2.3
The Economic Role of Commercial Nonprofits: The Evolution of the U.S. Savings Bank Industry*

Henry Hansmann

1. Introduction

In previous work, I and others have argued that nonprofit firms in general tend to arise as a response to problems of asymmetric information facing consumers — or, as I termed it in my earlier work, “contract failure” (Hansmann 1980). The notion, simply put, is that individuals prefer to deal with nonprofit firms rather than for-profit firms when they have difficulty policing the quality or quantity of the goods or services that the firm offers or provides. Under such circumstances, nonprofit firms — which operate under a “non-distribution constraint” that prohibits the distribution of the firm’s net earnings to anyone who exercises control over it (such as its directors, officers, or members) — hold the promise of behaving less opportunistically than would for-profit firms toward the individuals who patronize them, since those who control the firms cannot profit directly from opportunism and thus have less incentive to engage in it.

This theory is persuasive, and indeed seems to be widely accepted, for donative nonprofits — that is, nonprofits that rely upon gifts or grants for a significant portion of their income. Here the “customers” who have a contract failure problem are the organization’s donors, and the nonprofit form is undoubtedly employed in large part to provide them with a degree of fiduciary protection. The theory is more controversial, however, in the case of those nonprofits that I have elsewhere termed “commercial” nonprofits (Hansmann 1980). These are nonprofit organizations that receive little or no income from donations, but rather derive all or nearly all of their income from prices they charge for the goods and services they produce and sell. Such commercial nonprofits are becoming increasingly common in the United States today. For example, they account for most of the nation's hospital care, and also have large shares of other important service industries such as nursing care for the elderly, day care for children, and primary medical

* I wish to thank Eric Rasmusen for helpful comments on an earlier draft.
care. It is not obvious that, in industries such as these, consumers are at such an informational disadvantage that the crude protection from opportunism afforded by the non-distribution constraint could be very important. As a result, many analysts, myself included, have wondered whether we should look elsewhere for an explanation of the development of commercial nonprofits (Hansmann 1987a).

One alternative theory is that commercial nonprofits are just historical anachronisms. This seems like a persuasive theory where hospitals are concerned, for example. Nonprofit firms first came to dominate the hospital industry in the United States when it was entirely charitable and all the nonprofit firms in it were donatively supported. Health technology and financing techniques have since changed, however, in ways that have now rendered donative funding of hospitals largely unnecessary. Yet the nonprofit firm has survived in this industry, perhaps just through institutional inertia; hospitals that were formerly donatively supported institutions have simply evolved into commercial nonprofits.

Another theory is that commercial nonprofits are often just a response to tax exemption and other implicit and explicit subsidies that give them artificial cost advantages over their for-profit competitors: take away the subsidies, and eventually the commercial nonprofits would largely disappear.

There have been some efforts by economists in recent years to test these different theories empirically. For example, in an effort to provide a direct test of the asymmetric information theory, Burton Weisbrod and his students have sought to determine whether there are discernible differences in the quality of services provided by nonprofit and for-profit firms in service industries containing both types of firms (e.g. Weisbrod and Schlesinger 1986). Similarly, there have been efforts to determine empirically the extent to which commercial nonprofits in particular industries are simply a response to tax exemption (Hansmann 1987b). But, in general, this work has not yet provided us with a clear answer as to whether, and to what degree, the asymmetric information theory helps explain the role of commercial nonprofits in many, or even any, industries.

I now believe, however, that we can discern at least one industry in the United States in which commercial nonprofits clearly arose, from the beginning, primarily as a response to contract failure. And that is the savings bank industry. I shall explore the evolution of that industry in some detail

---

2. The Origins of the Savings Bank Industry

Nonprofit firms appear in the banking industry in the United States in the form of so-called mutual savings banks. Although the term “mutual” suggests that these banks are consumer cooperatives that are owned by their depositors, this is not the case. The depositors in a mutual savings bank have no voting rights or other means of exercising direct control over the organization, and thus are not members or owners in any proper sense. Instead, control over mutual savings banks lies in the hands of a self-perpetuating board of directors that holds the bank’s assets in trust for its depositors. “The term ‘mutual’ only indicates that all distributed earnings must be shared by the depositors” (Teck 1968: 13). (In addition, depositors in mutual savings banks arguably have the right, upon dissolution of the organization, to share among themselves the organization’s accumulated surplus.) Thus, mutual savings banks are appropriately classified as true nonprofit organizations rather than as cooperatives. In this respect, they should not be confused with mutual savings and loan associations, which are also common in American banking and — as discussed further below — are (at least formally) true consumer cooperatives that are owned by their depositors collectively.

