

RECORD OF SOCIETY OF ACTUARIES

1994 VOL. 20 NO. 2

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

Moderator: ROBERT W. BURGESS
Panelists: JOSEPH J. BUFF
EDWIN P. MCC AUSLAND, JR. *
MICHAEL J. MILLETTE †
Recorder: G. KEVIN GOUGH

This session will discuss the work of an industry group that is in the process of developing total return benchmarks and universes for the measurement of insurance company investment performance.

MR. ROBERT W. BURGESS: This should be an interesting panel discussion on a subject that's been evolving over the last year or so: investment performance measurement for life and property and casualty general account portfolios. We want to bring you up-to-date on all the work that has been happening in this area.

We're not talking about separate accounts or mutual funds, where the performance measurement processes are pretty much routine. Our focus is on the general account, where problems of investment performance measurement are well known. Interest is growing in finding ways to solve these problems, including the introduction of total return as the performance measure of preference.

All of the panelists have played key roles in the work done to date to address the general account investment performance issue. The panelists will take us through time—the past, the present, and the future. The shepherd of the past is Win McCausland. Win's a CFA and a vice president and managing director in Mass Mutual's Investment Management division. He's been with Mass Mutual since 1989. He's primarily responsible for working with line-of-business personnel and developing investment policies and investment strategies. Win's going to address problems he encountered at Mass Mutual in trying to solve the investment performance issue. The frustrations that arose from that experience led him to form a team with a consultant and representatives from three other insurance companies to launch the new industry survey on investment performance, which we'll discuss today.

Joe Buff, the shepherd of the present, is from Tillinghast. Joe is an FSA and is a principal in the New York office of Tillinghast; he's the firm's national practice leader for asset/liability management. Prior to joining Tillinghast in 1987, Joe spent two years at Morgan Stanley and six years at Guardian Life. Joe was the consultant brought in to help with the start-up of the project and to formulate the survey. Joe also has an ongoing role. He will talk about his role in this process and, more importantly, what's contained in the survey and what's happening with the survey as we speak.

*Mr. McCausland, not a member of the sponsoring organizations, is Vice President and Managing Director with Massachusetts Mutual Life Insurance in Springfield, MA.

†Mr. Millette, not a member of the sponsoring organizations, is an Investment Policy Officer with John Hancock Financial Services in Boston, MA.

RECORD, VOLUME 20

Mike, our shepherd of the future, is an investment policy officer at John Hancock, is responsible for asset allocation, asset/liability matching, hedging, and investment policy matters for a number of Hancock's portfolios. He also manages the stochastic modeling and risk analysis team for John Hancock. He has published a number of articles, and prior to John Hancock, was an account officer for Citicorp. Mike will, among other things, describe how his company plans to use the data generated by the survey, once they are available.

MR. EDWIN P. MCCAUSLAND, JR.: As an investment professional, I share a concern that each of you have for the meaningful measurement of investment performance within the insurance industry.

At Mass Mutual, we embarked on an odyssey about two years ago that has taken us in some new directions with some new partners. I'm pleased to share with you a synopsis of the journey that has brought us to this place in benchmarking general account investment performance.

Obviously, it takes some serious motivation to undertake a task such as this. Given prior, unsuccessful attempts to understand the relative "goodness" of one's investment performance within the insurance-industry context, this project had been described by some as a "fool's errand." Not knowing whether I was the fool or the errand boy, but understanding that our top management was frustrated with the inability to get relative investment performance data that were meaningful, we embarked on this exercise.

Everyone recognized that there were significant data problems, which I will review shortly. I'm sure that everyone in this room recognizes that a major element in improving the relative performance of general account products is improving the underlying investment performance on a relative basis. Regardless of the future inroads that variable products make in diverting new general account business, a meaningful portion of our assets still drive the policy returns for a block of business that we hope to retain for many years.

The question that none of us could answer, which was the motivating factor for this survey was, how good is your investment performance versus other companies?

We felt very good about the quality of reporting for variable products and separate accounts. Through the use of Lipper, Piper, Morningstar, and others, we knew how well each of these variable accounts stacked up versus the competition. We also knew how well they were performing versus an appropriate market benchmark, whether it was the Standard & Poor's 500, the Russell 2000, or some other index. All the pieces were in place to answer this basic question, except for the general account. For this account, we had the ability to measure performance versus market indexes, but not versus other companies.

Table 1 illustrates some of the differences between general account information and that available for other portfolios that we manage. All of this is old news to most of you; however, it is worth reviewing because it forms the basis for an approach that we believe has some merit.

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

TABLE 1
INSURANCE GENERAL ACCOUNT VERSUS OTHER INVESTMENT PORTFOLIOS

	Insurance General Account	Other Portfolios
Types of Liabilities	Liability mix	Similar liabilities
Management Objectives	Varying objectives	Similar objectives
Performance Index	Custom indexes can be constructed	Indexes readily available
Performance Measurement	Comparative performance cannot be measured from available data	Peer-group performance data readily available
Financial Reporting	Statutory reporting: -timing differences -book return	Market-value reporting: -time weighted -total return

We spent a great deal of time reviewing the body of publicly available data on insurance companies in an effort to develop some meaningful, comparative performance data. Annual statements, annual reports, rating-agency reports, research reports, etc. were all part of the review. In all of this research, there were few, if any, meaningful correlations between mix of business and asset mix, yield, duration, etc. What was missing? The liabilities and their link to the assets.

The liabilities are key elements in this puzzle. What are they? What is their duration? What asset-risk profile does a company want to assume for participating policies? Is it different for nonparticipating products? What was the timing when they were added to the company's balance sheet? (For book-value accounting, the timing means a lot with respect to the results you get.) Does the company engage in active trading or buy-and-hold strategies? The answer to each of these questions, plus others, is important in the understanding of statutory data. For each of our nongeneral account portfolios, "liabilities" are stated in terms of an investment objective. We have the ability, in each case, to eliminate the time factor from the analysis, because all returns are reported on a time-weighted basis.

