TRANSACTIONS OF SOCIETY OF ACTUARIES
1979 VOL. 31

A BAYESIAN APPROACH TO PERSISTENCY IN THE PROJECTION OF RETIREMENT COSTS

ARNOLD F. SHAPIRO

ABSTRACT

This paper uses a Bayesian approach to persistency to explore retirement cost projection variability. The analysis is divided into two parts. In the first part a subjective population projection model is developed, based on a beta-binomial distribution for the number of participants persisting through a given age. Probability statements regarding such items as the number of participants at a given age and the distribution of retirees are developed. The second part shows how the model might be used to generate confidence intervals for retirement cost projections. Both the single-entry-age case and the multiple-entry-age case are considered.

It is almost inconceivable that anybody could be in the position of having no *a priori* knowledge whatever regarding mortality.—
E. T. WHITTAKER

INTRODUCTION

PENSION actuaries long have realized that pension cost projections provide valuable insight into the cash-flow characteristics of pension plans.[1] Because of this, a pension cost projection typically is appended to the more elaborate pension plan proposals and valuations. Depending on the size of the plan, these range from simple projections that assume a closed group with no terminations other than for retirement[2] to more sophisticated models that introduce the full spectrum of pension plan parameters.[3]

[1] One of the first published accounts of the growth of a pension fund was contained in a paper by James J. M'Lauchlan, "The Fundamental Principles of Pension Funds," *TFA*, IV (1908), 195–227. In that paper M'Lauchlan illustrated the necessity of accumulating large investment funds during the early years of a fund's existence in order to provide for the heavy liability that ultimately will be maturing for payment.

[2] This type of projection is most commonly associated with the valuation of small pension plans. See, for example, *Calculating Auxiliary Fund Deposits for the Small Pension Plan* (Chicago, Ill.: A. A. Beaven & Co., Inc., 1975).

[3] Papers that discuss this type of projection include Robert J. Myers, "Some Considerations in Pension Fund Valuations," *TASA*, XLVI (1945), 51–58; A. M. Niessen,

While there is much to be said in favor of this practice, it suffers from at least two serious shortcomings. First, it provides no mechanism for incorporating the actuary's feelings regarding his confidence in the underlying assumptions. Perhaps the most important attribute of an experienced pension actuary is his intuitive notion of what should be. Ideally, there should be some vehicle for injecting this intuition into pension cost projections. Second, since projections invariably are based on expected value models, they provide no mechanism for introducing credibility.[4] Attached to any estimate of projected pension costs should be a statement of the actuary's confidence in that estimate.

It might be argued that this degree of refinement of pension cost projections is not warranted. Proponents of this view reason that pension costs are funded sequentially over a number of years and that periodic actuarial valuations will uncover underfunding problems before they can affect materially the solvency of a plan. The implication is that ex ante pension cost projections should be viewed strictly as rough (albeit best) estimates of ultimate pension plan costs. The fact that such projections may not convey an accurate picture of ultimate cost is regarded as only marginally relevant.

This proposition, however, disregards the question of whether a particular plan or plan liberalization would have been introduced initially had the plan sponsor realized that actual cost might be considerably in excess of the projected cost. Furthermore, this view presumes that the plan sponsor will be able to fund any deficiencies that arise. These considerations have become increasingly important in light of the liability that ERISA imposes on plan sponsors.[5] Thus, while ex post reconciliation of pension cost estimates remains an important facet of pension cost funding, there are compelling arguments for developing techniques to measure the variability of ex ante pension cost projections.

"Projections—How to Make Them and How to Use Them," *TSA*, II (1950), 235–53; Charles L. Trowbridge, "Fundamentals of Pension Funding," *TSA*, IV (1952), 17–43; and Frank L. Griffin, Jr., "Concepts of Adequacy in Pension Plans," *TSA*, XVIII (1966), 46–63.

[4] A. Guy Shannon, Jr., remarked: "Invariably, projections are based on expected value models and seldom is there a *quantified* statement of the actuary's confidence in the projection. Ideally, the individual assumptions and the composite results of the valuation should be viewed as the mean of the universe from which the experience of that pension plan will be drawn. The measurement of liabilities would be accompanied by a set of confidence limits based on the combined effect of the entire set of assumptions" (*Pension Topics* [Society of Actuaries Study Note 71-22-76], p. 10).

[5] Under ERISA, sec. 4062(b), an employer's liability may be as high as 30 percent of its net worth.

These observations suggest the need for a stochastic model for projecting pension costs. A straightforward procedure would be to base such a model on direct or deductive probabilities. One could assume, for example, that the number of participants who succumb to a particular decrement is distributed binomially and is based upon a probability of decrement that is constant or is given by a degenerate distribution. This assumption of an underlying degenerate distribution, however, is questionable in actual practice. Probabilities of decrement are obtained either from intercompany experience, which, at best, may only approximate the actual experience of a particular firm, or else they are derived from the firm's own experience, which, for the majority of firms, is not very credible. Thus, what is needed is a model in which underlying parameters may take on probability distributions.

These additional considerations lead naturally to a Bayesian approach to stochastic pension cost projections. Under this approach, not only are pension cost determinants (such as the number of decrements due to a given cause and the fund accumulation factor) assumed to be distributed stochastically, but the parameters upon which these determinants depend are themselves assumed to be distributed stochastically.

In this paper a Bayesian approach to persistency is used to explore retirement cost projection variability.[6] The analysis is divided into two parts. In the first portion the specifications of the model are developed. The second portion shows how the model might be used to generate confidence intervals for pension cost projections.

The paper ends with a comment on the use of stochastic models and suggestions for further study.

STOCHASTIC PENSION COST MODELS

Very few papers have dealt specifically with the development of a stochastic model of pension costs. Stone[7] investigated the impact of mortality fluctuations on pensions paid to pensioners. The main thrust of that study was the use of probability generating functions to develop probabilities, at various durations after employees had begun to retire, that the actual total pension payments would differ from the expected

[6] Persistency is, of course, not the only source of variability in pension cost projections. Deviations resulting from such factors as shifts in the distribution of salaries or returns on assets are also extremely important sources of variation. For the purpose of the present study, however, factors unrelated to persistency are assumed to be invariant.

[7] David G. Stone, "Actuarial Note: Mortality Fluctuations in Small Self-insured Pension Plans," *TASA*, XLIX (1948), 82–91.

total payments. Taylor[8] investigated the size of the contingency reserve needed to ensure that, with a given probability, the funds on hand would be sufficient to pay all promised pensions. Both of these studies dealt exclusively with the retired population, under the assumption that the number of retirees was known.

Papers that considered variability in pension cost estimates for active plan participants include the studies of Seal,[9] Knopf,[10] and Shapiro.[11] Seal investigated the impact of death benefits in a trusteed plan, using a normal approximation to the binomial distribution to introduce variance minimization into the design of pension plans. Knopf investigated the feasibility of full trusting of small pension plans using a simplified Monte Carlo approach. Shapiro considered the credibility of projected pension costs using a model based on the direct application of a conditional Bernoulli process.

To the extent that pensions may be regarded as annuities, a large number of other studies may be considered relevant. Piper,[12] for example, developed contingency reserves for life annuities based on the mean and variance associated with those annuities. Menge,[13] and later Hickman,[14] elaborated on the Piper paper—Menge using discrete functions and Hickman using continuous functions. Hickman's paper, in addition, extended the development to include loss functions and a probabilistic consideration of multiple decrement theory. The latter is directly applicable to pension populations.

Although it is clear that the number of lives that persist to a given age from an initial group of lives is generated by a Bernoulli process, the complexity of this process has resulted in the development of various approximation methods. Thus, Piper assumed a large group of lives and used a normal distribution, as did Seal; Taylor suggested fitting a Pearson Type III distribution to the total present value of life annuity costs;

[8] Robert H. Taylor, "The Probability Distribution of Life Annuity Reserves and Its Application to a Pension System," *PCAPP*, II (1953), 100–150.

[9] Hilary L. Seal, "The Mathematical Risk of Lump-Sum Death Benefits in a Trusteed Pension Plan," *TSA*, V (1953), 135–42.

[10] Myrna Knopf, "A Practical Demonstration of the Risk Run by a Very Small Company with a Trusteed Pension Plan," *PCAPP*, VI (1957), 230–43.

[11] Arnold Shapiro, "The Relevance of Expected Persistency Rates when Projecting Pension Costs," *JRI*, XLIV, No. 4 (December, 1977), 623–38.

[12] Kenneth B. Piper, "Contingency Reserves for Life Annuities," *TASA*, XXXIV (1933), 240–49.

[13] W. O. Menge, "A Statistical Treatment of Actuarial Functions," *RAIA*, XXVI (1937), 65–88.

[14] James C. Hickman, "A Statistical Approach to Premiums and Reserves in Multiple Decrement Theory," *TSA*, XVI (1964), 1–16.

Boermeester[15] applied a Monte Carlo approach to the problem, as did Knopf; Fretwell and Hickman[16] investigated upper bounds for the cost, using the inequalities of Chebyshëv and Uspensky; and Bowers[17] investigated the use of the Cornish-Fisher expansion to develop probabilities of sufficient reserves, based on correction factors applied to a standard normal table.

These studies relied generally on distributions whose underlying parameters were given. This study explores the use of a less constrained distribution.

PROBABILITY OF A GIVEN NUMBER OF PARTICIPANTS AT EACH AGE

Except at the entry age, the number of participants at age x in a pension plan may be regarded as a random variable, say $\tilde{l}_x^{aa}$, that depends on the number of participants at the previous age, $\tilde{l}_{x-1}^{aa}$, which is also a random variable. Let

$$^{k}l_x^{aa} = (^{k}l_{x-t}^{aa}\colon t = 0, 1, \ldots, x-a)' , \tag{1}$$

where a is the entry age, denote a vector of l_{x-t}^{aa} values consistent with a final value of l_x^{aa}, and call this vector a feasible l_x^{aa} array. If, for example, the number of entrants at age 20 were equal to 100, then the feasible arrays consistent with 98 participants at age 22 would be (100, 98, 98), (100, 99, 98), and (100, 100, 98).

Assuming that there are K distinct feasible l_x^{aa} arrays, the probability that $\tilde{l}_x^{aa}$ takes on some particular value is given by

$$\Pr\{\tilde{l}_x^{aa} = l_x^{aa}\} = \sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f(^{k}l_{x-t}^{aa} \mid l_x^{aa}) , \tag{2}$$

where f denotes the probability that exactly l_{x-t}^{aa} participants will persist through age $x - t - 1$, consistent with l_x^{aa} participants persisting through age x.[18]

[15] J. M. Boermeester, "Frequency Distribution of Mortality Costs," *TSA*, VIII (1956), 1–9.

