A Discretionary Wealth Approach for Investment Policy

JARROD WILCOX AND FRANK J. FABOZZI

JARROD WILCOX
is the president of Wilcox Investment, Inc., in Newton, MA.
jwilcox@wilcoxinvest.com

FRANK J. FABOZZI
is a professor in the practice of finance at the Yale School of Management in New Haven, CT.
frank.fabozzi@yale.edu

In this article, we suggest a paradigm for broad investment policy. Our main contribution is the development of a more robust model that can work well outside the contexts in which Markowitz mean-variance optimization is effective. We first briefly review a point-estimate model based on growing discretionary wealth. We then improve that approach using Bayesian logic. Risk in the outcomes of saving and spending plans is treated in parallel to risk in the investment returns and can be fully comprehended by a Bayesian approach. This is particularly useful, for example, in retirement planning where longevity risk is an issue. For many other cases in which investment funding and spending outcomes are nearly certain, the investment implications of a fully Bayesian model are more subtle and its application may not produce big effects. There appears to be material potential improvement, however, when long–short positions, high leverage, long time periods between rebalancing, or option positions are present. We will demonstrate this in the case of long–short portfolios.

The original discretionary wealth approach of Wilcox [2003] is a repeated single-period model built from an extended balance sheet. Implied assets include time-discounted values of predicted future cash flows to the investment portfolio. Implied liabilities include time-discounted foreseeable cash withdrawals from the investment portfolio. Discretionary wealth is the surplus of assets over liabilities, including these implied augmentations. As an accounting concept, it is akin to intangible book equity. The investment objective assumed each period is to maximize the expected log return of discretionary wealth.

Under conditions in which successive innovations in logarithmic (log) discretionary wealth are bounded and independent, the Feller–Lindeberg Central Limit Theorem applies even when returns (and innovations in saving and spending plans) follow different probability distributions in different time periods. Consequently, one would maximize not only the mean but also the median log return for the long run, and thus maximize the median discretionary wealth achieved without interim shortfalls. In practice, these conditions are not met perfectly; sufficient leverage may make log discretionary returns unbounded, the number of periods is not infinite, trading costs and autocorrelation of investment risk exist, and investment results in one period affect discretionary wealth in the next period, consequently affecting appropriate risk aversion and subsequent investment allocations and results. Nevertheless, in comparison with conventional Markowitz mean-variance optimization, especially based on subjectively chosen risk aversion, this approach appears to improve median discretionary wealth after many periods, while reducing the probability of interim shortfalls.

The discretionary wealth approach has several first cousins in the investment literature. In attempting to provide a more objective source for investor risk aversion, it shares goals with the application of constraints on probability of loss, as in Stutzer [2000, 2003]. In maximizing growth of discretionary wealth rather than of the investment portfolio per se, it shares antecedents with "fractional" Kelly strategies (see MacLean and Ziemba [2006] and Thorp [2006]). Its use of variable risk aversion is also related to the Black and Perold [1992] and Perold and Sharpe [1995] formulation of constant proportion portfolio insurance (CPPI). Its use of an extended balance sheet is related to asset–liability management in general, as in Ibbotson et al. [2007], but does not describe human capital as such or give it stock or bond attributes. It is formally similar to the work of Rubinstein [1976] in his description of generalized logarithmic utility, the generalization being the inclusion of an offsetting quantity to the investor's wealth.

Although the discretionary wealth approach is not useful in some circumstances, it appears to offer interesting advantages. Factors that favor this conclusion are transparency and easy customization, as well as its ability to provide an appropriate risk-aversion parameter absent from Markowitz's model and its potential to include higher-moment investment risks not captured by return variance. However, we believe the discretionary wealth approach can be improved considerably through the incorporation of Bayesian logic. This addition takes into account imprecision in specifying both investor and investment return characteristics and provides a path to more accurate and robust calculations where higher moments are concerned. The financial literature, for example, summarized in Fabozzi et al. [2007], has noted the advantage of Bayesian priors and estimation methods in making the inputs to investment decisions more robust. A few pioneers have gone further by making use of Bayesian methods for utility-based optimization, as in Harvey, Liechty, and Liechty [2008]. We applied a similar technique to the extension of the discretionary wealth model, but generalized it to include risk in investment portfolio funding and spending plans, risk that makes the appropriate application of risk aversion probabilistic.

In its current state, the model we present in this article appears suitable for broad asset allocation policy, but not yet for detailed management of securities. As in the original Markowitz model, it requires modification to limit the impact of trading costs over multiple periods. Further, the exercise of Bayesian logic, though conceptually simple, is expensive in computation and has not been thoroughly demonstrated for allocation problems with many assets.

The scope of this article is restricted to the consideration of tools with which to explore better broad investment policies given what is known of current saving and spending plans; that is, the investment planning process may be a single step in a more comprehensive iterative process needed to bring consumption plans into effective coordination with investment decisions. Accordingly, we assumed in each period that any revisions of spending or savings plans had already taken place before we addressed investment questions, understanding that the investment solutions will often be provisional because they depend on consumption plans that may prove suboptimal in this larger context. Note also that we are not trying to describe investors as they are, as we would from a behavioral finance viewpoint, but as they ought to be to achieve a specified objective. Though we recognize the applicability of behavioral finance in practice, we do not address the compromises a well-informed investment manager may find advantageous in providing services to a naïve investor, as stated by Brunel [2006]. We also do not confront agency problems. Finally, we applied in our examples simple Monte Carlo simulation rather than more advanced Markov chain Monte Carlo (MCMC) methods.

The article is organized as follows. In the next section, we illustrate a simple investment policy problem for a specific investor, first categorizing the investment opportunity set as two assets (cash and a bond-equity balanced fund), then expanding the alternatives to include cash, a bond, and two equity assets. We show how the original discretionary wealth approach based on point estimates can be applied to provide the risk-aversion parameter unspecified in the original Markowitz approach. Then we point out several areas in which practical problems remain to be overcome.

The implications of allowing decision inputs and subsequent intermediate calculations to involve Bayesian probability distributions rather than compressing them to point estimates and thereby losing information, especially when higher moments are subject to nonlinear transformations, are then discussed. This allows us to use Bayesian logic to construct a probability distribution of log return on discretionary wealth as a function of portfolio weights before selecting the best portfolio. Having mastered that practice, we can easily extend it by allowing appropriate risk aversion dependent on saving and spending plans to also be represented as a probability distribution. This will allow us to take into account imprecise estimates of future sources and uses of investment funds such as those derived, for example, from longevity risk.

We then employ our Bayesian version of the discretionary wealth approach to examine an asset allocation problem similar to that described in the next section under the more realistic conditions of imprecision in the knowledge of investment return parameters and also of the investor's extended balance sheet as subject to longevity risk. We compare the results with the original Markowitz mean-variance optimization under circumstances favorable to conventional optimization, that is, where higher moment effects are minimized. Although we have biased the test designs to favor conventional mean-variance optimization, even in these cases, the Bayesian approach is shown to be meaningfully superior as we extend problems to include uncertain future spending and long–short portfolios.

