubc.ca

CMPO is funded by the Leverhulme Trust.

1 Introduction

Existing theories suggest that nonprofit firms should arise in sectors where it is difficult for the purchaser to contract with the provider of the good or service sold.[1] Nonprofits are, indeed, usually found in the provision of services, where this type of contract is difficult to specify. But most service sectors are exclusively provided by for-profit firms, whose concern for their reputations ensures delivery of quality service. Existing theories of nonprofits based on contracting purchase difficulties can not satisfactorily explain why reputations do not similarly work for for-profit firms in sectors where nonprofit firms are widespread.

This paper is motivated by the observation that nonprofits are highly concentrated in the delivery of caring services where the service provided often has a public good component. Examples of such services are childcare, medical care, education and care for the aged. For multi-country evidence on this see Rose-Ackerman (1996, Figure 2). In such sectors, service providing workers, as well as purchasers, often have a civic-minded interest in service and consider the level or quality of service important. It is posited here that this is an important distinction between these services and the vast majority of others such as: financial services, insurance, and professional services etc., where much less civic-minded care exists.

The paper demonstrates that nonprofit firms are more effective than for-profit firms in obtaining care motivated effort, in addition to pecuniarily motivated effort, from workers. Their not-for-profit status provides a commitment that donated labor effort, which caring workers give in addition to that given if only pecuniarily motivated, actually affects service and will not be converted by the firm into profit. Because nonprofit managers have no (or at least greatly weakened) pecuniary incentives to cut costs elsewhere, nonprofit employees know that only modest crowding occurs due to their donated effort. In a for-profit firm, however, residual claimancy provides owners with strong incentives to reduce other inputs in order to profit from the worker's donated contribution. Knowing this, workers who care only about the service provided and not employer's profits, donate less labour effort than for a nonprofit employer.

The explanation provided for nonprofits is similar to the well-known "labour donation hypothesis" (see Rose-Ackerman 1996, p. 719, and Preston 1989, for examples). Like the present explanation, it is posited there that concerned workers may donate labour if they believe in the firm's mission. However, as here, such donations are subject to a free-rider problem common to all public goods, which previous statements of this hypothesis do not seriously consider. Since workers do not benefit by the very act of donating effort, but instead by the effects of effort on output, they would be better off if, instead of themselves working for low wages in such firms, someone else were to do so. This free-riding problem implies that the labour donation hypothesis,

[1] This is the motivation for non-profits explored by Hansmann (1980) and Weisbrod (1988). Glaeser and Shleifer (2001) formalizes this using a variant of Hart, et. al's (1997) model.

in its simplest form, cannot serve as a theory of nonprofits, because rational individuals will never work below the available market wage (or participation inducing wage) paid by firms elsewhere.

There are, however, situations when such labour donations can come about; namely, when the incentive compatible wage strictly exceeds the participation inducing wage. This typically occurs when labour effort cannot be directly contracted over, and the model here develops a relational contract solution to such formal contracting difficulties. In these, the worker is paid a wage strictly in excess of that required to induce participation. In order to keep their job, and wage premium, the worker supplies the non-contracted effort. Importantly, since such jobs pay a wage premium, the free-rider problem endemic to provision of the public good is solved: workers strictly prefer to work at firms offering a wage premium. Furthermore, for-profit firms are out-competed because the nonprofits' ability to commit to non-interference allows them to induce worker effort with a strictly smaller premium.

Since previous formulations of the labour donations hypothesis have not taken into account the free-rider problem, they have led to the conclusion that nonprofit firms should pay *lower* wages than for-profit firms; for example, this is the interpretation in Mocan and Tekin (2000). The equilibrium here predicts the opposite: within a sector, wages of identical workers at nonprofit firms should be *higher* than those of for-profits, because nonprofit firms' advantages only arise under a high-wage technology. The next section demonstrates this using a stripped-down model of for- and nonprofit competition.

