The Journal of Behavioral Finance & Economics Volume 3, Issue 1, Spring 2013
**1-20** 
ISSN: 1551-9570

# Misperceptions of Long-Term Investment Performance: Insights from an Experiment [1]

Michael Stutzer
*University of Colorado*

Susan Jung Grant
*Boston University*

## Abstract

Previous experimental studies have posed repeated fixed dollar gambles as a simplified model for studying long-term investment behavior. But long-term investments are more realistically modeled using fixed percentage returns rather than dollar returns. So in our experiments we made this substitution. Our experimental findings show that many subjects – including ones who have authored papers in the earlier experimental literature – would make long-term investments in a volatile asset with high expected return, but regret doing so once they are shown the actual distribution of long-term returns. We subsequently posed a well-diversified portfolio of two such investments to the same (now better informed) subjects. Many rejected the portfolio, but still changed their minds after being shown its long-term returns.

## 1. Introduction

Paul Samuelson (1963) used a simple binary fixed dollar gamble (i.e. one either wins a fixed amount or loses a fixed amount) to illustrate the propensity of individuals to accept the cumulative outcome of repeated gambles with positive expected value per repetition. The Samuelson gamble has been the focus of much discussion by noted academics (e.g. see Tversky and Bar-Hillel (1983), Chew and Epstein (1988), and Ross (1997)), so it is not surprising that Benartzi and Thaler (1999) tested the propensity of subjects to accept the Samuelson repeated gamble, hypothesizing that the most decision-relevant features of the probability distribution of cumulative wealth might be misperceived. They conducted experiments that supported this, i.e. many of their subjects opposed to the repeated gamble became favorably disposed to doing so after being shown the probability distribution of cumulative wealth resulting from it. Klos, et al. (2005) inferred from their own related experimental findings that “Computing, showing, and discussing aggregated distributions may have the potential to avoid utility losses in asset allocation decisions or other decisions involving repeated gambles.” This view reflects the seemingly sensible notion that the cumulative, long-term outcome of asset allocation and/or other investment decisions are affected by uncertain shocks to future returns that are analogous to the successive coin-tosses in a repeated fixed dollar gamble. Researchers thus hope that behavioral findings from experiments that use repeated gambles can serve as a straightforward and replicable experimental model for understanding the propensity of investors to make long-term investments.

But there is an important *qualitative* difference between a repeated fixed dollar gamble and a realistic investment, even when the investment’s returns are modeled to be binary and IID across periods, like the aforementioned fixed dollar gambles. In repeated gambles, one stands to make or lose fixed *amounts* of money in each repetition, so the outcome of repeated gambles is the *sum* of the outcomes of the separate gambles. But to more realistically model investments, financial analysts follow Samuelson (1965) in assuming that one stands to make or lose *percentages* of money, so the outcome of investment over time (i.e. the cumulative return) is the *product* of the sequence of future gross rates of return.

There is good reason to believe that this distinction will result in important behavioral differences. Chen and Rao (2007) thoroughly document, in both an experimental and a market setting, that many people make simple arithmetic errors in situations where one percentage change follows another. For example, they documented that many subjects overestimate the cumulative return arising from a $40\%$ increase in the value of a mutual fund over a six month period, followed by a $25\%$ decrease over the next six month period. The actual cumulative gross return over the year is $1.40\times 0.75=1.05$, i.e. a $5\%$ net cumulative return. But Chen and Rao (op.cit.) reported that “many participants made the computational error of adding up multiple percentages”, which in this case resulted in the erroneous belief that the net return would be $40\% - 25\% = 15\%$, rather than the actual 5% net return. They documented that errors of this sort also occurred when subjects evaluated other patterns of multi-period percentage changes. For example, when told that one mutual fund had a 62.5% return over a year, and that another mutual fund had a 30% return for the first six months of that year followed by a 25% return over the next six months of the year, moresubjects preferred the former fund – despite the fact that both had the same cumulative return over the year ($1.3\times 1.25=1.625$).[2]