BRIEF REVIEW OF THE DISCRETIONARY WEALTH APPROACH

We define discretionary wealth in terms of an extended accounting balance sheet, as in the example shown in Exhibit 1. On the left side of the example in the exhibit are investment assets of \$10 million plus the time-discounted value of foreseen financial contributions to the portfolio of \$1.5 million. This latter present value is an implied asset, that is, an asset implied by expectations of planned future savings before retirement. On the right side are current investment-related debts of \$2.2 million plus a \$5.3 million present value of foreseen financial commitments that must be satisfied by withdrawals, in this case, future withdrawals after retirement. In this case, both present values are time discounted using real, after-tax time discount rates. We term *discretionary wealth* the residual or surplus of \$4 million on the right side of the extended balance sheet after subtracting from total assets the current and implied (expected) liabilities. In practice, lack of certainty produces a fuzzy boundary between liabilities implied by expectations of future cash flows and what remains as discretionary surplus. In this section, the example amounts are sharply defined. In the next two sections, we will allow balance sheet quantities implied from expected future cash flows to have probability distributions.

EXHIBIT 1
Example of an Extended Balance Sheet (millions of dollars)

Investment Portfolio	10.0	Investment Debt	2.2
Implied Assets	1.5	Implied Liabilities	5.3
		Discretionary Wealth	4.0
TOTAL	11.5	TOTAL	11.5
Implied Leverage	2.50		

Implied Leverage = *L* = (Investment Portfolio/Discretionary Wealth) = 10/4 = 2.5

The ratio of the value of the investment portfolio to discretionary wealth we term *implied leverage*, abbreviated to leverage, L. Accordingly, the discretionary wealth approach criterion that we want to maximize is the expectation of $\ln(1 + Lr)$, where r is the single-period return on the investment portfolio, L is leverage, and ln is the natural logarithm function.

An important qualification of the approach is that in cases where there is an extremely minute probability of a 100% loss each period, the investor with a limited lifetime may appropriately disregard its full implications for investors with infinite lifetimes. This may be taken into account, for example, by censoring such losses to be no greater than 99%, which will be sufficient to cause avoidance for all but the tiniest probabilities.

The discretionary wealth approach typically provides a mild form of dynamic hedging that reduces, though it does not eliminate, the probability of discretionary wealth becoming negative as the result of investment returns. In practice, this feature should be used with limits on trading, for example, as specified by Donohue and Yip [2003], to avoid excessive trading costs. The improved Bayesian model in the next section further reduces the probability of negative discretionary wealth because a wider spectrum of potential return characteristics is taken into account.

Quantifying the Impact of Higher Return Moments

In Exhibit 1, L, the ratio of the value of the investment portfolio to that of the surplus or discretionary wealth, is 2.5. That is, in this example if we have a 10% loss in the investment portfolio, it results in a 25% loss in our discretionary wealth. Both the mean and the standard deviation of the arithmetic return are multiplied by the implied leverage L when we remap them from the investment portfolio to discretionary wealth. This means that variance (denoted by V) is multiplied by L squared; the third moment, involving skewness (denoted by S), is multiplied by L cubed; the fourth moment, involving kurtosis (denoted by K), is multiplied by L raised to the fourth power. Assuming that leveraged log return is bounded and that L is fixed, a Taylor series for the expected log return on discretionary wealth is shown as

$$\text{Expected}\ln(1+Lr)\cong\ln(1+LE)-\frac{L^2V}{2(1+LE)^2}+\frac{SL^3V^{\frac{3}{2}}}{3(1+LE)^3}-\frac{KL^4V^2}{4(1+LE)^4}+\ldots \quad (1)$$

Equation (1) makes clear the functional relationship of expected compound growth to leverage. Begin with variance V considerably smaller than expected return E. As implied leverage L is increased, the initial impact is to improve compound return, but with further leverage increases this relationship eventually peaks and turns downward. With still further increases in leverage, expected log return for discretionary wealth becomes negative and ensures eventual shortfall.

The investment literature contains investigations of portfolio optimization using the first three or four return moments of a Taylor series approximation for investment return.[1] The discretionary wealth approach, however, goes further by introducing an investor's implied leverage, L, and, therefore, through Equation (1) supplies an appropriate risk aversion for each return moment. That is, dividing through by L to put ourselves in the Markowitz context, the appropriate risk aversion for the second return moment is approximately $L/2$, the appropriate risk aversion for the third moment is approximately $L^2/3$, and the appropriate risk aversion for the fourth, or kurtosis, moment is about $L^3/4$.

Equation (1) also highlights the combination of high leverage, high variance, skewness, and kurtosis as a *joint* determinant of the relative strength of higher moments. This expands the usual research interest in return skewness and kurtosis to a broader list of interacting factors, making clear the relevance of higher moments to households with marginally adequate investment assets to support their future spending commitments. Though the investor can frequently limit the relative impact of higher return moments by shortening the length of periods between investment decisions, this ability may not always be practical, especially for the household investor and in cases where transaction costs are material. At the other end of the spectrum, investors who can afford to invest in illiquid investments, such as limited partnerships, greatly amplify their exposure to return negative skewness and fat-tailed kurtosis.

Relationship to Markowitz Optimization

The exploration of appropriate risk aversion previously noted can be translated directly to the Markowitz framework with minor approximation. Continuing to assume that L is fixed, suppose that LE and L^2V are small. Then, an approximation of Equation (1) implies that

$$\text{Expected}\ \ln(1 + Lr) \sim LE - L^2V/2 \quad (2)$$

Whatever investment portfolio optimizes the value on the right-hand side of Equation (2) will also maximize the value on the right-hand side of Equation (3),

$$(1/L)\ [\text{Expected}\ \ln(1 + Lr)] \sim E - LV/2 \quad (3)$$

This approximation of the discretionary wealth approach leads to the objective function of Markowitz mean-variance optimization, with $L/2$ providing the risk-aversion coefficient. The discretionary wealth approach specifies where the investor should be on Markowitz's efficient frontier of best return for a given degree of risk.

Simple Examples Using Point Estimates

Consider the investor whose extended balance sheet is represented by Exhibit 1. Suppose the investment opportunity set is a safe fund with practically no risk and a 1% annual return after tax, together with a riskier fund with an annual after-tax risk of 15% and an expected one-year after-tax return of 6%. The covariance in returns between the two funds is approximately zero. From Equation (2), expanded to deal with two assets and simplified to fit these facts, we obtain an optimum allocation for the risky fund,

$$W_R = (E_R - E_S)/(LV_R) \quad (4)$$

Given the investor's implied leverage of 2.5, we can then calculate a weight of $(0.06 - 0.01)/(2.5 \times 0.0225) = 89\%$, leaving 11% for the safe fund.

Now, we introduce a slightly more detailed example. A 55-year-old woman with an after-tax cash inheritance of \$10 million desires a \$300,000 after-tax income for her remaining lifetime in constant dollars after inflation. An after-tax, after-inflation interest rate is estimated at 2.5%. Her actuarial life expectancy is 27.7 years. A present value calculation of this implied liability as an ordinary annuity is about \$5.95 million. Assuming no other investment assets or liabilities, her discretionary wealth calculated on the basis of mean lifetime would be \$4.05 million and her implied leverage would be 2.47.