Given the varying objectives that companies may have for their general accounts, each driven by answers to the above questions, the statutory data are not very useful. Clearly we can construct custom indexes against which to measure the performance of segments of our general account for internal purposes. These will give us a clear picture of how well we are doing versus "the market." However, they don't do much for us in terms of seeing how well we are doing versus our peers.

This led us to the conclusion that we needed to get inside the liabilities and find a way that would enable us to look at asset allocation by liability segment as well as asset-type performance. We thought that by working alone we would be unable to answer management's question. So what did we do?

RECORD, VOLUME 20

The initial step in the process was to retain an outside consultant, Tillinghast, to help us find a solution to the problem. We then recruited three additional companies, Aetna, CIGNA and John Hancock to work with us in developing a method to gather and evaluate the necessary information to find the answer to our question, how good is your investment performance?

At our first meeting, we spent a considerable amount of our time discussing why previous efforts had fallen short of the mark. The conclusion was that, in general, they were too detailed and died of their own weight. The amount of data sought was either too great or too detailed to be meaningful. In addition, the efforts frequently focused on book returns.

We came away from that meeting with the goal of using total rate-of-return measures, recognizing that this might be a new or difficult concept for some participants to adopt. However, it seemed to us to be the only way to break out of the trap of book returns. Our second, and perhaps more meaningful conclusion was to "work around the edges." We did not seek to answer every question, but rather to gather data that would be meaningful to participants as they began to seek answers to questions about their investment performance. Asset allocation, asset-type returns, duration mismatch, and credit risk are among the areas in which we have attempted to gather new information from all participants. Much of this information is broken down by liability segment. We believe that by building on some early successes, we can, over time, create an even more meaningful study.

The project we will be describing to you has been and will continue to be viewed as work in process. It is not a final, definitive study by any means. Many changes have already been made in response to new participant needs and suggestions. However, as stewards of a process, we have been committed to several principles that have served as useful guides. The four overarching principles are: convergence, use of industry standards, high materiality threshold, and resistance to permutations.

CONVERGENCE

The concept here is that it would be nice in an ideal world to have all participants use the same methodologies from the start, but that would be a barrier to entry for many. We agreed on a range of acceptable methodologies, with the idea that we would seek convergence on common methodologies over time.

USE INDUSTRY STANDARDS

Wherever possible, we have avoided developing new methodologies and have used industry standards such as Frank Russell/National Council of Real Estate Investment Fiduciaries (NCREIF) and the Association for Investment Management Research (AIMR) performance presentation standards.

HIGH MATERIALITY THRESHOLD

For the first year, we have attempted to keep data generation from becoming burdensome by keeping asset/liability groupings rather large. In the future, we expect to move toward finer distinctions.

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

RESISTANCE TO PERMUTATIONS

Embodied in this concept is what was discussed earlier as working around the edges. In the future, more boxes will be filled in. The goal is not to fill in every box. Rather, it is to add data when they are both available and valuable to participants.

I hope that you find this presentation of interest. We welcome your questions, comments, suggestions, and your participation.

MR. JOSEPH J. BUFF: As Win has already started to explain, in many areas in the insurance industry it's important to understand as much as possible what sort of strategies are being taken by different entities and what kind of results they're producing. It's particularly important to be able to somehow position that information for yourself relative to the industry as a whole. Having more knowledge in that area is going to enable actions that produce more added value to the company.

The 1993 Intercompany Investment Performance Survey applies to both life and property/casualty (P&C) insurers and is intended to provide an additional tool for insurance companies to compare their investment results with other insurers' results. The survey will provide information to help companies review practices and results among participants, identify their own strengths and weaknesses, and do performance attribution relative to a universe of institutional investors with similar liability constraints and regulations. Portfolio performance attribution, for those of you who don't practice in the area, means taking a look at overall results, breaking them down into pieces, and understanding why we got the result that we did, relative to some sort of universe or benchmark, very much like the traditional actuarial analysis of gains and losses by source. The reports are highly confidential—they are only available to companies that actually participate in the survey by submitting data, and the reports may only be used by each company for internal management purposes.

The scope is general account investments, but we are also including separate account assets for fixed-guarantee liabilities, such as traditional guaranteed investment contracts (GICs). We are only looking at "managed" invested assets. In other words, we are excluding some things considered invested assets for statutory purposes but where the insurance company does not really have buy/sell-type management control, such as policy loans and home office real estate. We are including assets managed by outside asset management firms, if they are for the general account or fixed-guarantee separate account liabilities and the outside firm exercises buy/sell control for the company.

One key topic of the survey is asset allocations. Asset allocation is a big determinant of overall investment performance. An understanding of mixes by asset class is important in analyzing a company's results relative to other companies. Asset-allocation data are being gathered for the whole general account and also are broken down by line of business, on both a book-value and market-value basis.

We started talking about what kind of investment measures to use for this survey and very quickly settled on total return. Total returns are defined as accrued income plus change in price, both realized and unrealized. We are gathering this information for each asset class for the general account and also in the aggregate for each line of

RECORD, VOLUME 20

business. We are not, for this first survey year of 1993, attempting to gather total returns on asset class within line of business; this may come in later years.

We chose total return because this measure captures the key factors of economic performance in an integrated way, with appropriate weighting to different things, such as new money investing and credit experience in proportion to their financial impact on the company. Total return, unlike "yield" statistics based on amortized cost, can be readily compared among different companies. Unlike statutory investment income, it is not subject to the vagaries of the time of purchase of individual assets.

In addition to asset allocations and total returns, other information is needed to properly analyze investment outcomes and compare results for different investors. The survey is collecting information on a number of items, including quality, duration, and portfolio roll-over (active trading) rates, which form key elements of performance attribution. We also want to get an understanding of the durations of the liabilities, when those numbers are meaningful, and how they compare with assets.

Another thing that we've introduced ties in with the convergence, commonality, and data validation processes that Win mentioned. There are a series of short-answer questions in addition to the actual data that are submitted by each of the participants. These explain what they did to get their numbers and where there were exceptions in their numbers to the basic standards that we had established as a group. This information will be reviewed by Tillinghast, and an overview will be provided with the survey reports. Individual data items that are not reasonably consistent with the standards established in the survey formal documentation will be excluded from the report. This information is going to be delivered back to the participants along with the numerical data in the reports.