In the following section 2, we review the investment-relevant, qualitative differences between the repeated fixed dollar gambles used in the prior experimental literature and their closest investment analog, i.e. the standard textbook, fixed percentage “binomial” price tree that exhibits an investment’s possible cumulative returns. We dub this a *binary investment* for ease of exposition. [3] Specifically, the distribution of cumulative wealth resulting from a long-term investment (binary or otherwise) is highly positively skewed, and hence its median is both lowerand a more accurate measure of central tendency than the mean (i.e. expected) cumulative return. This realistic feature of cumulative investment returns does not appear in the cumulative returns from binary fixed dollar gambles used in the aforementioned experimental literature, and hence is a focus of our experimental design. In section 3, we substitute a volatile binary investment for the volatile binary fixed dollar gamble used in the aforementioned experiments, to determine whether or not misperception of this phenomenon would significantly change subjects’ long-run decisions. Because the phenomenon is motivated by asset and portfolio investment, we followed Benartzi and Thaler (op.cit.) by adding college business students (specifically, students in investment classes) to the survey pool. Moreover, we went beyond Benartzi and Thaler (op.cit.) in surveying a uniquely qualified group of experts: academics who have co-authored articles referenced in the aforementioned experimental literature that utilized binary fixed dollar gambles.

As summarized in the conclusions section 4, the experimental results indicate that many experimental subjects who were at first willing to make a volatile repeated binary investment with a high expected return, changed their minds after being shown the cumulative wealth distribution (with its lower median) – including many of the aforementioned academics. That is, the subjects would have overinvested in it had they not been shown the distribution of cumulative future wealth associated with their decision. Moreover, simple diversification among such volatile investments will not change the expected return, but will significantly improve the (more typically realized) median cumulative return. Even after viewing the cumulative future wealth distribution resulting from investment in a single volatile asset, many of the same experimental subjects who at-first rejected investment in the diversified portfolio, again changed their minds (i.e. wanted to invest in the diversified portfolio) *after* being allowed to view the distribution of cumulative future wealth associated with investment in the portfolio. Because real-world investors typically also do not know the distribution of cumulative future wealth associated with prospective investments [4], the experimental results suggest that some may overinvest in highly volatile assets that (as is often the case) also have high expected returns, and/or some may underinvest in diversified portfolios of highly volatile assets with high expected returns.

## 2. Binary Gambles vs. Binary Investments

Let $W_T$ denote the decision maker’s wealth after $T$ repetitions of a gamble, and let $X_t$ denote the payoff from Samuelson’s (1963) binary gamble’s repetition $t$, i.e. either $X_t = +200$ with probability $\frac{1}{2}$ or $X_t = -100$ with probability $\frac{1}{2}$. That is, the gambler either makes 200 dollars or loses 100 dollars with each repetition, depending on a fair coin toss. The expected dollar return per bet is 50 dollars with a standard deviation of 150 – *the expected return per bet is high, and this is accompanied by a high volatility per bet.* After $T$ repetitions:

$$W_T = \sum_{t=1}^{T} X_t = (\frac{1}{T}\sum_{t=1}^{T} X_t)T \quad (1)$$

The mathematical expectation $E[\ ]$ of wealth in (1) is

$$E[W_T] = \sum_{t=1}^{T} E[X_t] = [\frac{1}{2}200 + \frac{1}{2}(-100)]T = 50T \quad (2)$$