A more relevant point estimate, however, is that for the mean implied leverage rather than the implied leverage calculated from mean lifetime. Taking into account the probabilities for each lifetime from actuarial tables and using those to calculate a probability distribution for implied leverage, the leverage distribution is skewed with a long tail toward higher leverage as discretionary wealth is reduced. The mean implied leverage is consequently a somewhat larger ratio of 2.72. This refinement demonstrates the effect of transformations in the shape of probability distributions as they pass through nonlinear calculations, in this case from length of life to implied leverage. We will return to this topic in the next section.

The investment policy decision the client faces is what weight to give to four assets: a cash equivalent, a bond, and two similar equity assets. No short positions are allowed and transaction costs are ignored. The expected after-tax returns are 2%, 3%, 6%, and 6%, respectively, and the after-tax standard deviations are 1%, 5%, 20%, and 20%, respectively. Cash and bond returns do not co-vary with any other asset's returns, and the two stock returns have a 0.9 correlation. Given the Markowitz objective combined with the discretionary wealth-supplied risk aversion, we maximize expected portfolio return less half the implied leverage of 2.72 times the portfolio variance, as in Equation (3). Then the best asset weights revealed by an optimizer, such as one based on quadratic programming, are about 0%, 66.7%, 16.7%, and 16.7%, for the cash equivalent, bond, and two similar equity assets, respectively. (Moving beyond point estimates to Bayesian logic, this example will be modified and extended later in this article to illustrate the impact of knowledge imprecision.)

Opportunities for Improvement

Although it seems an important advance to have a method such as the discretionary wealth approach for deriving a more objectively based risk-aversion parameter, the examples just described otherwise carry all the practical challenges of the Markowitz model. The absence of trading costs can be partially cured in the same way as for Markowitz optimization, but there are other fundamental issues.

Without the deepening of the model as will be described in the next two sections, we would have made no advance on the most important problem that has plagued practitioners attempting to use conventional optimizers for asset allocation; that is, even though deterministic calculation of mean-variance portfolio optimization can be exquisitely sensitive to small changes in inputs, the basic framework makes no allowance for errors in input estimates and must be bound by various ad hoc constraints to produce useful practical results. In addition, the structure of the covariance matrix often requires simplification by factor analysis; still, in practice, the risk of solutions is often underestimated. Quasi-Bayesian procedures, such as mean and covariance shrinkage toward common values for "robustifying" the optimization to errors, have been invented and are valuable, but they are built upon a foundation that is unfriendly to imprecision.

In addition, even though the discretionary wealth approach provides an objective function that includes the impact of return higher moments involving skewness and kurtosis, it does not, in itself, provide a practical method for doing the optimization, and such methods are not trivial. Although this may not be an enormous problem for many conventional portfolio allocations, it is fundamental to the management of leveraged funds and funds that use options.

Finally and very importantly for individual investors and for collective investors, such as pension funds, conventional asset allocation assumes that we know the appropriate risk aversion. The discretionary wealth approach provides greater assurance of a sensible answer for those aspects of planned funding and withdrawals under investor control. But appropriate risk aversion also depends on future factors that are not under the control of the investor and consequently cannot be known with precision. Longevity risk affecting the financial commitment for retirement spending is just one example. Other important sources of uncertain investor circumstances include unknown future monetary inflation rates, uncertain taxes on currently unrealized capital gains, and the possibility of losing the ability to save because of a change in employment or marital circumstances.

DISCRETIONARY WEALTH MEETS BAYESIAN LOGIC

The Bayesian revolution in probabilistic reasoning seems to be gradually diffusing into financial research.[2] This is not limited to methods for more robust derivation of point estimate return parameters for use in conventional Markowitz mean-variance optimization. With the retention of full probability distributions for intermediate calculations, the subsequent asset allocation process itself can be upgraded.[3]

As before, we seek to maximize expected $\ln(1 + Lr)$, where L is implied leverage relating the investment portfolio to discretionary wealth and r is the portfolio return. But now we recognize that, to do so, we must derive a probability distribution for $\ln(1 + Lr)$.

In the Bayesian framework, probability is simply a framework for mapping the truth of assertions into real numbers bounded by zero (false) and one (true) that satisfy axioms that govern the calculation of compound assertions. The axioms lead directly to Bayes Law, which gives us a prescription of the proper evolution of our beliefs, expressed as probabilities, after exposure to data. This is often expressed as a posterior probability distribution being proportional to the product of a prior probability distribution and a data likelihood distribution. More generally, the axioms prescribe the appropriate combination of beliefs. The important example here is the combination of beliefs about the probability distributions of investment returns and extended balance sheet items to form a probability distribution of log returns of discretionary wealth, so that we may properly maximize its mean.

Nonlinear Transformations Require Entire Probability Distributions

An important insight useful for obtaining better broad investment decisions through Bayesian logic is to recognize that we cannot condense prediction distributions for decision inputs to a few parameters such as mean and variance without subjecting ourselves to potentially misleading chained inferences. Whenever there are logically subsequent nonlinear transformations of probability distributions, we benefit from retaining full probability distributions until we reach the point where an appropriate objective function for comparing decision outcomes can be calculated.[4]

In contrast, conventional practice with Markowitz mean-variance optimization collapses information to point estimates at the beginning, that is, a functional transformation of information from higher dimension to lower dimension takes place well before calculations of the optimal security weights. Whether the discarded information is relevant or irrelevant to this evaluation is greatly influenced by the degree of nonlinearity of the relationships between decision inputs and the output of interest.

Consider the following examples of potential mappings for even as simple a statistic as the mean as we go from one probability distribution to another. The mean of the product of two distributions is not the product of the component means. The mean of the distribution for $1/X$ is not the reciprocal of the mean for X. The means of the distributions for each of the elements of a covariance matrix's inverse cannot be determined from the means of the distributions for each element of the original matrix. The mean of a constrained variable's distribution cannot be obtained by constraining the mean of its original distribution. In all these nonlinear cases, an inference as to the mean of the function's results is distorted by attempting to infer it by applying the function to the mean(s) of the input distribution(s). The presence of intervening nonlinear relationships before calculating a preference index calls into question any decision based on point estimates of inputs with dispersed probability distributions.

Prediction Distribution Inputs to the Decision Process

We could improve conventional Markowitz mean-variance optimization by using Bayesian methods even if they were employed merely to derive better point estimates for input parameters. The well-known shrinkage models for expected return and covariance are, in spirit, highly specialized Bayesian models adapted to derive better point estimates for input parameters.[5] A more general approach to decision robustness is possible, however.

We need to recognize that our knowledge in the form of posterior distributions has a differently shaped probability distribution from the data likelihood distribution we are trying to model, and that the former is the appropriate starting place for decision making. When our posterior distributions concern parameters for another probability distribution, the next step is the construction of a *prediction distribution* for that variable of real interest. For example, if our posterior distributions are for the mean and variance of an investment return, the next step is to form a prediction distribution for returns, if that is what we care about.

Let us illustrate. Prior to the development of more capable MCMC methods in the 1980s, Bayesian estimation was typically done within *conjugate* probability distributions; that is, knowledge is represented with prior and posterior distributions, where both are assumed to be from the same function family, differing only in parameters that increment by specific formulae as more data are added. Commonly used likelihood distributions are each associated with their own conjugate prior-posterior distribution family.