[16] Robert L. Fretwell and James C. Hickman, "Approximate Probability Statements about Life Annuity Costs," *TSA*, XVI (1964), 55–60.

[17] N. L. Bowers, "An Approximation to the Distribution of Annuity Costs," *TSA*, XIX (1968), 295–309.

[18] Since this study deals strictly with retirement benefits, an alternative approach would be to restrict consideration to a model of the form $\tilde{l}_r^{aa} = {}_{r-a}\tilde{p}_a^{aa} l_a^{aa}$, as did Knopf, op. cit. This alternative approach has not been used, since this paper seeks to develop a model that is general enough to accommodate other benefits, that allows subjective judgment at each relevant age, and that provides a vehicle for tracing the progression of pension populations. The more general model used in this paper is necessary in order to accomplish these ends.

The Appendix gives a brief description of the programming logic for calculating probabilities over feasible arrays.

A Conditional Probability Distribution Function for l_x^{aa}

In order to implement equation (2), it is necessary to specify the probability distribution function. Given the assumptions of mutual stochastic independence and identical distribution, the number of employees who persist through a given age may be thought of as being generated by a Bernoulli process under which employees either persist as active members or leave the active group.[19] It follows that a conditional distribution of l_{x-t}^{aa} is specified by the binomial mass function[20]

$$f_b(l_{x-t}^{aa} \mid p_{x-t-1}^{aa}, l_{x-t-1}^{aa}) \propto (p_{x-t-1}^{aa})^u (1 - p_{x-t-1}^{aa})^v , \quad (3)$$

where $u = l_{x-t}^{aa}$ and $v = l_{x-t-1}^{aa} - l_{x-t}^{aa}$, and p_{x-t-1}^{aa} is the probability that an employee aged $x - t - 1$ will persist through that age.

In Figure 1 the binomial distribution is used to project the distribution of the number of plan participants at each age through age 65, assuming that there are 100 entrants at age 20. The probabilities of persisting are based on mortality rates from the 1971 Group Annuity Mortality Table,[21] disability rates used in the 1970 civil service pension valuation, and Turnover Table III presented by McGinn.[22] This data base, which is used for illustrative purposes, will be referred to subsequently as "the decrement data." The curve to the far right in Figure 1 represents the distribution of participants at age 21. The curve to the far left represents the distribution of the number of participants at age 65, who will be retiring. The intermediate curves are associated with participants at intermediate ages.

It is apparent from these curves that, even under conditions of perfect information, the actual number of participants at a given age may vary considerably from the best estimate of the number of participants. While this is not surprising, the considerable disparity that is likely to occur is

[19] Howard Raiffa and Robert Schlaifer, *Applied Statistical Decision Theory* (Boston: MIT Press, 1968), chap. 9.

[20] Ibid., p. 213. For an application of this distribution to the problem of projecting pension cost see Shapiro, op. cit.

[21] See Harold R. Greenlee, Jr., and Alfonso D. Keh, "The 1971 Group Annuity Mortality Table," *TSA*, XXIII (1972), 583–84.

[22] Daniel F. McGinn, "Indices to the Cost of Vested Pension Benefits," *TSA*, XVIII (1967), 235–36. If vesting were the topic of this study, bodily shifts in the rates of withdrawal subsequent to a vesting liberalization would be an important additional source of variation. Since this study deals solely with retirement benefits, however, this complication is not introduced.

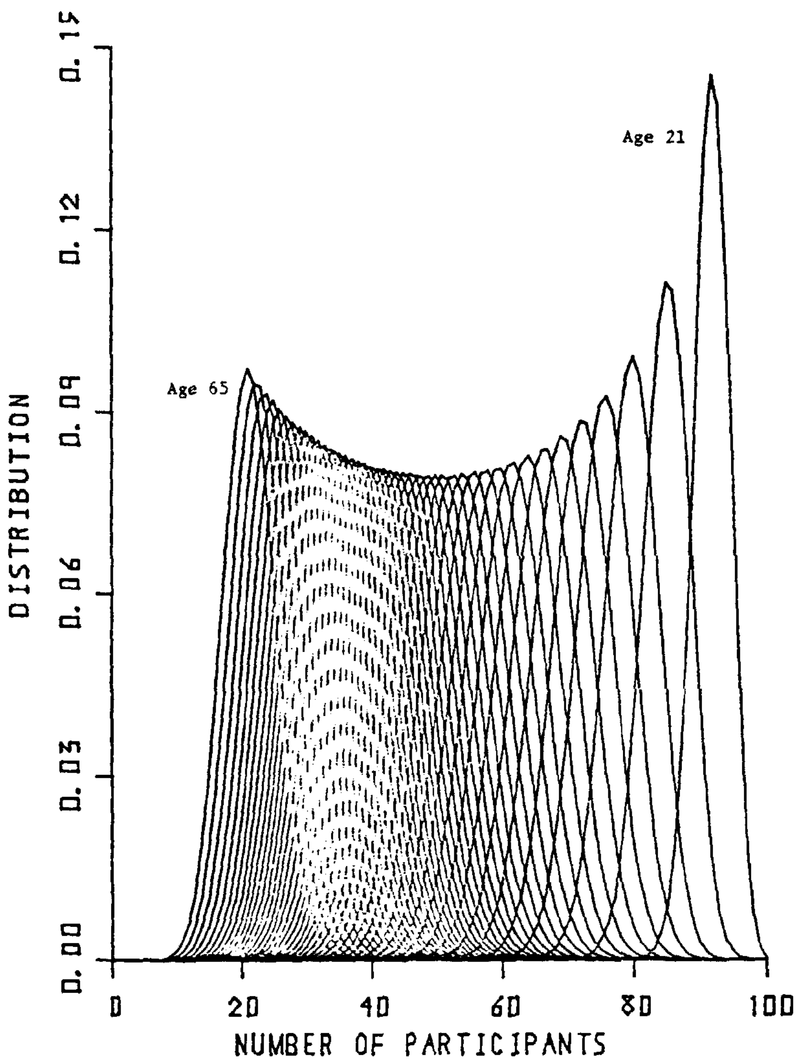

FIG. 1.—Distribution of number of participants at ages 21–65, given that they are distributed binomially. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

interesting. In the graph the locus of the modes of the distribution of participants is convex. The age at which the locus attains a minimum value represents the age at which the distribution of participants is most nearly symmetrical. Below this age the distribution of participants is negatively skewed and above this age the distribution of participants is positively skewed.

It is important to recognize that the binomial mass function is appropriate only under the assumption that the exact probabilities of persisting are known. This assumption, however, is generally not valid. Although estimates of p_x^{aa} are often available, the estimates may or may not be valid for the particular pension plan under consideration. Furthermore, the binomial mass function provides no mechanism for the actuary to indicate the intensity with which he views the credibility of the estimated value p_x^{aa}. These criticisms suggest the need for a more general probability distribution function. What is needed is a distribution that is not conditional upon a degenerate p_x^{aa}—that is, an unconditional distribution.

An Unconditional Probability Distribution for l_x^{aa}

Bayes's theorem may be used to transform the conditional probability of l_x^{aa} individuals persisting to an unconditional probability.[23] According to Bayes's theorem, if y has the probability density function $f_1(y)$, and the conditional probability distribution function of x, given y, is $h(x|y)$, then the joint distribution of x and y, $f(x, y)$, is given by

$$f(x, y) = h(x|y) f_1(y) . \tag{4}$$

If y has a continuous distribution, it follows that the marginal distribution of x, $f_2(x)$, is

$$\begin{aligned} f_2(x) &= \int f(x, y) dy \\ &= \int h(x|y) f_1(y) dy , \end{aligned} \tag{5}$$

which is independent of y.[24]

[23] For discussions of Bayesian analysis with insurance applications see Arthur L. Bailey, "Credibility Procedures," *PCAS*, XXXVII (1950), 7–23; Andrew R. Davidson and A. R. Reed, "On the Calculation of Rates of Mortality," *TFA*, XI (1927), 183–212; James C. Hickman and Robert B. Miller, "Notes on Bayesian Graduation," *TSA*, XXIX (1977), 7–21; Donald A. Jones, "Bayesian Statistics," *TSA*, XVII (1965), 33–57; Allen L. Mayerson, "A Bayesian View of Credibility," *PCAS*, LI (1964), 85–104; Wilfred Perks, "Some Observations on Inverse Probabilities including a New Indifference Rule," *JIA*, LXXIII (1947), 285–310; and E. T. Whittaker, "On Some Disputed Questions of Probability," *TFA*, VIII (1920), 163–206.

[24] A more general formulation would give the marginal distribution of x in terms of a generalized Riemann-Stieltjes integral with respect to y. However, since the distribution of the probability of persisting is continuous, this complication need not be introduced.

From the previous section we know the conditional distribution of exactly l_x lives persisting through age x. If we can assume that the probability of persisting, p_x^{aa}, is a random variable from a nondegenerate distribution, the unconditional probability of l_x^{aa} lives persisting can be determined.

Certain properties of p_x^{aa} seem evident. First, $0 \leq p_x^{aa} \leq 1$, so that the distribution from which p_x^{aa} is drawn, generally referred to as the "prior" distribution, must be distributed over this range. Second, the probability of persisting may take any value in this domain, so that p_x^{aa} has a continuous distribution. Finally, for any given age, the probability of persisting may be concentrated at no more than one value, so that the distribution of p_x^{aa} has a single mode.[25] It is assumed that any probability density function that is chosen to represent the probability of persisting must exhibit these properties.

In addition to the empirical properties mentioned above, another desirable property stems from the fact that it may be impossible to specify the distribution of p_x^{aa} exactly, owing to a scarcity of relevant data. The distribution that is used to characterize the probability of persisting should lend itself to updating as more sample information becomes available.[26]

A convenient choice for the prior distribution of p_x^{aa}, from an updating point of view, is the beta distribution in the form

$$f_\beta(p \mid r, n) = p^{r-1}(1 - p)^{n-r-1}/B(r, n - r) , \qquad (6)$$

where

$$B(r, n - r) = \Gamma(r)\Gamma(n - r)/\Gamma(n) .$$

This follows, since the updated, or "posterior," distribution also would be a beta distribution.[27]

[25] While this constraint seems generally appropriate, it has been argued that it may not be a necessary or desirable one. For example, G. E. Lidstone, in his discussion of Whittaker, op. cit., p. 196, suggested the possibility of using U-shaped curves in those instances where high probabilities occur in the upper or lower bounds of the distribution. Such a prior distribution subsequently was developed for the binomial distribution by Perks, op. cit., based on the hypothesis that $p_x dx \propto dx/\sigma_x$, where σ_x is the large sample standard error of x, a parameter in a probability law.

