Defined contribution pension plans have become an increasingly important component of many people’s retirement plans. More than 50 million Americans participated in these plans as of 1998 (PSCA 1999), and their assets exceed $2 trillion (EBRI 1999). A distinguishing characteristic of defined contribution plans is that participants have much greater responsibility for determining both the contribution level and the investment allocation, as compared to the norm in defined benefit plans. In view of this increase in the need for individual responsibility for retirement saving and investing, it is of interest to explore whether workers are financially literate enough, and sufficiently prepared, to be do an effective job of saving for retirement.

The news is not particularly encouraging. In a recent survey on retirement readiness, EBRI (1999) found that only 8 percent of workers were doing a “very good” job of retirement preparation. The same survey found that only about half of workers had ever even assessed their retirement needs. Other studies have uncovered similarly discouraging findings: Bernheim et al. (this volume) report that households need roughly to triple their retirement assets to secure a comfortable retirement; Moore and Mitchell (2000) conclude that the typical older household needs to save an additional 16 percent of income to maintain its preretirement standard of living; and Warshawsky and Ameriks (2000) estimate that more than half of U.S. households will fail to fund their retirement sufficiently.

Given this apparent lack of readiness, the question becomes: why are so many people underprepared for retirement? One answer may be that they are already saving as much as they possibly can. However, Moore and Mitchell (2000) concluded that this is unlikely for all but the poorest of households studied. A more likely answer may be that the difficulty and enormity of the computations simply overwhelm people, and certainly the
problem is an extremely complex one to solve. For instance, simply assessing how much one might need in retirement is difficult, and if one then considers the complexity of projecting investment returns under uncertainty, many will feel that the problem becomes unmanageable. This chapter shows how Monte Carlo simulation techniques can be used in the context of outcomes-based investing to help people make retirement decisions. Three key questions are considered:

- What is outcomes-based investing, and how does it relate to Monte Carlo simulation?
- What are the potential benefits of outcomes-based investing?
- What difficulties arise when implementing an outcomes-based approach?

Special emphasis is placed on the third question because addressing the difficulties is crucial to implementing a workable solution.

**Monte Carlo Simulation and Outcomes-Based Investing**

Outcomes-based investing simply refers to the process of evaluating possible outcomes associated with different portfolio allocation decisions, to determine what investment path is most consistent with the decision-maker’s goals. In fact, anyone seeking to evaluate possible outcomes prior to making a decision under uncertainty is employing a variant of “outcomes-based decision making.” Typically this requires the development of simulation methods, and indeed simulations in the context of financial decision-making have been practiced in the financial community for decades. For instance, large pension funds have employed Monte Carlo-based simulation approaches to evaluate investment choices, asking which investment policy minimizes the probability of exhausting pension assets given a particular contribution schedule. In the context of a defined contribution pension, an individual participant might want to know which investment portfolio maximizes the chance of achieving a particular retirement goal.

Though the concept of outcomes-based investment is not new, the complexity and cost of creating reasonable simulation models to help make investment-decisions has limited the availability of this approach to a relatively limited segment of the financial services industry. In recent years, however, two technological developments have made it possible to extend Monte Carlo-based methods to the individual retirement decision-making problem. The first has been the huge fall in the cost of computation. Monte Carlo methods are inherently computationally intensive, and fast, cheap computers now make implementing them more cost effective than before. The second has been the rise of the Internet, which has dramatically decreased distribution costs for software-based investment services. Together, these innovations make it much more likely that simulation based meth-
ods will become accessible to individuals seeking help in the retirement decision-making process.

Consider the problem an individual faces when seeking to accumulate retirement assets. If the target amount of wealth needed is $W^*$ by retirement, the saver may ask whether, given a current portfolio and planned saving level, he or she reach the goal. One way to answer this question is to turn to one of the widely available retirement calculators on the Internet. These calculators typically assume that the investor’s portfolio grows at a prespecified annual rate of return, and that inflation erodes a constant fraction of the portfolio each year. They may then generate an expected shortfall or “grade” associated with the saver’s plan, deeming it adequate or inadequate to meet the goal.