Because the cumulative wealth (1) depends on ($T$ multiplied by) an *average* whose distribution approaches normality (due to the Central Limit Theorem), the cumulative wealth (1) becomes approximately normally distributed and hence approximately symmetric as $T$ increases. As such, the expected value (2) of the cumulative wealth is close to its median, denoted $Median[W_T]$. So the expected value (2) is a useful measure of the central tendency of outcomes from repeated binary fixed dollar gambles. For example, with $T = 100$ repeated gambles, as posed by Samuelson (1963) and used in the Benartzi and Thaler (op. cit.) experiment, the expected value is $E[W_{100}] = 5000$ and the standard deviation is $\sqrt{Var[W_{100}]} = 1500$. Because the probability of a loss is more than three standard deviations from the mean, Samuelson (1963, p.109) concluded that "By the usual binomial calculation and normal approximation, this probability of making a gain is found to be very large, $P_{100} = .99+$". Moreover, Rabin and Thaler (2001, p.223) note that $W_{100}$:

...has an expected return of $\$5000$, with only a $\frac{1}{2,300}$ chance of losing any money and merely a $\frac{1}{62,000}$ chance of losing more than $\$1000$. A good lawyer could have you declared legally insane for turning down this gamble.

Reasoning that many experimental subjects will be unable to intuit the nature of cumulative outcome probabilities like that, or even to quantitatively calculate them if given time and resources, Benartzi and Thaler (op. cit., p. 370) formulated and experimentally tested their "Hypothesis 2" ,i.e.

**Hypothesis H2**: Subjects will find repetition of a binary fixed dollar gamble with positive expected return more acceptable after they are presented with the distribution of final outcomes.

Benartzi and Thaler (op. cit.) found experimental evidence that generally supported this hypothesis, by posing Samuelson's gamble and related gambles to a variety of students and coffee shop customers. Specifically, some subjects who initially refused to make $T = 100$ repeated bets decided to do so *after* being shown the cumulative wealth distribution from $T = 100$ repetitions[5]. However, in a different experiment intended to model individuals' asset allocations in 401K plans, Sundali and Guerrero (2009) found that subjects only slightly increased their allocations to stocks with high expected returns after being shown the expected long-run *cumulative* return. They conjectured that this was due to subjects' subjective considerations about the risks (higher volatility) being takenover the long-run.

This suggests the possibility that repeated fixed-dollar binary gambles posed in Samuelson (1963, op.cit.) and used in the aforementioned literature may not be the best simplified experimental set-up for studying long-term investment. Surely Samuelson (who neither devised nor conducted the aforementioned survey experiments) would have agreed with this, for he argued (Samuelson (1965, op.cit.)) that random *percentage* gains and losses made more sense when modeling investments. The simplest implementation of this is the canonical investment textbook model, called the *binomial tree* model. In that model, an asset's price either goes up by some percentage with some probability, or goes down by some different percentage with the complementary probability. The gain/loss percentages and their respective probabilities are repeated each period that the investment is held. Hence, this model is analogous to a repeated binary fixed dollar gamble, after replacing its fixed dollar amounts with fixed percentage amounts. To differentiate this from the literature's ubiquitous binary fixed dollar gambles, let us call this a *binary investment*.

Analogous to Samuelson's binary gamble with its high positive expected return and volatility per repetition, consider a binary investment with high expected return and volatility per repetition: the investment either increases $80\%$ in value with probability $\frac{1}{2}$, or decreases in value by $50\%$ with probability $\frac{1}{2}$, and is repeated in each subsequent period that the investment is held. *The two possible outcomes are intentionally starkly different, as are the two outcomes in Samuelson's aforementioned and oft-analyzed binary fixed dollar gamble (i.e. +200 and -100), to facilitate comparison of the binary gambling experiments with our binary investment experiment.* Starting with $100 placed in the binary investment, the expected wealth after the first repetition (i.e. after $T=1$) is $\frac{1}{2}\$180+\frac{1}{2}\$50=\$115$, with high standard deviation of $\$65$. While both its expectation and standard deviation are high, the corresponding numbers for the oft-studied Samuelson binary gamble are higher still – the expected wealth after its first repetition is $\$150$ ($\frac{1}{2}\$300+\frac{1}{2}\$0=\$150$) with standard deviation of $\$150$.[6]

Like the binary gamble, the binary investment has a highly positive expected return per repetition, so one might expect to observe experimental results in accord with the aforementioned Hypothesis 2 of Benartzi and Thaler (op.cit.), i.e. subjects will find $T=100$ repetitions of it more acceptable after they are shown the distribution of cumulative future wealth resulting from that. However, as shown in Section 3, our experimental findings are the opposite: many subjects who initially wanted to make $T=100$ repetitions decided not to do so *after* being shown the distribution of cumulative wealth resulting from $T=100$ repetitions.