[26] It is important from the point of view of pension plan valuations, that is, the going-concern analysis, to be able to update estimates of pension population parameters as more data become available. The development of a stochastic pension valuation model that incorporates this facility is currently under investigation by the author and will form the basis for a sequel to the present study.

[27] See Raiffa and Schlaifer, op. cit., p. 263. It is interesting to note that Sir G. F. Hardy alluded to this distribution in the form $x^r(1 - x)^s$ in correspondence regarding a Bayesian approach to mortality (*Insurance Record*, XXVII [October, 1889], 433 ff.). It is not clear, however, whether Hardy recommended the method for practical use. See Lidstone's discussion of Davidson and Reed, op. cit., p. 225.

The beta distribution satisfies all the empirical requirements mentioned above except that it is not necessarily unimodal. This requirement is met, however, if r and n are restricted to positive values and max $\{r, n - r\}$ exceeds unity.[28]

Given that p_x^{aa} has a beta distribution as specified above, the unconditional distribution of the number of employees who persist through age $x - 1$ is given by

$$\int_0^1 f_b(l_x^{aa} | p_{x-1}^{aa}, l_{x-1}^{aa}) f_\beta(p_{x-1}^{aa} | r_{x-1}, n_{x-1}) dp_{x-1}^{aa}$$

$$= \binom{l_{x-1}^{aa}}{l_x^{aa}} \frac{B(l_x^{aa} + r_{x-1}, l_{x-1}^{aa} - l_x^{aa} + n_{x-1} - r_{x-1})}{B(r_{x-1}, n_{x-1} - r_{x-1})}, \qquad (7)$$

$$l_x^{aa} = 0, 1, \ldots, l_{x-1}^{aa},$$

$$n_{x-1} > r_{x-1} > 0,$$

$$\max \{r_{x-1}, n_{x-1} - r_{x-1}\} > 1.$$

This distribution appropriately is called the beta-binomial distribution, and its probability distribution function is given by $f_{\beta b}(l_x^{aa} | r_{x-1}, n_{x-1}, l_{x-1}^{aa})$.[29]

From the foregoing, it follows that an unconditional probability of exactly l_x^{aa} employees persisting to age x is

$$\sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f_{\beta b}({}^k l_{x-t}^{aa} | l_x^{aa}, r_{x-t-1}, n_{x-t-1}, {}^k l_{x-t-1}^{aa}). \qquad (8)$$

The remainder of this paper assumes that the beta distribution describes adequately the distribution of the probability of persisting, and that the beta-binomial distribution describes appropriately the distribution of the number of participants at a given age.

ESTIMATING THE PARAMETERS r AND n

In order to utilize the beta-binomial distribution, we must either know or estimate the parameters r and n. In practice it is unlikely that the exact values of these parameters are known, so it is necessary to estimate them. In this section the method of moments in conjunction with subjective judgment is used to develop an approach for estimating these parameters.[30]

[28] The beta distribution is bimodal if max $\{r, n - r\}$ is less than unity, in which case it has a U-shape. It was this distribution, in the form $n = 2r = 1$, that was developed by Perks, op. cit., p. 298.

[29] Raiffa and Schlaifer, op. cit., p. 237.

[30] There are, of course, more sophisticated methods for developing sample estimates of r and n, such as the method of maximum likelihood. See, for example, S. W. Dhar-

Let the tabular probability of persistency, for some particular age, be denoted by $\hat{p}_x^{aa}$. Since $E(p_x^{aa}) = r_x/n_x$,[31] it follows that, if one assumes that the tabular probability of persisting is approximately equal to the probability of persisting, then

$$r_x \doteq n_x \hat{p}_x^{aa} . \tag{9}$$

Furthermore, since the variance of the beta distribution is[32]

$$V(p_x^{aa}) = E(p_x^{aa})E(1 - p_x^{aa})/(n_x + 1) , \tag{10}$$

it follows that

$$V(p_x^{aa}) \doteq \hat{p}_x^{aa}(1 - \hat{p}_x^{aa})/(n_x + 1) . \tag{11}$$

It is clear that the estimated variance of the prior distribution of the probability of persisting will be inversely proportional to the size of the n_x parameter that is chosen. In this sense, n_x may be regarded as a precision parameter. The greater the confidence in the tabular persistency rate, the greater the value of n_x that should be chosen, that is, the smaller should be the estimated variance. Once an appropriate n_x is chosen, r_x is determined by solving equation (9). Denoting by $\hat{r}_x$ and $\hat{n}_x$ the estimated parameters of the prior distribution, the unconditional distribution of l_{x-t}^{aa} becomes

$$f_{\beta b}(l_{x-t}^{aa} \mid \hat{r}_{x-t-1}, \hat{n}_{x-t-1}, l_{x-t-1}^{aa}) . \tag{12}$$

Implementation of the Foregoing Procedure

Table 1 and Figures 2 and 3 exemplify the mechanics of the foregoing procedure by showing how one might determine a subjective prior distribution for p_{20}^{aa}. Given the decrement data, the tabular value of $\hat{p}_{20}^{aa}$ is 0.918247. Table 1, which was developed by substituting this value in equation (11), shows the trend of the estimated variance for various choices of $\hat{n}_{20}$. As the table indicates, an actuary who feels extremely confident in the tabular persistency rate might choose an $\hat{n}_{20}$ of 100 or more. This would result in a prior distribution for p_{20}^{aa} that has a variance of 0.0007507 or less. This distribution may, for all intents and purposes, be degenerate, and a binomial distribution might be used in this case. On the other hand, an actuary may be satisfied that 0.918247 represents a good estimate of the mean of the prior distribution but may, at the

---

madhikari, "A Simple Modification of the Binomial Distribution," *JIASS*, XV (1960), 436–44. However, the simplicity of the approach used in the text is a strong argument in its favor, particularly as a method of forming initial estimates.

[31] See Raiffa and Schlaifer, op. cit.

[32] Ibid., p. 213.

same time, feel that there is a considerable possibility that $\bar{p}_{20}^{aa}$ may take on some other value. In an extreme situation of this kind, the actuary may be so uncertain of the outcome that he chooses to introduce substantial variability. This could be done by choosing an $\hat{n}_{20}$ equal to 2; a distribution with a mean of 0.918247 and a variance of approximately 0.025 would result. Any distribution between these two extremes also would be available.

Figure 2 shows the impact of various choices of $\hat{n}_{20}$ on the prior beta distribution of p_{20}^{aa}. A choice of $\hat{n}_{20}$ equal to 1,000 results in an almost symmetric distribution about the mean. At the other extreme, a choice of $\hat{n}_{20}$ equal to 2 results in a distribution for p_{20}^{aa} that is highly skewed toward the origin and has a maximum value at unity.

Figure 3 shows the probability that a given number of participants

TABLE 1

IMPACT OF n ON VARIANCE OF THE PRIOR BETA DISTRIBUTION, GIVEN A MEAN OF 0.918

Prior Value of n_{20}	Variance	Prior Value of n_{20}	Variance
1*	0.0359350	500	0.0001501
2	0.0250231	10,000	0.0000075
100	0.0007507	∞	0.0000000

* The actual value of n is 1.089+, which is the smallest value of n that is consistent with a unimodal beta distribution.

will persist to age 21, given 100 entrants at age 20, based on some of the distributions given in Figure 2. The expected number of participants at age 21 is 91.82. It is apparent that, as the variance of the distribution of the probability of persisting approaches zero, the distribution of the number of employees approaches its limiting distribution, the curve labeled $n = \infty$, which is based on a binomial mass function. This is as expected, since, in the limit, the beta-binomial distribution approaches the binomial distribution.[33] Once again, if $\hat{n}_{20}$ is equal to 2, a hyperbolic curve results.

Implications of the Choice of the Parameters r and n

Before proceeding, it is appropriate to mention the implications of different choices for the parameters r and n. The choice of a small prior

[33] Intuitively, the fact that, in the limit, the beta-binomial distribution approaches the binomial distribution follows from the observation that the binomial distribution results when the prior distribution becomes degenerate.

n is tantamount to the assumption that the estimated probability of persisting at a given age, although the best available estimate, is questionable. There is a considerable chance, based on the subjective judgment of the actuary, that the probability of persisting will take some value more or less than the best available estimate. On the other hand, the choice of a large n is tantamount to the assumption that the actuary's

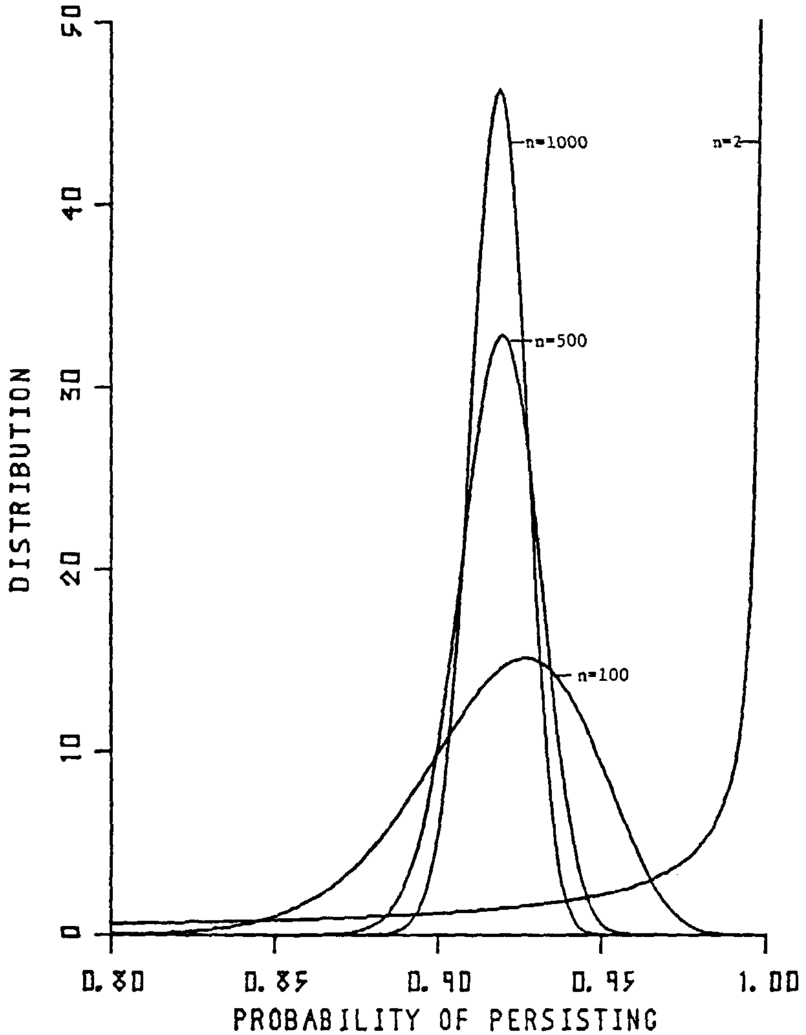

FIG. 2.—Effect of the choice of n on the beta probability density function, given a mean of 0.918.

