Choosing Outcomes versus Choosing Products: Consumer-Focused Retirement Investment Advice

DANIEL G. GOLDSTEIN
ERIC J. JOHNSON
WILLIAM F. SHARPE*

Investing for retirement is one of the most consequential yet daunting decisions consumers face. We present a way to both aid and understand consumers as they construct preferences for retirement income. The method enables consumers to build desired probability distributions of wealth constrained by market forces and the amount invested. We collect desired wealth distributions from a sample of working adults, provide evidence of the technique's reliability and predictive validity, characterize individual- and cluster-level differences, and estimate parameters of risk aversion and loss aversion. We discuss how such an interactive method might help people construct more informed preferences.

In this article, we present an approach for gaining insight about consumers' preferences for investments. Because it is based on both normative and behavioral views of investment decision making, the approach may also help financial services firms better meet consumers' needs. Unlike current practice, the approach does not attempt to match people to products or to reduce risk to a unidimensional construct. The key idea is enabling consumers to explore probability distributions of prospective outcomes, which are constrained by market forces and the amount invested.

To begin, imagine an employee on the first day of a new job. He or she is taken down the hall to Human Resources, is told about the retirement plan, and is shown a catalog of investments consisting of stocks, bonds, and money market funds.

At this juncture, the retirement investment decision is a product-selection task. The products are funds. The prices are the fees, which are usually 1%–3%, with lifetime costs that can reach tens of thousands of dollars. The product attributes are the funds' performance history, composition, brand image, and so on (Capon, Fitzsimons, and Prince 1996). To help make this decision, investment management firms provide brochures describing product attributes and so-called risk-tolerance quizzes. While commonly used, these quizzes do not appear to be the result of extensive evaluation; we are aware of only one published study on them, which shows that the various quizzes used by different fund providers have at best modest intercorrelations (Yook and Everett 2003).

There are many potential sources of information. Employees may take the provided quiz, do their own research over the course of several days, or talk to an advisor. However, there is some evidence that people do not do much research. One study of university employees found that the majority spent less than an hour deciding and did not consult with anyone other than family members (Benartzi and Thaler 1999). The short time spent on the initial allocation is not troubling if employees later revise their choice. But for many, this allocation rarely changes (Agnew, Balduzzi, and Annika 2003). An early study found that 100% of employees did not change their initial allocation in their entire tenure of employment (Samuelson and Zeckhauser 1988), and a

*Daniel G. Goldstein is an assistant professor of marketing, London Business School, London, NW1 4SA, United Kingdom (dgoldstein@london.edu). Eric J. Johnson is Norman Eig Professor of Business, Columbia University, New York, NY 10027 (ejj3@columbia.edu). William F. Sharpe is the STANCO 25 Professor of Finance, Emeritus, Stanford University Graduate School of Business, 518 Memorial Way, Stanford, CA 94305-5015 (wfsharpe@stanford.edu). This research has been supported by the Columbia Center for Excellence in E-Business and by National Science Foundation grant SES-0352062 to Eric J. Johnson. The authors thank the editors and reviewers, in addition to the following people, for help in developing this article: Beatrice Belizaire, Shlomo Benartzi, Phil Blythe, Noel Capon, Isaac Dinner, Emel Filiz, Wendy Garrido, Dominique Goldstein, John Hauser, Bruce Hardie, Peggy Hu, Raghu Iyengar, Anja Lambrecht, John Payne, Margaret Pierson, Drazen Prelec, Duncan Simester, Richard Thaler, Elke Weber, Martijn Willemsen, and Besir Wrayet.

John Deighton served as editor and Ann McGill served as associate editor for this article.

Electronically published August 21, 2008

 0093-5301/2008/3503-0010$10.00. DOI: 10.1086/589562

later one found that 87% made one or zero reallocations (Ameriks and Zeldes 2004). Thus, retirement investing has similarities with the purchase of subscription services in which initial decisions exhibit inertia.

The decision space faced by the consumer is large. A recent study of hundreds of retirement plans showed that they offered as many as 59 funds, with most offering between six and 22 (Huberman and Jiang 2006). Assuming a plan presents 10 funds, our employee must make three main decisions: how much to contribute, which funds to invest in, and how much to allocate to each of the funds.

If we use the United States as an example, the first decision—contribution level (the amount of money to invest each year)—can range from $0 to a maximum of $16,000. While this is an important decision, it is not the main focus of this article, in part because various companies and countries are experimenting with the introduction of auto-enrollment, linking retirement investing to the marketing literature on defaults (e.g., Madrian and Shea 2001). We focus on the second and third decisions.

The second question—how many different funds to select—could be simple. An investor might dedicate 100% of his contribution to a single fund. However, if he seeks further diversification, this is followed by a third decision—how to allocate the contribution across funds (asset allocation). Allocating between funds A and B, he could split the contribution many different ways: 5% to fund A and 95% to fund B, 10% to A and 90% to B, and so on. To compute the size of the decision space, we can assume that a person may choose a portfolio of between 1 and 10 funds—this would cover nearly all the plan participants studied by Huberman and Jiang (2006)—and make allocations in units no smaller than five percentage points. Assuming 16 possible contribution levels from $1,000 to $16,000 per year, this yields a choice set over 160 million possibilities.[1] In practice, the decision space is much larger, since employees may be presented with dozens of funds and can deal in units as small as 1% and $1. Though the options are numerous, somehow they are whittled down to one in as little as an hour. These rapid decisions can lead to outcomes that differ by large amounts over the course of a lifetime. This is obvious for the case of saving something versus saving nothing, but even a reallocation of a few hundred dollars per year from a fund returning 7% to a fund returning 3% can make tens of thousands of dollars of difference over the years.

This decision is also marked by uncertainty at multiple levels. For instance, there are at least two levels of risk at play in fund investing: risk between portfolios and risk within portfolios. Choosing a portfolio, as shown above, is risky because it involves choosing one option from a large and varied set. Once a portfolio is selected, however, each component fund or stock is a risk on its own, taking on many possible values over time. The risk of a portfolio also involves changing interfund correlations, interest and inflation rate fluctuations, currency risk, and beyond. In defined-contribution plans, a person is not signing up for a guaranteed level of income in retirement, as in defined-benefit plans. Rather, the income in retirement is essentially a random draw from a probability distribution of wealth. The three decisions we have discussed thus far affect the distribution, not the outcome. While marketers have modeled uncertainty (see Woodruff [1972] for an early example), it is usually uncertainty in the beliefs of the consumer about product attributes, not uncertainty in the realized outcome.

Forecasting a portfolio's outcome is difficult. Experts do so with Monte Carlo simulations and still miss the mark when fundamental assumptions are wrong. Research from behavioral finance suggests that people either do not know or do not like how their investments will translate into probability distributions of wealth in retirement. Benartzi and Thaler (2002) used financial techniques to forecast the retirement wealth distributions resulting from people's own investment choices and those resulting from two alternatives (based on the mean and median choices of other plan participants). Only 20% of people preferred the outcome distributions arising from their own investment choice over those of the median plan participant, which suggests that people could not forecast how their choices would translate into outcomes or that people had not rebalanced in a while, or both. These findings suggest that consumers have significant room for improvement when making retirement investment decisions.

Economic theory (Hakansson 1970; Merton 1969, 1971; Samuelson 1969), investment advice (Perold and Sharpe 1988), and the marketing collateral of mutual fund vendors often recommend that people divide their contributions between a risky and risk-free (or very low-risk) asset and perform periodic rebalancing to fixed proportions by dollar value. These recommendations are based on the assumption that people's risk preferences are described by constant relative risk aversion (CRRA). If returns are uncorrelated over time, such an investment strategy yields a roughly log-normal distribution of long-term returns.[2] If consumers follow this strategy, their wealth at retirement will effectively be drawn from a log-normal distribution. This raises a key question: How many investors prefer a log-normal distribution of wealth? Is a log-normal distribution the most suitable alternative given most people's stated preferences? Some investors may desire more downside protection or more upside potential. If it turns out that a substantial number of people are loss averse (Benartzi and Thaler 1995; Kahneman and Tversky 1979) and reference dependent, they would prefer non-log-normal distributions. Following standard investment advice would produce outcomes inconsistent with such nontraditional stated preferences.