For example, suppose that a return-generating process is believed to be normal, but with unknown mean and unknown variance. The relevant data for this distribution are limited. Regrettably, the normal distribution as prior and posterior is not conjugate to a normal likelihood distribution. A different distribution—a normal mean conditional on a scaled gamma distribution for the inverse of the variance—is conjugate.[6] It leads to a Student's t distribution for the marginal distribution of the mean—which can produce fat tails. Also, the marginal distribution for variance will be skewed, with a fat tail extending to higher values. That is, no matter what variance we have observed, there is always a possibility that these observations were drawn from a likelihood process with much larger variance, whereas it is impossible for the likelihood variance to go as low as zero if we have observed variation. Ignoring these two facts by focusing just on observed sample variance, along with the usual omission of the variance of the mean, leads to both variance and kurtosis underestimation. In contrast, following Bayesian logic by first generating a prediction distribution for returns based on random draws from posterior distributions for mean and variance, before quantifying from it the return means and variances to use in decision making, avoids these sources of bias.

Similarly, we have prior knowledge that correlations of returns cannot be greater than +1 nor less than −1, which should cause our posterior probability distribution for correlations to be skewed away from extremes, consequently affecting our posterior beliefs in the form of probability distributions about return covariances. From these posteriors, along with posteriors for means, we can generate prediction distributions for multiple asset investment returns.

Prediction distributions for security returns, together with a prediction distribution for implied leverage, give us an improved set of inputs for our investment decision.

Integrating Uncertainty in Appropriate Risk Aversion

Since Markowitz's 1952 contribution, financial research has paid a great deal of attention to risk regarding investment returns. In contrast, far less attention has been given to the inevitable, and often quite large, risk concerning the investor's future sources and uses of investment funds. The present values of these future cash flows are important determinants of current discretionary wealth and, in turn, determine the appropriate level of single-period risk aversion. Consequently, appropriate risk aversion should not be considered as a point estimate, but as a probability distribution. When we consider appropriate risk aversion as a probability distribution, we discover that the shortcut Markowitz objective function obtained by dividing Equation (2) by L, the implicit leverage, to arrive at the simpler objective shown as Equation (3) is often inadequate for broad investment policy. Other factors being equal, when the mean of L squared is not very close to the square of the mean of L, the Markowitz objective will tend to produce overly risky portfolios. This effect tends to be exaggerated for those whose discretionary wealth is slim. Bayesian decision making applied to the discretionary wealth approach gives us a conceptually simple way of bringing these additional considerations into the investment policy decision, including not only the variance of leverage but higher moments as well. It applies equally to households and to pension funds.

The complexity involved in risk regarding sources and uses of investment funds can be summarized in a single predictive distribution for discretionary wealth and, consequently, leverage. However, the relationships are often highly nonlinear, as is their cross product with investment returns, making it especially important that we address funding and spending uncertainty. Imprecision in appropriate risk aversion does not prevent us from portfolio optimization; however, it is a further reason for investors facing uncertain needs to exercise great caution in interpreting the results of conventional mean-variance optimized portfolios for broad investment policy.

Nonlinear Mappings in Portfolio Optimization

Two sequences of mappings are involved in comparing conventional Markowitz optimization to Bayesian optimization as a method of optimizing mean log return on discretionary wealth. They are each highly nonlinear, but in different ways, and we should not be surprised that they give different answers.

Consider the Bayesian sequence of mappings of probability distributions from knowledge of return parameter inputs to the probability distribution of $\ln(1 + Lr)$ before maximizing its mean. We are concerned, first, with the distribution of the *product* L times r, where both L and r have their own posterior distributions, and, second, with the logarithmic version of this product. The second calculation implicitly gives a positive weight to the mean of that product, a negative weight to the variance of the product, and alternating positive and negative weights to higher powers of the product. Finally, we adjust asset weights so as to maximize the mean of that distribution. Any nonlinearities in this last calculation are harmless, because the final probability distribution has already been calculated.

Consider, in contrast, the conventional use of Markowitz mean-variance optimization. To begin, consider only the version of the problem presented in Equation (3) in which the terms of the objective are divided through by L. The calculation process runs from implicit, rarely explicit, prediction distributions for return and risk aversion characteristics to input point estimates of mean return vector, return covariance matrix, and risk-aversion parameter, and thus to a preference ordering that leads to a proposed optimal set of weights. This sequence entirely omits the construction of a prediction distribution for the objective.

For Markowitz mean-variance optimization, we maximize

$$\text{Objective} = \max[w'E - (L/2)w'Vw], \text{ subject to constraints} \tag{5}$$

In Equation (5), w is the vector of security weights and is subject to at least a budget constraint, and possibly to other constraints; w' is the transpose of w; E is the vector of point-estimate mean security returns; and V is the matrix of point estimates of return covariances. $L/2$ is a risk-aversion parameter. Suppose the only constraint is the budget constraint; the sum of the weights, positive or negative, must be unity. We can solve for optimal weights using a Lagrangean multiplier for this constraint. Augmenting V to V'' and E to E'' to take into account the multiplier, we produce the optimal weights for Equation (5), augmented by a value for the multiplier, to form w'' given by

$$w'' = V''^{-1}(E''/L) \tag{6}$$

The determination of weights in Equation (6), whether explicitly accomplished as we do here or implicitly arrived at through the use of penalty functions or some other method, involves the inverse of the augmented covariance matrix, a highly nonlinear operation. This is multiplied (another nonlinear operation) by mean returns and divided (another nonlinear operation) by the implied leverage determining appropriate risk aversion. At first glance, there might seem to be no harm because no explicit probability distributions are involved, but this evaluation would be incorrect because we have ignored the error distribution implicit in the posterior distributions for means and covariances and even leverage. In contrast to our Bayesian procedure, the imposition of nonlinearities in the optimization or preference-ordering step is not harmless because it takes place too early in the process.

The situation for conventional Markowitz optimization gets worse when we consider more fully the impact of a dispersed probability distribution for L. Returning to Equation (2), we see that to approximate our objective of maximizing expected log return on discretionary wealth, the appropriate objective in a mean-variance scheme is to maximize expected $\ln(1 + Lr) \sim LE - L^2V/2$. Especially troublesome is the squared leverage multiplier for the covariance distributions. In other words, taking into account moderate probabilities of higher-than-expected leverage greatly amplifies the predicted degradation of the objective caused by investment return dispersion. We could make some improvement in the conventional process here by explicitly using separate point estimates for L and L squared and optimizing $< L > E - < L^2 > V/2$, but we still would not have adequately accounted for instances when unusually large leverages interact with unusually poor returns.

Comparing the two mapping sequences, a conventional point-estimate-based sequence of nonlinear calculations cannot be expected to closely track the result of compounding probability distributions through the intervening nonlinear mappings. Those who use such conventional optimizations must employ a variety of ad hoc strategies to produce acceptable practical results. For example, adding position constraints, such as the requirement that no securities take negative weights and that no security can take more than a modest maximum weight, can reduce the biases introduced by the point estimate approximation. Markowitz [1983] did something similar when he advocated allowing only constrained short positions. At best, however, constraints reduce errors in sub-optimal fashion. Additional strategies, such as 1) reducing the number of securities considered, 2) assuming a simpler factor structure for V (Jacobs, Levy, and Markowitz [2005]), or 3) Bayesian shrinkage of the input V toward a simpler structure (Ledoit and Wolf [2004]), can all reduce the scope for distortion by reducing the nonlinearities produced by matrix inversion. Although these strategies reduce the symptoms, they do not truly cure the disease of relying on point estimates for a problem with intrinsically nonlinear mappings of probability distributions. In any case, the disease reappears as soon as we include imprecise risk aversion in our problem.