During the survey design phase in 1993, the working group (Aetna, CIGNA, Mass Mutual, and John Hancock) working with Tillinghast defined a series of classifications for assets and liabilities for the survey. These are consistent with general practices and financial reporting rules in the insurance industry. In some cases, we suggested a preferred approach where we knew company practices vary. The basic asset classes being used in the survey are:

- Public bonds: treasuries versus corporate
- Private placements
- Tax-exempts
- Securitized assets
- Commercial mortgages
- Foreclosed real estate
- Investment real estate
- Common equity
- Cash and short-term assets
- Other managed assets

I don't think these basic asset classes are too mysterious! They're very similar to the standard practices in the industry. The preferred approaches to asset classification are given in the formal documentation, which is supplied as part of the survey process. We recognize that practices and some specific microareas will vary from company to

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

company, such as convertible bonds. Do you consider them to be fixed income or do you consider them to be equity? Some companies do one, some companies do the other. Yes, ideally, you'd want everybody to treat them the same way. We're willing in the first year to live with having companies do whichever they prefer, depending on what's supported by their current data systems, and to tell us what they did so that we can report it back to the participant group.

Another current example is our treatment of tax-exempts. We're asking that they be put in with the results as is; not be grossed up or otherwise adjusted for their tax-exempt status. That, however, is a very important issue, especially on the P&C side. P&C companies are taxed differently from life companies, and we're going to be paying more attention to that as we redesign the survey this summer to go into the second year.

On the liabilities side, we also have had to define standard classifications. We can't look just at numbers for the whole general account and think that they tell us very much. We need to break it down by major characteristic of the liabilities. Here again we must compromise. When you want to study something, the more you divide it into smaller and smaller sample cells that are more and more homogeneous and comparable with each other, the less meaningful data you have in each cell. We actuaries are used to this, going back decades, for instance, in developing mortality tables. The same problem comes up all the time in investment-performance measurement. So for the first year, we've defined about a half dozen lines of business for life and the very basic ones for P&C, commercial lines versus personal lines. In addition, we've established another "line of business" (maybe segment is a better word for it) for surplus for those companies that do maintain surplus separately with a special investment strategy for it. The lines of business are shown below:

- Individual life/health
- Individual deferred annuities
- Nonparticipating annuities: GIC versus "other"
- Participating Pension
- Group life/health
- Surplus funds—life
- Property & casualty commercial lines
- Property & casualty personal lines
- Surplus funds—property & casualty

Data from each participating company are submitted for the survey on diskette. The questionnaire spreadsheet is organized by topic. The first section covers company-wide data for life, the second for P&C, and the third section covers the breakouts by line of business. Shown in Table 2 is the top half of the first page of the questionnaire, which gathers book-value, market-value, and total-return data for the life general account major asset classes:

RECORD, VOLUME 20

TABLE 2
 FORMAT OF FIRST PAGE OF QUESTIONNAIRE

Total Invested Assets Life	Book Value (\$000s)	Market Value (\$000s)	Total Return (%)
Public Bonds	0	\$0	0.00%
Private Placements	0	0	0.00
Tax Exempts	0	0	0.00
Securitized Assets	0	0	0.00
Commercial Mortgages	0	0	0.00
Foreclosed Real Estate	0	0	0.00
Investment Real Estate	0	0	0.00
Common Equity	0	0	0.00
Cash and Short-Term (yield)	0	0	0.00
Other Managed Assets	0	0	0.00
Total Managed Invested Assets	0	0	0.00

We want companies to be able to join the survey without facing hurdles due to the limitations of their current systems and data-tracking procedures. After careful discussion, the working group decided to make some of the data input items "optional," at least for the 1993 survey. Optional items include total return on mortgages and real estate, bond total return broken down by duration and quality, and other items as indicated in the survey questionnaire. For 1993, all participants will receive the full report even if they cannot supply all the optional input data.

An important part of the utility of the survey is a set of procedures to ensure maximum consistency and validity of the data submitted. When you talk about numbers such as total return or durations, there are a number of different approaches to calculating them. When you look at the details—the practical aspects of systems administration, classification of cash flows to different lines of business within a company and its accounting systems, calculations of market value, how often you calculate them, whether you treat cash flows as actual timing or centered in the month or at the beginning of the month or at the end of the month, what assumptions you use for volatility of interest rates to price options, etc.—many different judgments go into the numbers at any one company.

The design phase prepared a white paper and questionnaire guidebook explaining the reasoning behind some basic decisions we had to make last year. In particular, the guidebook lists a series of "required" and "recommended" standards for asset-allocation, market-value, total-return, and duration statistics. Wherever possible we have followed existing industry standards. These include AIMR formulas for reporting total return, Society of Actuaries standards for reflecting credit quality in fixed-income

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

assets, and the Russell-NCREIF definitions of total return for real estate. Again, we view the survey project as a vehicle for discussion of standards and methodologies. We hope that over time approaches will converge to enhance the quality of investment-performance data for everyone. The working group has shared information with the ACLI about investment-survey procedures. We are looking forward to the results of industry initiatives to standardize pricing and total-return data for private placements and real estate. These should become good sources for some of the inputs for the Intercompany Investment Performance Survey's coverage of overall asset allocations, total returns, and supporting information.

For 1993, we are using time-weighted total returns which ignore the incidence of liability cash flows and the relative magnitude of investable funds during the year. This facilitates comparisons among companies. In the future, we may also look at dollar-weighted returns, which are a bit more complicated to develop, but they better reflect the dollar amount of investment return to the bottom line during the period.