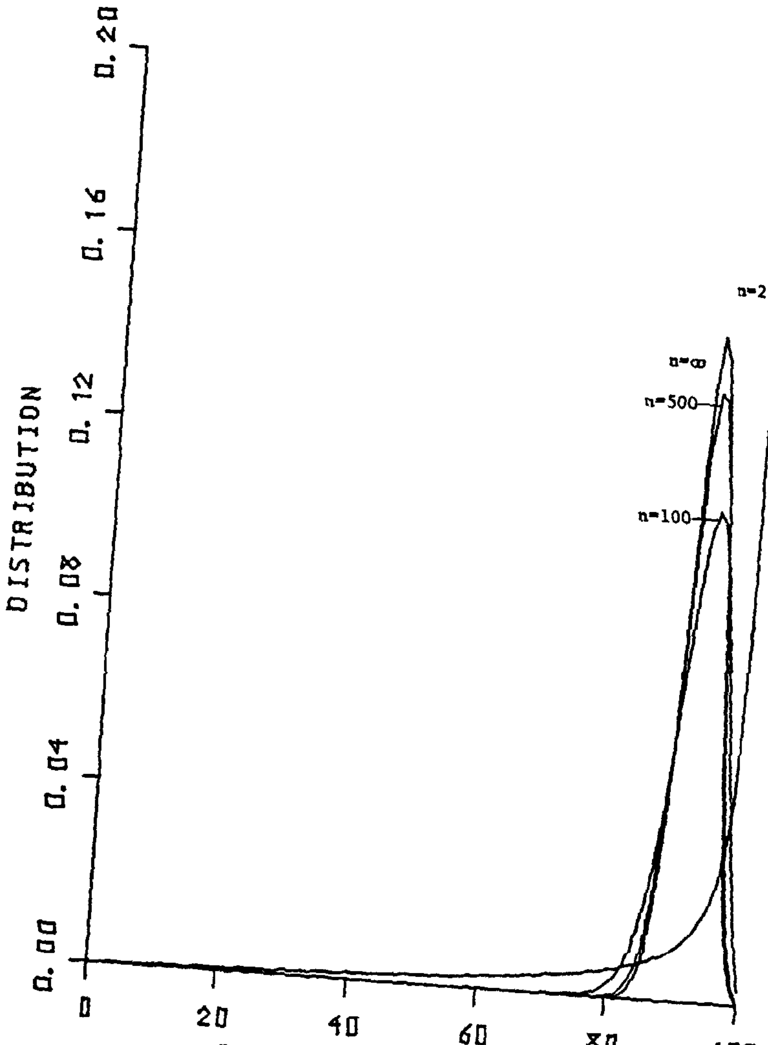

FIG. 3.—Effect of the choice of n on the distribution of employees at age 21, given a probability of 0.918 of persisting. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

subjective evaluation of what that probability should be is the best available estimate of the probability of persisting.

It also should be noted that the variance associated with the distribution of p_x^{aa} need not be the same for each age. The variance may, for example, be somewhat larger for the ages in the vicinity of the initial or full vesting ages, where an actuary might be unsure of his best estimate of p_x^{aa}. For other ages, where the impact of vesting might be slight, an actuary may have considerable confidence in his estimate and may choose a somewhat smaller variance for the distribution of p_x^{aa}.

PROBABILITY OF THE PROJECTED NUMBER OF PARTICIPANTS

The expected number of participants at a given age is the same regardless of the assumption concerning the prior distribution of the probability of persisting, but this is not the case for the probability that the expected number of participants will occur. In fact, the probability that the actual number of participants at a given age is equal to the projected number of participants is very much a function of the distribution of p_x^{aa}. Figure 4, which was derived by using equation (8), exemplifies this characteristic for 100 entrants aged 20. The values shown are interpolated values, since in most cases the projected number of participants is not integral. The curves are convex because the distribution of the number of participants is more compact at the extreme ages.

As the prior distribution tends toward degeneracy the probability of the expected number of participants occurring increases. For example, if the prior distribution of p_x^{aa} is degenerate at its mean, that is, if n_x is infinitely large, the probability that the projected number of retirees at age 65 will be equal to the actual number is 0.0962404. This is more than five times the probability of 0.0183497 obtained using an n_x equal to 2. Note, however, that even with a degenerate distribution, that is, with perfect information, the probability that the expected number of participants will occur at any age is relatively small.

While the probability of the projected number of participants at age x varies directly as the size of the n parameter of the prior distribution, it varies inversely as the size of the active population. Other things being equal, the larger the size of the population, the smaller the likelihood that the actual number of participants at a given age will equal the projected number of participants at that age. This follows directly from equation (2) if one considers the impact as the number of entrants approaches zero.

Figure 5 exemplifies the foregoing observation by showing the probability that the projected number of participants occurs at each age, given

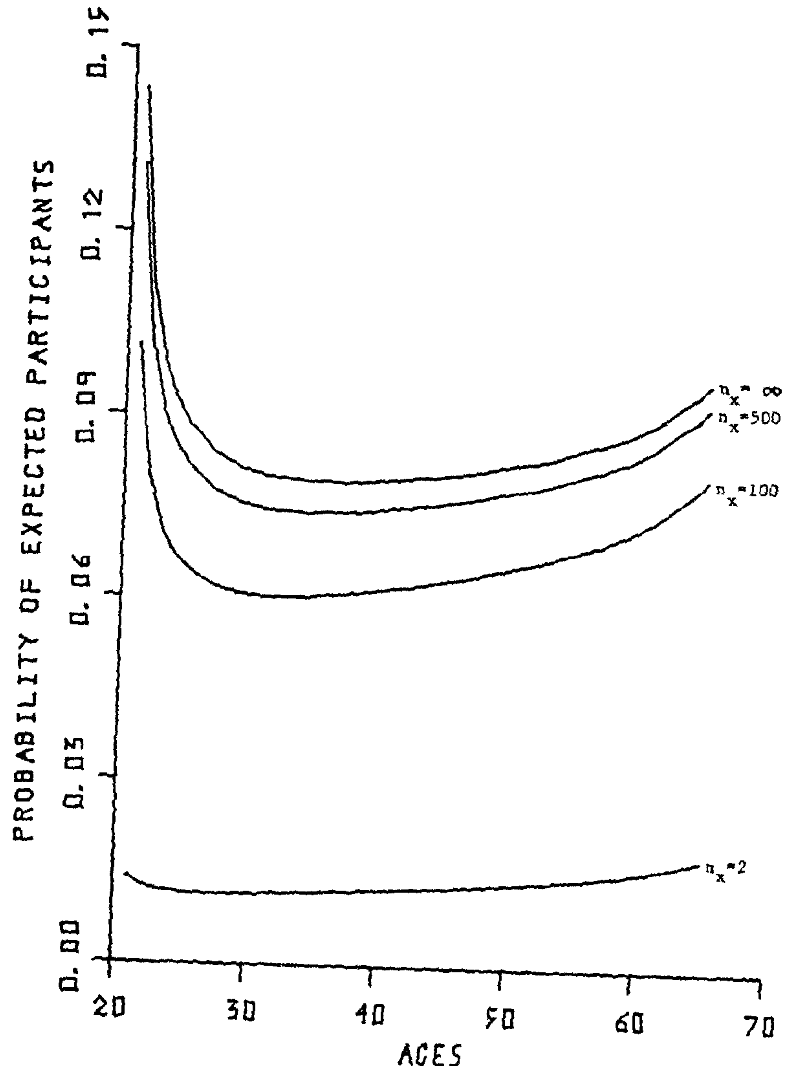

FIG. 4.—Probability of the expected number of participants, given 100 entrants at age 20 and various values of *n*. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

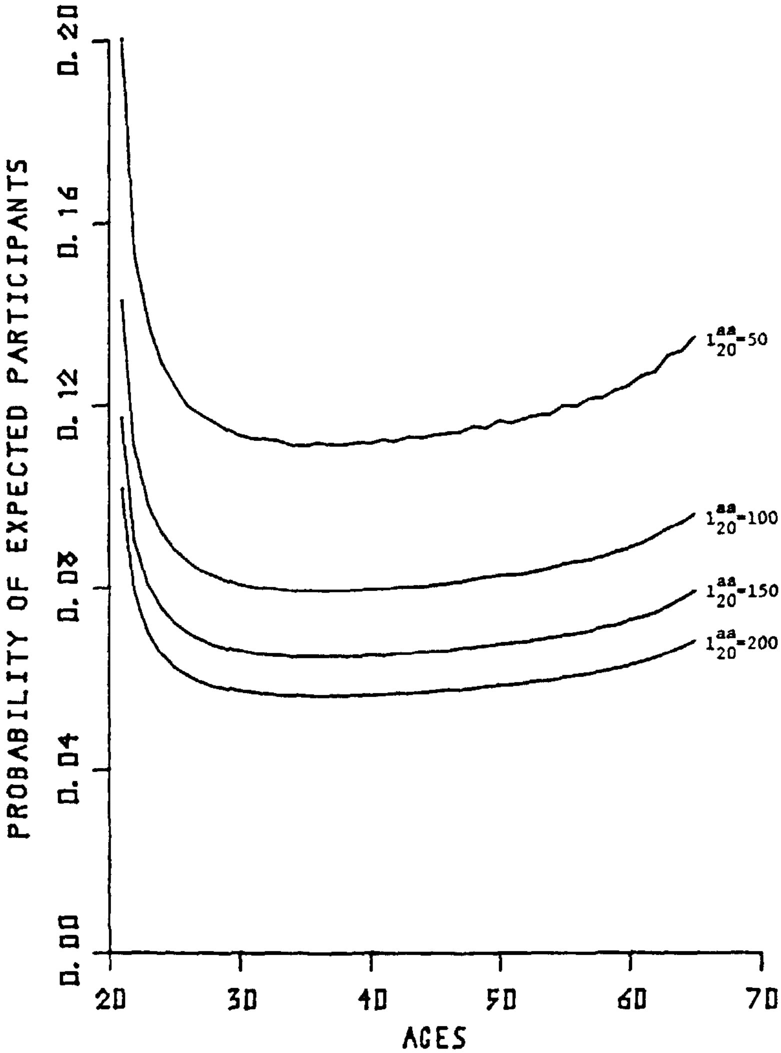

FIG. 5.—Probability of the expected number of participants, given a degenerate prior distribution and from 50 to 200 entrants at age 20. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

various numbers of entrants at age 20 and a degenerate prior distribution. Of particular note is the result that the probability that the number that retire will be equal to the projected number is approximately doubled if there are 50 entrants (0.1348059) as compared with 200 entrants (0.0684788). Thus, although a larger data base acts to increase the credibility associated with probabilities of decrement, one must not make the mistake of attributing a higher confidence to a larger exposure estimate of the expected number of participants at a given age.