Bayesian-Optimized Asset Allocation

Conceptually, the Bayesian asset allocation process is simple, though it may involve a large number of calculations. We need to begin with a postulated likelihood process for a distribution of returns, which can be as simple or as complex as we wish. A very simple one, used in the next section to give the Markowitz model its best opportunity for favorable comparison, is multivariate normal.[7] In that case, suppose we already have available joint posterior probability distributions for return mean vector E and covariance matrix V.[8] (We may postulate that E and V distributions are not independent, if we wish, although for simplicity we will not do so in the next section.) We then randomly draw a large sample of E, V pairs. These pairs are then substituted into our likelihood model to produce a large predictive sample for returns. We do the same for implied leverages given discretionary wealth uncertainty in order to produce an independent prediction distribution for the risk-aversion coefficient. These return and appropriate risk aversion scenarios are combined to form a more comprehensive joint probability distribution, winsorized so that returns on discretionary wealth are no worse than −99%. This winsorization allows us to apply a log-optimal criterion for discretionary wealth in our objective without complication. Finally, we optimize security weights over this large sample according to our objective function of the mean log return on discretionary wealth, taking into account any further constraints. These constraints may include not only a budget constraint but a prohibition against short positions, as well as any other constraints imposed by the investor's or the investment manager's situation. Dimensionality of input information is thus preserved until the final ranking.

Again, within this general framework, we may choose from a variety of likelihood models, of greater or lesser complexity, for returns and implied leverage. Derivative securities may be included. We can also cope with more complex solution constraints than were addressable when Markowitz first formulated the portfolio optimization problem.

In essence, we are employing a form of the old technique of scenario analysis—although used differently than by Markowitz and Perold [1981]—by using the calculation speed of the computer to range over thousands of scenarios.

ILLUSTRATIVE TESTS

How much do we gain in moving from using the discretionary wealth approach to supply the absent risk-aversion parameter to using a Bayesian version? In this section, we compare, in both long-only and long–short examples, the log returns on discretionary wealth obtained by conventional Markowitz optimization versus a Bayesian approach. We do so in a manner that favors conventional approaches by assuming that the likelihood process for return generation is multivariate normal, and we further reduce the impact of higher moments by rebalancing for only a one-month interval rather than a one-year interval. (See Equation (1) for the impact of periodicity on the relative influence of higher moments.) Nevertheless, higher-moment effects are important, not only because plausible posterior distributions of return will be fat-tailed as a result of imprecision in variance estimates, but because of previously noted nonlinearities in subsequent calculations as well as non-normal distribution of implied leverage.

The examples are small enough to be easily simulated without MCMC, but contain the principles that lead to distortion and inaccuracy when attempting to maximize expected log return on discretionary wealth using a conventional Markowitz mean-variance optimization. For example, to highlight the impact of poorly conditioned covariance matrices, we employ a very high return correlation between two assets to simulate the effect that may result from more moderate return correlations among a larger number of assets.

We observed two distinct benefits. First, the asset weights derived from limited evidence using log returns on discretionary wealth are more robust than those derived from conventional mean-variance optimizations when they are tested out of sample. A second benefit occurs in accuracy even when our input distributions are essentially fully informed with respect to our test distribution of returns and leverages. These benefits are merely moderate for long-only portfolios, perhaps because we employ few assets and none with marked higher return moments such as would result from options, but they are still worthwhile, especially for preventing very poor results. In comparison, when the Bayesian approach is applied to long–short portfolios, the benefits are dramatic.

Test Structure

One test set explores robustness with respect to small-sample error (see Exhibit 2) and the other explores accuracy given such a large sample that it is essentially full information regarding the shape of the joint distribution governing investment returns and leverages (see Exhibit 3). Within each set, we have subtests—one constrained to only long positions and the other permitting short positions as well. Each test compares conventional Markowitz mean-variance optimization to a Bayesian discretionary wealth model. We include only four assets: cash, a bond, and two stocks.

Scenario returns are constructed as follows. For each scenario, annual asset return means are drawn from a posterior Student's t distribution (the conjugate prior for the mean of a normal) with five degrees of freedom and means of 2%, 3%, 6.5%, and 6% for the cash equivalent, bond, and two stocks, respectively. These means are known with standard deviations calibrated at 0.1%, 0.5%, 1.5%, and 1.5%, respectively. Asset return covariances are independently drawn from an inverse Wishart distribution (the conjugate prior for a multivariate normal) with eight degrees of freedom, targeted at covariances based on standard deviations of 1%, 5%, 20%, and 20%, respectively, and with zero intercorrelations, except that the two stocks have expected return correlations of 0.99. Finally, for each scenario, the means and covariances drawn are used to generate a set of four normally distributed returns, winsorized to prevent returns of less than –99%. Note that although each scenario is drawn from a multivariate normal, the collection of return scenarios represents a mixture of different normal distributions.

A probability distribution for the remaining lifetime of our example of a 55-year-old woman appears in Exhibit 4. Note the transformation in shape that the probability distribution takes when it is mapped from lifespan into leverage, L, as shown in Exhibit 5. The skew is reversed as a longer lifetime translates to increasingly larger leverage and, consequently, greater appropriate risk aversion.

To further reduce the impact of higher return moments, the joint leverage and return samples are scaled to represent monthly rather than annual allocations, though tabled results are scaled up to annualized figures for presentation. Finally, winsorizing again, no combination of leverages and returns is permitted to go below –99%. (This constraint was required in less than 1 in 100,000 cases.)

Procedure and Results

For our robustness test, we first constructed 1,000 samples, each containing 100 joint instances of implied leverages and asset returns. These instances should be regarded as representing bits of evidence rather than a

EXHIBIT 2
Robustness Out of Sample

		TESTED MEAN LOG RETURN ON D.W.		
		Cross-Learning Sample Statistics		
Example	Treatment	Mean	Median	Std. Deviation
1	Markowitz mean-variance	0.0573	0.0751	0.4313
	Bayesian mean log on DW	0.0684	0.0772	0.3542
2	Markowitz mean-variance	0.0053	0.0418	0.8484
	Bayesian mean log on DW	0.0387	0.0565	0.6761

Notes: For 1,000 learning samples of 100 observations.

Example 1—Short positions not allowed; Example 2—Short positions allowed; the test sample was 900,000 observations. D.W. refers to discretionary wealth.

Exhibit 3
Accuracy Given Full Information

Example	Treatment	Cash	Weights: Bond	Weights: Stock 1	Weights: Stock 2	Fitted Log Return on D.W.: Mean	Fitted Log Return on D.W.: Std. Dev.
1	Markowitz mean-variance	0	0.633	0.367	0	0.08106	0.3199
	Bayesian mean log on DW	0	0.729	0.274	0	0.08377	0.2391
2	Markowitz mean-variance	-0.827	1.430	1.931	-1.534	0.07769	0.5726
	Bayesian mean log on DW	-0.213	0.962	1.322	-1.069	0.09058	0.3084

Notes: For a sample of 1,000,000 observations.
Example 1—Short positions not allowed; Example 2—Short positions allowed D.W. refers to discretionary wealth.

