

## EVALUATING A NATIONAL PENSION PLAN DATA BASE WHILE TESTING FOR ECONOMIES OF SCALE

Emily S. Andrews, Social Security Administration  
Olivia S. Mitchell, Cornell University

### I. INTRODUCTION

Recent theoretical research on private pension plans can be grouped into two general categories. The first focuses on the role of pensions in the labor market: how pensions fit into firms' employment strategies, how employees react to pension incentives, how costly it is to provide pensions, and the importance of pensions in post-retirement income.<sup>1</sup> The second group of analysts studying pensions is composed of finance specialists, whose primary concern is an understanding of how pensions fit into a firm's overall investment strategy.<sup>2</sup> Because of limitations on available data, both groups of researchers have moved slowly into empirical testing of theoretical models. The purpose of the present paper is to evaluate a new data source on private pensions in light of empirical research requirements. Since many of our conclusions are based on an in-depth study of pension costs,<sup>3</sup> we present the basic analysis and extensions in the next section. Possible biases due to data problems are highlighted and suggestions made for future researchers intending to use this data source.

### II. PENSION PLAN COSTS AND PENSION PLAN SIZE

#### A. The Model

For many years, pension administrators have claimed that large pension plans are cheaper to operate than are smaller ones. We wished to test this proposition empirically, to see if higher levels of pension plan output can be achieved with less than a one-for-one increase in administrative expenses. From a government perspective, such information can help policymakers determine whether plan consolidation should be encouraged. From a private sector viewpoint, our analysis could help employers and workers determine whether their pension monies are being allocated most effectively, or whether plan growth and merger might be more cost effective.

In order to test for scale economies in pension plan operation, we used a recently released sample of private pension reports filed with the U.S. Department of Labor, reports known as "5500 Forms." These are financial accounts of private pension plans filed pursuant to the Employee Retirement Income Security Act (ERISA) of 1974. Until recently, only the 1975 cross-section of plans was publicly available; this edited file contains reliable information on plans with 100 or more people. Files for 1977 have now been released including a greater number of plans.

The 1975 data are suitable for our purposes because they report on a set of private pension plan characteristics not available previously. First, they contain information on plan finances including the level of assets at the end of 1975, and a fairly detailed breakdown of asset holdings by type. No nationally representative sample of private pension plans had reported such a wealth of data. Nonetheless, there are some drawbacks. For instance, rates of return cannot be completed since many companies reported the book value of their holdings at the beginning of the year and

market value at the end of the year. Another useful feature of the data is that they provide some information on pension plan participants; for instance, the participant breakdown into retired and active members was never available in the past. No actuarial data on age, sex, retirement or turnover rates, however, were presented in the 1975 reports. Some of the implications of these shortcomings will be detailed below. Nevertheless, since this is still a most detailed and representative source of plan data, it remains a valuable tool for pension research.

A conventional Cobb-Douglas cost function was utilized in our analysis of scale economies for pension plan expenses, and related (the natural log of) expenses to (the natural log of) pension plan output.<sup>4</sup> Strictly speaking, the output of a pension plan is the payment of promised benefits to plan participants on retirement. But promised benefits are not reported in the 5500 Form survey. Therefore, we identified two intermediate outputs related to plan costs which could be measured: service to the participant, and investment performance of the pension plan's funds. Service to the participants is measured by the number of total participants in the plan (PARTS), since records must be maintained for all active and retired beneficiaries. Of course, the paperwork involved will vary with retirement patterns, participant counselling and other features of the plan. To control for this sort of variation we incorporated a second variable, the proportion of retired participants (RET). The disbursement of benefit checks and record-keeping for this group may produce different expense patterns than those for active workers.

Investment performance would normally be measured by some rate of return yardstick (income plus unrealized capital gains), but as noted the 5500 data do not permit a determination of such a figure. Instead, the value of assets held by the plan was used as a proxy (ASSETS); the proportion of assets held in mutual or pooled funds is also incorporated (POOL) to control for lower portfolio management expenses.

The model is summarized in the following equation, where all variables are specified in natural log form:

$$(1) \text{ EXPENSES} = a + b_1 \text{ PARTS} + b_2 \text{ ASSETS} \\ + c_1 \text{ RET} + c_2 \text{ POOL} + e.$$

It is assumed for hypothesis testing that the disturbance term  $e$  is randomly and normally distributed. If either  $b_1$  or  $b_2$  is less than one, it suggests that a one percent increase in plan output is associated with a less than one percent increase in expenses *ceteris paribus*; thus scale economies are said to exist. With participants held constant, an increase in assets implies that benefits per participant may be higher. Thus, the coefficient on the asset term indicates the relationship between increases in pension plan expenses and the level of benefits likely to be

available to retirees. Finally, the sum of  $b_1$  and  $b_2$  suggests what would happen to pension plan expenses if a new participant were added to the plan with pension benefits similar to those promised to current plan members holding other things constant.