A difficulty with this type of simplified calculator model is that it can easily output wrong information regarding expected portfolio values at retirement. For example, if the user has extremely optimistic beliefs regarding likely future investment performance, and supplies the calculator with an expected annual return of ≤20 percent, he or she will likely leave the exercise confident in achieving the retirement objective. With a more reasonable return assumption the analysis would indicate that the worker was woefully unprepared for retirement. In practice, we find that people are often surprised that they have a relatively small chance of achieving the wealth implied even by average historical returns. Even if historical returns are predictive of future returns (possibly an over-optimistic assumption), the average return is difficult to achieve. One reason is that returns are reduced by administrative and other costs including management fees, taxes, and transaction costs. Another reason is that the volatility of market returns implies that average returns overestimate cumulative returns. Since these calculator models do not address the issue of investment uncertainty directly, users will tend to construct retirement plans based on insufficient or faulty information. Even worse, people may believe and act as if they were on track for a “dream” retirement, when a much less happy outcome could reasonably occur.

The major drawback with a calculator approach to wealth projection is that portfolio returns are uncertain. Because of this complexity, the likely range of wealth outcomes at retirement cannot easily be derived analytically, even with relatively simple assumptions for portfolio returns. This is where Monte Carlo simulation is useful. Specifically, Monte Carlo simulation involves constructing an economic model capable of capturing many of the important characteristics of investment returns. Using this model, it becomes possible to simulate thousands of potential paths an asset portfolio may take over time, as people move toward retirement. The user can then examine numerous potential scenarios to determine what range of outcomes is most reasonable, what outcomes are unusually optimistic, and what the range of quite pessimistic outcomes might be. Instead of determining
the distribution of retirement wealth analytically, Monte Carlo simulation generates an empirical estimate of the final wealth distribution. This step is helpful in producing informative statistics such as a saver’s chances of reaching a target wealth level. As explained below, the number of simulations evaluated determines the quality of the approximation to the final retirement wealth distribution.

In short, outcomes-based investing means evaluating different investment plans based on the likely outcomes the plan can deliver. Monte Carlo simulation allows extremely complex stochastic distributions to be evaluated without knowledge of the actual analytic distribution of outcomes.

**Potential Benefits of Outcomes-Based Investing**

At its core, outcomes-based investing—or more generally, outcomes-based planning—helps one make better decisions through improved information. By providing individuals with better information about the likely impact of portfolio risk, savings level, investment horizon, and other factors, the expectation is that the saver can make decisions that ultimately lead to improved outcomes.

To assess the benefits of outcomes-based investing, it is helpful to examine a prominent alternative, namely the results of basing investment decisions on deterministic returns. This latter deserves scrutiny because of the potential for investors to ignore portfolio risk. With deterministic returns, a given portfolio generates an anticipated average return. As described above, this average return can be used to calculate a single value of projected retirement wealth. Typically this final wealth level is compared to the wealth target, and the resulting shortfall (or surplus) is reported.\(^2\) A saving plan is deemed “successful” if a surplus is identified, while a projected shortfall indicates the need to adapt the saving path and investment decisions. A fundamental problem occurs, however, when an investor compares portfolios having different risk levels. If, as is typically the case, the higher risk portfolio generates a higher expected return, then the higher risk portfolio would appear to dominate the lower risk portfolio. In short, the saver would be encouraged to think that higher risk is better because higher risk generates a higher average return. This conundrum is termed by Sharpe (1997) as the “return/return trade-off.” As he notes, the loss potential of the portfolio is usually described in the discussion of the decision process, but usually it does not influence the calculator’s surplus/shortfall calculation. So the key measure of a saving plan’s success depends only on the average return rather than the risk and return of the portfolio, which can lead to inappropriate investment decisions.

Avoiding the return/return tradeoff is a specific illustration of how outcomes-based investing can improve decision-making. The more general point is that simulation models do a better job at helping people understand
risk. Outcomes-based investing allows users to determine their risk tolerance in the context of the impact of risk on their own particular outcomes. For example, someone who is sensitive to losses over the short term, absent any context, might be slotted into a conservative portfolio by a risk questionnaire. On the other hand, minimizing short-term risk might mean that this individual has little chance of achieving his long-term objective. This might indeed turn out to be the appropriate portfolio, but the individual might arrive at a different portfolio if the full range of risk/return trade-offs were illustrated. Casting the decision as a choice between low short-term risk with minimal chance of reaching long-term objectives, versus moderate short-term risk and a better long-term chance of hitting the target, would allow this user to make a more informed decision.