Exhibit 4
Lifespan Distribution

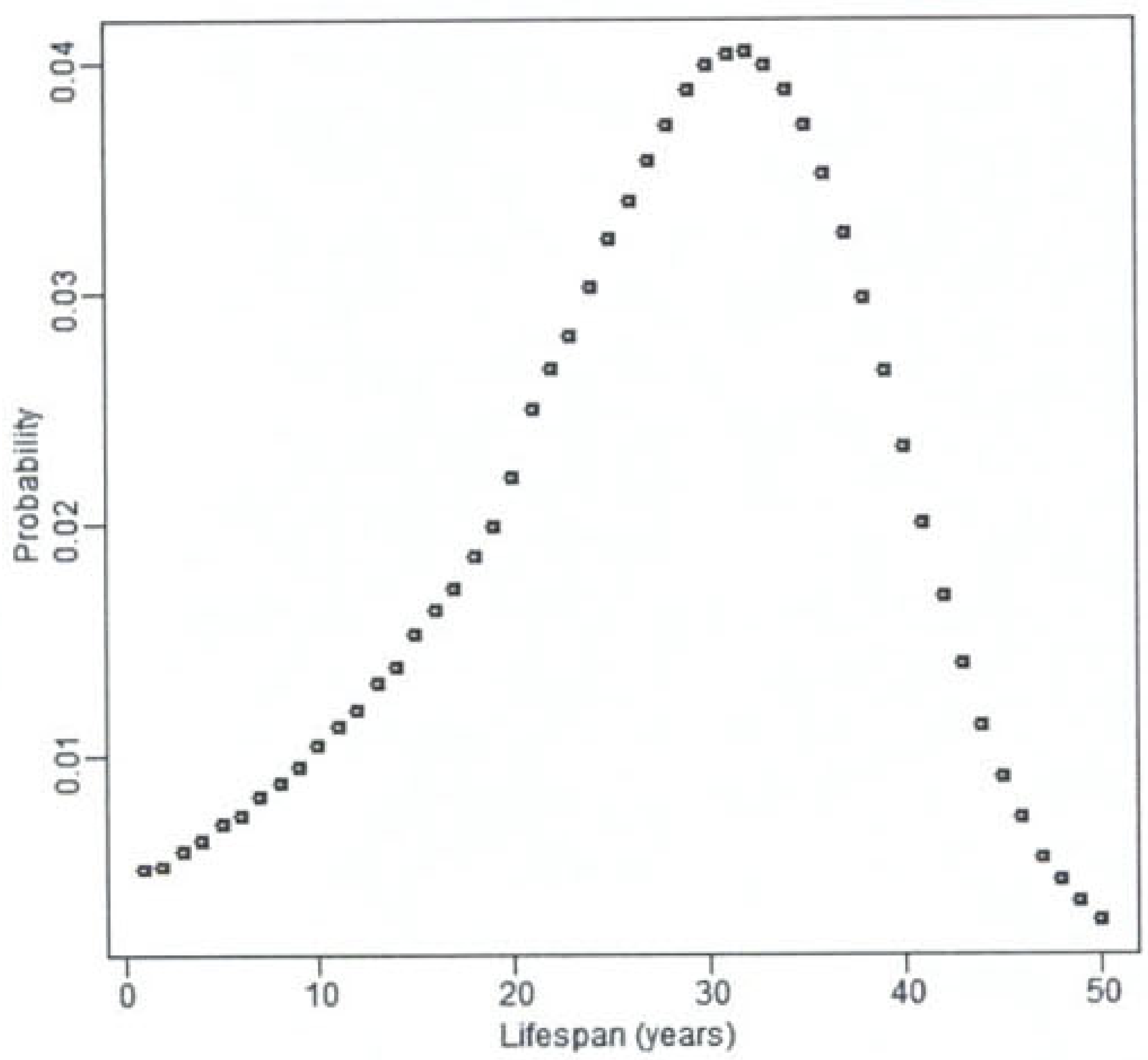
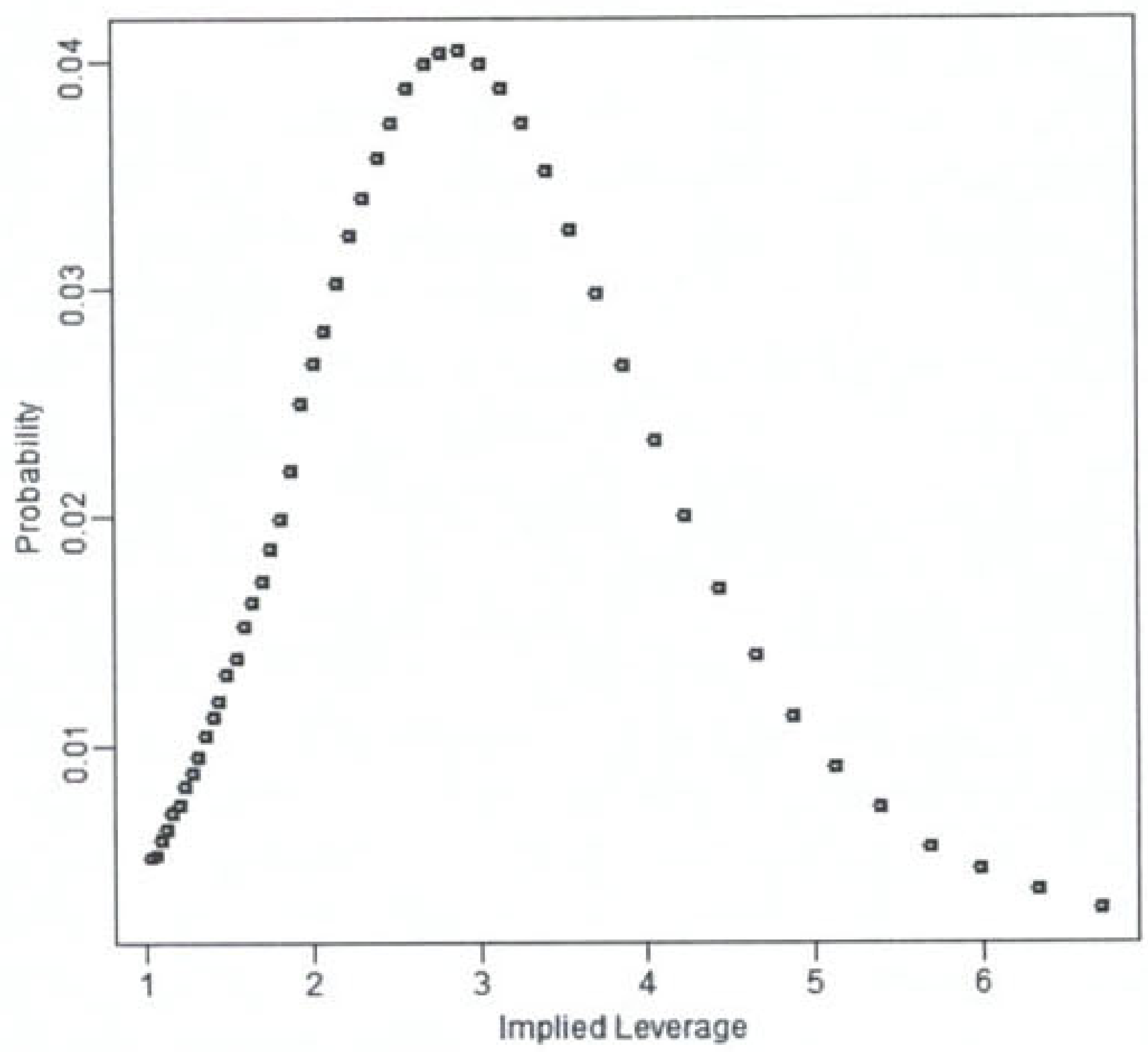