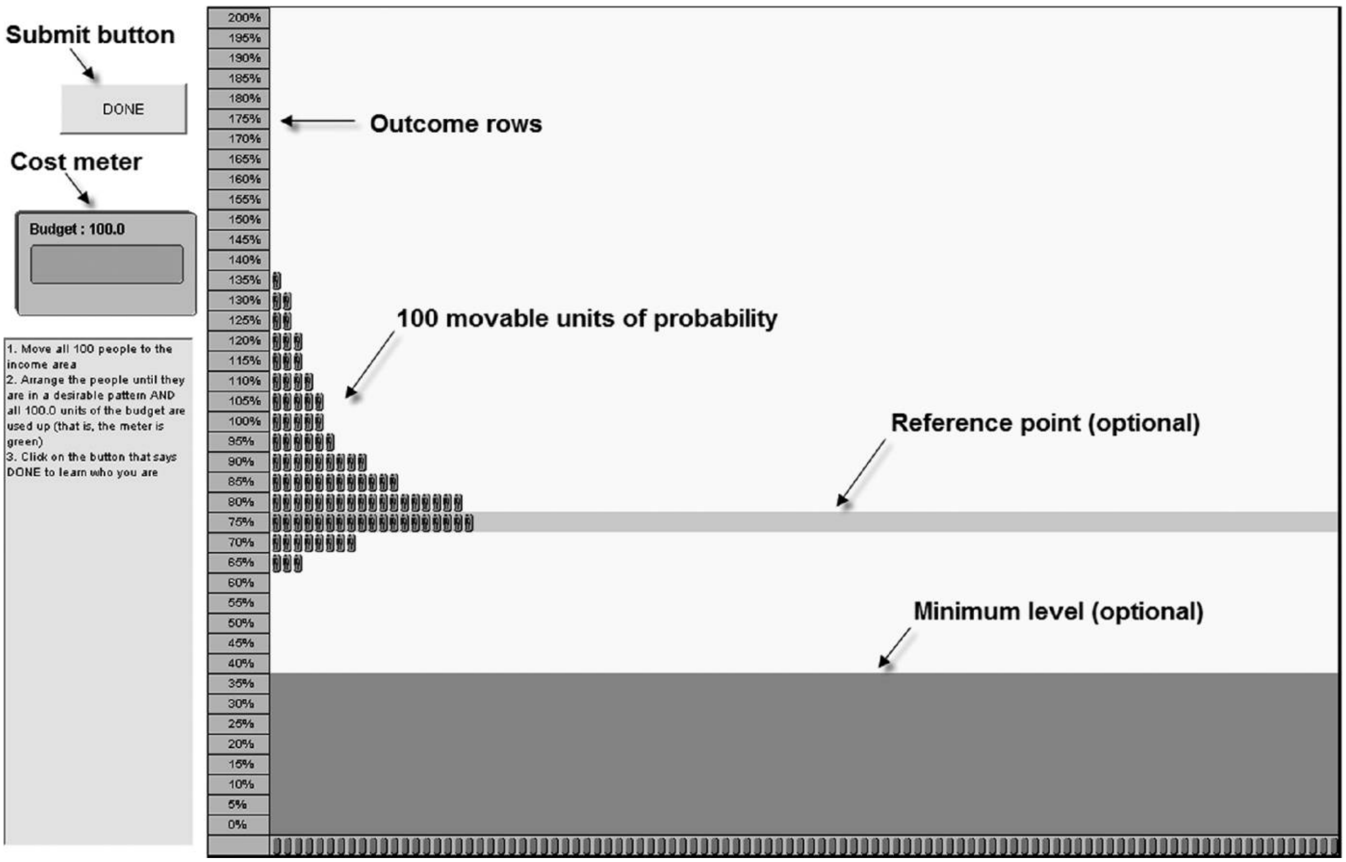
In what follows, we describe a method for gaining insight

[1]There are 20 units of 5% to be allocated and 10 possible funds in which to place them, and there are $\binom{N + r - 1}{r - 1}$ ways to partition N objects into r categories.

[2]If returns are independent, the ending value for such a strategy will converge to log-normality as the number of periods increases. If r is the rate of return, the ending value is the product of $1 + r$, its log is the sum of the logs of $1 + r$, and the central limit theorem holds that the sum of uncorrelated random variables converges to normality.

FIGURE 1

THE DISTRIBUTION BUILDER INTERFACE



NOTE.—Using movable units of probability, participants can create arbitrarily shaped discrete probability distributions over numerous outcomes (on the vertical axis). Between two and 40 outcomes and one and 100 units of probability can easily be displayed on a standard-size monitor. The 40-outcome/100-unit case provides over 10^{34} unique distributions to choose among. A cost meter (upper left) can be used to constrict the space of allowable distributions, for example, to those that have a particular risk-return relationship. The cost meter functions by not allowing one to submit a distribution (using the submit button on the upper left) until it satisfies an arbitrary cost function. Users can see how every change to the distribution affects the cost meter numerically and graphically. All movements are seen in the context of their effects on the system as a whole. Color version and video demonstration available as online enhancements.

into investor preferences and use it to elicit desired wealth distributions from a sample of working adults who have been saving for retirement for 5–30 years. The method passes tests of reliability and validation and allows us to characterize individual differences in both risk attitude and loss aversion. We conclude by discussing how this method—exploring probability distributions of potential outcomes—might serve as a market research tool and might benefit consumer welfare by helping people construct, as opposed to state, preferences.

THE DISTRIBUTION BUILDER

The Distribution Builder (DB) is a market research tool that has been used to estimate the coefficient of constant relative risk aversion (CRRA parameter) from user-constructed probability distributions of desired outcomes (Sharpe 2006; Sharpe, Goldstein, and Blythe 2000). In this article, we describe a version of the tool created to elicit preferred probability distributions of retirement income.

Figure 1 shows the interface of the DB. On the vertical axis are outcome rows, percentages ranging from 0% to 200%. These represent income in retirement expressed as a percentage of preretirement income. If a person earned \$100,000 in the year before retirement, the 75% row would represent \$75,000 annual income in retirement. In the experiments described below, participants were told that 75% of preretirement income is a typical goal for income in retirement, and the 75% row was highlighted as a reference point. Indeed, popular investment advice, such as that given in *Ernst and Young's Retirement Planning Guide*, lists income-replacement ratios ranging from 68% to 80% (Arnone et al. 2002).

The interface has 100 markers forming a probability distribution against its left vertical axis. At the beginning of a session, the markers sit at the bottom of the screen and are dragged up into the square exploration area with the mouse. Clicking and dragging one marker moves it and all the markers on its right. Each of the 100 markers has an image of a person drawn on it. The participant is told that just one of the 100 markers represents her and that the others are just placeholders. Importantly, she is told that the marker representing her was chosen at random by the computer, and there is no way to tell which one it is before submitting a distribution. The participant arranges the markers into a probability distribution of her pleasing—that is, one she would be happy to have determine her own wealth in retirement. When finished, she clicks the "Done" button to simulate observing a randomly determined investment outcome. Clicking "Done" causes 99 randomly chosen markers to disappear slowly, one by one, until only one marker is left standing on its respective row. This relates a draw from a probability distribution to the amount of wealth one might have in retirement. For a demonstration of the DB, see video 1 (6.4 MB) in the online edition of the *Journal of Consumer Research*.

One obvious question is why, if higher rows represent more retirement wealth, do participants not move all 100 markers up to the top row, giving a 100% chance of enjoying 200% of preretirement income in retirement? The answer is that distributions are cost constrained, and participants may only submit a distribution of a given cost. The cost of each distribution is represented by the cost meter (the "Budget" reading on the left in fig. 1). As markers are moved upward, the cost goes up, and as markers are moved downward, the cost goes down. In retirement decision making, the DB uses a cost function that estimates the cost of obtaining each distribution in financial markets. The details of the cost function are beyond the scope of this article but may be found in Sharpe et al. (2000) and in Sharpe (2006). This type of cost function works by weighting each of the 100 markers by a unique state price and only allowing the submission of distributions in which the sum of the 100 chosen outcome rows (each weighted by its state price) is roughly equal to a given budget. In intuitive terms, with the DB, the participant is given a budget and state prices (weights) and is asked to create a distribution of outcome values that satisfies a cost constraint. We use the distribution of markers to make inferences about the person's utility for outcome levels, as will be shown.

In our studies, participants can only submit distributions that cost between 99 and 100 units of a hypothetical budget. We used a budget that was large enough such that participants could obtain 75% of their preretirement income risk free. That is, they could put all 100 markers on the 75% row and satisfy the cost meter. The 75% row is important because it serves as a reference point for three reasons: First, participants are told that 75% is a typical goal. Second, it stands out from the background. Third, it is the risk-free alternative that participants can discover, although they are not told about it. Figure 2 shows that the risk-free alternative satisfies the cost meter (i.e., all 100 units of the budget are used).

In the studies that follow, we have configured the DB so that it allows a user to trade off risk and return. People can decide to accept a certain amount of downside loss to provide an even greater possible amount of upside gain. For instance, in figure 1, a person taking 11 chances in 100 of ending up below 75% is rewarded with 70 chances in 100 of ending up above 75%.

At this point, we briefly provide technical details of how the DB allows for the estimation of utility functions. Readers less interested in these specifics might wish to skip ahead to the description of the experiments.

USING COST-CONSTRAINED DISTRIBUTIONS TO MAKE INFERENCES ABOUT UTILITY FUNCTIONS

Our goal is to use the cost-constrained placement of markers on the DB interface to estimate utility functions and parameters at an individual level. As a practical consumer concern, we can test whether participants' utility is consistent with advice often given to retirement investors.

Distributions are constrained by cost according to the function $\sum_{i=1}^{N} o_i s_i = B$, where the budget B is the exact enough amount of money needed to retire at 75% of preretirement income if only a risk-free investment is used, o_i is the outcome row associated with the ith marker, and s_i is the Arrow-Debreu state price (or weight) of the ith marker (Arrow 1964; Debreu 1959). The outcome rows are indexed in increasing (or nondecreasing) order, but the state prices are indexed in decreasing (or nonincreasing) order. The least-cost way to obtain a distribution of outcomes is thus $\sum_{i=1}^{N} o_i s_i$, which is the cost associated with the distribution and is shown as a percent of the budget (B) on the cost meter.