To help understand why, we turn to a mathematical derivation of the probability distribution of cumulative wealth after $T=100$ repetitions. The gross return $X_t$ of each dollar invested is either 1.80 with probability $\frac{1}{2}$ or is 0.50 with probability $\frac{1}{2}$. Starting from an initial investment denoted $W_0$, (e.g. $W_0=\$100$) the invested wealth after $T$ periods is:

$$W_T = W_0 \prod_{t=1}^{T} X_t . \quad (3)$$

which attains the value

$$W_T = W_0 1.8^u 0.5^{T-u}$$

with binomial probability

$$\frac{T!}{u!(T-u)!}\left(\frac{1}{2}\right)^{u}\left(\frac{1}{2}\right)^{T-u}$$

where $u$ denotes the number of up moves (i.e. +80% increases) in the sequence of $T$ repetitions. Figure 1 illustrates the frequency distribution of cumulative wealth at $T=100$, after starting with $W_0=\$100$. Note that the median outcome is to have only around $50$ *cents* left of the initial \$100, despite the 15% expected return per repetition! How can this be? Now the plot thickens. The mathematical expectation $E[\ ]$ of cumulative wealth from the IID process (3) is the product of the individual period expectations, i.e.

$$E[W_T]=W_0[1.15]^T \tag{4}$$

Hence (4) shows that the *expected* cumulative wealth $E[W_T]$ compounds geometrically to infinity as $T$ increases, qualitatively similar to, but eventually outstripping, the linear increase (2) of the fixed dollar gamble's expected cumulative value. For the case $T=100$ and $W_0=\$100$ used both in the aforementioned binary gamble experiments and in our survey instrument, $E[W_T]=100*1.15^{100}=\$117,431,345$. And that would sufficiently characterize the outcome of a $\$100$ bank account growing at a constant compound interest rate of 15%. But unlike the cumulative outcome (1) of repeated fixed dollar gambles, the cumulative outcome (3) of repeated binary investment does not depend on an average of the IID random variables, so the Central Limit Theorem cannot be used to infer that it has an approximate normal (and hence symmetric) distribution with similar expected and median cumulative values of wealth. In (3), the limited liability of typical (e.g. stock or bond) investments precludes $X_t<0$, so $\prod_t X_t \geq 0$, and always $W_T \geq 0$. That is, just like a real investment in stocks or bonds, you can't lose more than your initial investment.[7] As shown in Figure 1, the distribution of $W_T$ in (3) becomes extremely positively skewed (i.e. long right-hand tail and short left-hand tail) for large $T$, so the expected cumulative value $E[W_T]$ in (4) is much higher than the *median* cumulative value of wealth. When investment returns per period are IID (as our example is), the short mathematical appendix describes a good, large $T$ approximation for the median value. Applying that approximation to our example yields:

$$Median[W_T]\approx W_0 e^{E[\log X_1]T}=W_0 e^{[\frac{1}{2}\log 1.8+\frac{1}{2}\log 0.5]T}=W_0 e^{-.0527T} \tag{5}$$

Again starting from an initial investment of just $W_0=\$100$, (5) shows that the median cumulative value of wealth will be only about $\$100e^{-5.27}\approx\$0.50$, consistent with the frequencies depicted in Figure 1, despite the expected cumulative value of wealth of $\$117,431,345$ computed from (4)! This type of disparity is greatest with highly volatile investments, and does not occur at all in the aforementioned binary fixed dollar gambles, regardless of their volatility. *Hence this is the focus of our experimental design: do subjects anticipate some decision-relevant, qualitative difference between expected and median values?*