Exhibit 5
Leverage Distribution



sequence of observations. The asset returns were scaled down to monthly equivalents. For mean-variance optimization, optimal weights are based on mean sample leverage together with return means and covariances formed within each sample. For optimization of log return on discretionary wealth, optimal weights were formed based on maximizing the mean of the log-leveraged returns within each sample of 100 instances.

Each 100 instance sample's resulting set of weights forms a "rule." The 1,000 rules are tested on a much larger sample of 900,000 observations. The mean of log-leveraged portfolio return was collected for each rule. Exhibit 2 shows the single-period grand mean, median mean, and the standard deviation of mean log return on discretionary wealth across samples; the last may help assess speed of convergence toward best long-term compounded return. Note that the log returns shown in Exhibits 2 and 3 are for discretionary wealth. They must be converted to arithmetic returns and divided by leverage to measure returns on the investment portfolio.[9]

In the long-only case, the mean deficit experienced by the Markowitz approach as compared to the Bayesian discretionary wealth approach was 111 basis points (bps) (684 – 573). Although the median result is not the relevant one for judging long-term compounded median results, we show it as well, with the Markowitz approach having a median deficit of 21 bps (751 – 772). The 90 bps difference between mean and median deficits highlights that much of the superiority of the Bayesian approach is a consequence of avoiding very poor outcomes. One explanation is that the conventional models did not sufficiently avoid instances in which the combination of longer lifetime and poor investment results was present.

When short positions were allowed, the results were a greatly increased degradation of relative performance for Markowitz mean-variance optimization. The mean deficit was 334 bps, which is consistent with Markowitz's presidential address (Markowitz [1983]) to the American Finance Association warning against excessive short positions in the portfolio. To the extent that reliability is an issue, because, after all, we do not face an infinite number of periods, standard deviations are shown as well, and there is greater variation in mean log returns on discretionary wealth among rules learned by conventional Markowitz methods. Again, the difference between mean and median deficits illustrates the benefit of the Bayesian approach in avoiding the worst outcomes.

Our accuracy test is based on a single combined sample of one million instances from which to form estimates of mean and variance and, in the case of our Bayesian approach, other relevant distribution attributes. In it, we see that the Markowitz-optimized mean log return on discretionary wealth in the long-only case was only 27 bps worse than that of the Bayesian discretionary wealth approach. In comparison, the difference for the long-only portfolios in our robustness test was 110 bps. But when short sales were allowed, even though the Markowitz approach had essentially full information as to the distribution of both returns and leverages, the deficit expanded to 129 bps.

Exhibit 3 also provides the asset allocation selected by the two approaches, given information on one million instances. In the long-only case, the Bayesian discretionary wealth approach took a less aggressive stance, consistent with higher-moment effects stemming in part from the long tail on leverage. In the case allowing short positions, the Bayesian discretionary wealth approach also made less use of short positions. Finally, as in Exhibit 2, it showed less variation in log returns, signifying more rapid convergence potential.

Further Discussion

We cannot conclude from limited examples that the addition of full Bayesian logic is always worthwhile; it appears likely, however, that for major investment policy decisions the extra effort is justified. The enhancement is marked when knowledge of the parameters of the extended balance sheet and investment returns is imprecise. And investors who are also contemplating long–short investment strategies have an even greater motivation to try this approach.

Turning to speculation beyond the results of our tests, the structure of the mathematical relationships is such that, other things being equal, the greatest improvement should be for those investors with high implicit leverage. Since it is the common observation that most household retirement plans are underfunded, this observation carries considerable implication for practice. In other words, the benefits of the full Bayesian analysis can be material, with even very small numbers of assets, for several different types of investors. One is the household investor of proportionately lower discretionary wealth and consequently high implied leverage. For such investors, uncertainty with regard to future cash flows, especially because of longevity risk, is likely to create a strongly skewed distribution of possible implied leverages, as well as making higher return moments of greater relative interest because of greater volatility of investment returns on discretionary wealth; this conclusion could also be extended to many pension funds. Another investor who may benefit from a full Bayesian analysis is the professional investment manager who may have high leverage, both implicit and explicit, and who may take on short positions and use derivatives with strong high-moment effects.

CONCLUSIONS

The model we have presented in this article flows naturally from a paradigm shift built on four ideas. The first two ideas comprise the original discretionary wealth approach. First, an augmented balance sheet encompassing the present values of future savings and spending plans to quantify discretionary wealth provides a superior description of an investor's position. Second, long-run maximization of median wealth without interim shortfalls is a worthy goal, and, although we have not proven an optimum strategy for its attainment, we are satisfied that an approach based on maximizing expected growth of discretionary wealth in the current period is a marked improvement from conventional practice. This provides the normative risk aversion absent from Markowitz mean-variance optimization, as well as allows consideration of higher return moments. We find worthwhile improvement for application to broad asset allocation decisions from extending the discretionary wealth approach through adoption of Bayesian logic.

The third idea is to use full probability distributions to preserve all relevant information, which reduces the fragility of point-estimate-based optimizing methods and reduces the need for ad hoc constraints to produce usable results. The fourth, and perhaps most important, idea—the effective integration of uncertainty in saving and spending plans—is comparable in importance to dealing with return risks in setting investment policy and may be done straightforwardly through a Bayesian approach.

We tested this paradigm on illustrative asset allocation problems. The full application of Bayesian logic appears to provide better results for expected growth in discretionary wealth than simply using the discretionary wealth approach to create the risk-aversion parameter needed for conventional point-estimate-based Markowitz mean-variance optimization. The advantages appear notable not only in the case of long–short portfolios, but also when there is dispersion in estimates of future investment funding and withdrawals, as with longevity risk.

Our analysis raises issues of more general interest. These include theoretical questions of how close the approach comes to optimality of long-term median results, and how it may be modified to take into account trading costs, essential if we are to better understand implications for more dynamic policy than the usual annual or multi-year rebalancing of strategic allocations. From an implementation perspective, we know that using MCMC rather than naïve Monte Carlo simulation would enable the practical solution of problems with a larger number of assets, but we do not know how far we can go in this direction. As regards broad investment policy, our analysis questions the usefulness of lifecycle mutual funds that are insufficiently customized to the investor and insufficiently flexible as circumstances change. Perhaps the most critical policy issue for many investors, including pension funds, is the urgency of taking into account uncertainty in future financial needs. We would like to know if the approach we have demonstrated in this article could materially improve investor welfare.