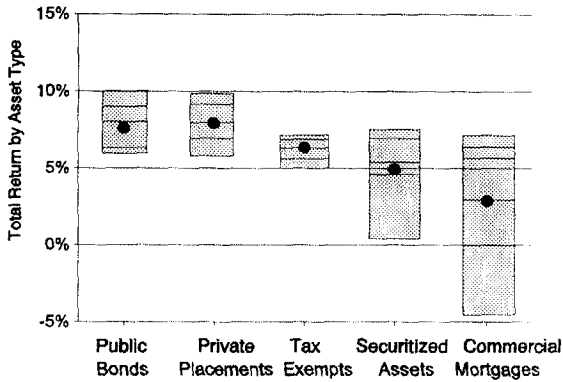
The work on the 1993 survey is going on during the first half of 1994. We have already sent participants a "flash report" overviewing some of the methods being used by the companies to prepare the input data. We will review each company's submittal for reasonable compliance with the survey standards, and we will query and resolve any important issues that arise. Each company is required to submit documentation to Tillinghast identifying areas where it has departed from our established standards.

We will be working hard to maintain a sense of "materiality" over practices that may differ among companies but don't really matter. An example already identified in the flash report is that some companies use notional transfers of assets or investment income on assets among lines of business. We would like this to be consistent, but we do not believe it's a material concern. Again, data entries that do not properly comply may be excluded from the statistics in the reports.

The reports will include commentary by Tillinghast about data standards issues and will suggest how to use the reports as a management tool. If there are enough valid entries for a particular data item, we will publish the range of results in the form of percentiles: 10th percentile, median, 90th percentile, etc. For the 1993 survey, "enough" means ten entries. This will be after excluding entries that are not valid or for which exposure is not deemed material. To protect the confidentiality of individual companies' entries, outliers beyond the 10th or 90th percentiles will not be published. We will also provide various summary statistics, such as rankings, sample sizes, means, standard deviations, etc., even if there are not enough entries to publish ranges. In addition, we will publish scatter diagrams among key items, such as total return and duration, but with individual data entries grouped to avoid disclosing values for particular companies. Each company will receive its own special version of the report, which will plot its own entries against the ranges in the survey universe. Other than this, no entries will be identified by name. In fact, to preserve confidentiality, each company's report is identified by a "secret code number" rather than by the name of the company printed on each page. Each report will include graphs generated by Microsoft® Excel, along with tables of the pivotal points of the graphs, summary statistics, and the participants' own values for the items it submits. A sample of the outputs is shown in Chart 1, Table 3 and Table 4.

RECORD, VOLUME 20

CHART 1
SAMPLE OF REPORT FROM THE SURVEY*



*Display of total return data by asset class. The circle represents the specific company's return.

TABLE 3
TOTAL RETURN BY ASSET TYPE—SURVEY GROUP STATISTICAL SUMMARY *

	Public Bonds	Private Placements	Tax Exempts	Securitized Assets	Commercial Mortgages
90th Percentile	9.96%	9.59%	7.31%	7.66%	7.11%
75th Percentile	9.00	8.63	7.10	6.89	6.22
50th Percentile	7.70	7.67	6.30	5.36	5.33
25th Percentile	6.40	6.71	5.50	4.60	3.00
10th Percentile	6.00	5.75	4.94	0.40	(4.60)
Average	7.43	7.67	6.37	4.82	3.15
Standard deviation	2.78	1.38	1.12	3.07	6.57
Standard error	0.56	0.28	0.22	0.61	1.31
Number of values	25	25	25	25	25

*Table of summary statistics for all participating companies. Percentiles correspond to bars in (a).

TABLE 4
TOTAL RETURN BY ASSET TYPE—COMPANY NUMBERS *

	Public Bonds	Private Placements	Tax Exempts	Securitized Assets	Commercial Mortgages
Company value	8.80%	6.71%	7.10%	4.60%	3.00%
Percentile	70.80%	20.80%	70.80%	20.80%	20.80%
Rank	8	16	7	16	19

*Specific company data and ranking. Percentiles correspond to dots in (a).

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

Twenty-two companies are participating in the survey for 1993—all for life and eleven for P&C. We believe we have covered a broad spectrum in terms of stock or mutual, product mix, size, and geographic location.

Here are the participating life companies; additional P&C participation is noted in parentheses: Aetna (P&C), Allianz (P&C), Allstate (P&C), Central of Iowa, CIGNA (P&C), Farmers Group (P&C), Fortis (P&C), General American, Great West, Hartford Group (P&C), IDS, International Nederlanden (P&C), John Hancock (P&C), Lincoln National (P&C), Lutheran Brotherhood, Mass Mutual, Merrill Lynch Life, Mutual of America, Northwestern Mutual, Northwestern National, Paul Revere, and Prudential (P&C).

The consultant plays several roles in the survey. For one thing, we have participated in the activities of the design phase and worked hard to recruit companies to join the survey. In particular, we serve as an intermediary for individual companies' data, to preserve the confidentiality of the numbers and to avoid antitrust problems. Attorneys from Tillinghast and the working group reviewed the project at each stage and will continue to do so. A letter of agreement covering data ownership, confidentiality, and other issues is signed by all participants after review by their own attorneys. Each company owns its own data. The working group owns the "copyright" on the survey, having cofunded the design phase with Tillinghast in 1993. As already mentioned, Tillinghast is responsible for preparing procedural and questionnaire documentation, reviewing and validating inputs, helping each company fill out the questionnaire, preparing the flash report and the final reports from the input data, checking audit letters and reviewing answers to the supplemental/methodology questions, conducting the user seminars, interfacing with the working group, etc. Consequently, Tillinghast and the working group think the survey should be viewed more as a service than as a product. The 1993 cost to each participant for Tillinghast's work is \$16,000, based on 22 participants. This cost would decline slightly if more companies joined the 1993 survey.

Now let me just say a little bit more about the confidentiality and antitrust procedures. For the most part, the data that we're getting are not public. Whatever the trends are toward mark to market and any changes in GAAP or statutory accounting, much of the data will remain nonpublic for the foreseeable future. Some of it are confidential and companies think that it's very sensitive. Consequently, we can't give out results that say Mass Mutual has this or Aetna has that. In fact, we cannot even say Company A, Company B, Company C. Attorneys were present at each of the working group meetings, and it was their unanimous opinion after studying antitrust and confidentiality issues that we had to thoroughly disguise the identity of any individual company.