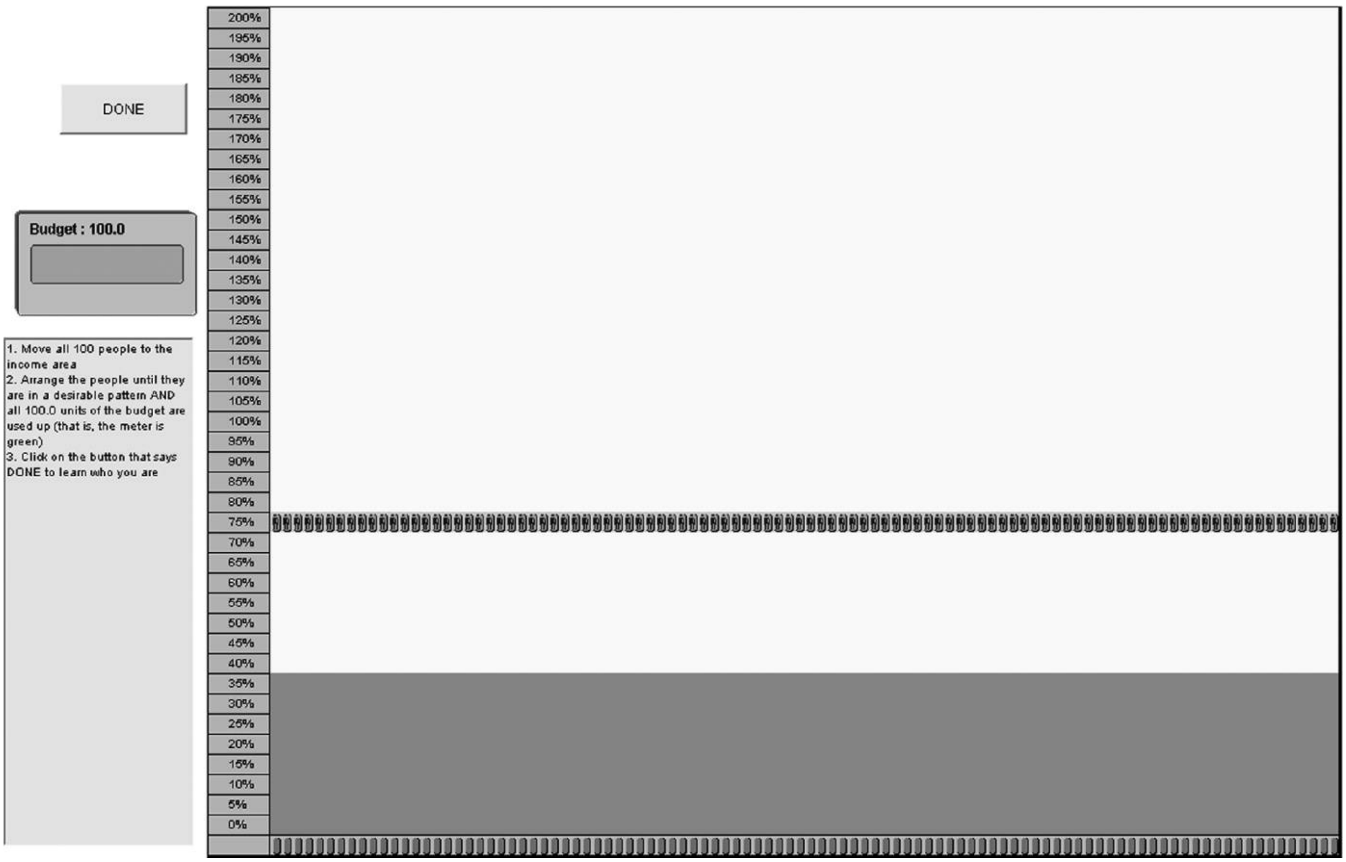
A key attribute of this procedure is that taking on downside risk allows for proportionally more upside gain, as in real markets. The intuition for this is as follows. The markers at the bottom of the distribution are more heavily weighted than those at the top. Each time a marker is moved, the state prices are reassigned so that this relative weighting holds. Since lower markers have more impact on the cost function than do higher markers, the result is a risk-return trade-off, which can be seen by comparing figure 1 with figure 2. Both figures satisfy the cost meter, but figure 1 clearly has the higher expected value.

While any set of unique state prices could provide useful information about utility functions, it is desirable to choose trade-offs that are similar to those in actual capital markets. We provided the DB with a set of state prices that, given assumptions, made the budget meter readout proportional to the cost of obtaining the distribution with a least-cost investment strategy (Dybvig 1988; Sharpe et al. 2000).

Inferring utility: The DB can be used to infer utility functions from distributions instead of from choices between simple two-alternative gambles, as was the dominant paradigm in the twentieth century. In order to do this, we assume that a person building a distribution is maximizing

FIGURE 2

LOWEST-RISK ALTERNATIVE



NOTE.—The lowest-variance distribution that satisfies the cost meter has all 100 markers on the 75% row. Notice that figure 1 has roughly the same cost but a visibly higher expected value, reflecting the risk-return trade-off of real markets. Color version available as an online enhancement.

$\sum_{i=1}^{N} p_i u(o_i)$, where p_i is the probability of marker i (all 1/100) and $u(o_i)$ is the utility associated with the outcome row chosen. This maximization is subject to the cost constraint $\sum_{i=1}^{N} o_i s_i = B$. Solving the maximization problem results in the family of equations $p_i u'(o_i) = k s_i$ for each marker i, where k is a constant. If we let K be any of the identical p_i divided by k, this becomes

$$Ku'(o_i) = s_i. \qquad \textbf{(1)}$$

That is, the slope of the utility function at marker i is proportional to the ith state price. This key result can be used to make inferences about individual utility functions and to assess reference dependence (loss aversion).

Predictions of CRRA: A common classical assumption is that investors' risk preferences possess a property called constant relative risk aversion, or CRRA (Arrow 1970; Barberis 2000; Pratt 1964; Safra and Segal 1998). The three most commonly used utility functions in financial economic models are the quadratic, the constant absolute risk aversion, and the CRRA. The quadratic function implies that an investor will put fewer dollars in riskless assets as he or she becomes wealthier. The constant absolute risk aversion function implies that an investor will put the same number of dollars in riskless assets as wealth increases. The CRRA function implies that the proportion of wealth invested in riskless assets is invariant with respect to changes in initial wealth level. Of the three functions, CRRA provides implications closest to observed behavior on the part of most investors (Arrow 1970; Pratt 1964).

People with CRRA utility functions should prefer a constant mix of assets if returns are uncorrelated over time, resulting in log-normally shaped distributions of terminal wealth. A typical CRRA utility function is the power utility function $u(o) = (o^{1-\alpha})/(1-\alpha)$. An investor with such a function will have a marginal utility of $u'(o) = o^{-\alpha}$. Combining this with equation 1 and taking logarithms gives

$$\ln(s_i) = \ln(K) - \alpha \ln(o_i). \qquad \textbf{(2)}$$

In the finance literature, this is commonly referred to as the coefficient of CRRA. If most people's risk preferences are well fit by CRRA, standard investment advice should help them select investment products. If not, then many people who follow the standard advice could be holding portfolios that violate their expressed preferences. In the experiment that follows, we fit equation 2 at an individual level by plotting the log state prices of the markers against the log outcome row levels and computing the regression line. If the relationship is not linear (low R^2), it would suggest that investor preferences may be better described by a model other than CRRA. If the relationship is linear, we assume that CRRA fits well and would expect estimates of the coefficient of relative risk aversion α to fall in the range commonly found by other means.

Alternatives to the traditional CRRA model are loss-averse utility functions such as that in prospect theory (Kahneman and Tversky 1979), which have received recent attention in the marketing domain (Ariely, Huber, and Wertenbroch 2005; Camerer 2005; Novemsky and Kahneman 2005). In loss-averse utility functions, losses have a higher impact than gains due to a loss-aversion parameter λ, which applies on one side of a reference point perceived to be the border between gains and losses. One means of estimating λ is taking the ratio of the slopes of the utility function on opposite sides of this border (Köbberling and Wakker 2005). The loss-aversion parameter has been empirically estimated to be around 2.25 (Tversky and Kahneman 1992). Note that we do not concern ourselves with fitting various parameters of prospect theory in this article; we only estimate the loss-aversion parameter.

Predictions of loss aversion: As mentioned, the participants in our experiment had a reference point: the 75% row. Recall that from equation 1, the state price s_i of the ith marker equals the slope of the utility function times a constant—that is $Ku'(o_i) = s_i$. The kink in the utility function parameterized by λ can thus be estimated by the ratio of state prices of the markers on opposite sides of the reference point (loss side divided by gain side). From the distributions submitted, we will estimate λ at an individual and group level.

Experiment 1

The purpose of the first experiment is to estimate the parameters of risk aversion and loss aversion for preferences concerning wealth in retirement and to obtain a descriptive overview of data gathered with DB.

Method and Participants. Experimentation was carried out online. Participants were 152 geographically diverse U.S. citizens who were randomly selected from our participant pool on the basis of age and citizenship and who were paid for their participation. Participants had an average age of 42 years with a standard deviation of 8 years. All have been saving for retirement for 5–30 years. Of the participants, 76% were married, 12% were single, and 12% were divorced or widowed; 71% were female, in line with the gender ratio of the entire subject pool at the time. Median income was $50,000 with a standard deviation of about $33,000. Average net worth was estimated to be about $200,000, and average amount saved toward retirement was about $110,000.

Participants were instructed to think about income in retirement as a percentage of preretirement income. They were told that 75% of preretirement income was a typically recommended goal. An extensive training session explained how to interact with the DB. Participants created one practice distribution that did not have any budget constraint (cost meter) but that did demonstrate the one-by-one random selection of markers once the distribution was submitted. After this, users were instructed on the role of the cost meter. They were told their task was to find a pattern of markers that they would like to have apply to their own income in retirement and that uses between 99% and 100% of the budget. The importance of treating the task as if it concerned their own income in retirement was emphasized. Participants were strongly advised against taking a chance of going below 25% of preretirement income, and, for this reason, as seen in figures 1 and 2, the bottom rows of the DB are shaded. Participants created two distributions, one after the other, and then went on to answer a survey containing demographic questions.