PROJECTED NUMBER OF RETIREES

Consider now the application of the beta-binomial mass function to the problem of projecting the distribution of retirees. Figure 6 shows the distribution of participants at age 65 resulting from 100 entrants at age 20, given the decrement data and various precision parameters. The projected number of retirees is 21.38. It is apparent that, as the probability of persisting at each age tends to degeneracy, the distribution of retirees approaches its limiting distribution. It also should be noted that, the less credible the prior distribution of the probability of persisting, the greater the probability that the projected number of retirees will exceed the actual number of retirees.

Under a condition of considerable uncertainty, that is, a precision parameter equal to 2 for all ages, the probability that the actual number of retirements will be less than or equal to the projected number is 59.88 percent. This is because of the extremely skewed nature of the distribution of retirements under a condition of high uncertainty. Attributing a high uncertainty to the estimated value of the probability of persisting is tantamount to assuming that the probabilities of decrement may be higher than the best estimates indicate. Thus, there is considerable likelihood that the actual number of retirees will be exceeded by the estimated number of retirees. On the other hand, under a condition of high certainty, that is, a precision parameter approaching infinity for all ages, the probability that the actual number of retirements will be less than or equal to the projected number is 52.09 percent.

PROJECTED RETIREMENT COSTS

We turn now to the development of projected retirement costs.[34] This development proceeds in three stages. First, the concept of a select group is extended, both to generalize the model and to simplify notation. Next,

[34] Ultimately, studies of the stochastic nature of pension costs will encompass all elements of those costs, including such items as vesting and early retirement. These refinements are beyond the scope of this paper.

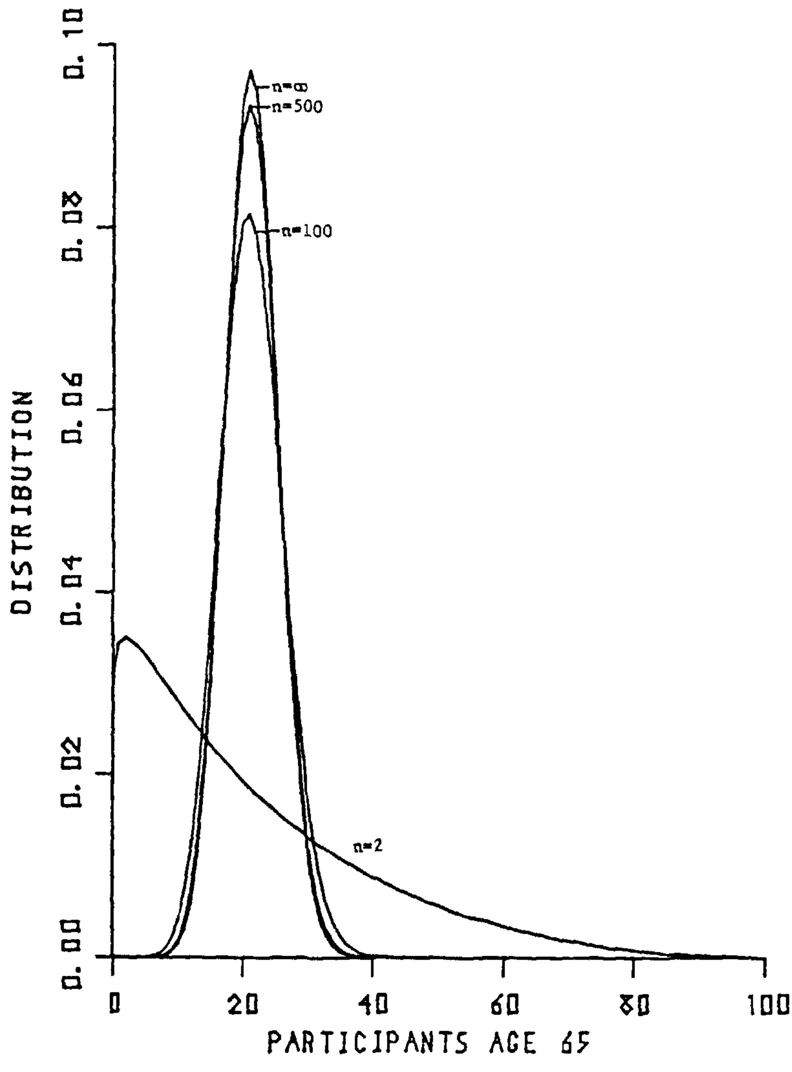

FIG. 6.—Effect of the choice of *n* on the distribution of retirees at age 65. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

the probability that projected retirement costs exceed actual retirement costs is developed. Finally, a contingency charge is introduced.

Total Attribute Groups

To facilitate the development of the model it is convenient to segregate pension populations by qualification ages. To accomplish this, first the types of qualification ages are isolated. Hence, all possible entry ages are grouped, all possible initial vesting ages are grouped, and so on. Then, each type of qualification age is partitioned by age. For example, there may be ten different entry ages, ten different initial vesting ages, and so on. Given this classification scheme, a plan participant can be assigned to a unique group on the basis of the ages when he or she qualifies under each plan provision. Let each such group of participants be defined as a *total attribute* group, that is, a group having all qualification ages in common, and let C denote the set of all total attribute groups.[35]

An example of a total attribute group is a group of active participants with an entry age of 25, an initial vesting age and initial disability qualification age of 30, an early retirement age of 55, a normal retirement age of 65, and a mandatory retirement age of 70. This particular total attribute group would be denoted by (25, 30, 30, 55, 65, 70).

The concept of a total attribute group has been implemented generously in the pension literature, albeit in a somewhat disguised form. Many papers, for example, refer to "select" groups, where the common attribute is the entry age. However, in most of the papers the other qualification ages are the same for each member of a select group, which means that each select group is in fact a total attribute group.

Probability that Projected Retirement Costs Exceed Actual Retirement Costs

The retirement cost associated with any particular total attribute group is

$$ {}^{c}\tilde{l}_{r}^{aa} \; {}^{cB}\bar{a}_{r}^{rr} , \qquad c \in C , \tag{13} $$

where ${}^{cB}\bar{a}_{r}^{rr}$ represents the present value, at the retirement age r, of the pension benefits; ${}^{cB}\bar{a}_{r}^{rr}$ is assumed to be given.[36] The probability that the

[35] It should be mentioned that the "uniqueness" of the qualification ages is to be interpreted in a computational sense. For example, two employees whose entry age nearest birthday is 20 might both be given an entry age of 20 for computational purposes, even though in fact they may not be the exact same age. Another computational convenience that often is used is to classify entry ages into quinquennial age groupings. Under this procedure each quinquennial age constitutes a unique entry age.

[36] The retired life annuity also could be regarded as a random variable. See Piper, op. cit. For the purpose of this study, however, annuities are assumed to be purchased at an annuity purchase rate of $\bar{a}_{r}^{rr}$.

projected retirement cost for this group exceeds the actual pension cost is equal to

$$\Pr \{ {}^{c}\bar{l}_r^{aa} \; {}^{cB}\bar{a}_r^{rr} \geq {}^{c}l_r^{aa} \; {}^{cB}\bar{a}_r^{rr} \} , \tag{14}$$

where a bar over a function indicates the expected value of the function. This reduces to[37]

$$\Pr \{ {}^{c}\bar{l}_r^{aa} \geq {}^{c}l_r^{aa} \} . \tag{15}$$

In view of the integral properties of the function ${}^{c}l_r^{rr}$, this latter probability becomes[38]

$$\sum_{{}^{c}l_r^{aa}=0}^{[{}^{c}\bar{l}_r^{aa}]} \Pr \{ {}^{c}l_r^{aa} \} . \tag{16}$$

Most pension plans, of course, have entrants at more than one age, so it is appropriate to extend the foregoing analysis to recognize this situation.

In general, the probability that the total projected retirement cost exceeds the total actual retirement cost is

$$\Pr \left\{ \sum_c {}^{c}\bar{l}_r^{aa} \; {}^{cB}\bar{a}_r^{aa} > \sum_c {}^{c}l_r^{aa} \; {}^{cB}\bar{a}_r^{rr} \right\} . \tag{17}$$

The solution to equation (17) is facilitated by defining two arrays: a retirement benefit array and a feasible retirement array. Let

$${}^{B}\bar{\mathbf{a}}_r^{rr} = ({}^{cB}\bar{a}_r^{rr} \mid c \in C)' \tag{18}$$

be defined as the retirement benefit array associated with the pension plan under consideration, that is, the array whose elements are the present values, at retirement, of the retirement benefits associated with each total attribute group. Additionally, let

$${}^{n}\mathbf{l}_r^{aa} = ({}^{nc}l_r^{aa} \mid c \in C)' \tag{19}$$

be defined as a feasible retirement array. The elements of this array are composed of possible numbers of participants from each total attribute group who reach normal retirement age and satisfy the condition

$$({}^{n}\mathbf{l}_r^{aa})^T \; {}^{B}\bar{\mathbf{a}}_r^{rr} \leq \sum_c {}^{c}\bar{l}_r^{aa} \; {}^{cB}\bar{a}_r^{rr} . \tag{20}$$

Assuming that there are N distinct feasible retirement arrays, it follows that the probability that the total expected retirement cost will

[37] Note that, for a given total attribute group, the probability that the expected pension cost will be adequate is independent of the benefit function defined by the plan.

[38] The function $[m]$ represents the largest integer in m.

exceed the total actual retirement cost is

$$\sum_{n=1}^{N} \Pr \{{}^{n}l_r^{aa}\} . \tag{21}$$

However, since the probability of a given feasible retirement array is simply the product of the probabilities of the joint occurrence of each element of the array, the probability that the total projected cost will exceed the total actual cost becomes

$$\sum_{n=1}^{N} \prod_{c} \Pr \{{}^{nc}l_r^{aa}\} . \tag{22}$$

Contingency Charge

The determination of the contingency charge needed to increase the probability of adequate funds to a given level follows immediately from the foregoing analysis. The only change is that, instead of defining a feasible array in terms of projected cost, one would define a contingent feasible array in terms of some multiple of the projected cost. Thus, one might define a contingent feasible retirement array as a feasible retirement array that satisfies the condition

$$({}^{n}l_r^{aa})^T \; {}^{B}\bar{a}_r^{rr} \leq (1 + m) \sum_{c} {}^{c}\bar{l}_r^{aa} \; {}^{cB}\bar{a}_r^{rr} , \tag{23}$$

where the factor $(1 + m)$ defines the multiple of the projected cost that is to be funded, and where the product of m and the projected cost represents the contingency charge.