Antitrust is related to confidentiality, but it's a little bit different in thrust. The point is to make sure that we don't, through the forum of this survey, begin to collaborate or reach agreements on strategies going forward that can reduce competition, restrain trade, or influence the way we price. Such a strategy might be to boycott certain markets for assets/liabilities. All of that is forbidden, and currently we're in a climate where we have to be very sensitive to antitrust issues. We want to make sure this survey is as immune as possible to these concerns, which is why we worked carefully to make sure that people don't learn of each other's data.

I should again stress that the only people who can see any of the reports are the companies that participated. They're not available under any circumstances to any other entities. The formal agreement letters that everybody has signed restrict a company's individual use purely to internal-management purposes. They cannot be shown to rating agencies. They can't be talked about to agents. They can't be used to drive advertising material. They can be used in discussion with the board of directors, including outside directors, and in discussion with nonaffiliated asset managers, but that's as far as we're letting the information be used outside the confines of the company.

As mentioned, the working group of Aetna, CIGNA, John Hancock, and Mass Mutual jointly "own" the survey. They will continue as an "outside board of directors" to review basic policy issues and ensure the quality of Tillinghast's work. We are going to gradually refine the standards over time and also raise the minimum standards for participation. Some companies expressed interest in joining in the survey but we had to ask them not to because they could not provide certain minimum required data. We're going to be taking a look later this year at what sort of minimum standards and minimum data support we need from each company and whether we will raise the hurdle a little for 1994 over 1993. It may indeed add some value to do that, although we'd be very hesitant to ask any companies that have participated this year to step out.

We all hope that the survey becomes a recurring annual process. We have received a number of suggestions for how to improve the survey for 1994 and will give them careful consideration. We welcome additional suggestions from companies regardless of whether they join the survey.

MR. MICHAEL J. MILLETTE: I'm going to talk about what we're planning on doing with this information once we get it at John Hancock. There are three main areas where we hope to use the information from this study: performance evaluation, performance monitoring, and asset allocation. Under the heading of performance evaluation, I'm going to discuss some alternative approaches, specifically benchmarks versus universes and total return versus other performance statistics.

Under performance evaluation, we hope to use the survey in two different ways: retrospectively and prescriptively. We plan to use it retrospectively to understand how well we did. Once we understand how well we did, we'll use it as a tool in incentive compensation. Also, prescriptively, we'd like to look through our investment operations and diagnose problems and opportunities.

Retrospectively, we look forward to using these data to reintroduce total return into our incentive compensation system for our fixed-income managers and analysts. We used to include a total-return measure, calculated by using the modified Dietz method. The formula is $TTR = (MV_{Ending} - MV_{Beginning} - CF) / (MV_{Beginning} + CF_{Weighted})$ where $MV_{Beginning}$ is the beginning market value, MV_{Ending} is the ending market value at the end of the period, CF is intraperiod cash flows, and $CF_{Weighted}$ is intraperiod cash flow weighted by the percentage of days in the portfolio. The modified Dietz and two alternative calculations—daily valuation method and modified Bank Administration Institute (BAI) Method, internal rate of return (IRR) based—are discussed in the AIMR's *Performance Presentation Standards* (AIMR; Charlottesville, VA, 1993).

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

The problem was, once we calculated it, we had no idea what it was telling us about relative performance by our managers. We compared the total return with a quality and duration-adjusted benchmark developed from Lehman bond market data, but still could not understand whether 25 basis points of outperformance made us superstars or just decent performers in the pack. Also, we were generally managing private placements and were comparing their performance with public-bond benchmarks, even though it was clear to us that the private market was following its own drummer. We used a contorted procedure to transmogrify our return into one not affected by liability constraints and insert it into the SEI public bond-manager universe, and after much effort, succeeded in creating a performance measure that no one really believed. We threw it out of our incentive-compensation plan last year.

The Intercompany Investment Performance Project will get us over many of these hurdles. First, a total-return universe will allow us to see performance by individual qualities and durations. We will be able to weight together returns to develop a baseline portfolio return. We will be able to measure the relative as well as the absolute performance of our managers. We will be able to compare privates with privates, mortgages with mortgages, and so on. Also, we will be able to compare our performance with that of other managers who share the constraints we live with. In future years, as the study moves on to break down returns by lines of business, we will be able to gain a more precise evaluation of relative manager performance under various specific-liability constraints.

The data do not provide us with all the data points in the first year (for example, we do not have total returns for qualities broken down by durations), but we believe that we will be able to approximate rankings with fair precision, and we look forward to achieving greater precision in the future.

You may be asking yourselves how we're going to use something that comes out in July and that has all sorts of measurement problems in incentive compensation. The answer is that this is a long-term investment. We're hoping to get results earlier each year (the long-term target is the end of January) and, in the future the study will have greater precision. Down the road, we think that this is going to be an important tool for measuring eligibility for incentive compensation for our managers.

In addition to incentive compensation uses of this data, we expect that it will be quite valuable in assessing our capabilities in various asset classes. This is the "prescriptive" use of the data. This study will tell us whether an investment manager is performing weakly. It is important for us to know this so that we can assess the reasons. It may be because our manager is involved in a strategy that is different from the universe, and that results will move dramatically relative to the universe as the vector of the market turns around. If that is the case, we would assess the likelihood of such a turnaround. If it is likely, we may rest content with the strategy. Alternatively, we may decide either that the turnaround is unlikely, or that the manager needs some remedial work. We currently use whatever data floats our way from the ACLI, Frank Russell, public market universes, and ad hoc surveys to make such assessments. The study will provide us with a systematic framework. Likewise, the survey will help us to better understand where we have competitive advantages that should be exploited more fully.

Almost needless to say, these data must be interpreted with care. Some of the data will provide only very rough comparability due to problems with data quality at other companies or due to differences in calculation measures. Much of this noise should work itself out during the first few years as we converge on standards and audit away measurement problems. Also, as mentioned, some strategies could result in performance that is orthogonal to the relevant asset-type basket, even at the level of fineness that we are seeking in our divisions of asset classes (there will be more than 70 different bond baskets). For example, interest-only (IO) and principal only (PO) bonds will be in the same basket. Obviously, a pure IO manager will be an outlier in a distribution of results of pure PO managers. The orthogonal strategy may be a great strategy, and a few years of bad relative performance may be a great idea in the long run.