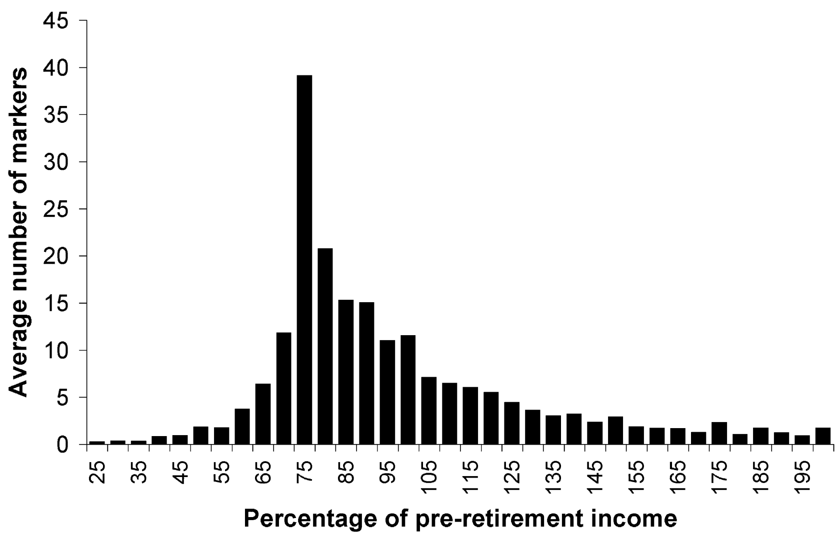
One year after submitting their first two distributions with DB, 85 participants from the above described set and 73 participants from a similar DB experiment completed a follow-up study, for a total of 158 year 2 respondents. This other experiment, which we do not analyze here, was the same as the reported one, except that the budget constraint had been set such that the risk-free level was 60% instead of 75% in the first session. Jumping ahead, there is no significant difference in year 2 risk-tolerance estimates between participants from two different year 1 experiments (t ratio $-.4688$, 144 df). Demographics of this group are naturally very close to the first-year set: average age 41 years, median income $50,000, 68% female.

As in year 1, participants submitted two distributions concerning desired income in retirement. To provide criteria for validation, respondents were then presented a DB on which they were asked to play two small-stakes gambles for a gain of up to $1.25 or a loss of up to $.75, with a risk-free alternative of $0 and outcome row increments of $.05. The stakes were real: money was added or subtracted to the participants' payments on the outcome of these gambles. Outside of the axis labeling, the mechanics of the DB were exactly the same as in the retirement scenarios. After submitting the two retirement and two gamble distributions, participants engaged in additional validation tasks, described in experiment 2.

Distributions. Figure 3 shows the aggregate distribution based on the number of markers at each wealth level from both distributions, averaged across all participants in the first year. Note that the composite distribution is roughly log-normal in shape, as CRRA would predict, but exhibits a peak at the reference point at 75%, congruent with loss-averse preferences, as will be discussed. Later, we will break apart this aggregate distribution into the clusters it com-

FIGURE 3

AGGREGATE DISTRIBUTION



NOTE.—The distribution of the average investor, created by combining the number of markers at each wealth level in both distributions and averaging across all participants. The distribution is right skewed and shows a peak of twice the next highest level at the reference point. The 75% row served as a reference point because participants were told it is a typically recommended goal level for income in retirement and because 75% was the outcome that could be obtained without taking any risk (placing all 100 markers at 75% would satisfy the budget meter).

prises. Looking across the retirement distributions individually, the most common modes of the distributions are the 75%, 80%, 85%, 90%, and 100% wealth levels, which describe 55%, 11%, 7%, 6%, and 3% of distributions, respectively. Clearly, the intended reference point of 75% was a popular choice. Since we could have chosen 70% or 80% as the reference point, and distributions would presumably have shifted accordingly, these data show the reference-dependent nature of preferences. For instance, in the previously mentioned experiment in which the risk-free level was 60% but the recommended level was 75%, we observed aggregate distributions that peaked in both places.

Aggregate Model Fit. We fit a power function $u(o) = (o^{1-\alpha})/(1-\alpha)$, which is a CRRA utility function, to each participant's first and second distributions. Since the state prices (s_i values) were given and the participants chose a distribution of outcome (o_i) values, we treated the logarithm of the former as the independent variable and the logarithm of the latter as the dependent variable in our regressions. Rearranging equation 2 shows that the slope in such a regression provides an estimate of $-1/\alpha$, which can be easily transformed into the coefficient of relative risk aversion α.

The loss-aversion parameter λ was estimated by taking a ratio of the state prices (which are multiples of the slope of the utility curve) of the markers on either side of the reference point, namely the 75% row. Since this row is a bin that could be seen as ranging from 72.5 to 77.5, one cannot say which markers would sit just above and below 75. We take as a proxy the two most differing state prices in this row, which should reflect the change in slope experienced when moving across the reference row. Because of this issue of granularity, and because this estimate of loss aversion is not independent of risk aversion (someone who places many markers at 75 may simply be expressing a preference for a low-variance distribution), we consider our estimates to be somewhat rough approximations. We have tested a more complex model-based method to estimate loss aversion while accounting for risk aversion and find its estimates to correlate highly (>.9) with the estimates we provide here. For the sake of simplicity, we stick with the straightforward method.

Table 1 provides estimates of α, λ, and R^2 for the six individual distributions and for all retirement distributions together. We exclude cases in which CRRA and loss-aversion parameters cannot be estimated because no markers were placed below the reference point. This excludes 50 of the 620 retirement distributions we collected, a sizable exclusion that in the future can be avoided by altering the tool's submission constraints. Apart from these cases, the CRRA model tended to fit the data well, with median R^2 values around .9. The difference between the medians and means, as well as figure 4, shows skewness in both parameters, with most individual distributions having rather low risk and loss aversion. When the tool was used to engage in two small-stakes gambles for real money, risk aversion was less than when retirement was contemplated, as would be expected for smaller stakes. The first distribution from the first session stands out with its low risk- and loss-aversion estimates, possibly due to learning effects.

TABLE 1

PARAMETER ESTIMATES

	Risk aversion		Loss aversion		R^2	
	Median	Mean	Median	Mean	Median	Mean
Year 1 distribution 1	4.3	5.8	1.3	2.4	.91	.88
Year 1 distribution 2	6.1	10.6	1.8	4.7	.89	.85
Year 2 distribution 1	6.3	8.8	1.8	3.9	.90	.86
Year 2 distribution 2	8.3	10.7	2.2	4.9	.88	.84
Year 2 gamble distribution 1	4.9	8.5	1.4	4.4	.88	.82
Year 2 gamble distribution 2	4.1	6.8	1.5	4.6	.88	.81
All retirement distributions	6.1	9.0	1.8	4.0	.90	.86

NOTE.—Estimates of coefficient of constant relative risk aversion α and loss-aversion parameter λ, as well as R^2, for the six types of distributions and for all retirement distributions together.

For all retirement distributions, median α is 6.1, consistent with economic estimates in the range from 1 to 10 or higher (Blake 1996; Brav, Constantinides, and Geczy 2002; Campbell 1996; Friend and Blume 1975; Mehra and Prescott 1985). The median λ is 1.8, which is consistent with many other estimates (Tversky and Kahneman 1992). While this might be considered to indicate widespread occurrences of kinked utility curves (i.e., loss aversion), it is important to recall the limitations of our estimate of λ given earlier and to recall that figure 4 shows that many individuals are not loss averse.

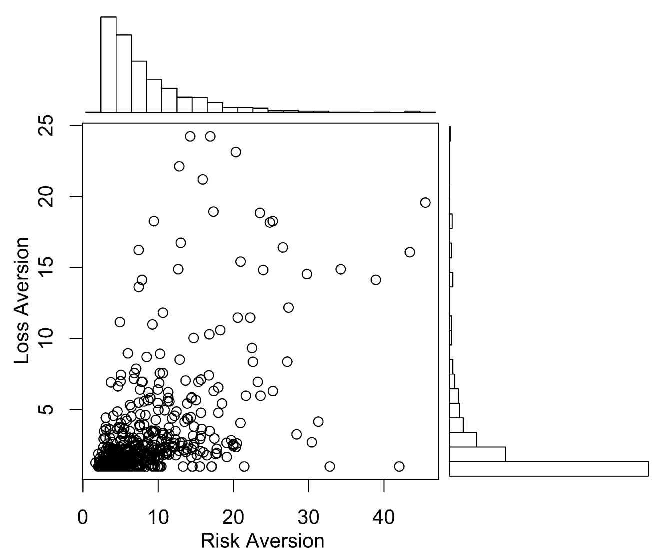
Individual Differences. Do these aggregate parameter estimates mask variation across people? In other words, are there subpopulations of people with different risk preferences? Are participants who are not well fit by CRRA loss averse, or do they deviate in some other way from the CRRA model?