In practice the factor $(1 + m)$ would be determined so that the probability of adequate funds attains some desirable level.

WORKING FORMULAS FOR DETERMINING THE PROBABLE ADEQUACY OF PROJECTED RETIREMENT COSTS

Given that the number of participants at a given age has a specified distribution, it is a simple matter to set down a working formula for the probable adequacy of the projected retirement cost. For a specific total attribute group, the probability that the projected retirement cost exceeds the actual retirement cost is

$$\sum_{{}^{c}l_r^{aa}=0}^{[{}^{c}\bar{l}_r^{aa}]} \sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f({}^{k}l_{r-t}^{aa} \,|\, {}^{c}l_r^{aa}) , \tag{24}$$

obtained by substituting equation (2) in equation (16). To incorporate a beta-binomial distribution, $f({}^{k}l_{r-t}^{aa} \,|\, {}^{c}l_r^{aa})$ is replaced by equation (7). On the other hand, the probability that the total projected retirement cost exceeds the total actual retirement cost is

$$\sum_{n=1}^{N} \prod_{c} \sum_{k=1}^{K(c)} \prod_{t=0}^{{}^{c}r-{}^{c}a-1} f({}^{k}l_{r-t}^{aa} \mid {}^{nc}l_{r}^{aa}) \,, \tag{25}$$

obtained by substituting equation (2) in equation (22). A working formula for the probable adequacy of some multiple of the projected cost is defined similarly.

The final section of this paper deals with the application of these formulas.

ESTIMATING THE ADEQUACY OF PROJECTED PENSION COSTS

As a first example of the estimation of the adequacy of projected retirement costs, consider the probability that the projected cost will exceed the actual cost, given a specific total attribute group. Figure 7 shows this probability for 100 entrants at age 20 and various degrees of confidence in the decrement data. These probabilities take the form of step functions, since increasing the amount of funds has an impact only at the point at which an additional retiree can be accommodated.

If the accumulated funds are less than the projected cost, the greater the variability assumed for the probability of persisting the greater the probability that the projected funds will be adequate. This is emphasized by the curve labeled $n = 2$. The opposite is true if the funds are greater than the projected cost, as shown by the curve labeled $n = \infty$. Once again, this is caused by the skewness of the beta-binomial distribution under a condition of uncertainty.

As a second example of the implementation of the beta-binomial distribution, consider the determination of the contingency charge for a plan as a whole. For the purpose of illustration, assume that entry takes place quinquennially from age 20 through age 50, inclusive, with the proportions of entrants at each age being 0.28, 0.24, 0.18, 0.12, 0.08, 0.05, and 0.05, respectively. Assume also that the total number of entrants is chosen so that, if entry were to take place annually, an ultimate population of approximately 2,000 employees would result. In addition, the benefit function is based on 2 percent of final salary for each year of service, using Salary Scale S-3 of the *Actuary's Pension Handbook.*

Figure 8 shows the results of this analysis. Once again, the probability that the projected cost will be adequate is greatest under a condition of high uncertainty regarding the probability of persisting, shown by the curve labeled $n = 2$. However, as a contingency charge is added, its impact is directly proportional to the confidence in the prior distribution of the persistency rate. The greater the degeneracy of the prior distribu-

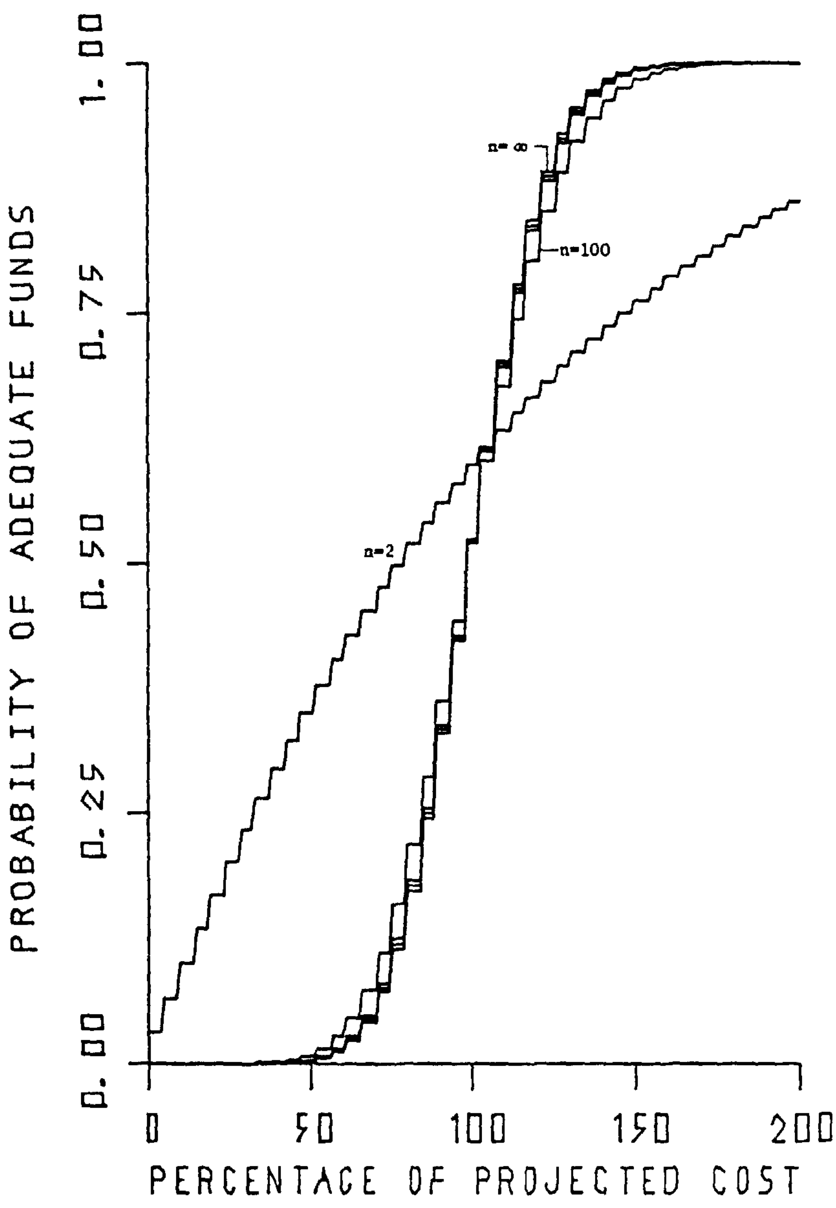

FIG. 7.—Probability of adequate funds—single-entry-age case. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, and McGinn's Turnover Table III.

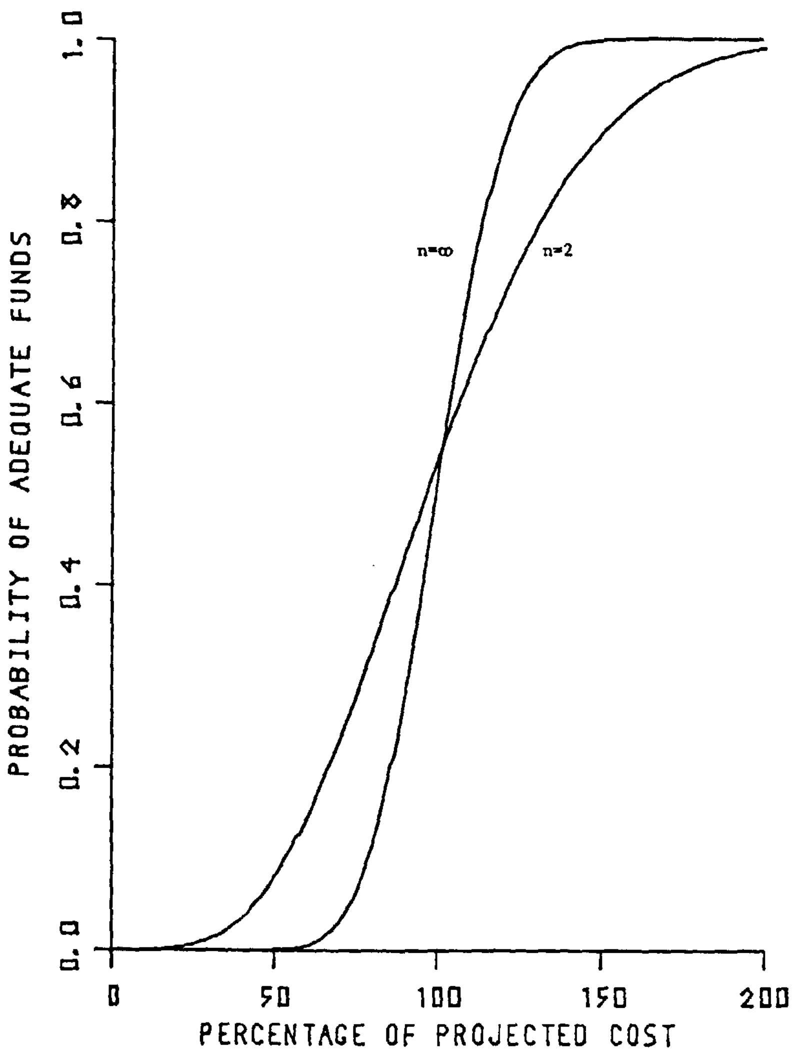

FIG. 8.—Probability of adequate funds—multiple-entry-age case. Data base: 1971 Group Annuity Mortality Table, disability rates from the 1970 valuation of the civil service retirement system, McGinn's Turnover Table III, and Salary Scale S-3 of the *Actuary's Pension Handbook*.

tion, the smaller the contingency charge needed to obtain a given probability of adequate funds. This is shown by the curve labeled $n = \infty$.

Note that, on the basis of the decrement data, even with perfect information the probability that the projected cost will be adequate to fund the actual cost approaches one-half, and that a contingency charge of 40 percent of projected cost would be required to attain a 99 percent probability of adequate funds.[39] Note also that under a condition of high uncertainty a contingency charge of 100 percent of projected cost would be needed to raise the probability of adequate funds to 99 percent.

COMMENT

While it is clear that a stochastic study of pension cost is superior to a deterministic study, in the sense that it introduces credibility into the pension cost estimates, there have been very few stochastic studies of pension costs. This fact leads one to inquire into the reason for this situation. It is clear that the technology is available; the present paper shows that such a study is possible. Hence, something other than the lack of technology has caused the lack of stochastic pension cost models. Two obvious possibilities are lack of interest and lack of resources. The first of these possibilities can be eliminated easily. For many years actuaries have been concerned with the limitations imposed by projected cost estimates. Rather than use stochastic models, however, many investigators have chosen to provide a number of cost projections based upon a spectrum of assumptions. Whether these researchers would have used a stochastic model if it were available is a matter of conjecture. However, there appears to be no question that the limitations imposed by a deterministic model are of significant concern, and it seems reasonable to rule out a lack of interest as a reason for not employing stochastic models.