Before presenting results, two sample restrictions should be noted, and our reasons for imposing them. First, our analysis includes only plans which are managed by a pension trust, omitting pensions handled by insurance companies. This is necessary because ERISA does not require that assets handled by insurance companies be included in a pension plan's annual report. Therefore, insured plans probably substantially underreport assets. Insured plans may also underreport retired participants, since retirees can be carried on an insurance company's rolls rather than on a pension plan's roster of participants.

Second, the analysis below focuses only on defined benefit plans, or those pensions which guarantee specific benefits to retirees. The majority of plans with over 100 persons are of this type. The other major type of plan, the defined contribution plan, frequently provides benefits to supplement those of a defined benefit plan. Moreover, when pension coverage is provided solely through a defined contribution plan, management by an insurance carrier is not unusual. Because such complicating features in the defined contribution case restrict sample size and produce incomplete asset and participant data, these plans are not analyzed.

A further concern prompts us to break the sample into two subsets--those plans operated by a single employer, and those operated by a multi-

employer group. This approach is motivated by the recognition that single employer plans tend to have their sponsoring corporation absorb some or all of the pension's administrative expenses, rather than identifying individual services provided to the plan. Administrative personnel required to run the plan may, for instance, be regular firm employees and only spend some percentage of their time on the plan; these expenses are probably not reported. Multi-employer plans, on the other hand, are much less likely to underreport expenses since a central pension trust usually manages the plan and allocates expenses to each participating employer. Differences in single and multi-employer plans' pension expenses are investigated below.

#### B. Empirical Findings

Table 1 indicates significant scale economies in the administration of both single and multi-employer pension plans. For both types of plans, a percentage rise in participants raises costs by around one-half of one percent. Holding participants constant, a percentage rise in assets increased total expenses by a little over one-fourth of a percent, indicating substantial economies in the investment function. The sum of the two coefficients implies that pension expenses rise less than proportionally as participants increase, given the same benefit per participant.

The coefficients on the multi-employer equation are somewhat more statistically significant and the proportion of variance explained higher. These results are consistent with the suspected tendency of single employers not to charge expenses to the plan. The hypothesis that the sin-

TABLE 1

Regression Estimates of the Impact of Pension Plan Characteristics on Pension Plan Expenses: 1975<sup>a</sup> (t statistics in parentheses)

Item	Single Employer Plans	Multi-Employer Plans	Both
Constant	1.44	2.49	0.900
PARTS	0.446 (6.68)	0.557 (17.06)	0.780 (15.75)
ASSETS	0.270 (5.84)	0.270 (10.11)	0.162 (4.39)
RET	0.109 (2.57)	0.086 (3.86)	0.138 (4.13)
POOL	-.018 (3.02)	-0.037 (3.19)	-0.118 (7.95)
R <sup>2</sup>	.19	.69	.31
N	1499	745	2244

<sup>a</sup>The dependent variable in each column is the natural log of total administrative expenses, which includes fees and commissions, salaries and allowances, insurance premiums and other administrative expenses.

gle and multi-employer equations are identical for all coefficients is rejected at the one-percent level ( $F=126$ ). However, we cannot reject the hypothesis that the coefficients on the ASSETS and PARTS terms are identical across plan types. Thus, it appears that the major difference between the single and multi-employer equations is the constant term.

This evidence is consistent with the hypothesis that single employers may underreport pension expenses in a very simple way. Consider how single and multi-employer plan expenses might differ if multi-employer plans report all expenses but singles always understate true expenses by some fixed fraction  $d$  (less than one). For the purposes of the argument, collapse all right hand side variables in equation (1) to a vector  $X$ , and all coefficients  $b_i$  and  $c_i$  into a vector  $\beta$ . In this event reported expenses for singles ( $E^r$ ) would be related to true expenses ( $E^*$ ) as follows (assuming  $u$  is independently distributed):

$$(2) \quad E^r = d E^* + u$$

Substituting (2) into equation (1) produces an equation for expenses reported by single employer pension plans of the following form (assuming that  $\text{EXPENSES}^r$  is the natural log of  $E^r$ , defined above):

$$(3) \quad \text{EXPENSES}^r = (a - \ln d) + \beta X + (e - \ln u).$$

Thus if (2) holds, only the constant term in the equation for single employer pensions expenses is biased, but other coefficients remain unchanged. On the other hand, if underreporting by single employer pension plans took some other form, presumably the entire set of coefficients  $\beta$  could also be affected. The results in Table 1 indicate only a difference in the constant term between the two types of plans and imply that single employer pensions report only about 35% of their expenses. Naturally, this figure should be compared to evidence from other data sources for confirmation, but it suggests that understatement of expenses may be quite serious for single employer plans.