The final area investigated is the important, possibly counterintuitive effect that investment diversification has on the cumulative return distribution through its median. The most common long-term investment advice is to allocate savings across assets that are not too highly correlated, periodically *rebalancing* the portfolio back to its original asset weights. Indeed, a professor invested in TIAA-CREF can elect to have her/his pension portfolio automatically rebalanced to weights of her/his choosing.[8] Hence we added diversification questions to our survey, which were the simplest and most transparent extensions of our prior questions concerning a single binary investment. We wished to see whether or not subjects would choose to put half of their $100 initial stake in one binary investment, and the other half in an *identical* second binary investment whose return is independent of the first asset's return, with the provision that a computer will automatically continue to split the invested wealth between the two before each subsequent repetition. Thus, this is an experimental model of an equally value-weighted rebalanced portfolio of two independently and identically distributed assets. We asked the subjects whether or not they would accept this, both before and after showing them the diversified *portfolio's* distribution of cumulative future wealth.

One might intuit that the portfolio will have future results similar to the single asset – aren't two half-losers equivalent to one whole loser? Calculation of the *expected value* of the diversified portfolio seems to support that heuristic reasoning, because the equally-weighted rebalanced portfolio's expected return is $\frac{1}{2}15\% + \frac{1}{2}15\% = 15\%$, i.e. the same expected return as each of the separate investments that comprise it. As a result, the portfolio will have the same enormous expected cumulative wealth as the single investment (i.e. in excess of $100 million), and our subjects might intuit that the portfolio's future wealth distribution will also look similar to the single asset's future wealth distribution. But that intuition is wrong. To see this, do the following simple math using the same approximation:

$$Median[W_T] \approx W_0 e^{E[\log(\frac{1}{2}X_1 + \frac{1}{2}X_2)]T} = W_0 e^{(\frac{1}{4}\log 1.8 + \frac{1}{2}\log 1.15 + \frac{1}{4}\log 0.5)T} = W_0 e^{+.0435T} \quad (6)$$

In contrast to the typical almost-complete ruin stemming from the single binary investment, (6) shows that an equally-weighted rebalanced portfolio of two of them will typically *make* money as $T$ increases. Starting from an initial investment of $100, (6) shows that the diversified portfolio's median cumulative wealth after $T = 100$ repetitions will be around $7750, in contrast to the $0.50 median cumulative wealth from investing everything in (either) single investment used to form the portfolio. The frequencies depicted in Figure 2 are consistent with this.[9] Because of diversification, the intuition that 'half of two identical losers is the same as one loser' is extremely misleading when projecting the typical (e.g. the median) cumulative wealth, because the median is determined by the expected *log gross* return per repetition (i.e. -5.27% for the single investment and +4.35% for the equally-weighted rebalanced portfolio of two independent ones).

In summary, intuition about the central tendency of future wealth resulting from both the single investment, and the well-diversified portfolio of two independent investments, will be misled by focus on expected returns, which will not be changed by diversification.

We hypothesize that subjects will be heavily influenced by knowledge of the central tendency of the cumulative wealth distribution, which in binary gambles is determined by the expected value and in binary (and real-world) investments is determined by a significantly different median value. Because the investment's median depends on the expected log gross return per repetition, we formulate the following hypothesis as a more investment-relevant alternative hypothesis to Benartzi and Thaler's Hypothesis 2 (stated earlier in this paper):

**Hypothesis H2'**: Subjects will find repetition of a binary investment with positive expected *log gross* return more acceptable *after* they are presented with the distribution of final outcomes. The opposite will occur for repetition of a binary investment with negative expected log gross return, even when the expected net return is positive.