Mutual savings banks arose in the United States early in the nineteenth century, following earlier English models. The first mutual savings bank was chartered in Massachusetts in 1816; by 1849, 87 mutual savings banks were

---

1 In primary medical care, nonprofits appear in the form of firms of doctors organized as group practices — commonly termed health maintenance organizations — that sell their services on a prepaid basis. Such firms are also often organized on a for-profit basis.

2 After writing an earlier draft of this paper I discovered a paper by Eric Rasmusen, since published (1988), that makes much the same argument about the historical role of mutual savings banks.

3 See Teck (1968: 13—14). The term “arguably” is used here, because depositors’ rights to the distribution of surplus upon dissolution are a bit unclear; see In re Dissolution of Cleveland Savings Society, Ohio Ct. Com. Pls. (1961); Morristown Institute for Savings v. Roberts, 42 N.J. Eq. 496, 8 A. 315 (1887).

4 In classifying nonprofits, I have elsewhere (Hansmann 1980) distinguished between “mutual” nonprofits, in which control of the organization is in the hands of the class of patrons for which the organization is a fiduciary, and “entrepreneurial” nonprofits, in which control is vested in a board of directors that is self-perpetuating or appointed by third parties. Within this scheme, mutual savings banks, despite their name, are appropriately classified as entrepreneurial rather than mutual nonprofits.
in operation, primarily in urban centers in the northeastern and mid-Atlantic states (Teck 1968: 8, 16). They were typically founded as philanthropic institutions, with their initial capital donated by wealthy businessmen. The founders’ motivation, it is said, was to help prevent pauperism, and relieve the burden on public charity, by encouraging thrift among the working class (Welfling 1968: 17).

Given this early history, one might conclude that mutual savings banks were established simply as a vehicle whereby the rich could provide charitable services to the poor, in the form of subsidized interest on the latter’s savings. Indeed, this is the conventional view. These banks would then have assumed the nonprofit form, rather than being established as proprietary organizations, for the same reasons of contract failure that lead donative institutions in general to be formed almost universally as nonprofits: in order to provide some degree of fiduciary protection for the organization’s donors, who otherwise would have little assurance that their contributions were being used for the purposes they intended rather than simply going into the pockets of the organization’s proprietors.

Yet this theory seems unsatisfying for several reasons. To begin with, although hard data seem to be lacking, the amounts of capital contributed by the founders were probably inadequate to yield more than a trifling subsidy per individual depositor. Thus, it seems implausible that mutual savings banks were established merely as charitable intermediaries through which the rich could redistribute some of their income to the poor, or even to the frugal poor. Likewise, the subsidy per investor must surely have been too small to provide any important incentive in itself for saving, and thus to encourage noticeable changes in the savings behavior of the working classes.

Further, and more important, commercial banks at the time did not serve as places where individuals could deposit their savings. Although commercial banks were numerous in the early nineteenth century — there were over 300 in the United States in 1820 (Polakoff 1970: 68) —, they did not accept small deposits from individuals (Gup 1980: 137). Rather, they generally dealt only with businesses. They served primarily a monetary function, creating money in the form of bank notes, which were then the principal circulating currency (Polakoff 1970: 17). These bank notes were issued in exchange for notes from merchants, which the bank purchased at a discount.

“... neither the merchant, nor the saver, [of the early nineteenth century] thought of banks in connection with deposits. A bank ... meant primarily a place of discount for his notes. He owed bills ..., but [his] own note would not suffice to pay those bills, even though his credit were excellent, because it would not pass acceptably from hand to hand. But if he exchanged it for the note of some bank, paying for the privilege, through a discount, he would obtain something which would pass acceptably” (Bennett 1924: 20—21).

Commercial banks obtained their working capital not through deposits, but through sale of stock in the banks. And the savings that went to purchase this stock were, presumably, large sums accumulated by wealthy merchants, not the meager weekly savings of the working class.

In the early nineteenth century, then, there was no convenient vehicle through which persons of modest income could invest. Thus, the mutual savings banks were not founded simply to provide a place where the working poor could get a more attractive rate of return on their savings; they were founded as the only place where such people could deposit their savings at all. To understand the role of the early mutual savings banks, therefore, we have to understand why it was that there were no commercial savings banks until well into the nineteenth century. That is, why did the commercial banks not take deposits from individuals, and consequently obviate the need for the mutual savings banks?