Outcomes-based investing also has the potential to allow investors to view their overall consumption/saving/investment process in the context of their own long-term objectives. For example, analyzing the myriad financial products on the market becomes easier when the answer to the question “Is this financial product right for me?” turns into an evaluation of whether the product improves the individual’s likely outcomes. Other questions also become more manageable in an outcomes-based framework. For example, this framework can answer the following question: “The DOW dropped 500 points, what does this mean to me?” Ultimately, having a single framework to analyze financial uncertainty and make financial decisions could result in improved decision-making.

**Challenges in Implementing an Outcomes-Based Approach**

Some problems arise in the process of implementing an outcomes-based approach, many of which we have solved in the process of devising a new simulation model offered by Financial Engines. In developing the program, the most significant challenges were (a) to summarize vast amount of information generated by the simulation, and (b) to create a Monte Carlo simulation with sufficient precision to facilitate decision making summarizing all the information generated is challenging because at each point in time, the situation model generates a full wealth distribution. Determining which characteristics of the simulation are most relevant to making decisions ultimately depends on the preferences of the person making the decisions. It is possible, although unlikely, that someone thinking of retiring in thirty years worries most about the 75th percentile of the wealth distribution ten years prior to retirement. Likewise, some other individual might care deeply about a different point on the outcomes distribution, at some other point in time. Implementing an outcomes-based system requires determining which pieces of information are relevant to individual decision making.
To help determine which pieces of information to display, Financial Engines relies on behavioral finance research. This literature is summarized by Shefrin (2000), who describes some of the key elements to investor decision-making as follows:

According to folklore, greed and fear drive financial markets. But this is only partly correct. While fear does play a role, most investors react less to greed and more to hope. Fear induces an investor to focus on events that are especially unfavorable, while hope induces him or her to focus on events that are favorable. In addition to hope and fear, that apply generally, investors have specific goals to which they aspire.

This research, along with user testing, leads us to summarize the Monte Carlo information generated using five summary statistics. The most important one relates to the chances of achieving a specific goal. In particular, the user is shown the probability that he will have enough wealth to achieve the retirement income objective. Given that many people are unaccustomed to seeing and interpreting probabilities, Financial Engines uses a weather analogy to help communicate the financial forecast (sunny, cloudy, stormy, etc.).

The second highest priority statistic deals with investor fear regarding the short-term loss potential, defined as the 5th percentile of portfolio wealth distribution over the next year. Highlighting these two pieces of information allows the user to trade off hope and fear. A higher risk portfolio may improve one’s chance of long-term success, but it also implies larger short-term volatility.

In addition to measures focused on hope and fear, Financial Engines models also report other key information regarding the user’s retirement prospects. In particular, the program generates estimates of upside, median, and downside wealth outcomes at the specified investment horizon, where upside, median, and downside are defined as the 95th, 50th, and 5th wealth percentiles respectively.

While determining which statistics to show can be a difficult task, presenting these statistics is also challenging, since many people do not handle probabilities comfortably. To address this concern, each of the five statistics is reported in an accessible way. For example, the loss associated with the 5th percentile of year 1 wealth is reported as an amount the user might lose over the short term, if markets perform poorly. Similarly, the long-term 95th, 50th, and 5th percentiles are reported as the long-term upside, median, and downside scenarios. Finally, the probability of achieving the long-term objective is reported using a graphical weather metaphor in addition to reporting the actual statistic. Irrespective of the implementation, it is important to recognize these types of challenges facing any outcomes-based approach to investing.

The second important challenge in implementing an outcomes-based model is the size of the computational requirements for the simulation approach. An advantage of Monte Carlo simulation is it yields approximate
distributions even in cases where analytic distributions are unknown. But its
disadvantage is that substantial simulation runs might be needed before the
results are precise enough to form the basis for decision-making. As an
example, we consider the problem of estimating the probability of achiev-
ing a particular wealth goal at retirement. Irrespective of the complexity of
the stochastic process, the wealth associated with each scenario will either be
deemed sufficient or falling short of the goal. In this sense, each scenario
can be characterized as Bernoulli random variable that takes a value of 1
when the goal is met or exceeded and a value of zero otherwise. To express
this statistically, we let:

\[ X_i = \begin{cases} 
1, & W_{T,i} > W^*, \\
0, & \text{otherwise},
\end{cases} \]

where \( X_i \) refers to the scenario \( i \) indicator variable (=1 if goal achieved); \( W_{T,i} \)
refers to scenario \( i \) simulated wealth at investment horizon \( T \); and \( W^* \) is the
retirement wealth objective.