## 2.1 What Can Be Learned From Our Experiment?

We emphasize here that neither us, nor the authors of the previously published articles (cited earlier) making use of Samuelson's (1963) repeated binary gamble, were attempting to discover whether or not experimental subjects can compute cumulative wealth distributions in the course of filling out their respective survey questionnaires. Few (if any) subjects faced with accepting or rejecting $T = 100$ repeated Samuelson gambles could compute the expected value and standard deviation of the approximately normal distribution of cumulative wealth resulting from that, as we calculated and reported early in Section 2. Nor could the subjects in earlier experiments compute expected utilities or other decision-theoretic techniques for this and related repeated gambles that were posited and analyzed by the eminent decision theorists Tversky and Bar-Hillel (op.cit.), Chew and Epstein (op.cit.), and Ross (op.cit.). Likewise, we did not believe that our subjects could make the analogously difficult computations required to calculate the cumulative wealth distribution resulting from our repeated binary investment. Fortunately for the earlier published authors and for us, orthodox economic theory does *not* posit that rational decision makers actually perform such calculations – even when given adequate time and computational resources – but rather that their decisions are at least qualitatively the same *as-if* they had made such calculations.

Hence the purpose of the experimental literature is to uncover aspects of decision problems that impede investors' abilities to make such "as-if" decisions. Our experimental results indicate that many subjects – even those with some college-level statistical skills – do not intuit situations where median cumulative values behave *qualitatively* differently than expected cumulative values, even when the subjects receive more quantitative information about the return generating process (the specific binary investment) than real-world investors typically have.[10] Subjects' failures to anticipate the positively skewed nature of the distribution of investments' cumulative future values impedes the subjects' ability to behave in accord with the aforementioned "as-if" hypothesis, as witnessed by changes in their decisions upon being shown the actual distributions.

## 3. The Survey

Because our main contact with prior literature is the robustness of findings in the repeated additive gamble experiment conducted by Benartzi and Thaler (op.cit.) that made use of Samuelson's (1963) repeated additive gamble, we substituted our previous section's binary investment with its analogously high expected return and volatility per repetition. Question 1 of the survey instrument (see the appendix) models the decision to invest $100 in a binary investment. This serves to prime the subjects for the important Question 2 that models the decision to buy-and-hold that investment over $T = 100$ repetitions. The use of both a $W_0 = \$100$ investment and $T = 100$ repetitions matches the initial bet and number of repetitions in the Samuelson (1963) repeated binary gamble and in the Benarzi and Thaler (op.cit.) experiment, *so were adopted here in order to isolate effects that arise from substituting the binary investment for the repeated binary gamble.* To help ensure that the subjects' answers reflect their respective internal decision processes, rather than the influence of normative opinions from finance professors, investment advisors and financial media figures, the phrases "investment" and "buy and hold" were not used in the wording of the survey questions. Doing otherwise would have biased our comparisons to the Benartzi and Thaler (op.cit.) results. Question 3 again inquires about accepting $T = 100$ repetitions of the binary investment, but is posed *after* the subjects are shown the probability distribution of the cumulative wealth after 100 repetitions (Table 1 in the appendix). Again, this parallels the development in the Benartzi and Thaler (op.cit.) experiment. Question 4 models the decision to invest in the equally-weighted, diversified portfolio of two of these assets, when they are independently distributed. Again, we did not use normative investment buzzwords, e.g. "diversified", "portfolio" or "rebalanced", in order to avoid potential bias from prior normative investment advice that subjects may have heard or read. The question is somewhat longer than the others, because of the need to properly describe the de-facto portfolio return generating and de-facto rebalancing processes. Question 5 inquires similarly, but is posed *after* the subjects are shown the probability distribution of the cumulative portfolio wealth after 100 repetitions (Table 2 in the appendix). Taken together, Questions 4 and 5 naturally extend the approach of the first three questions in order to investigate perceptions of diversification effects, in a way that Benartzi and Thaler (op.cit.) did not.