2 The Model

To explicitly contrast with existing explanations, it is assumed here that the service produced is fully contractible. There is a downward sloping market demand curve for the service produced:

$$G = G(p), \tag{1}$$

where $G(p)$ is aggregate demand for service at price p and $\frac{dG}{dp} < 0$. Purchasers of this service may be the government, private agents, or a combination of both.[2]

Time is discrete, a measure N of agents are alive each period. Agents, who are all identical, remain alive with probability $\beta < 1$ for another period, and die with probability $1 - \beta$, to be replaced by identical agents. Time subscripts are suppressed and β also acts as a quasi-discount factor. Agents have the following within period preferences defined over their consumption of a numeraire good, y, the level of service, G, and the effort they exert at work e, all of which are defined on the non-negative real line:

$$u(y, g, e) = y + \gamma G - e. \tag{2}$$

[2] In a slightly different setting, Francois (2000) has examined the implications on organizational form of allowing the service purchaser to design more sophisticated contracts in a tendering situation. Allowing such complications here will not affect the basic results.

We do not carefully model the utility benefit derived by direct purchase of the service, as represented implicitly in equation (1) , as it is orthogonal to the focus here. We are instead concerned with the public good component, G, which enters each individual's utility function separately. The strength of their concern for this good or service is captured by the parameter $\gamma > 0$, which is equivalent for all individuals.

Production is undertaken by firms with access to a common production technology. A firm is composed of one worker and one manager. The production function converts worker effort into service according to $g(e)$, with $g' > 0, g'' \leq 0$. The actions required of the manager vary with the technology chosen and are specified below. Firms are price takers. The level of service is fully contractible so that each unit produced is sold at its market price, p. Thus the total revenue of a firm producing $g(e)$ is $pg(e)$, and aggregate service level simply adds the output of all producing firms $G = \Sigma g$.

Though the production mapping from effort to service level is common across all firms, the means of obtaining worker effort may vary. Third party verifiability of contracts between worker and firm is ruled out. Thus payments as a function of effort cannot be enforced. Moreover the worker cannot be the owner of the firm. This rules out the trivial solution where self-employed workers provide all output.[3]

Supervision Technology

There are two solutions to the non-contractibility problem here. The first is for the firm to use a supervision technology which involves a fixed cost, F, incurred on a per period basis. This has the advantage of allowing direct monitoring of the worker and thus solving the moral hazard problem.

Relational Contracting

Alternatively, the firm can dispense with the need for supervision and use a self-enforcing relational contract that exploits the potential repetition of interaction to provide incentives for the worker. The firm pays the worker a total amount we^* at the start of the period in return for a promised level of effort e^*. If e^* is provided, the firm re-employs the worker under the same terms in the next period. If effort is below e^*, the worker is not re-hired. For w high enough, it is well known that such an arrangement can be incentive compatible; see Macleod and Malcomson (1989) for an early formalization.

Care Does Not Affect Participation Constraints

Since G enters each individual's utility function, it might be conjectured that workers could be motivated by their care to volunteer some of their labor effort, and work for a wage that is below their cost of effort, one. But this cannot happen. Individuals correctly anticipate that any

[3]More realistically, since production generally requires other inputs and many workers, self-employment will not solve the non-contractibility problem. It needs to be ruled out here by construction because the simplifying assumption that production requires only labor effort and the effort of a single worker, makes the self-employment option a possibility.

wage they would work for would also have been accepted by someone else. Thus, their accepting work does not affect the level of service provision; i.e., in deciding whether to participate they take G as fixed. There is thus a form of free-riding problem here which implies that the direct concern for output cannot be used to make individuals accept wages below that necessary to compensate for their disutility of effort (equal to one).

Shirking

Unlike an on-the-spot supervisor, the manager in a relational contract detects shirking only with a lag and can then make up for this only imperfectly - and not by immediate dismissal and re-hiring. The shirker's poor performance implies the level of contracted upon service is not properly provided; e and thus $g(e)$ falls. Recall that since service provision is fully contractible, this should mean the firm loses an amount p times the amount of decline in g. To cover for the worker, and perhaps only partially offset the service loss, the manager may provide extra effort herself. This requires taking actions that are not usually part of the manager's job until the shirker is dismissed and a replacement hired (the next period). Since such actions are all departures from the usual (optimally chosen) labor allocation of a non-shirking employee, an efficiency loss is entailed. We model this extra cost by assuming that remedial managerial effort used in production costs $\alpha > 1$, and thus exceeds the worker's cost of effort.