The principal reason, it seems, is that commercial banks were too untrustworthy to serve as a repository for the savings of persons of modest means. That is, the problem was probably not on the supply side but on the demand side. If individuals had been willing to entrust their savings to commercial banks, the latter might well have taken them; but willing depositors were probably too few to make the activity worthwhile.

The reason that commercial banks were so untrustworthy in the early nineteenth century is that they were then largely unregulated; they did not have to maintain minimum levels of reserves, and there were no restrictions on the ways in which they could invest their assets. Commercial banks therefore had both the incentive and the opportunity to behave opportunistically toward their depositors. In particular, they had an incentive to invest depositors’ savings in highly speculative ventures that would pay off handsomely if successful, but that ran a substantial risk of not paying off at all. If the bank was lucky in such investments, it would earn a large profit. And if it was unlucky, it would go bankrupt, leaving its depositors to bear most of the losses. Further, commercial banks had an incentive to maintain only very small reserves. That way, if the bank’s investments went sour, only a minimum of the owners’ assets would be exposed; most of the losses would fall on the depositors.

Depositors would, of course, have had an incentive to try to bind banks by contract to maintain adequate reserves and not to undertake excessively speculative investments with their savings (and banks, in turn, would have

5 Similar incentive problems are now well recognized as a limitation on the feasible debt/equity ratios for business corporations (Jensen and Meckling 1976). And they also go far toward explaining why mutual firms dominated the life insurance business in its early days in the first half of the nineteenth century (Hansmann 1985).
had an incentive to bind themselves in this way in order to attract more business). But, under the circumstances, it was probably impossible to draft a contract of this type that was both effective and enforceable.

In short, consumer deposit banking was characterized by a high degree of asymmetric information (contract failure) in its early years: depositors could not know, or control, what commercial banks were doing with their funds.

Indeed, the early commercial banks often engaged in speculation, and not infrequently behaved opportunistically toward holders of their notes for example, by making it difficult for them to be redeemed in specie (Scoggs 1924). And they were, in fact, highly risky ventures: nearly 50% of all commercial banks formed between 1810 and 1820 closed by 1825, and the same proportion of banks formed between 1830 and 1840 failed before 1845 (Trescott 1963: 19). For these reasons, commercial banks were popularly viewed with distrust during the first half of the nineteenth century (Clain- Stefanelli 1975: 51). Individuals would, with reason, have been very hesitant to permit such institutions to hold their life savings for any length of time. Thus the commercial banks confined themselves to short-term transactional services, such as discounting notes, that exposed their customers to only limited losses in case the bank failed.

There was, consequently, a demand for savings banks that would provide a higher degree of fiduciary protection for depositors than commercial banks could offer. This demand grew particularly strong in the early nineteenth century, when for the first time there was a large class of urban industrial workers who received their income in the form of wages rather than in kind, and who lacked the traditional supports of the farm communities to tide them through periods of unemployment (Welting 1968: 5). The mutual savings banks met this need. They obtained their seed capital—a problem for nonprofits in general, since they cannot obtain equity capital—from wealthy philanthropists. The mutual (nonprofit) form thus served the useful purpose of providing a degree of fiduciary protection both to the depositors and to the donors.6 But, unlike other types of donative nonprofits (such as

6 Although commercial banks did not take consumer savings deposits until the middle of the nineteenth century, there did exist before then a number of commercial trust companies that administered private and charitable trusts. Why did these commercial trust companies develop and survive in this period, while commercial savings banks did not? The answer presumably lies in the size of the individual trusts and the method of remuneration devised for trust managers. The trust companies, then as now, took a percentage of the total assets as their form of compensation. This reduced their incentive to behave opportunistically; indeed, it essentially made each individual trust into a small nonprofit firm. The trust managers could not keep any fraction of the gains from speculate irresponsibly with the trust assets, so they had little incentive to engage in such activity. In a savings

traditional redistributive charities), mutual savings banks were evidently not established as nonprofit rather than proprietary firms primarily to protect donors. Rather, the mutual savings banks sought donative financing because they were nonprofit, and they were nonprofit to protect their "commercial" customers, the depositors. Confirming this, donative financing seems to have been largely confined to the initial formation of mutual savings banks; once established, they tended to become purely commercial nonprofits.

3. The Development of Commercial Savings Banks

The mutual savings banks grew rapidly throughout the nineteenth century, reaching their peak in number of banks around 1900, when there were 652 such banks with a total of $2.1 billion in deposits (Teck 1968: 13). By the turn of the century, however, commercial banks had begun actively entering the savings field. In 1900, commercial banks held only $600 million in savings deposits; by 1915, this had increased to $3 billion (Welting 1968: 58–59). By 1983, total deposits at commercial banks were roughly ten times those at mutual savings banks (FDIC 1983: Table 1).