The first set of performance-evaluation alternatives that I'd like to talk about is benchmarks versus universes. What we are producing here is a manager universe for insurance investment managers analogous to the Lipper universe for mutual funds or the SEI universe of public bond managers. There will be no tracking of securities in this study—only of managers. A proper benchmark is a passive portfolio composed of a list of securities defined in advance. The S&P 500 and the Lehman government-corporate index are very broad benchmarks. Benchmarks have many virtues. They provide a basis from which managers may deviate. They are investable, meaning that managers may choose to invest in the benchmark (or an approximation that has low tracking error) rather than in a portfolio of their own selection (this practice is known as indexing if a broad-based benchmark called an index is used). They can be constructed with unambiguous lists and weights of securities whose performance is measurable by managers on a frequent basis. They can be customized to reflect the precise style and bias of the manager.

Those who follow this line of thought will be quite happy to learn that Salomon Brothers and John Levy from Barrons have developed a total-return benchmark for commercial mortgages and the ACLI is working with a group of large companies (including Hancock and Mass Mutual) to develop a benchmark for private placements.

The problem with benchmarks is that they do not tell you where your investment manager stands. If you are beating your benchmark by 10 basis points, but the average manager is beating it by 30, you have below-average performance. If a below-average level of performance relative to other managers is sustained, you may want to change managers or reorganize your existing manager to get better performance because it is clearly available. In perfectly efficient markets, the average manager should perform worse than a passive benchmark over the long haul because of transaction costs, but we are dealing with many inefficient markets in the insurance industry. Moreover, it is no longer clear that any markets are efficient, even in a weak form. It is quite possible that the majority of managers could add value to a passive benchmark, and so such a benchmark may not be very helpful in identifying the profit opportunities available to your company.

Many people may question why we are producing a total-return universe as opposed to some other type of investment-performance measurement universe. Quoting from the March 1991 survey on total return by the ACLI: "Most general-account investment is an asset/liability match/spread exercise, not a total rate-of-return game. We

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

measure investment performance by looking at actual investment spreads versus target spread over benchmark curves."

Why not a universe of acquisition rates on fixed-income assets, because insurance is really a spread business these days and total return has a lot of irrelevant noise? The short and unsatisfying answer is that such a universe already exists. It is produced every month by the ACLI in composite form (private-placement commitment rates), and is available on a no-names basis in universe form as well.

The longer and more satisfying answer is that such a measure doesn't cut it. Related investment measures, such as net investment income and capital gains and losses, don't cut it either. They give your managers perverse incentives. They provide incomplete and often misleading information. Total return provides a comprehensive picture of economic value added or deducted by investment management over a period.

Let me give you an example. Many companies use acquisition spread-based incentive compensation systems. This makes sense because we are primarily in investment businesses and the spread is what matters, right? Well, the first problem is that asking analysts to beat a single-spread bogey will cause them to drop quality and increase duration. This is easily resolved by generating a matrix bogey by duration and rating. What if the ratings are wrong? Add in a clawback that reduces their excess spread if defaults are in excess of those priced for in the matrix. What if defaulted deals eventually turn around (or get worse)? Track them until they leave the company. This means tracking mortgages through foreclosure into real estate and on to sale. What about call or extension risk? Ask that the spreads be option adjusted. What if the option adjustments turn out to be optimistic? Add in a clawback for excess option losses or gains. This means tracking streams of option-based cash flows on IOs, Resolution Trust Company securitizations, sticky jump Z's, LYONS, and so on. What if the investment department sits on cash and cherry picks deals? You must develop a way to track the opportunity cost of holding cash. This must take account of desired levels of liquidity, the moral hazards presented to your allocators, and any hedging activity, etc. Essentially, by the time you have made all these adjustments you will have created a total-return measure.

The result of this sort of process will be to create two new businesses within your company, neither of which will earn any money. The first business creates and tracks these numbers. This will be an expensive business because these are complicated issues and will require the attention of very up-to-date professionals. It will require this attention because of the second business—the business of finding and exploiting methodological gaps. Some of your very best investment-management energy will be wasted at this, and I suspect that this is the more expensive part of the mix. Some people reject the possibility that their investment managers will indulge in this sort of exercise. I question whether we should expect this. We pay investment managers to find and exploit complex situations and inefficiencies in the market. Their compensation plan should guide them toward such opportunities, and the designer of the plan is at fault if it instead guides them toward loopholes. An alternative approach to performance evaluation is liability based, involving cost of funds at a simple level and dollar-weighted total return at a complex level. Again quoting from the ACLI survey: "In my opinion, any procedure that might distract

management attention from accurate and proper definition of liabilities and appropriate comparison of a company's liability profile with its asset profile would be destructive."

You need liability-based measures to understand whether your business is profitable. They are less useful as investment-performance evaluation measures because they don't measure what your investment people are really doing—competing in investment markets to achieve superior returns. It is true that they are competing on a constrained basis, but this study and more refined iterations of this study will take that into account. It is possible that your investment people could be doing great against liability benchmarks, because your liability people are great at issuing low-cost funds or else because they are underwriting well against persistency (which is the same as saying that they are issuing low-cost funds on an option-adjusted basis). The investment folks should not be compensated for that, and you do not gain any understanding of your asset business from such a measure. You would not know whether you could make more profit on the asset side.

Total return is not a substitute for asset/liability management for general-account insurance products with guaranteed-principal, and interest. It is a valuable and increasingly necessary supplement.

Total return is the best measure of investment performance because it measures investment performance on an investment basis; it cuts the Gordian knot and provides a comprehensive measurement of all economic value added or subtracted; total return-based management produces the highest economic long-run returns; and market-value accounting is the future.

I have already alluded to the first three issues. The fourth goes to the heart of many traditional objections in the industry to total return. The Financial Accounting Standards Board is clearly moving toward market-value accounting with *SFAS 115*. In such a world, total-return management is quite consistent with asset/liability management, and we will finally view our liabilities as books of securities threaded with exotic options.