A natural estimator for the probability of achieving the wealth objective is
to determine the fraction of scenarios that achieve the goal. Given the
definition of \( X \) above, this estimator \( Y \) is simply expressed as follows:

\[ Y = \frac{\sum_{i=1}^{N} X_i}{N}, \]

where \( N \) = number of scenarios in the given simulation. For example, if
there were 100 scenarios in the simulation \( (N=100) \), and 50 of these sce-
narios met or exceeded the wealth objective, then the estimate for the
probability of achieving the wealth objective would be 50 percent \( (50/100) \).

Nevertheless the estimated probability of achieving one’s wealth objective
is a noisy estimate. In the example above, if another 100 scenarios were
generated, only 40 might achieve the goal. In this case, the estimated
chance of success is 40 percent even though the true chance of success is 50
percent. This dependency of the estimated chance of success on the particu-
lar sample of scenarios is called “sampling error.” And as is well known,
sampling error with only a few runs (e.g., 100 scenarios) can be so large as to
invalidate decision making.

To determine the uncertainty associated with any particular estimate, the
variance of the estimator must be calculated (Appendix). It turns out that
the reasonable range of outcomes from a given Monte Carlo simulation is
approximately \( \pm 1/N^{1/2} \), where \( N \) is the number of scenarios simulated. For
example, if \( N \) is 100, then the error bound is approximately \( +/− 1/10 \) or
10 percent! Put another way, a 100-scenario Monte Carlo simulation would
likely yield an estimated probability of achieving a particular goal of be-
Table 1. Potential Monte Carlo Results (Sample Size = 100) (percent)

<table>
<thead>
<tr>
<th>Portfolio Risk Level</th>
<th>True Probability of Achieving Goal (P)</th>
<th>Likely Range for Monte Carlo Simulation</th>
<th>Potential Monte Carlo Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46</td>
<td>36–56</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>38–58</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>40–60</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>42–62</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>54</td>
<td>44–64</td>
<td>44</td>
</tr>
</tbody>
</table>

Source: Author’s calculations.

between 40 and 60 percent, when the true underlying probability was 50 percent.

To put this result in context, consider an individual hoping to use Monte Carlo simulation to make outcomes-based investment decisions. Suppose this investor would like to evaluate five different portfolios, each generating a different probability of achieving his wealth objective. Table 1 illustrates the precision problem faced when relying on a small number of Monte Carlo scenarios (in this case, 100 runs). With so few scenarios, only a very coarse assessment can be devised. What is interesting is that even when the scenario count rises to 400, the uncertainty is still large (±5 percent).

To distinguish consistently the possible portfolios described in Table 1, research has shown that the Monte Carlo model would require on the order of 10,000 runs (±1 percent error). Depending on the complexity of the simulation, this number of scenarios may not be computationally feasible. Fortunately, sampling techniques exist with the potential to decrease significantly the number of computations required to achieve a low-variance estimate. The best variance reduction technique depends on the particular problem. Based on our experiments, we find that stratification techniques seem to work well with the wealth accumulation problem. This refers to a technique for reducing the variance of a Monte Carlo simulation. The draws used for the simulation are not chosen completely at random. Rather, specific statistics of the sample are required to match their theoretical distributions exactly.

As an example of stratified sampling, consider a Monte Carlo experiment to find the average elevation of Massachusetts (hypothetically deemed to be a rectangle for our purposes). Suppose we can choose 12 points within the state, and at each of these points we measure the elevation, taking the mean of these measurements as an estimate of the average state elevation. The question arises as to how the 12 points should be selected. If we were to select 12 points at random, there is some chance that they might cluster about the low-lying Cape Cod seashore. A more scientific approach would be to divide Massachusetts into 12 equal areas and choose one point ran-
domly from each area (Figure 1). Theoretically, each of these areas should be sampled with equal frequency, and this stratification scheme enforces this result. This is an example of two-dimensional stratification.