When fitting the CRRA model, median R^2 in the experiment centers around .9. If we perform a median split and categorize participants with $R^2 \geq .9$ as CRRA and $R^2 < .9$ as non-CRRA, the difference in the distributions between the two subpopulations of participants is easily visually observed. Figure 5 shows the composite distributions of high and low R^2 participants. The average R^2 was computed per participant on the basis of his or her two submitted distributions. Both distributions went into either the high or low R^2 composite histogram depending on this average. The distribution of the high R^2 group is roughly log-normal in shape, skewed to the right. The low R^2 distribution deviates from this smooth form in a very specific way, consistent with reference dependence. To understand why this shape suggests loss aversion, recall that the loss-aversion parameter is estimated by comparing the state prices of the two most differing markers at the reference row. As the number of markers at the reference row increases, so does this loss-aversion ratio. Note that this relationship between R^2 and loss aversion is not obvious. It is possible to create low R^2 distributions with low loss aversion, as is the case with a bimodal distribution or a distribution with probability massed around a point other than the reference point.

Summarizing experiment 1, we have estimated risk- and loss-aversion parameters on six distributions, obtaining estimates that fall into the expected ranges. The DB technique captures 100 points of a distribution that can be used to infer the slopes on a utility curve and can be analyzed by various means to gain insight into individual differences. While this may be a useful tool, practical applications depend upon establishing the reliability and validity of the method. Our data also suggest that there are important individual differences, which we explore later. In experiment 2 we go beyond face validity to demonstrate reliability and examine predictive validity.

FIGURE 4

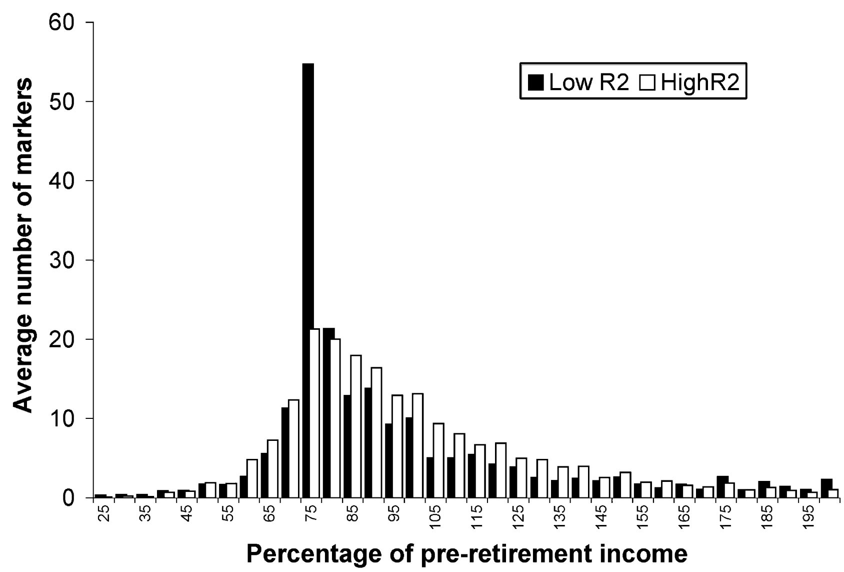
RISK-AVERSION AND LOSS-AVERSION PARAMETER ESTIMATES



NOTE.—Estimated risk-aversion parameter α and loss-aversion parameter λ computed for all retirement distributions submitted in year 1 and year 2. Marginal histograms show that the distribution of both estimates is strongly skewed, with most points having low risk and loss aversion. For scale, the chart omits 11 points with loss aversion greater than 25.

FIGURE 5

POPULATION SPLIT BY FIT TO CRRA MODEL



NOTE.—Composite distribution of participants who are well (high R^2) and poorly (low R^2) fit by the constant relative risk aversion (CRRA) model. The low R^2 people do not deviate arbitrarily, but rather in a systematic way, from CRRA (log-normality). The massing of probability at the reference point (here 75%) is a property of a loss-averse distribution. The bars of each color sum to 200 because each participant submitted two 100-marker distributions.

Experiment 2

Method. In order to assess individuals on covariate measures of risk preference, participants from the year 1 study completed the following validation tasks after creating their two distributions.

Outcome Preferences Task. In this task, participants were shown histograms representing distributions of returns resulting from portfolios holding 0%, 10%, 20%, 30%, 40%, 50%, or 60% in stock and the rest in a risk-free asset and were asked which pattern of investment results they would like to apply to their own retirement. Graphs were made with the same assumptions about budget, holding period, and stock returns as those used for the DB. Histograms of returns were chosen as a validating measure for four reasons: they have been shown to facilitate estimates of volatility (Ibrekk and Morgan 1987), they have been favored by people assessing risks (Thompson and Bloom 2000), they are commonly used in financial prospectuses and studies of financial risk perception (Siebenmorgen, Weber, and Weber 2000), and, finally, unlike asset allocation tasks, they preclude application of the 1/*N* heuristic (Benartzi and Thaler 2001), which can bias responses. Because the histograms we present consist of a constant mix of stocks and a risk-free asset, they are log-normal in shape and thus approximate the options available to people who follow the constant-mix investment advice.

Gamble Choice Task. To validate DB in comparison with a dominant form of risk-assessment task, participants were presented with three choices between sure amounts and gambles. The first item was "Which would you prefer if offered right now 1) $4.50 for certain 2) A 50% chance of getting $1 and a 50% chance of getting $15." The choices in the other items were "1) $1 for certain 2) A 10% chance of getting $12 and a 90% chance of getting nothing" and "1) $8 for certain 2) A 90% chance of getting $10 and a 10% chance of getting nothing." The number of risky choices (0–3) was computed for each participant.

Risk-Tolerance Scales. Participants completed two risk-tolerance scales. The first was the gambling and investment subscale of the Weber-Blais-Betz Domain Specific Risk Attitude Scale (Weber, Blais, and Betz 2002). Participants also completed a reworded variant of an actual risk-tolerance self-assessment scale used by one of the world's largest providers of retirement investment products.

Reliability, Validation, and Predictive Validity

Reliability. To assess the reliability of the risk measures collected with DB, we computed the reliability coefficient as the Pearson correlation between parameters estimated from two distributions submitted one after the other, as shown in table 2. Within the first session in year 1, the two measurements of the transformed CRRA model-based risk parameter $-1/\alpha$ have a Pearson correlation of .700 (Spearman .702). In the year 2 session, this correlation reached .803 (Spearman .793). Note that the respondents started with a cleared screen for each distribution, so these high corre-

TABLE 2

TEST-RETEST RELIABILITY

Component	Pearson correlation	Spearman correlation	Count
Within year 1	.700	.702	152
Within year 2	.803	.793	148
Between year 1 and year 2:			
Averages:			
Uncorrected	.431	.454	75
With correction for attenuation	.575		75
With ≥$25,000 savings:			
Uncorrected	.583	.559	57
With correction for attenuation	.778		57

NOTE.—Test-retest correlations of the constant relative risk aversion (CRRA) parameter computed within years, between years, with and without correction for test-retest attenuation, and with the exclusion of those who have little or nothing saved toward retirement.

lations are not due to merely resubmitting the previous response. To look at reliability over time, the average $-1/\alpha$ value from year 1 is correlated with its corresponding value in year 2, giving a Pearson correlation of .431 (Spearman .454). As Ghiselli, Campbell, and Zedek (1981) point out, the between-year reliability measure $r_{12(\mathrm{true})}$ must take into account the attenuation due to within-year reliabilities and is $r_{12}/\sqrt{r11r12}$, where r_{11} and r_{22} are the respective reliabilities within year 1 and year 2. Using the Pearson correlations as reliabilities, here $r_{12(\mathrm{true})}$ is equal to .575. Reasoning that participants with little or no experience in investing for retirement may provide less reliable data, we reran the analysis, excluding those with the lowest level of retirement savings ($0–$24,999). The uncorrected across-year correlations were higher at .583 (Spearman .559), and with the correction they reached .778.

Validations

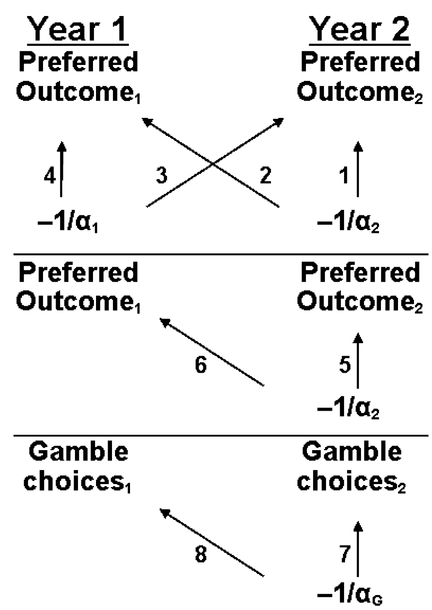
Because the validations span two dependent variables (outcome preferences and gamble choice) and two types of estimates of $-1/\alpha$ (from retirement distributions and small-stakes gamble distributions) across two years, figure 6 is provided to serve as a visual guide. The goal of the validations is to show that the risk-aversion parameter, as estimated by DB, is a valid predictor of preferences even in the presence of numerous traditional predictors of risk attitude. In addition to the psychological and industry scales described, we look at the demographic variables of age, income, and gender, which have long been studied as covariates of risk preference. In general, younger people, wealthier people, and males are found to exhibit less risk aversion (for a review, see Bajtelsmit and Bernasek [2001]).

Predicting Outcome Preferences. The variance of the distribution of returns in the outcome preferences task is a common measure of risk attitude: greater variance implies greater risk tolerance. We modeled the standard deviation of the histogram chosen in the outcome preference task with six predictors: (1) $-1/\alpha$, the average transformed coefficient of relative risk aversion as estimated by DB, thus transformed to make it amenable to regression analysis; (2) age in years; (3) income, log-transformed; (4) gender, coded 1 for male and 2 for female; (5) industrial risk profile; and (6) psychological risk profile, the participant's score on the gambling and investment subscales of the Weber-Blais-Betz Domain Specific Risk Attitude Scale (Weber et al. 2002).