Why did the commercial banks begin entering the savings account business late in the nineteenth century? A likely explanation is that the advent of account, the depositor receives a fixed rate of return on his savings, and the bank keeps all profits (or absorbs all losses) that result from its investment of these funds. Thus the incentive for the bank to behave opportunistically is much higher than in the case of a trust account.

Of course, the banks could have arranged a method of remuneration for savings accounts that looked more like that of trusts. An individual savings account is generally too small to permit a bank to segregate and account separately for the investments it makes with the amounts deposited in the account; the funds from a number of such accounts must be pooled for efficient administration. Nevertheless, a bank could simply have confined itself to a fixed rate of compensation for the entire pool of savings, such as a percentage of the total assets. All earnings beyond this would be returned pro rata to the depositors as interest on their accounts. Yet this approach would essentially turn the bank into a nonprofit entity. Such a method of compensation makes the bank a trustee of the depositors' funds. The pool of assets administered by the bank would be held by it in trust for the beneficial owners, who are the depositors. In effect, this is the type of contractual relationship that was established between the managers of the mutual savings banks and their depositors. (Alternatively, the pooled assets could be owned by the depositors as a group not just beneficially, but directly; acting as a group, they would then simply hire the bank's management. This is the arrangement employed in the mutual savings and loan associations.)
state and federal banking regulation did what private contractual mechanisms could not — namely, make commercial banks a relatively safe place for members of the general public to deposit their savings. Prior to 1860, there was relatively little regulation of banks in general, and the regulation that existed was directed almost exclusively at protecting holders of the banks' notes rather than depositors. A number of states passed legislation during this period that required banks to maintain reserves of some kind.

Massachusetts was evidently the first to act, in 1829, and the movement toward such legislation accelerated rapidly after the banking crisis of 1837 (Dewey 1915: 155; Sharp 1970). Typically, this legislation limited note issues by a bank to a stated multiple of the amount of specie or other reserves held by the bank. Absent from most of this legislation, however, was any provision for a reserve requirement against deposits, whether demand deposits or time deposits. Louisiana was the first state to enact a reserve requirement against deposits as well as notes, in 1842. Prior to 1860, it was followed in this only by Massachusetts, which enacted a reserve requirement covering both deposits and notes in 1858 (Sharp 1970: 112–113).

In 1863 and 1864, however, the federal government, in response to the financial pressures created by the Civil War, passed legislation providing for federally chartered banks. This legislation required that banks chartered under it maintain a specie reserve of 25% against both notes and deposits. Many states copied this chartering system after the Civil War, and thus laws requiring reserves against deposits as well as notes became common (Rodkey 1934). This legislation, by limiting the ability of banks to act opportunistically toward their depositors, was probably a critical precondition for the increasingly strong role that the commercial banks played in savings banking in the latter part of the nineteenth century. On the other hand, such legislation still provided something less than complete protection to depositors. The reserve requirements were often rather lax; most states, for example, permitted banks to keep a substantial portion of their reserves in the form of demand deposits at other banks (Rodkey 1934: 393). Thus it is not surprising that the mutual savings banks continued to grow during this period, and at the end of the century still held in aggregate far more savings deposits than did commercial banks.7

A further decisive step in banking regulation took place in 1933, however, when the federal government passed legislation establishing federal deposit insurance that provided complete insurance for savings deposits at commercial banks (and mutual savings banks as well). This insurance essentially

---

7 It should be noted, however, that other factors, such as governmental regulation of the types of investments that could be made by commercial banks and mutual savings banks, respectively, probably also contributed to the relative shares of the savings deposit business held by these two types of institutions.

4. Mutual Savings and Loan Associations

Mutual savings banks have a close cousin in the form of mutual savings and loan associations (MSLAs), which have also played an important role in savings banking. Unlike the mutual savings banks, however, MSLAs are not nonprofits, subject to a strict non-distribution constraint, but rather are true cooperatives: their depositors have formal voting control over the organization as well as the sole claim to residual earnings. Although space precludes extensive discussion of these institutions here, a few words about their role may be appropriate, for purposes of comparison with mutual savings banks.