All this notwithstanding, insurers are not pure total-return managers. We have accounting constraints that stem from the distinction between realized and unrealized gains. We have liability constraints on fund management. These are reasons why universes of insurance asset managers for total return are the right universes.

Total-return measures are difficult to "game" because the direction of maximum returns is not clear. In an acquisition-rate-based incentive system, there is always an incentive to overrate deals because hurdle matrices are lower at higher qualities. With total return, however, it is never clear in advance whether higher or lower quality assets will outperform.

In addition to evaluating performance, we hope to use this study to flag and explain anomalies, that is, for performance monitoring. A company that finds itself an outlier on a measure has two choices. It may believe either that there is something about its measurement or about its management that needs explaining. Such flags are valuable management tools, especially for investment-policy and asset/liability management professionals.

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

Likewise, the study may help us to explain situations that appear to be anomalous but really aren't. For example, your commercial-mortgage delinquency rates may look high relative to published measures. Your mortgage managers might argue that this is a result of an aggressive foreclosure policy and that they are producing lower loss rates than the industry. Their total return performance relative to that of the industry may help to corroborate this.

We have already found that the study has helped us to improve performance monitoring on a microlevel. Because of the number of breakdowns of the bond portfolio, we decided to move to seriatim total returns for bonds. Once you are calculating returns asset by asset, it becomes easy to flag and rectify measurement problems by grouping similar assets together and looking for outliers. Such analysis can also help you to identify winning and losing patterns in your investing.

The third thing we'd like to do with this study is improve our asset allocation. We look forward to having the line-of-business duration and asset-allocation information that this study will generate. Aside from occasional ad hoc studies, we have little knowledge of how people are investing against specific products in this industry. By knowing the distributions of asset durations and perceived mismatches across the industry for various products, we will have available the first systematic picture of how people perceive the length and optionality of our liabilities. By knowing the asset-type distribution, we will have a better picture of how people perceive product liquidity needs. A company that finds itself on the fringes on either score will have to either explain to itself why its liabilities are different from everyone else's or else reconsider its asset/liability management.

The very distribution of strategies will confront companies with alternatives that may challenge their traditional assumptions. I look forward to being surprised and to reconsidering the way we do business.

Once we develop a track record of total returns for asset classes in a constrained, insurance company investment environment, we will better be able to construct and implement optimal asset/liability management strategies. We will have a better understanding of expected returns for various asset classes.

We will also be able to quantify the cost of insurance constraints by calculating the difference between median total return for asset classes within this study and median total return for unconstrained asset managers, perhaps in the SEI universe. Armed with these shortfall numbers, we will be able to adjust expected-return figures produced in the marketplace for unconstrained asset management and use them in developing our own optimizations (or at least in reality testing our optimizations).

Also, over time, we will develop two-dimensional volatility measures by asset class and type. First, we will have the standard deviation of manager performance over time. Tillinghast will be able to calculate this from individual company data. Second, we will have standard deviations across asset managers within each period. These statistics will be helpful in establishing the efficient frontier.

We will also be able to do performance evaluation of our asset allocation. By line of business, we will use the median asset allocation in conjunction with some of the

RECORD, VOLUME 20

data on duration and asset-class performance to develop a baseline return. We will compare our own line of business return with this figure and grade ourselves on asset allocation.

In conclusion, we're going to put these data to work doing retrospective performance evaluation, prescriptive performance evaluation, performance monitoring, efficient asset allocation and asset/liability management. We have felt the need for these tools at Hancock during the past few years, and we look forward to using them to make money in the future.

MR. BURGESS: You've heard where we've been, you've heard where we are now, and you've heard what one company plans to do with its results. Could I have your questions and comments please?

MR. JOHN D. MURRAY: I think the work that you're doing to enable you to measure investment performance against other insurance companies, instead of external indexes, which is currently a downfall, makes a lot of sense. I am concerned, though, within our company and others about the recognition of the liabilities. Like it or not, the liabilities are at book value, and like it or not, we do have obligations to pay at certain points in time. You mentioned, early on, Joe, that liability data were being gathered. The only characteristic that I think you talked about was duration measure. I also know that there's a lot of hedging that can be done to keep that in place. Could you speak more to what kind of detail you're getting on liabilities and how you're going to recognize these obligations?

MR. BUFF: There's a limit to what we can do in the first year, but underway are some plans that go back to the earliest discussions of this, which we think we will phase in gradually. Again, the issue has been to get at something useful this year, concentrating on investment performance and then to gradually build and mature the survey as we go forward.

One idea that we have had is to take a look at one or more of the techniques for comparing investment results to some sort of liability bogey, or liability-based index, or liability cost of funds, whether it's on a spread or more of a market-value perspective. This way you can track the impact on bottom-line profitability and wealth accumulation for the company by looking at the things that are happening on both sides of the balance sheet. One reason that total return is certainly not a perfect measure is it looks at all of the changes in price due to various factors and treats them all as being equally relevant and important to any particular investor or insurance company. However, there are some aspects of what makes total return what it is that matter to different companies more than to others. For instance, the buy-and-hold company probably doesn't care too much about changes in price due to general changes in treasury rates.

However, a company that does actively trade does care. We make the case that total return is really not a perfect measure any more than statutory income is. But having both available solves some of the weaknesses of each and enables companies to have more tools to do the job. For instance, in the area of thinking of a company as a spread manager, you can do that on the basis of yield versus credited rate, and you can say that's all there is to the picture. Or you could think more about the

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

underlying economic behavior based on the concept that assets and liabilities both have future cash flows, which imply that they both have market values and embedded options. Even if those things don't show in current statutory or GAAP reporting, we all know they're there. We all know that they're going to have an effect on the company. If you ignore things such as option risk or credit risk until losses have already occurred, then that's subtracting value for the company, not adding value. We've created a framework to address that issue more keenly in future surveys. The framework has been started in the 1993 survey by looking at asset allocations and total returns separately on different lives of business.