Regressions 1–4 in table 3 show the results of predicting preferred outcomes. Note that N varies from regression to regression because not all participants in a given year are able to be predicted with data from another year. For instance, only 85 participants from experiment 1 participated

FIGURE 6

GUIDE TO THE EIGHT VALIDATION REGRESSIONS



NOTE.—Guide to the eight validation regressions. α_1 and α_2 are based on the retirement distributions in years 1 and 2. α_G is based on the small-stakes gamble distribution, which was made in year 2 only.

TABLE 3

VALIDATIONS

	Year 2 retirement distributions[a]						Year 1 retirement distributions[a]					
	Year 2 outcome preference[b] (1)			Year 1 outcome preference[b] (2)			Year 2 outcome preference[b] (3)			Year 1 outcome preference[b] (4)		
	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio
Intercept	−4.072	9.841	−.41	−5.484	14.856	−.37	−7.230	10.891	−.66	−8.537	12.281	−.7
$-1/\alpha$	−52.490	8.252	−6.36***	−50.613	13.818	−3.66***	−35.643	9.029	−3.95***	−19.390	10.104	−1.92*
Age	−.069	.112	−.62	.046	.163	.28	−.058	.129	−.45	.079	.142	.56
Income	.965	1.634	.59	1.840	2.519	.73	1.420	1.815	.78	.551	2.008	.27
Gender	.359	.961	.37	3.043	1.463	2.08**	.669	1.089	.61	1.653	1.212	1.36
Industry risk scale	.380	.390	.97	.526	.617	.85	.648	.431	1.5	.932	.512	1.82*
Psychological risk scale	.352	.134	2.62**	−.032	.189	−.17	.336	.161	2.09**	.270	.160	1.68*
R^2		.322			.224			.199			.115	
% of explained variance due to $-1/\alpha$		61.68			64.55			46.83			22.08	
No. of observations		145			79			141			136	

	Year 2 gamble distributions[a]											
	Year 2 outcome preference[b] (5)			Year 1 outcome preference[b] (6)			Year 2 gamble choice[b] (7)			Year 1 gamble choice[b] (8)		
	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio	Estimate	SE	*t*-ratio
Intercept	12.028	9.703	1.24	−2.575	14.075	−.18	1.844	.708	2.61**	1.692	.739	2.29**
$-1/\alpha$	−21.622	6.536	−3.31***	−16.514	9.249	−1.79*	−.977	.477	−2.05**	−.939	.498	−1.89*
Age	−.130	.128	−1.02	−.009	.174	−.05	.002	.009	.26	−.025	.010	−2.6**
Income	1.926	1.878	1.03	3.950	2.749	1.44	−.142	.137	−1.04	.104	.143	.73
Gender	1.150	1.097	1.05	1.699	1.513	1.12	.295	.080	3.69***	.093	.083	1.11
R^2		.124			.117			.120			.109	
% of explained variance due to $-1/\alpha$		56.40			33.05			22.53			21.32	
No. of observations		142			78			142			142	

NOTE.—Predicting preferred outcome distributions and gamble choices based on estimates of the constant relative risk aversion (CRRA) parameter obtained with the DB method. The parameter accounts for 22%–65% of explained variance, even when predicting ahead or back one year in time. Income is log of income in U.S. dollars. α is the coefficient of CRRA, a transformation of which is used to improve suitability for regression analysis. Outcome preference is the standard deviation of the participant's choice in a task that presented participants with seven log-normal distributions of terminal wealth corresponding to a 0%, 10%, . . . , 60% investment in stock and the rest in a risk-free asset.

[a]Source of $-1/\alpha$.

[b]Dependent variable.

*$p<.1$.

**$p<.05$.

***$p<.01$.

in experiment 2, and 79 of these participants provided usable parameter estimates by submitting two non-zero-variance distributions in year 2. The transformed risk-aversion parameter α accounts for 22%–65% of variance explained by the various models and is particularly predictive in tests that involve year 2. Remarkably, the DB estimate of risk aversion predicts outcome preferences expressed one year later and one year earlier than the time it is measured, and it does so in the presence of five explanatory variables.

Regressions 5 and 6 provide a validation across tasks (gambles versus retirement distributions) and years. Since the gamble distributions were presented only in year 2, there are two regressions of this type instead of four. In both cases, the risk-aversion parameter is significant and accounts for roughly 56% and 33% of explained variance. No other predictor is significant.

Predicting Gamble Choices. Regressions 7 and 8 provide a validation entirely in the much-investigated domain of small-stakes gambles. Here, the risk-aversion parameter predicts the number of risky gambles chosen both within year 2 and from year 2 to year 1 and accounts for 23% and 21% of explained variance, respectively. It is interesting to note that when regressions 1 and 4 are compared, R^2 is higher within year 2 than within year 1, and, similarly, table 2 shows greater reliability within year 2, which is suggestive of practice or learning effects.

In regressions 5–8, models were built with four predictors: $-1/\alpha$, age, log-transformed income, and gender. We ran four additional regressions numbered 9–12, which correspond to models 5–8 but with two additional predictors: the industrial and psychological scales. In models 10, 11, and 12 (but not

TABLE 4

DEMOGRAPHICS AND RISK AVERSION

Condition and variable	Pearson correlation	p	Spearman correlation	p	N
All participants:					
Years 1 and 2 α:					
Age	.286	.013**	.261	.024**	75
Income	−.215	.066*	−.109	.354	74
Gender	−.013	.909	−.001	.991	75
Year 2 α:					
Age	.182	.028**	.200	.016**	146
Income	−.062	.456	−.081	.335	145
Gender	.022	.794	.040	.630	146
Year 2 gamble α:					
Age	.238	.004***	.211	.011**	143
Income	−.170	.042**	−.123	.145	142
Gender	.071	.397	.040	.639	143
Year 1 α:					
Age	.175	.040**	.154	.070*	139
Income	−.174	.043**	−.159	.065*	136
Gender	.056	.512	.074	.389	139
≥$25,000 retirement savings:					
Years 1 and 2 α:					
Age	.301	.030**	.261	.062*	52
Income	−.333	.016**	−.269	.054*	52
Gender	.042	.769	.094	.506	52
Year 2 α:					
Age	.243	.015**	.238	.017**	100
Income	−.190	.059*	−.214	.033**	100
Gender	.135	.181	.147	.144	100
Year 2 gamble α:					
Age	.313	.002***	.268	.008***	98
Income	−.195	.055*	−.186	.067*	98
Gender	.179	.079*	.170	.094*	98
Year 1 α:					
Age	.131	.208	.127	.222	94
Income	−.249	.017**	−.234	.025**	92
Gender	.065	.534	.076	.467	94

NOTE.—Relationship between transformed risk-aversion parameter and demographic covariates age, income, and gender.
*$p < .1$.
**$p < .05$.
***$p < .01$.

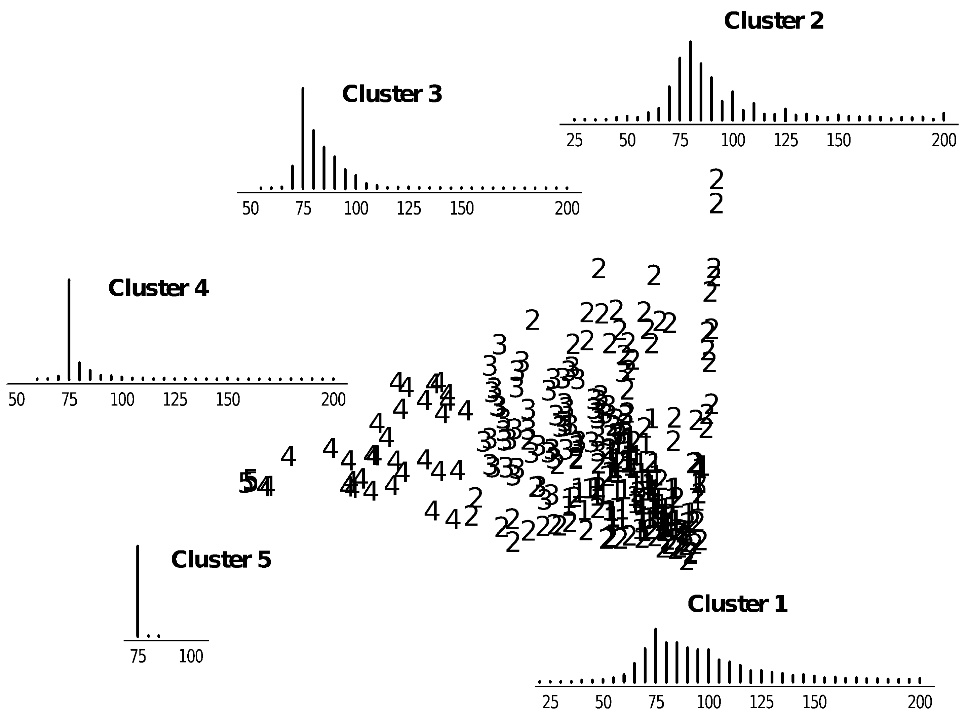
9) the coefficient on $-1/\alpha$ falls below conventional levels of significance with t-values of −1.6, −1.48, and −1.11. Thus, in models with many predictors, parameters estimated from DB gamble distributions seem to be less predictive of retirement outcomes and of two-alternative gambles. The industrial scale turns out to predict the small-stakes two-alternative gambles quite well—its t-values in models 11 and 12 are 3.15 and 5.00, respectively.