A more feasible reason may be a lack of resources. In this respect, the

[39] It might be argued that a 99 percent confidence interval is too high to be realistic. However, a comment by Piper in his discussion of Menge, op. cit., p. 609, bears repeating. Piper, in discussing a 99.9 percent confidence interval, observed:

"It would be possible for fluctuations during the early years of observation to exhaust the contingency reserve and compel borrowing from some undefined R.F.C. which is assumed to be ready to lend its funds at 4 percent interest. A satisfactory answer in statistical terms to the retention question would require, it seems to me, a calculation of the contingency reserve, which would at no time exhaust the available funds.

"It is to be hoped that a further investigation can be made along this line, for a rational solution to the problem of retention limits is extremely important to companies of small to medium size."

relatively sophisticated computer program and the execution time required may be critical limitations. In addition, depending on the scope of the model—whether it includes multiple entry ages, the cost of vesting, and so forth—the size of the pension population that can be accommodated may be a limitation.

To the extent that these are the reasons stochastic models have not been employed in the pension area, they are becoming less of a constraint. Sophisticated programs have become commonplace in the insurance business. Computer core has become less of a problem, primarily as a result of virtual memory facilities. Of course, it takes time to develop an efficient model, and the development of a "most efficient" model is an area for future study. All things considered, lack of resources should prove to be less of a reason for not employing stochastic models than it has been in the past.

It is hoped that the model developed in this paper will be instrumental in stimulating both theoretical and empirical research into the stochastic nature of pension costs. With regard to the former, there are many refinements that might be incorporated into the model, including such items as a stochastic accumulation of funds and a stochastic retirement annuity. With regard to the latter, the results presented in this paper are intended primarily as examples of the implementation of the model and therefore are far from exhaustive. Future researchers should find the empirical study of the stochastic nature of pension costs a fruitful area for exploration, particularly if they have at their disposal an accommodating computer facility.

ACKNOWLEDGMENTS

The author is grateful to the Office of the Vice-President for Research and Graduate Studies at the Pennsylvania State University and to the College of Business Administration for research assistance.

APPENDIX

A flowchart that describes the calculation of probabilities over feasible arrays is shown in Figure 9. The following definitions are used:

EA = Entry age;

$F(K|M) = f_{\beta b}(K|r_{x-1}, n_{x-1}, M)$;

N = Number of entrants;

NRA = Normal retirement age;

$\underline{P} = 2 \times (N + 1)$ matrix of probabilities of feasible arrays;

X = Attained age.

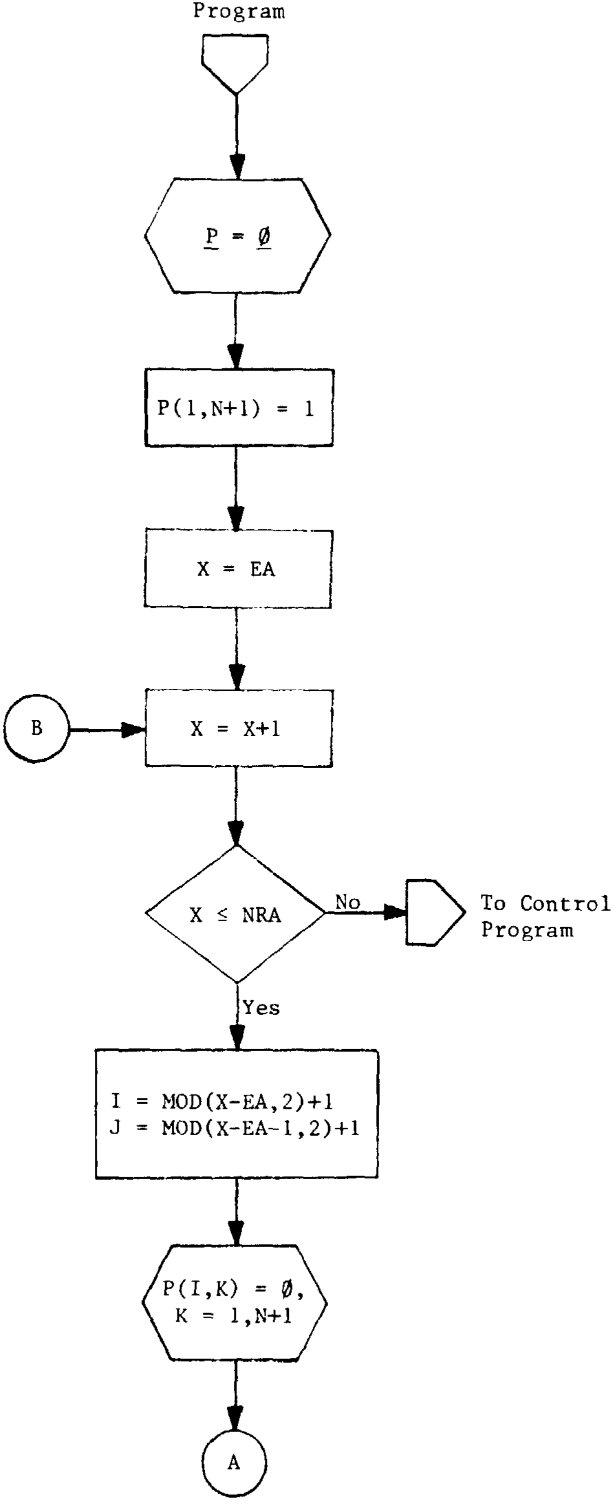

FIG. 9.—Flowchart for calculation of probabilities over feasible arrays

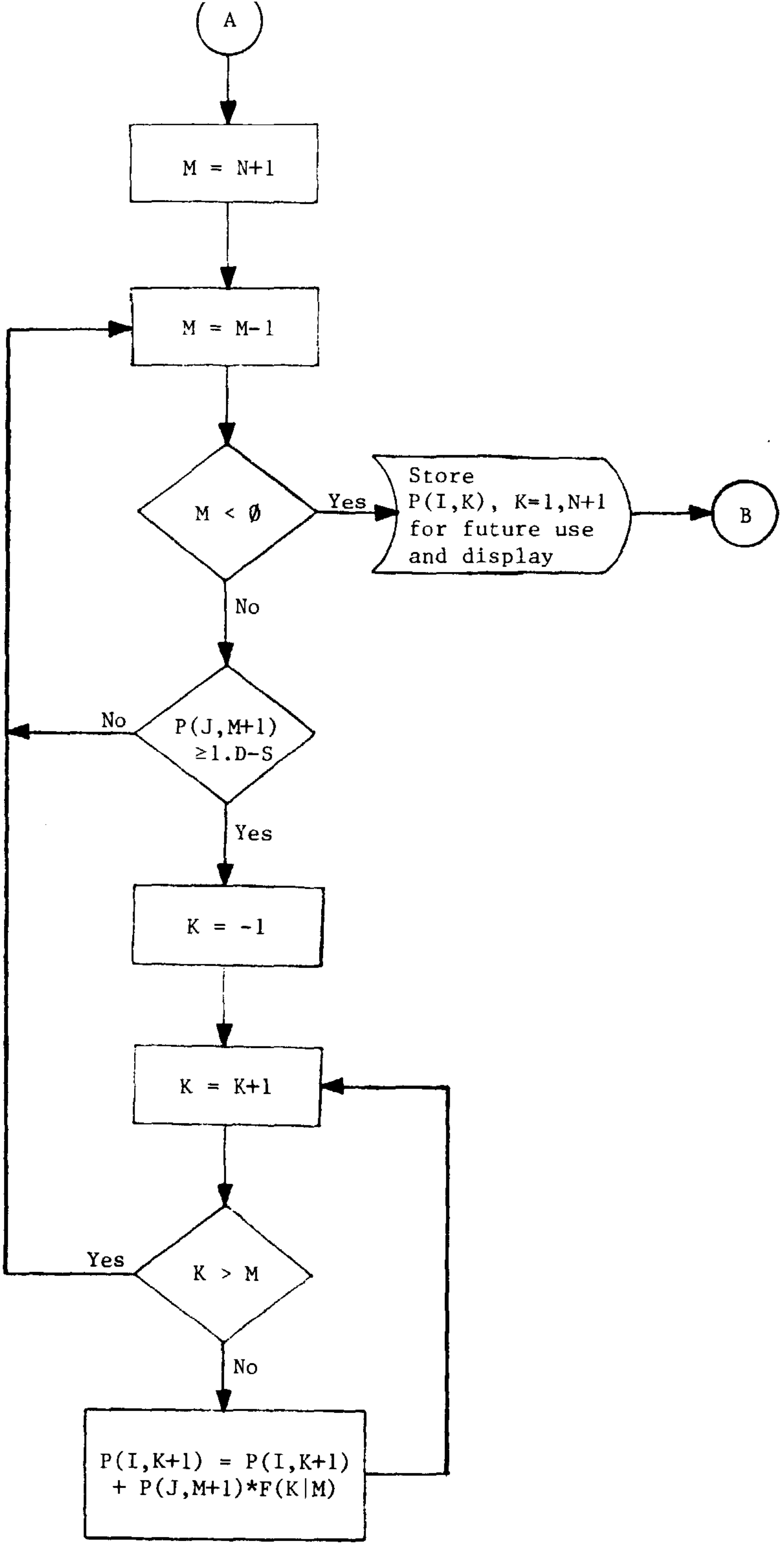

FIG. 9.—*Continued*

DISCUSSION OF PRECEDING PAPER

STUART J. KINGSTON:

It is quite natural for an employer responsible for paying the cost of a defined benefit pension plan to be concerned about the reliability of the contribution estimates. However, the very large bibliography in Mr. Shapiro's paper, and his paper itself, indicate to me that every actuary who has addressed this topic has barked up the wrong tree. For example, "The Mathematical Risk of Lump-Sum Death Benefits in a Trusteed Pension Plan" by Hilary L. Seal (*TSA*, V [1953], 135–42) contains some overambitious language and some very advanced mathematics but nevertheless succeeds in barking up the wrong tree. The overambitious language appears on page 135, where Mr. Seal says: "The purpose of this note is to investigate whether, *for a given level of annual contribution rate*, the introduction of a lump-sum death benefit in lieu of part of the pension otherwise payable, can result in a reduction of the mathematical risk due to chance deviations from expectation. The mathematical derivation of the results is given in some generality."