Stratified sampling tends to reduce variance primarily because the stratified sampling avoids the potential clumping of samples. For example, Figure 2 illustrates a potential sample using standard Monte Carlo. This particular sample overweights the western portion of the state, and thus the result could deviate substantially from the true average elevation. Stratified sampling explicitly rules out this type of clustering, thus increasing the likelihood that an estimate closer to the true value is achieved.

Returning to the wealth accumulation problem, we posit that stratification could reduce the variance of the terminal wealth distribution and thus reduce the noise around a statistic derived from the wealth distribution. Suppose the objective was to create an estimate of the probability of achieving a particular goal that was precise to within ±1 percent. As described above, a Monte Carlo simulation with 10,000 samples would ensure this level of precision. An interesting question is, how much improvement could stratification provide? To simplify the example, let us suppose that the wealth accumulation process is as follows (ignoring future contributions):

\[ W_T = W_0 \prod_{j=1}^{T} (1 + R_j) \]

In other words, ending wealth is simply starting wealth accumulated at the appropriate (stochastic) rate of return. Ideally, retirement wealth could
itself be stratified. In fact, if retirement wealth could be stratified, the resulting precision would be approximately $\pm 1/N$. In this case, our research shows that only 100 samples would yield the desired precision!

Unfortunately if retirement wealth could itself be stratified, this would imply knowledge of the analytic retirement wealth distribution. In this event, there would be no need to estimate the probability, since it could be calculated exactly. But since retirement wealth cannot be stratified, another alternative is to stratify the individual return random variables. The problem with stratifying the returns is that the dimensionality of the stratification becomes too high. Recalling the Massachusetts example, the stratification occurred along two dimensions: horizontal and vertical bands. This two-dimensional delineation required at least twelve observations ($3 \times 4$) to appropriately sample the state. A full stratification of the returns would require a $T$-dimensional stratification (a dimension for each return random variable). Even if only two bands were considered per random variable, the resulting requirement for complete coverage would be at least $2^T$ scenarios. If returns were annual and the projection period was 50 years, the number of scenarios is clearly seen to be prohibitive.

One strategy might be to ignore some dimensions and simply stratify a few of the returns. Another desirable approach would be to stratify using some measure highly correlated with retirement wealth, if not wealth itself. The intuition is that if something highly correlated with retirement wealth is required to have minimal clustering, this should decrease the clustering properties of retirement wealth. In order to motivate a quantity correlated with retirement wealth consider the following approximation:
This approximation indicates that retirement wealth should be highly correlated with the sum of the returns. If the returns in this problem were normal, then the sum of the returns is also normal. This fact can be exploited to stratify the sum of the individual returns.

The benefit of stratification in this simple case is clear from Table 2. For this simple problem, stratifying the sum of the returns decreases the number of scenarios required to achieve the objective precision by over an order of magnitude. In more complicated problems, the improvement tends to degrade because the correlation between the stratified quantity and the terminal wealth degrades. Even in this case where degradation occurs, we find that stratification can yield significant improvement over standard Monte Carlo simulations.

**Conclusion**

Improvement retirement planning requires better, more sophisticated, outcomes-based investing and Monte Carlo analysis. Simulation modeling can now permit evaluation of complex outcomes and better retirement decision-making. An outcomes-based approach has several benefits. One is the avoidance of the return/return trade-off problem, arising when riskier portfolios are perceived to be superior to less risky ones, solely due to their higher average expected return. Outcomes-based approaches also offer improved risk assessment by quantifying the risk of a particular plan in terms of the range of potential outcomes. This quantification allows trade-offs between short-term sensitivity to losses and long-term goal achievement. Outcomes-based approaches also can offer a single framework for interpreting new products and events.

We have also identified some solutions to many of the challenges faced by developers of outcomes-based models. Displaying key statistics in a way people can understand proves to be a nontrivial problem, which motivates the statistics shown by Financial Engines’ outcomes-based approach. We also
Table 2. Stratification Results for Terminal Wealth Problem

<table>
<thead>
<tr>
<th>Return Distribution</th>
<th>Simulation Technique</th>
<th>Samples Required to Ensure ± 1% Goal Probability Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_j \sim \text{Normal}(0.1, 0.2) )</td>
<td>Simple Monte Carlo</td>
<td>10,000</td>
</tr>
<tr>
<td>( R_j' \sim \text{Normal}(0.1, 0.2) )</td>
<td>Stratification</td>
<td>900</td>
</tr>
</tbody>
</table>

Source: Author’s calculations.

show how to assess the probability that a particular retirement plan achieves a wealth objective, and the difficulty that arises with small numbers of simulations. We suggest that stratification is a method to reduce the variance of probability estimates.