ENDNOTES

We thank the referee, Russell Fogler, for his thoughtful and helpful comments. In addition, we are grateful to Stan Beckers, John Campbell, Campbell Harvey, Haim Levy, André Perold, Rodney Sullivan, Meir Statman, and Guofu Zhou for their helpful comments on an earlier version of this article.

[1]See Harvey et al. [2002], for example.

[2]For a review of basic Bayesian concepts and methodologies, see Gelman et al. [2004] and Rachev et al. [2008].

[3]The literature on this subject has become too large to do justice to all the relevant research. Meucci [2005] and Fabozzi et al. [2007] reviewed a wide span of applications. Recent contributions incorporating a Bayesian decision include Greyserman, Jones, and Strawderman [2006], Kan and Zhou [2007], and Harvey, Liechty, and Liechty [2008].

[4]This fully Bayesian approach should not be confused with a resampling method that deals with uncertainty by building a portfolio based on an average of security weights derived through repeated conventional mean-variance optimizations given perturbed input parameter point estimates. Michaud [2001] has shown that averaging classically optimized portfolios indicated by perturbed input point estimates can improve results. This appeared to Markowitz and Usmen [2003] to be better than a Bayesian approach in a relatively benign case, though with a rejoinder by Harvey, Liechty, and Liechty [2008].

[5]This thread includes, for example, Jorion [1986], Black and Litterman [1992], and Ledoit and Wolf [2004].

[6] See Gelman et al. [2004].

[7]Although log-normal or Student's *t* distributions for returns are common in the literature, neither these nor a normal return process is a precise model for realistic security return generation. A multivariate normal likelihood distribution will be used in the next section in our test examples, because we want a best case for Markowitz mean-variance optimization before comparing it with our fully Bayesian model.

[8]We omit estimation of the input parameter posterior distributions because this topic is already well covered in the investment literature. The key difference here is that we retain the full distribution, rather than obtain a point estimate from it.

[9]Although a basis point (one hundredth of a percent) does not strictly apply to logarithms, it is convenient terminology in the absence of better terminology.

REFERENCES

Black, F., and R. Litterman. "Global Portfolio Optimization." *Financial Analysts Journal*, Vol. 48, No. 5 (1992), pp. 28–43.

Black, F., and A.F. Perold. "Theory of Constant Proportion Portfolio Insurance." *Journal of Economic Dynamics and Control*, Vol. 16, No. 3–4 (1992), pp. 403–426.

Brunel, J. "How Sub-Optimal—If at All—Is Goal-Based Asset Allocation?" *Journal of Wealth Management*, Fall 2006, pp. 19–34.

Donohue, C., and K. Yip. "Optimal Portfolio Rebalancing with Transaction Costs." *Journal of Portfolio Management*, Vol. 29, No. 4 (2003), pp. 49–63.

Fabozzi, F.J., P.N. Kolm, D. Pachamanova, and S.M. Focardi. *Robust Portfolio Optimization and Management*. Hoboken, NJ: John Wiley and Sons, 2007.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*, 2nd ed. London: Chapman and Hall, 2004.

Greyserman, A., D.H. Jones, and W.E. Strawderman. "Portfolio Selection Using Hierarchical Bayesian Analysis and MCMC Methods." *Journal of Banking and Finance*, Vol. 30, No. 2 (2006), pp. 669–678.

Harvey, C.R., J.C. Liechty, and M.W. Liechty. "Bayes vs. Resampling: A Rematch." *Journal of Investment Management*, Vol. 6, No. 1 (2008), pp. 1–17.

Harvey, C.R., J.C. Liechty, M.W. Liechty, and P. Müller. "Portfolio Selection with Higher Moments." Working Paper, Duke University, 2002.

Ibbotson, R.G., M.A. Milevsky, P. Chen, and K. Zhu. *Lifetime Financial Advice: Human Capital, Asset Allocation, and Insurance*. Charlottesville, VA: Research Foundation of CFA Institute, 2007.

Jacobs, B.I., K.N. Levy, and H.M. Markowitz. "Portfolio Optimization with Factors, Scenarios, and Realistic Short Positions." *Operations Research*, Vol. 53, No. 4 (2005), pp. 586–599.

Jorion, P. "Bayes-Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis*, Vol. 21, No. 3 (1986), pp. 279–292.

Kan, R., and G. Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3 (2007), pp. 621–656.

LeDoit, O., and M. Wolf. "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices." *Journal of Multivariate Analysis*, Vol. 88, No. 2 (2004), pp. 365–411.

MacLean, L.C., and W.T. Ziemba. "Capital Growth: Theory and Practice." In S.A. Zenios and W.T. Ziemba, eds., *Theory and Methodology*, Vol. 1 of *Handbook of Asset and Liability Management*, Amsterdam: Elsevier, 2006, pp. 429–469.

Markowitz, H.M. "Nonnegative or Not Nonnegative: A Question about CAPMs." *Journal of Finance*, Vol. 38, No. 2 (1983), pp. 283–295.

Markowitz, H.M., and A.F. Perold. "Portfolio Analysis with Factors and Scenarios." *Journal of Finance*, Vol. 36, No. 4 (1981), pp. 871–877.

Markowitz, H.M., and N. Usmen. "Resampled Frontiers versus Diffuse Bayes: An Experiment." *Journal of Investment Management*, Vol. 1, No. 4 (2003), pp. 9–25.

Meucci, A. *Risk and Asset Allocation*. New York: Springer Finance, 2005.

Michaud, R.O. *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. New York: Oxford University Press, 2001.

Perold, A.F., and W.F. Sharpe. "Dynamic Strategies for Asset Allocation." *Financial Analysts Journal*, Vol. 51, No. 1 (1995), pp. 149–160.

Rachev, S.T., S.J. Hsu, B. Bagasheva, and F.J. Fabozzi. *Bayesian Methods in Finance*. New York: John Wiley and Sons, 2008.

Rubinstein, M. "The Strong Case for the Generalized Logarithmic Utility Model as the Premier Model of Financial Markets." *Journal of Finance*, Vol. 31, No. 2 (1976), pp. 551–571.

Stutzer, M. "A Portfolio Performance Index." *Financial Analysts Journal*, Vol. 56, No. 3 (2000), pp. 52–61.

——. "Portfolio Choice with Endogenous Utility: A Large Deviations Approach." *Journal of Econometrics*, Vol. 116, No. 1/2 (2003), pp. 365–386.

Thorp, E.O. "Portfolio Choice and the Kelly Criterion." *Business and Economics Statistics*, Proceedings of the American Statistical Association (1971), pp. 215–224.

——. "The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market." In S.A. Zenios and W.T. Ziemba, eds., *Theory and Methodology*, Vol. 1 of *Handbook of Asset and Liability Management*, Amsterdam: Elsevier, 2006, pp. 385–428.

Tobin, J. "Liquidity Preference as Behavior toward Risk." *Review of Economic Studies*, 67 (February 1958), pp. 65–86.

Wilcox, J. "Harry Markowitz and the Discretionary Wealth Hypothesis." *Journal of Portfolio Management*, Vol. 29, No. 3 (2003), pp. 58–65.