The overambitiousness is attested to by Mr. Cecil J. Nesbitt's written discussion of the Seal paper. On page 318, Mr. Nesbitt points out that the lump-sum death benefit that minimizes chance deviations in cost is "one under which the sum insured is equal to the accumulated contributions." So Mr. Seal, with his undoubted mathematical prowess, arrived at the wrong answer ("7.8 times the annual pension rate") on page 141, whereas Mr. Nesbitt (using the "a priori knowledge" recognized by E. T. Whittaker as quoted in Mr. Shapiro's introduction, combined with no mathematics at all) arrived at the correct answer.

What is the wrong tree up which Mr. Shapiro and all his predecessors have been barking? Perhaps an extreme example will bring out the fundamental misconception common to all these actuaries, which wastes the elegance of their mathematical techniques.

Suppose that, to compute the cost of a defined benefit pension plan with a normal retirement age of 65, no early retirement, no disability, and no vesting, I choose a mortality table that ends before the normal retirement age. The estimated contribution would be zero. Were the same table used to establish the confidence level, it would be 100 percent or close to 100 percent, regardless of technique. But my a priori knowledge tells me that the confidence level is really zero, or very close to zero. This is because using the same probabilities (or even similar ones the appropriateness of which is equally bad) to measure confidence levels is a logical fallacy, called "reasoning in a circle." Mathematical power applied in the form of a logical fallacy is wasted.

Obviously, the most important step is to choose the "best" actuarial assumptions. The technique for finding the confidence level is relatively unimportant. If the "best" assumptions are used, the confidence level is very high.

The smaller the plan, the less past experience available to draw on and the more volatile the future—all causing great difficulty in choosing the "best" actuarial assumptions. We need articles on how to approach the goal of "best" assumptions, using methods that do not require either esoteric mathematical knowledge or excessive computer expense. I do not know how to do this, especially in the case of small plans, which often have erratic experience, particularly in the area of investments and salary changes, but maybe someone not prone to reasoning in a circle can provide some guidance. If the "best" assumptions are used, estimates of random fluctuations will be meaningful to the small employer. They will help him adopt a conservative plan, the basic cost of which, plus a probable random fluctuation (for the degree of confidence desired by the employer), is within the budget.

Perhaps these goals are unattainable; if so, it would help to know that, but it is harmful to develop ultrarefined techniques for the purpose of reasoning in a circle.

BARNET N. BERIN AND KEVIN CHESLACK-POSTAVA:*

To determine a unique pension cost, the actuary works with a defined group, a specific benefit plan formula, assets, a funding method, a set of actuarial assumptions, and an amortization period. All of these are variable over time. These elements may be viewed as having distributions that are characterized more by their differences than by their similarities. Individual variances are large. There may be a dozen, or more, different actuarial assumptions, each with its own distribution changing over time.

Pulling all these disparate elements together into one multivariate frequency distribution, to enable probability statements to be made and confidence intervals to be established, is a formidable task. The resulting confidence interval may be so large as to be less than helpful. To be able to make a probability statement such as "95 times out of 100 the cost of this pension plan will be between 1 percent and 18 percent of payroll, averaging about 8 percent of payroll," is not useful to a plan sponsor, particularly if the underlying frequency distribution is questionable.

The author of this paper has collapsed many of the individual elements

* Mr. Cheslack-Postava, not a member of the Society, is a candidate for Associateship.

that go into determining a unique pension cost into one beta distribution and has considered reserves accumulated at retirement age, while concentrating on the persistency of the group. Yet the result is much the same: too wide a confidence interval. In the "Comment" section, the suggestion is made that lack of interest and lack of resources are perhaps responsible for the limited investigation of this particular approach. We suggest a third possibility, namely, that the confidence interval is bound to be so loose, and therefore of such limited value, that further studies, while of theoretical interest, have limited practical value. This is consistent with the author's two cases: with "perfect information" a contingency charge of 40 percent is necessary to attain a 99 percent confidence level for adequate funds; and with "high uncertainty" a contingency charge of 100 percent is necessary to attain the same confidence level for adequate funds. The variance of the cost pattern is an area requiring considerable theoretical investigation.

The question of whether a pension valuation conveys an accurate picture of ultimate cost misunderstands the purpose of the valuation, namely, to determine a range of tax-deductible contributions at a point of time, and to test the continued appropriateness of the actuarial assumptions through the gain-and-loss analysis. Self-correcting action is a ready by-product of diligent attention to this analysis.

The competent pension actuary does not disregard the question of plan liberalizations and their effect on costs. In fact, to do so would be quite unprofessional.

The studies that sometimes supplement pension valuations, or are undertaken before instituting change in a pension program, do not address the actuary's confidence in the estimated cost produced by the valuation; rather, they attempt to answer a host of perfectly reasonable questions in the "what if" category. If enough possibilities are allowed for (for example, a broad range from low to high for each possible assumption), heuristic probability statements might be made, using geometric areas; but clearly these are only as good as the full range of assumptions that produce the cost curves.

The paper is interesting from a theoretical viewpoint. However, some of the author's statements are disappointing as to what can be expected from present practice as opposed to the use of the model introduced in this paper.

(AUTHOR'S REVIEW OF DISCUSSION)

ARNOLD F. SHAPIRO:

A common theme of the discussants is that a principal reason that pension cost projections may be invalid is the likely inappropriateness of many of the underlying assumptions. The author does not quarrel with this observation. On the contrary, the primary focus of the paper is to explore a model that helps resolve this difficulty.

Mr. Kingston, using rather vivid prose, makes the point that he would prefer articles that deal with approaches for developing "best estimates." He notes that a best estimate is often an illusive abstraction, particularly in the small-plan area, and he concludes that any discussant who presumes the availability of a best estimate is "reasoning in a circle" and "barking up the wrong tree."

In support of his contention, Mr. Kingston fabricates what he identifies as an extreme example. I agree that his example is extreme; I do not agree, however, that it is relevant. A fundamental premise of the paper is that the tabular value used represents the "best available estimate." If a priori knowledge suggests that tabular decrement rates are absurd, one certainly will not use those rates.

It could be that Mr. Kingston has not been fully informed regarding the characteristics of the plan. One possibility is that he is discussing a pension plan for football players. Any football player who attempted to play to age 65 would, in all likelihood, be dead before that time. This is a plan design problem, of course—there is nothing wrong with the assumption.

In a more serious vein, I sympathize with Mr. Kingston. The small-plan area is a perplexing one. His comments regarding the best estimate, however, lead me to suspect that he has some misconceptions. First, ERISA does not require the actuary to use "the" best estimate, but merely "his" best estimate. Guy Shannon alluded to this notion when he remarked: "I do not feel that a 'best' contribution figure exists. I cannot seriously argue that a change of 10 percent or even 20 percent in a best estimate makes it wrong or even less good."[1] Second, the frustration expressed by Mr. Kingston is not unique to the small-plan actuary. Paul Jackson captured the essence of this frustration when he lamented that "most of the things the actuary knows about really are not very important."[2] Finally, there are a number of articles that deal with assumptions in the small-plan area.[3] A reading of some of these articles may help Mr. Kingston develop his best estimates.

[1] See Avon Guy Shannon's discussion of Preston C. Bassett, "Accrued Benefit Level Cost Method," *PCAPP*, XXVI (1976–77), 103.

[2] Paul H. Jackson, "Panel: Funding Problems," *PCAPP*, XXV (1975–76), 229.

[3] See, for example, Arnold Shapiro, "A Survey of Post-ERISA Small Pension Plan Valuations," *PASPA*, 1977, pp. 75–99. Reprinted in *Advanced Pension Planning Study Guide* (Philadelphia, Pa.: American College, 1979), sec. 14.

I must disagree with Mr. Kingston's statement that "if the 'best' assumptions are used, the confidence level is very high." The remark, by implication, conveys the impression that if perfect information were available, a pension plan that was funded to the extent of its projected cost likely would have adequate funds to meet its actual cost. Momentary reflection, however, leads one to reject this conclusion.

Consider, for example, the simple experiment of tossing one hundred fair coins. The best estimate of the number of heads that will result is fifty, but the probability of obtaining exactly fifty heads is 12.56 to 1, and the probability of obtaining at least fifty heads is 0.5398. Insofar as pension cost projections are concerned, the implications of this example are clear. Even with perfect information, the probability that the actual cost of a pension plan will be arbitrarily close to the projected cost may be very small, and there may be considerable likelihood that a plan that is funded to the extent of its projected cost is inadequately funded.

This example also has implications from a valuation point of view. Messrs. Berin and Cheslack-Postava stress the importance of a tax-deductible contribution. However, contributions that fall within deductibility ranges may not be sufficient to guarantee plan solvency. Many actuaries have circumvented this problem by using "conservative" assumptions, where, as often as not, the degree of conservativeness has not been quantified. A direct approach would be to revise section 412 to allow for a contingency reserve equal to some percentage of expected cost.

Gain-and-loss analysis, as generally advocated, is not nearly as finely tuned as Messrs. Berin and Cheslack-Postava seem to suggest. The magnitude of the gain or loss no doubt will become the objective test for determining whether the reasonableness of the assumptions should be questioned, and it likely will serve as a basis for the IRS to consider assessing an excise tax, or for allowing a contribution as a deductible item. However, IRS staff will use audit guidelines to identify significant gains and losses, and at least some of the guidelines, perhaps most, will not have been statistically validated. There is a host of questions that have never been resolved. For example, how does one segregate random fluctuations from errors in judgment, and how does one quantify the difference? Only stochastic models can provide definitive responses to questions of this type.

Messrs. Berin and Cheslack-Postava suggest that business decisions can be based on traditional "what if" analysis. However, this traditional type of analysis, while helpful, does not provide sufficient information from which to make *optimal* business decisions, since the standard procedure is to make projections rather than forecasts. Projections are the numerical consequence of the assumptions chosen. The numbers obtained are conditional on the assumptions being fulfilled: if entry and termination rates move in a certain fashion, the total impact on costs will be such and such. The cost projections are correct beyond any test against a subsequent valuation. In fact, they can be incorrect only in the trivial sense that the actuary made an arithmetic error that prevented his final numbers from being consistent with his initial assumptions. Forecasts, on the other hand, require a *quantified* statement of the actuary's confidence in the projection.

Optimal decision making requires forecasts rather than projections. Human nature being what it is, regardless of the intentions of the actuary, actuarial analysis presented as an innocent, indeed tautological, projection is accepted in some sense as a forecast of the future. The bridge between these two points of view—the actuary who is right if his assumptions hold and the client who relies on his analysis—has to be the goal for which we strive. An important step toward attaining this goal is to recognize that the deterministic models upon which most, if not all, of our technology is based, ultimately will be superseded by stochastic models.

I would like to extend my thanks to the discussants. While I did not always agree with their observations and conclusions, I am indebted to them for pointing out several possible areas of future research.