FROM THE FLOOR: If we're looking at this as a measurement of investment performance, I haven't really heard anything about the risk involved. You're breaking it down by asset classes, but even within one asset class, let's say, private placements, we could have private placements of small companies versus larger companies. There may be an incentive for a manager to get a higher total return by taking more risk. What are you doing to address that problem?

MR. MILLETTE: That was just a function of the level of detail shown. Actually, for each of the asset classes that you saw, we have breakdowns by quality, by duration, or by other important characteristics. In the bond portfolio, there are more than 70 separate baskets, and most participating companies actually succeeded in providing information for the bulk of them. You will be able to compare total returns for BAA private placements with those of other BAA private placements. Or compare the returns of seven-year-duration bonds with the returns of other seven-year duration bonds. In the future, we'd like to produce even finer breakdowns, such as returns for specific qualities by durations and so on. We haven't asked for such statistics at this point because to have done so could have caused the survey to collapse under its own weight.

MR. BURGESS: I'd like to make a comment on the incentive angle that Mike talked about so forcibly in his comments. Speaking for at least one other company, CIGNA, we aren't ready and we have no plans, even a year from now, to turn this product into some kind of incentive compensation tool. We are still a buy-and-hold organization for our general-account monies. We are very interested, though, in total-return measurement, even if we're not doing total-return management for many of our monies, because as you've heard from the panelists, there is the economic bottom line that is represented by your total-return measurement. We'd like to see what our buy-and-hold approach is doing relative to that bottom-line economic measure. Having said that, I would expect that we'll have all kinds of motivation and perhaps some better tools to eventually move to more active management and some incentive compensation accordingly using this kind of measure.

MR. DAVID A. Y. CAMPBELL: I have two questions. First, where do immediate annuities fit in your business breakdown? Second, the list of companies that you showed participating in this survey struck me as being rather restricted. How are you planning to increase that sample during the coming years?

MR. BUFF: Our basic plan is to get more companies to join the survey. To explain a bit more what we did in the recruiting process, we first had a mailing campaign to the top 100 life and the top 100 P&C companies, as identified by data from *Best's* for

RECORD, VOLUME 20

the U.S. Once we sent out some formal documentation, we had a follow-up telephone and personal-visit campaign, where we worked to connect with the chief investment officer, at least over the telephone. The companies that came on board in the first year were the ones that felt a bit courageous about wanting to try to do the work. Most were interested in seeing what kinds of information they could get, maybe because we caught them at a particular point in their own process of addressing these issues.

It's interesting to see that there were some differences in why companies on the life side and on the P&C side said that they didn't want to join for 1993. On the life side, it was almost entirely because they didn't have the basic minimum data. Some companies could not give us good, robust, total-return data for the general account for life. However, several companies have privately given us firm commitments from their chief investment officers to join the survey for the calendar-year 1994. This is designed to be a very participative process and we hope that more companies will join. The more that come in, the more useful the information will be for everybody.

We're also following some other publicity processes, such as talking about this at conferences like this one and sending out more information in the fall.

MR. MILLETTE: Immediate annuities are a nonparticipating annuity; that bucket is subdivided into one bucket for GICs and another bucket for immediate annuities, structured-settlement annuities, and pension buyout annuities. All are long-dated, fairly certain, cash-flow liabilities.

I want to add something to what Joe was saying about the recruitment process. This sale process is important, not just to Tillinghast, which is actually the employee of the participating companies, but it's also important to all of those participating because an investment-manager universe is like a party. It only works if many guests show up. We need other companies to join the survey so that it will provide a statistically reliable picture of insurance company investment management. Only then will it be usable for the purposes that I described.

MR. BUFF: One last thought: the formal documentation that we developed during the project is available. If anybody is interested, this is public information. Just let one of us know and we'll see that you get a copy quickly. Chances are that your chief investment officer or somebody in his or her area did get this some time in the last few months.

MR. STEVEN H. C. LEE: I have a question about some of the information you mentioned, such as assessing performance on mortgage loans or something like that. It seemed like an extended time period might be valuable. I wonder if you saw, going forward, any possibility of chaining the results together for the companies that were involved over a longer period of time.

MR. MCCAUSLAND: The short answer is yes. At our very first meeting, we recognized that in the first year or two, certain valuation problems that might exist at a particular company might impact the returns that it reported. But we thought that its valuation problems would begin to disappear over time and become less meaningful. As you begin to link together multiyear returns, you would probably get better

MEASURING INSURANCE COMPANY INVESTMENT PERFORMANCE

data. That's why we think this becomes more valuable over time and why we're hoping to build the universe of participating companies.

MR. BURGESS: It's more valuable over time for a couple of obvious reasons. First, linking together multiple-year results will give you a better sense of the investment performance. Holdings are obviously going to behave differently at different points of the interest rate cycle or business cycle. Second, convergence, which was discussed along the way, is more operative during a longer time period.

MS. LINDA M. LANKOWSKI: Why is this only to be used for internal purposes? I understand why not to send it to the policyholders, but why can't we share this with rating agencies?

MR. MCCAUSLAND: We were concerned that companies might not participate if they thought the data were being used in any way outside the group of companies that were putting it together. We were also concerned about the control of information or giving information to rating agencies for some companies when other companies might not have it. Would you be criticized because you were giving it to them, or would you be criticized because you didn't participate? You should have people participating because they're genuinely interested and they're doing it for the right reason, which is to measure performance, not to make them look better or worse in a relative universe. We thought that if it was used for marketing purposes, it might distort or bias the figures.

MR. BURGESS: Having said that, would any of the panelists disagree that there will be a loosening of the restrictions on the use of the results over time, assuming the survey is successful?

MR. MILLETTE: I think that what we're trying to develop and what we hope this will become over the long term is an analog to the Lipper universe within the insurance industry. We hope this study will grow into that, and it will become a performance study for industry investing that everyone will participate in and for which there will be considerable moral suasion encouraging everyone to participate. We hope we won't have a situation in which companies that are doing poorly will feel free to drop out. When we feel as though we have reached such a level of moral suasion, the time will have come to make this much more public. But that point is some years down the road. Until then, the study must be confidential.