Differences in organizational type

The owner of a for-profit firm keeps profits, whereas the manager of a nonprofit firm is subject to a non-distribution constraint. Each has a straightforward objective; to maximize utility. The for-profit owner is subject to her own budgetary constraint, but since the manager of a nonprofit is not responsible for the firm's losses nor gains, her constraint is less immediate. We impose one here by assuming that the nonprofit manager is subject to a non-negativity requirement: the projected costs of production have to, at least, be met by expected revenues.

$$pg(e) \geq cost(e) \text{ for any chosen } e. \tag{3}$$

A more elaborate discussion of the behaviour of nonprofit managers can be found in Glaeser (2002).

2.1 Production With Supervision Technology

By paying the per-period fixed cost, F, of the supervision technology, the firm's owner or manager is able to directly observe and dismiss shirking workers, so that the agency problem is solved. If this monitoring is effortless, then both for-profit owners and nonprofit firm managers are equivalent. However, such efforts are rarely costless, and any $\varepsilon > 0$ effort cost to supervision will see the for-profit firm have an advantage due to residual claimancy. For brevity, the precise means by which residual claimancy provides a benefit will not be formally modeled here, since it is already well known; see Alchian and Demsetz (1972). The for-profit firm's actions are then

straightforward. The owner chooses e to:

$$\max_{e} pg(e) - e + \gamma g(e) - F, \tag{4}$$

where the workers are paid only to induce participation (one per unit of effort), so that wage payments are e.[4] Given the convexity of the problem, the solution is well defined, and moreover there exists a unique p such that the maximized value of profits is zero. Denote this price by p^f. If the supervision technology is available to all entering for-profit firms, a condition of equilibrium will be that any equilibrium price p satisfies:

$$p \leq p^f. \tag{5}$$

2.2 Production With Non-supervision Technology

This form of production requires the worker to be paid a wage in excess of her opportunity cost, one, but has the advantage of not requiring the costs of the supervision technology, F. This section compares the effectiveness of both non- and for-profit firms in the use of this technology.

2.2.1 The for-profit firm

The worker knows that if not contributing effort correctly she will be dismissed and forego the stream of future wage premia, until obtaining another relational contract. Additionally, output, about which she cares, may fall. The amount of this fall depends on any remedial actions taken by the owner.[5]

Remedial effort by the owner

Suppose that the worker has contributed a level of effort, e^w, the owner then chooses e to satisfy:

$$\begin{aligned} &\max_{e} pg(e^w + e) + \gamma g(e^w + e) - \alpha e \\ &\textit{subject to} \\ &e \geq 0. \end{aligned}$$

This yields first order condition:

$$\begin{aligned} &g'(e^w + e) \leq \frac{\alpha}{p + \gamma}, e \geq 0, \\ &\text{with at least one equality.} \end{aligned} \tag{6}$$

The solution is unique and denoted by $e^m(e^w)$, where the other variables are suppressed. Note also that it can be easily demonstrated that $e^m(0) > 0$.

[4] In this, and all remaining optimizations we consider only γg, i.e. the firm's own level of output, and not the aggregate level γG in evaluating the public good component since the actions of other firms are taken as given.

[5] Here we solve for the case when the manager is the firm's owner, i.e. the case of the neo-classical firm. Even if a much weaker concern for profit ensues, provided concern for profit is non-zero, results are identical.

Let $V^E(w,e)$ denote the discounted expected lifetime value to not shirking under a relational contract (w,e):

$$\begin{aligned} V^E(w,e) &= we + \gamma g(e) - e + \beta V^E(w,e) \\ &= \frac{(w-1)e + \gamma g(e)}{1-\beta}. \end{aligned} \tag{7}$$

Since optimal shirking implies $e = 0$, the net lifetime benefit of shirking is:

$$V^S = we + \gamma g(e^m(0)) + \beta V^U,$$

where V^U is the present net discounted value of losing one's job and entering into unemployment, solved below.[6]