Demographic Analysis. Correlating the DB estimate of risk aversion with age, income, and gender provides indirect validations. If DB provides sensible estimates, we would expect to find significant correlations (though typically low in the literature) and consistent directional relationships between these variables and the risk-aversion parameter. Table 4 shows these correlations for all estimates of $-1/\alpha$ discussed so far, in addition to an average of all four estimates. Looking at both Pearson and Spearman correlations, age and income are related to $-1/\alpha$ in 12 of 16 cases ($p < .1$) spanning the year of measurement and type of distribution (retirement or gamble). As with table 2, results are shown excluding the lowest retirement savings group, and slightly stronger relationships are seen (14 of 16 significant relationships). Notably, in all 32 cases involving age and income, the correlation is in the expected direction (age increases risk aversion; income decreases it).

Equally notable is that in 14 of 16 cases, there is no relationship between gender and risk aversion, and in two cases only modest relationships are found (Pearson correlation .179, Spearman .17). However, in 14 of 16 cases, the expected sign is observed (being female predicts greater risk aversion). Though many studies have found relationships between gender and risk aversion, few offer ideas on why this relationship might exist, with one exception of those in evolutionary psychology (Low 2000). For a review and

FIGURE 7

CLUSTERING ANALYSIS



NOTE.—Cluster membership and location of distributions in principal components space. Data are retirement distributions submitted in year 2. The horizontal and vertical axis frames, representing the first and second principal components, respectively, are omitted for clarity. The number of distributions in clusters 1 through 5 is 100, 98, 66, 32, and 20, respectively. The five inset histograms are the aggregated markers within each cluster, plotted with vertical axes of the same scale.

some pointers to a conceptual framework, see Bajtelsmit and Bernasek (2001).

Describing Individual Differences: Cluster Analysis

We have seen that when segmenting distributions by their fit to the classic CRRA model, the well-fitting distributions appeared log-normal, while those poorly fit by the model were peaked at the reference point. Just as the overall distribution in figure 3 masked the two interesting subpopulations in figure 5, insight into individual differences may be gained by segmenting even more finely. To avoid seeing the world only through the lens of utility models, we now cluster distributions in a model-free way. Clustering requires the specification of variables that will be used to judge similarity between histograms. Recall that the distributions consist of 100 markers placed in 41 bins (wealth levels 0–200). We clustered two ways, using either 41 variables (the number of markers in each bin) or 100 variables (the wealth level of the marker associated with each of the 100 markers). Since similar results were found with both approaches, we present the first here because we find it simpler and more intuitive. An optimal model-based clustering routine that uses the Bayesian Information Criterion to choose among models (Fraley and Raftery 2003) was applied to both retirement distributions built in the second year. The procedure was constrained to fit up to five spherical, variable-shaped, equal-sized mixture components after determining them to have the best Bayesian information criterion values of 10 candidate component shapes. The procedure suggested five clusters we refer to as 1–5, comprising 100, 98, 66, 32, and 20 distributions, respectively.

To visualize the relationships between clusters, a principal components analysis was performed on the same 41 variables. The first two components accounted for 74% of the total variance, allowing for the construction of an informative two-dimensional representation. In figure 7, the cluster label of each point is plotted at its location in principal components space. The five inset histograms show the aggregated number of markers in each bin of each distribution making up each cluster, plotted with a common vertical axis scale. The points fall into a triangular shape. At the left corner, clusters 4 and 5 comprise narrow distributions peaked at the reference point, 75. Moving right, cluster 3 has higher variance and still peaked at 75. Cluster 2 takes up more area than the others and, interestingly, peaks at 80, not 75. Finally, cluster 1, which is roughly log-normal and

TABLE 5

CLUSTER PROFILES

Cluster	Size	Age	Income	Preretirement income	% Female	% College
1	100	39.7	57,400	77,600	65.0	66
2	98	40.5	68,776	96,289	68.4	76
3	66	42.9	61,212	78,154	74.2	54
4	32	41.3	53,333	70,938	71.9	72
5	20	39.9	53,000	59,500	60.0	60
Overall	316	40.8	61,083	81,656	68.4	67

NOTE.—Mean values by cluster of age in years, present income in U.S. dollars, participant estimates of preretirement income, percentage of female participants, and percentage with a college education or beyond. The bottom row is the mean for all year 2 retirement distributions.

highest in spread, takes up a small area in the lower right corner but contains about as many distributions as cluster 2. When we inspect the loadings, the first principal component is most influenced by the 75 bin, and the second principal component mostly by the 80 and 85 bins. As a robustness check, we have repeated the clustering analysis on each of the four retirement distributions individually. The principal components plot maintains a similar triangular shape, and the inset histograms showing cluster composition bear a strong visual similarity to those presented here.

Table 5 shows mean demographics by the clusters in figure 7. Values were computed by averaging the demographic values of the person who built each distribution within the cluster. Since each participant made two distributions for this analysis, one person's demographics may appear in the means of one or two clusters. Unlike those in table 4, these data are thus not useful in significance testing. However, when this technique is applied to a large number of distributions, one per individual, it is a promising way to segment people. We repeated this analysis separately for each of the six distribution types, in which one person could belong to only one cluster. The two lowest-variance clusters in each analysis (e.g., clusters 4 and 5 in fig. 7) tended to have a lower income than the others. When collapsing these two low-variance clusters into one, it had the lowest mean income (and lowest mean estimated retirement income) in five of the six distribution types, a finding that accords with the correlation between income and risk aversion reported in table 4.

Summary of Experiments

In retirement investing, an important determinant of consumer welfare, we estimated the coefficient of relative risk aversion and the loss-aversion parameter, using the Distribution Builder, and plotted the marginal distributions of these parameters based on several hundred observations. Parameter estimates fell within ranges observed in the literature, providing a first level of validation of the tool. Two correlation metrics were calculated within sessions and between years showing significant long-term reliability, particularly for those actively saving more for retirement. Stronger validation tests used DB estimates of risk aversion to predict preferred investment outcomes both within and across years and in the presence of five additional predictors known to correlate with risk preferences. Crossing domains, a DB designed for a small-stakes gamble task provided risk-aversion estimates that significantly predicted both preferred investment outcomes and choices between gambles. Correlations with age and income provided an indirect validation of the method, while correlations with gender were for the most part not observed in these domains.

In addition to parameter estimation and validation, an interesting finding of this investigation was that the CRRA preferences, a common assumption in financial engineering and prescriptive advice, did not describe all participants' data well. Estimating a loss-aversion parameter, we found that participants who were not well fit by CRRA were more loss averse (reference dependent) than the rest. One group of investors that had CRRA preferences may be satisfied with the investment advice of maintaining a constant asset allocation between a risky and risk-free asset. The group with loss-averse preferences might be more concerned about the likelihood that their investments could go below a reference level. Interestingly, in recent years, investment and insurance firms have offered products that offer upside gain when the market goes up and absolute downside protection when the market goes down. The latter group of consumers might well be the intended audience for these products.

Two points are relevant in this connection. First, the investors in the experiments described here, as in the real world, could only obtain upside gain (results greater than 75%) by accepting some downside loss (results less than 75%). Hence they were limited to strategies that would provide the reference return over a range of market outcomes, with lower returns in very bad markets and higher returns in very good markets. The second point is that in real markets, for every investor with a strategy that essentially buys protection against falling below certain reference wealth levels, there must be one or more with a strategy that sells this insurance. There has been speculation that loss aversion addresses why stocks, which provide better returns over the long run, seem to be underpriced relative to bonds, an observation called the equity premium paradox (Benartzi and

Thaler 1995). For a discussion of equilibrium in a capital market with investors who have kinked utility curves, see the recent work of Sharpe (2006).

Segmenting the population of distributions more finely, a clustering analysis uncovered subpopulations of distributions differing on variance, location of mode, and peakedness consistent with reference-dependent preferences.

DISCUSSION

Financial Services Marketing

The presumption of efficient markets would seem to render marketing irrelevant to understanding financial markets. However, developments in behavioral finance suggest that marketing may be quite relevant. Recent research has started to emphasize differences in preferences for investments (Wilcox 2003) or trading style (Dhar and Zhu 2002). In addition, behavioral finance has started to characterize psychological differences among investors. By asking about differences in needs in investing, finance is asking the same question that motivates much of marketing research: how do consumer needs differ? We have begun to answer this question by identifying two main preference segments, one well described by standard theory, the other more reference dependent. An important next step would be to examine these differences more closely and determine the correlates and causes of the differences.

Beyond financial services, the basic DB framework could be used to study other consumer choices in which there is a risk-reward trade-off, including waiting times at health clinics (or on customer support lines), delivery times of packages, and overage charges for mobile phone plans. In a sense, this research extends Woodruff's (1972) investigations of eliciting distributions of beliefs about product attributes to eliciting preferences for distributions of prospective outcomes.

Constructing Constructive Preferences

Many tools in marketing research presume to measure existing preferences. The DB suggests a revised view: market research can help people construct preferences in a way that increases their welfare. We do not, for example, believe that consumers of investment products hold ideal probability distributions of retirement income in their heads. Income distributions are likely to be something they have never thought about before. However, each of the millions of employees who specify a fixed asset allocation for their 401(k) plan is indeed signing their name to a probability distribution of wealth, a distribution of which they may never even learn the mean or standard deviation. Since most investors must actively make decisions, consumers can benefit from a tool such as the DB to help them explore the costs and benefits of downside protection and upside gain. We suggest three aspects of the DB that might facilitate better preference construction.