MSLAs first began to be formed in the United States in the 1830s. They originally arose as institutions in which small groups of working people would pool their savings, and from which they would then take loans, by turns, with which to finance the construction or purchase of a house. In the early stages of their development, an MSLA would be dissolved once all of its original members had acquired a house; subsequently, they became perpetual...
organizations with fluid membership (see generally Teck 1968). Evidently the impetus for the formation of MSLAs was in large part the same as that described above for the formation of mutual savings banks: asymmetric information in the management of consumer deposits by commercial banks. Mutual ownership, like the nonprofit form of the mutual savings banks, mitigated the hazards of opportunistic conduct.

Part of the impetus for the formation of the MSLAs, however, evidently came as well from the fact that, at the time, commercial banks commonly refused to make loans for the purchase or construction of a house (Teck 1968: 18, 21). The advantage that the MSLAs had over the commercial banks here was presumably that they were better at dealing with adverse selection and moral hazard on the part of the borrowers. A group of workingmen undoubtedly had better information with which to determine which of their friends and fellow workers would be good risks, and thus should be permitted to join the mutual, than a commercial bank would have had. Further, when times are hard, a borrower is likely to be less inclined to default when he knows that his friends and neighbors will bear the loss than when he knows that the owners of a commercial bank will bear it.

Thus mutual savings and loan associations had the potential of solving two different problems, where the mutual savings banks only solved one. Whether for this or other reasons, mutual savings banks never took root in those sections of the country where MSLAs developed — which were essentially those parts of the country (the South and West) where development occurred primarily after the 1830s.8

5. Conclusion

Commercial nonprofits are the great puzzle of the nonprofit sector today. The historical experience with mutual savings banks throws some important light on the possible roles that commercial nonprofits can play, and on the patterns of evolution that characterize them.

To begin with, the experience with mutual savings banks shows that nonprofit firms can play an important role in the early stages of purely commercial industries that are characterized by severe problems of asymmetric information. In effect, they offer a form of consumer protection. But that experience also indicates that public regulation is likely, in the long run, to be more effective than the nonprofit form as a means of dealing with problems of asymmetric information in commercial enterprise. Regulation can make for-profit firms viable; and for-profit firms, with better access to capital than nonprofit firms, and better incentives for customer responsiveness and cost efficiency as well, are then likely to begin to take over the industry. Yet, nonprofit firms, once established, tend to become embedded and do not quickly leave an industry, even after the conditions to which they initially responded have long disappeared. And thus we see that mutual savings banks have survived for more than half a century after they became anachronistic.

Nursing care and day care are arguably two other industries that have evolved along paths similar to that followed by mutual savings banks. When, several decades ago, these industries were new, consumers might reasonably have been wary of trusting commercial firms to provide the sensitive services involved, and therefore preferred nonprofit providers. Yet, as public regulation of these industries became tighter, as consumers became more knowledgeable, and as for-profit firms developed stronger reputations, the need, and hence the special demand, for the nonprofit form of organization presumably diminished. Nevertheless, the already established nonprofits still occupy a substantial market share in these industries, and may continue to for some time to come.

References


8 For further discussion of the role of cooperative enterprise, and of the way in which it compares and contrasts with the role of nonprofit enterprise, see Hansmann (1988).
2.4 Nonprofit Organizations and Consumer Policy: The Swedish Model

Victor A. Pestoff

1. Introduction

Since 1940, the Swedish government has encouraged nonprofit organizations to participate actively in consumer policy, first as an interlocutor for private business, and then as a countervailing force against business. This development was facilitated primarily by the use of political, rather than monetary resources. The government gave the nonprofit organizations access to official consumer agencies by providing them with seats on the governing bodies of consumer authorities, and thereby endowed them with official status and influence, rather than providing them with direct financial assistance. As a result, consumers were organized and institutionally pitted against producers and business in order to resolve, through negotiations, their many conflicting interests.

This paper, therefore, explores the changing role of nonprofit organizations in the formulation and implementation of Swedish consumer policy since the end of World War II. How were traditional hurdles to consumer organizing overcome? What policy has the Swedish government pursued to augment the influence of consumers? What do recent governmental proposals for decentralizing consumer policy imply for the future success of this policy?

2. Theoretical Concerns and Systemic Considerations

Four theoretical concepts are important in order to understand Swedish consumer policy and the role attributed to nonprofit organizations. These are exit and voice, collective action, countervailing power, and integrated participation. These concepts, which are related to certain basic properties of Swedish markets and politics, have influenced the options facing policymakers in Sweden.

2.1 Exit and Voice

Hirschman (1970) analyses the typical economic and political responses to decline of firms, organizations, and states in terms of exit and voice. For