Let ρ denote the (endogenous, in equilibrium) probability that an individual will obtain work at one of these firms in any period. For simplicity, assume no other occupations generate such surpluses. Thus we have:

$$\begin{aligned} V^U &= 0 + \beta\rho V^E + \beta(1-\rho)V^U \\ &= \frac{\beta\rho V^E}{1-\beta+\beta\rho}. \end{aligned}$$

So that

$$V^S = we + \gamma g(e^m(0)) + \frac{\beta^2\rho V^E}{1-\beta+\beta\rho}, \tag{8}$$

and incentive compatibility implies:

$$V^E(w,e) \geq V^S. \tag{9}$$

Substituting from (7) and (8) and re-arranging, yields the incentive compatible wage:

$$w \geq \frac{e(1+\beta\rho) + \gamma\left[g(e^m(0))(1+\beta\rho-\beta) - g(e)(1+\beta\rho)\right]}{e\beta}. \tag{10}$$

If $\gamma = 0$, workers and managers do not care about output, then we have the usual relational contract $w \geq \frac{1+\beta\rho}{\beta} > 1$. It is also necessarily the case that $g(e) \geq g(e^m(0))$ because the cost of managerial effort, α, exceeds worker's costs so that $\frac{dw}{d\gamma} \leq 0$, i.e., an employee's concern lowers her incentive compatible wage.

2.2.2 The nonprofit firm

In computing the incentive compatible wage for the nonprofit employee we proceed as previously and compute the level of output resulting when a worker shirks. In that case, setting $e = 0$ again. The nonprofit firm manager chooses e to:

$$\max_e \gamma g(e) - \alpha e.$$

[6] It is standard in relational contracts that optimal shirking is to set $e = 0$. Though slightly more complicated in the present framework, since the worker has direct concern for output, this result still holds.

Which implies a solution for the manager's level of effort solving:

$$g'(e^a) = \frac{\alpha}{\gamma}. \tag{11}$$

Denote this value $e^a(0)$. The nonprofit manager only provides effort because she cares about output, in contrast to the for-profit firm owner who took into account profits as well as care. We thus have:

Lemma *In the event of worker shirking, a nonprofit firm manager provides less remedial effort than a for-profit firm owner/manager;* $e^a(0) < e^m(0)$

Proof: Immediate from (6) and (11).

This is the draw-back of residual claimancy. It provides a weaker commitment to workers that their effort matters, and, because of this, implies they are less willing to donate labor effort out of their concern for output. It thus immediately follows that nonprofit firms can pay their workers' lower incentive compatible wages than for-profit firms.[7]

Result *Given* ρ*, for any* e*, binding incentive compatible wages in nonprofit firms are*

$$w^n(e,\rho) = \frac{e(1+\beta\rho) + \gamma\left[g\left(e^a(0)\right)(1+\beta\rho-\beta) - g(e)(1+\beta\rho)\right]}{e\beta}, \tag{12}$$

where $e^a(0)$ *is solved from* (11)*. These wages are strictly lower than those that would be paid by a for-profit firm were it to use the supervision technology, solved from* (10)*.*

Finally, the no-entry condition for nonprofit firms is:

$$\pi^n(e,\rho) = pg(e) - w^n(e,\rho) \leq 0 \qquad \forall e. \tag{13}$$

An immediate implication of this result is that, in equilibrium, the only firms using a non-supervision technology are nonprofit firms, since free entry will ensure that for-profits cannot compete in its use. The next section characterizes the equilibrium outcome more fully.

2.3 Equilibrium

In reality, many sectors feature non- and for-profit firms simultaneously providing similar or identical services; see Rose-Ackerman (1996) Table 2. The model suggests that each organizational type has distinct advantages so that such a mixed equilibrium naturally occurs here.