#### C. Further Comments About Plan Service Variables

The record-keeping and other service functions of the pension plan were controlled for in Table 1 with only two variables, PARTS and RET. Thus, we have omitted a variety of other factors that might influence the quality and types of services provided. In work reported elsewhere,<sup>5</sup> the sample was further broken down to the industry level in order to determine whether the results differed along this dimension. Unfortunately, the data do not permit one to determine whether variation in coefficients is attributable to differences in management style or in the characteristics of workers. Furthermore, plan service requirements may be especially high for a workforce that is relatively young and mobile; the effects of a firm's workforce turnover and mobility on its pension plan have not been incorporated in the analysis. Better data are required on the characteristics of pension plan participants over time to determine whether these factors would affect our conclusions.

#### D. Further Comments About Financial Variables

In work reported elsewhere,<sup>6</sup> we have noted that our use of the level of assets rather than plan performance may lead to a biased estimate of scale economies. In particular, if portfolio performance is the desired measure, but the level of assets is all that is available, the size of the bias can be determined. We have shown that under certain simple assumptions scale economies will be overstated but that the magnitude appears small.

A second problem also arises with respect to the financial variables: the pension balance sheet neglects unfunded pension liabilities. These represent pension promises made for which the fund has not yet been provided contributions.<sup>7</sup> The way in which such liabilities fit into the structure of plan output is not clearly understood at present. Yet, with appropriate data the role of unfunded liabilities could begin to be investigated. Information required to determine the extent of unfunded liabilities is potentially available on Schedule B of the 5500 form. Data were not provided for the 1975 sample of plans, however, because of definitional and coding problems. To the extent that pension plans incur added expenses that depend on these liabilities, scale coefficients will be underestimated if they are omitted.

#### III. CONCLUSION

The next step would be to extend our findings from defined benefit plans managed by a trust to the full range of other types of pension plans; a number of hypotheses are obvious candidates for investigation. Such generalization is impossible, however, without certain data improvements. For instance, data on small plans (with fewer than 100 persons) would have to be available to determine whether economies of scale arising from small plan consolidation or merger are greater than those for larger plans. Similarly, in order to understand why some plans utilize insurance carriers while others do not, information on differences in expenses incurred by plan type would have to be available.

We have indicated that the evidence suggests that single employer plans underreport expenses in a relatively simple way. Yet, in order to confirm this hypothesis, financial and personnel data from the sponsoring enterprise would have to be added to the pension balance sheet. Finally, only full information on defined contribution plans would tell us whether these are sometimes less costly alternatives to defined benefit plans. Such research on insured plans, single employer plans and defined contribution plans cannot be undertaken without data linkages to private sector sources. Linkages are also needed in order to study the full range of research questions including labor market and capital market issues.

While private sector linkages need to be stressed, a number of deficiencies in our analysis can be solved through Federal reporting and disclosure files if data development plans proceed in the coming years. For instance, measures of actual pension portfolio performance could be developed using more recent 5500 forms. In particular, assets are reported in terms of market value at the beginning and end of each plan year

and, in addition, panel data spanning several plan years can be developed. Furthermore, starting with the 1980 filing, annual reports will include improved actuarial data with better information on unfunded liabilities. Such data are required for a variety of topics, including our study. Actuarial data can also be used to enhance the information available on benefit obligations and participant characteristics of particular plans. Finally, although full reporting for small plans (those with under 100 participants) is only required every third year, representative information could still be developed to investigate small plan behavior.

In conclusion, the 1975 pension plan filing sample provides a unique opportunity to investigate scale economies. While the empirical results are encouraging, a number of data shortcomings became apparent during the course of the analysis. Although some deficiencies are related to gaps in data availability for the initial year of ERISA reporting, others point to the needs for linkages with private sector information. These problems are likely to plague researchers interested in analyzing many aspects of the economies of pension plans. Consequently, the sample and variable deficiencies identified in our study are generally applicable, and suggest that a community of researchers would benefit from future data development.

#### FOOTNOTES

<sup>1</sup>For a review of many of the recent studies on

pensions and labor markets, see Mitchell, Olivia S., and Fields, Gary S., "The Effects of Pensions on Retirement: A Review Essay," Labor Economics Working Paper, Cornell University (Spring 1981).

<sup>2</sup>Examples of recent research on the financial aspects of pension plans include various studies issued as Working Papers by the National Bureau of Economic Research.

<sup>3</sup>The research reported in this study builds on the analysis appearing in Mitchell, Olivia S. and Andrews, Emily S., "Scale Economies in Private Multi-Employers Pension System," Industrial and Labor Relations Review, July 1981.

<sup>4</sup>Input prices would also be included, but the data we are discussing here did not provide them--namely, salary figures for actuaries, lawyers, accountants and financial advisors to the pension plan. These are likely to be similar across the Nation, since such labor markets are nationally competitive: thus, their omission does not seriously affect the results discussed below.

<sup>5</sup>See Mitchell and Andrews, op. cit.

<sup>6</sup>See Mitchell and Andrews, op. cit.

<sup>7</sup>Under ERISA this deficiency must be made up over a 30-year period.