Appendix

The variance for Bernoulli random variables is well-known, leading to the following result for our estimator, \( Y \):

\[
\text{Var}(Y) = \text{Var} \left( \frac{\sum_{i=1}^{N} X_i}{N} \right) \\
= \frac{\text{Var} \left( \sum_{i=1}^{N} X_i \right)}{N^2} \\
= \frac{\sum_{i=1}^{N} \text{Var}(X_i)}{N^2} \\
= \frac{N \cdot P \cdot (1 - P)}{N^2} \\
= \frac{P \cdot (1 - P)}{N}.
\]

In this derivation, \( P \) is the true probability of achieving the wealth objective and \( N \) is the number of scenarios created in the Monte Carlo simulation. The first equality above simply substitutes for \( Y \). The second equality follows from a property of how scalars impact variances. The third equality recognizes that each Bernoulli random variable is independent, implying that the variance of the sum equals the sum of the variances. The fourth equality utilizes the fact that the variance of a Bernoulli random variable is \( P^*(1 - P) \), where \( P \) is the probability of the Bernoulli random variable taking a value of 1. In this case, \( P \) is the probability of achieving the wealth objective. Finally, the last equality is simply algebra.
In addition, as $N$ becomes large, the estimator for the chance of reaching the goal becomes approximately normal. In other words, as $N$ becomes large:

$$Y \sim N(P, \operatorname{Var}(Y)),$$

where $P$ is the true probability of achieving the goal, and $\operatorname{Var}(Y)$ is defined above. Assuming this normal approximation holds, then a reasonable confidence interval around $Y$ would be $+/−$ two standard deviations. Using the variance above, two standard deviations is

$$2 \times \text{STD}(Y) = 2 \times \operatorname{Var}(Y)^{1/2} = 2 \times \left(\frac{P \times (1 - P)}{N}\right)^{1/2} = \frac{1}{N^{1/2}}, \quad P = \frac{1}{2}.$$

The last equality occurs when the true goal probability, $P$, is 50 percent. Since this is the point that corresponds to maximum variance, it is useful to consider this as a worst-case approximation.

**Notes**

1. For example, an investor who receives a 50 percent return one year followed by a loss of 20 percent had a cumulative return of close to 10 percent. However, the average of the two returns is 15 percent.
2. This latter is sometimes referred to as “gap analysis.”
3. This company offers a very sophisticated online adviser that “provides alternative portfolios based on four factors you are free to adjust: the year in which you choose to retire; how much income you expect in retirement; your savings rate; and the amount of risk you are willing to assume. To do ‘what if’ scenarios with one or more of these different variables, you simply adjust onscreen ‘sliders’ that resemble the bass and treble controls on a stereo. Adjust your retirement age, or your risk tolerance, for example, and the program quickly fashions an alternative portfolio” (Longman 2000). The program analyzes the user’s portfolio using many different scenarios, and outputs probabilities of achieving retirement asset targets. For more information see www.financialengines.com.
4. As other analysis in this volume points out, earnings (and hence contribution) streams and investment horizons are also stochastic (Davis and Willen, this volume; Bernheim et al., this volume), as is life expectancy in retirement (Brown et al. this volume). However, as of this writing, the models employed by Financial Engines take contributions and investment horizon as exogenously determined by the investor.
5. This number of scenarios is consistent with the typical defined benefit asset/liability study, in our experience.
6. This argument is a bit loose, in that $Y$ is actually degenerate as $N$ becomes large. More formally, $Y$ should be scaled up by the square root of $N$ in order to approximate a normal distribution. The test offers a useful shorthand for the discussion; see Amemiya (1985) for additional detail.
References


Brown, Jeffrey R., Olivia S. Mitchell, and James M. Poterba. This volume. “Mortality Risk, Inflation Risk, and Annuity Products.”

Davis, Steven J. and Paul Willen. This volume. “Risky Labor Returns and Portfolio Choice.”