Description of equilibrium

In a mixed equilibrium, nonprofits specialize in the non-supervision technology and for-profits in the supervision technology. Both earn zero expected profits and the market price must be at

[7] In a slightly different setting, Francois (2001) has shown that the commitment to non-interference advantage of nonprofit firms cannot be mimicked by the for-profit firm developing a reputation for non-interference. A similar result holds in the present framework.

p^f solved in equation (5). The equilibrating factor is entry of nonprofit firms. If their feasibility constraint under the non-supervision technology were not binding (equation (13) were slack) nonprofit firms would enter and thus push up the incentive compatible wage, by increasing ρ. Conversely if their constraint were violated, some would exit, ρ would fall, and so too would w^n. Moreover, whenever their feasibility constraint binds, that of for-profit firms fails, so that for-profits cannot also use the non-supervision technology.

A feature of the equilibrium is that nonprofit firms pay higher wages than for-profit firms. At first sight, the analysis would seem to suggest the opposite: recall the result of the previous section is that for-profits would have to pay higher wages if they were to use the supervision technology. The nonprofit premium occurs however because, in equilibrium, for-profit firms cannot profitably use the non-supervision technology and must resort to the supervision one. Consequently, since the nonprofit firms specialize in the non-supervision technology, they alone pay a premium to ensure incentive compatibility.

It is, however, also the case that nonprofit firms receive labor donations based on workers' concern for service, whereas for-profits do not. Even though for profit employees care just as much about the level of service as nonprofit ones (the term γ is common) this care does not motivate them to donate labor effort: for-profit workers' wages are independent of γ since they are working under the supervision technology.

3 Conclusion

Though able to account for the sectoral breakdown of nonprofits, the equilibrium implications of the present theory may seem strange. Nonprofit firms should be observed to pay more than for-profit firms for identical workers in identical sectors, yet, at the same time, workers at nonprofit firms should be seen to donate labour effort, whereas their for-profit counterparts do not. Mocan and Tekin (2000) using an unusually detailed worker/firm matched data set for the US childcare sector, have come up with precisely these findings. They found a significant nonprofit wage premium (controlling for characteristics of firms and workers). However, an unusual question in their data set (What is the main reason for your choice of working in the childcare industry?) also allowed, to my knowledge, the first direct test of the labour donations hypothesis. One of the option answers to this survey question was "I think this is an important job someone needs to do". Choosing the "important job" option had a significant downward effect on wages if working for a nonprofit firm, suggesting labour donations. In contrast, workers in for-profit firms who chose this option had either no, or a positive wage premium, suggesting no labour donations. Future work plans to more carefully investigate this and other implications of the present theory.

References

[1] Alchian, A.A., and Demsetz, H. (1972) Production, information costs, and economic organization, *American Economic Review*, Vol. 62, No. 5, pp. 777-795.

[2] Francois, P. (2000) Public service motivation as an argument for government provision, *Journal of Public Economics*, Vol. 78, pp. 277-299

[3] Francois, P. (2001) Employee care and the role of nonprofit organizations, *Journal of Institutional and Theoretical Economics*, Vol. 157, pp. 443-464.

[4] Francois, P. (2002) Why nonprofit firms are mainly in the caring sectors and why they motivate employees more effectively, mimeo Tilburg University, The Netherlands.

[5] Glaeser, E. (2002) The governance of not-for-profit firms, *NBER working paper* 8921.

[6] Glaeser, E and Shleifer, A. (2001) Not-for-profit entrepreneurs, *Journal of public economics,* Vol.81(1), pp. 99-115 .

[7] Hansman, H. (1980) The role of nonprofit enterprise, *Yale Law Journal*, Vol 89, pp 835-901.

[8] Hart, O., Shleifer, A. and Vishny, R., (1997) The proper scope of government: theory and an application to prisons, *Quarterly Journal of Economics*, Vol. 112(4), pp. 1127-61.

[9] Mocan, H.N. and Tekin, E. (2000) Nonprofit sector and part-time work: an analysis of employer-employee matched data of child care workers, *NBER working paper 7977*, forthcoming in *Review of Economics and Statistics.*

[10] Preston, A. (1989) The nonprofit worker in a for-profit world, *Journal of Labor Economics,* Vol 7, pp 438-63.

[11] Rose-Ackerman, S (1996) Altruism, nonprofits and economic theory, *Journal of Economic Literature*, Vol. 34, pp. 701-28.

[12] Weisbrod, B. (1988) *The nonprofit economy*, Cambridge MA.: Harvard University Press.