Focusing on Prospective Outcomes, Not Product Offerings. In the world of behavioral research, much attention is paid to choices between gambles, and not on the overall risk that results from considering all these gambles jointly. In the world of investment advice, the same is true—not about gambles, but about choices of multiple funds. Risk-tolerance questionnaires like the one we tested ask how people would feel about a single investment product that could lose varying percentages of its worth overnight, but not about shocks to overall net worth, which could affect well-being not just overnight, but over a lifetime. The actual investment products in a portfolio should largely be irrelevant to the consumer. What matters is how investments combine to give an overall distribution of possible outcomes. With DB we have an efficient method for constructing and exploring complex outcome distributions.

Simulating Experienced Outcomes. The DB can be run in a repeated mode to cycle through the process of building distribution and drawing from a distribution many times. The current study used only two iterations through this cycle, but in other research we are beginning to explore the effect of repeated trials. Early results suggest that extensive experience systematically changes the preferences people construct about risk. In this way we think that the DB can function like a flight simulator, allowing investors to explore the outcomes of their decisions with only virtual outcomes.

More Realistic Utility Functions. The DB can fit not only standard CRRA utility functions but also exotic variants incorporating, for example, loss aversion. We do not take a position on whether or not reference dependence is a mistake in retirement settings. After all, reference wealth does matter—one must continue to make mortgage payments, for instance. However, we do point out that the DB makes explicit the trade-off resulting from buying downside protection. It is possible that, with sufficient experience with the DB, the desired level of insurance against losses might decrease, suggesting an exciting direction for future research. The unfortunate alternative, which exists currently in financial services, is to sell overpriced products that cater to loss-averse consumers, offering full downside protection at the cost of significantly reduced upside potential.

CONCLUSION

Employees making retirement investment decisions are likely to have inchoate preferences over the dizzying number of possibilities before them. Since well-defined preferences are necessary to guide the reasoned selection of investments, it is not clear how employees choose products. Experimental evidence suggests that many could be unsatisfied with the probable outcomes of their choices (Benartzi and Thaler 2002). We suggest an alternative approach in which consumers can explore candidate wealth distributions and, in so doing, construct informed preferences for outcomes instead of products. The idea that marketing could help people construct better preferences is a shift away from the view that stable preferences exist and that it is the job of marketing to uncover them.

REFERENCES

Agnew, Julie, Pierluigi Balduzzi, and Sunden Annika (2003), "Portfolio Choice and Trading in a Large 401(K) Plan," *American Economic Review*, 93, 193–215.

Ameriks, John and Stephen Zeldes (2004), "How Do Household Portfolio Shares Vary with Age?" Working paper, Economics Department, Columbia University Graduate School of Business.

Ariely, Dan, Joel Huber, and Klaus Wertenbroch (2005), "When Do Losses Loom Larger than Gains?" *Journal of Marketing Research*, 42 (2), 134–38.

Arnone, William J., Freida Kavouras, Glenn M. Pape, Martin Nissenbaum, Charles L. Ratner, Kenneth R. Rouse, David C. Voss, and Patricia A. Wiley (2002), *Ernst and Young's Retirement Planning Guide*, New York: Wiley.

Arrow, Kenneth J. (1964), "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Review of Economic Studies*, 91–96.

——— (1970), *Essays in the Theory of Risk-Bearing*, Amsterdam: North-Holland.

Bajtelsmit, Vickie L. and Alexandra Bernasek (2001), *Risk Preferences and the Investment Decisions of Older Americans*, AARP Public Policy Institute report no. 2001–11, Washington DC: American Association of Retired Persons.

Barberis, Nicholas (2000), "Investing for the Long Run When Returns Are Predictable," *Journal of Finance*, 55 (1), 225–64.

Benartzi, Shlomo and Richard H. Thaler (1995), "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110 (1), 73–92.

——— (1999), "Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments," *Management Science*, 45 (3), 364–81.

——— (2001), "Naive Diversification Strategies in Retirement Saving Plans," *American Economic Review*, 91 (1), 79–98.

——— (2002), "How Much Is Investor Autonomy Worth?" *Journal of Finance*, 57 (4), 1593–1616.

Blake, David (1996), "Efficiency, Risk Aversion and Portfolio Insurance: An Analysis of Financial Asset Portfolios Held by Investors in the United Kingdom," *Economic Journal*, 106 (438), 1175–92.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy (2002), "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence," *Journal of Political Economy*, 110 (4), 793–824.

Camerer, Colin (2005), "Three Cheers—Psychological, Theoretical, Empirical—for Loss Aversion," *Journal of Marketing Research*, 42 (2), 129–33.

Campbell, John Y. (1996), "Understanding Risk and Return," *Journal of Political Economy*, 104 (2), 298–345.

Capon, Noel, Gavan J. Fitzsimons, and Russ Alan Prince (1996), "An Individual Level Analysis of the Mutual Fund Purchase Decision," *Journal of Financial Services Research*, 10 (1), 59–82.

Debreu, Gerard (1959), *Theory of Value*, New York: Wiley.

Dhar, Ravi and Ning Zhu (2002), "Up Close and Personal: Investor Sophistication and the Disposition Effect," *Management Science*, 52 (5), 726–40.

Dybvig, Philip H. (1988), "Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market," *Review of Financial Studies*, 1 (1), 67–88.

Fraley, Chris and Adrian E. Raftery (2003), "Enhanced Software for Model-Based Clustering, Discriminant Analysis, and Density Estimation: Mclust," *Journal of Classification*, 20 (2), 263–86.

Friend, Irwin and Marshall E. Blume (1975), "The Demand for Risky Assets," *American Economic Review*, 65 (December), 900–922.

Ghiselli, Edwin E., John P. Campbell, and Sheldon Zedeck (1981), *Measurement Theory for the Behavioral Sciences*, San Francisco: Freeman.

Hakansson, Nils H. (1970), "Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions," *Econometrica*, 38 (September), 587–607.

Huberman, Gur and Wei Jiang (2006), "Offering vs. Choice by 401(K) Plan Participants: Equity Exposure and Number of Funds," *Journal of Finance*, 61 (2), 763–801.

Ibrekk, Harald and M. Granger Morgan (1987), "Graphical Communication of Uncertain Quantities to Nontechnical People," *Risk Analysis*, 7 (4), 519–29.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decisions under Risk," *Econometrica*, 47 (2), 263–91.

Köbberling, Veronika and Peter P. Wakker (2005), "An Index of Loss Aversion," *Journal of Economic Theory*, 122 (1), 119–31.

Low, Bobbi S. (2000), *Why Sex Matters: A Darwinian Look at Human Behavior*, Princeton, NJ: Princeton University Press.

Madrian, Brigitte C. and Dennis F. Shea (2001), "The Power of Suggestion: Inertia in 401(K) Participation and Savings Behavior," *Quarterly Journal of Economics*, 116 (4), 1149–87.

Mehra, Rajnish and Edward C. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15 (2), 145–61.

Merton, Robert C. (1969), "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51, 247–57.

——— (1971), "Optimum Consumption and Portfolio Rules in a Continuous Time Model," *Journal of Economic Theory*, 3, 373–413.

Novemsky, Nathan and Daniel Kahneman (2005), "Boundaries of Loss Aversion," *Journal of Marketing Research*, 42 (2), 119–28.

Perold, Andre F. and William F. Sharpe (1988), "Dynamic Strategies for Asset Allocation," *Financial Analysts Journal*, 44 (1), 16.

Pratt, John W. (1964), "Risk Aversion in the Small and the Large," *Econometrica*, 32 (1/2), 122–36.

Safra, Zvi and Uzi Segal (1998), "Constant Risk Aversion," *Journal of Economic Theory*, 83 (1), 19–42.

Samuelson, Paul A. (1969), "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51 (3), 239–46.

Samuelson, William and Richard Zeckhauser (1988), "Status Quo Bias in Decision-Making," *Journal of Risk and Uncertainty*, 1 (1), 7–59.

Sharpe, William F. (2006), *Asset Prices and Portfolio Choice*, Princeton, NJ: Princeton University Press.

Sharpe, William F., Daniel G. Goldstein, and Phil W. Blythe (2000), "The Distribution Builder: A Tool for Inferring In-

vestor Preferences," Working paper at http://www.stanford.edu/˜wfsharpe/art/qpaper/qpaper.html.
Siebenmorgen, Niklas, Elke U. Weber, and Martin Weber (2000), "Communicating Asset Risk: How the Format of Historic Volatility Information Affects Risk Perception and Investment Decisions," Sonderforschungsbereich 504 Publication no. 00-38, University of Mannheim, Mannheim.
Thompson, Kimberly M. and Diane L. Bloom (2000), "Communication of Risk Assessment Information to Risk Managers," *Journal of Risk Research*, 3 (4), 333–52.
Tversky, Amos and Daniel Kahneman (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (4), 297–323.
Weber, Elke U., Ann-Renee Blais, and Nancy E. Betz (2002), "A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors," *Journal of Behavioral Decision Making*, 15 (4), 263–90.
Wilcox, Ronald T. (2003), "Bargain Hunting or Star Gazing: Investors' Preferences for Stock Mutual Funds," *Journal of Business*, 76 (4), 645–64.
Woodruff, Robert B. (1972), "Measurement of Consumers' Prior Brand Information," *Journal of Marketing Research*, 9 (August), 258–63.
Yook, Ken C. and Robert Everett (2003), "Assessing Risk Tolerance: Questioning the Questionnaire Method," *Journal of Financial Planning*, 18 (8), 48–55.