Again following Benartzi and Thaler (op.cit.), our subjects included both coffee shop customers and students (ours found near our university instead of theirs). The students were nearing the end of an undergraduate investment course. Hence not only had they already passed a college-level statistics course, but they also had been exposed to typical undergraduate investment coursework which surely includes use of percentage returns, expected returns, etc. *But unlike Benartzi and Thaler (op.cit.), we additionally surveyed a group of academic authors of very closely related literature.* The academic authors were drawn from (i) authors referenced in Benartzi and Thaler (op.cit.) and (ii) authors who subsequently cited Benartzi and Thaler (op.cit.), and (iii) the same two groups of authors generated from the Klos, et.al. (op.cit.) paper. We tasked a research assistant to seek the email address of at least one co-author from each reference and citation source just described. 22 of the 72 authors identified did respond to our emailed request to complete our online survey (a 30% response rate). *A priori*, one might think that subjects in this academic subsample would be unlikely to change their responses upon observing the cumulative wealth distributions, because (1) in the course of acquiring a Ph.D and publishing experimental papers, they have demonstrated knowledge of statistical theory and its applications, and (2) their research in the area would have helped them anticipate the *qualitative* differences between binary gambles and binary investments. Moreover, the academic subjects participated in an on-line survey, and hence could have used pencil-and-paper analytics or spreadsheet simulations to gain insight before answering the questions, had they felt it necessary.

Finally, note that all subjects were asked about their willingness to undertake the hypothetical investments, without having to actually do so, and without being rewarded proportionally to an actual, simulation-based outcome. Paying subjects proportional to the outcomes of their decisions may be appealing, but it would complicate comparison of our results to Benartzi and Thaler (op.cit.), who did *not* pay subjects in proportion to outcomes. Moreover, in his survey of the literature, Camerer (1995,p. 599) notes that *"Psychologists do not always motivate subjects financially – though many have and a few are adamant about doing so – because incentives usually complicate instructions and psychologists presume subjects are cooperative and intrinsically motivated to perform well."*

3.1 Experimental Results

We start by examining the robustness of Benartzi and Thaler's Hypothesis H2 (stated in Section 2 herein). Did our subjects find $T = 100$ repetitions of the single binary investment (with its expected cumulative value in excess of $100,000,000) more attractive *after* they were presented with the probability distribution of outcomes, i.e. after they were shown the survey instrument's Table 1 (see the appendix)? Comparing responses to our Questions 2 and 3 shows that they did *not* find it more attractive after seeing the distribution of outcomes, despite the positive expected return (15% per repetition) emphasized by Benartzi and Thaler. Pooled results for our three subsamples are shown in Table 1.

INSERT TABLE 1 HERE

Note that while 34 of those sampled rejected the repeated binary investment before viewing the distribution of cumulative outcomes, 62 subjects rejected *after* viewing the distribution of cumulative outcomes, despite its positive expected net return of 15% per repetition. Also note that while 49 of those sampled accepted it before viewing the distribution, only 21 accepted after viewing it. Hence they were not more likely to accept the binary investment after viewing the distribution. Because the investment has a negative expected *log gross* return, and hence a long-run median cumulative outcome near zero, the results support our alternative Hypothesis H2'.[11]

The separate subsamples all exhibited the pattern of *less* frequent acceptance of $T = 100$ repetitions after seeing the distribution of outcomes. 26% of coffee shop customers were willing to accept before viewing the distribution of cumulative outcomes, while only 16% were willing to do so after viewing. 83% of the investment students were willing to accept before viewing, while only 23% were willing to accept after viewing. Our academic panel was no different; 73% were willing to accept before viewing, while only 41% were willing to do so after viewing (statistically significant at the 5% level). Hence any prior knowledge possessed by either the investment students or the academics who authored related papers did not lead them to anticipate the decision-relevant information contained in the distribution of cumulative outcomes. Table 2 shows that the decision-relevant information contained in the distribution of cumulative outcomes led most subjects in each subsample to reject the opportunity of $T = 100$ repetitions after viewing distribution of cumulative outcomes. Most subjects acted *as if* the near-zero median value of cumulative wealth after $T = 100$ repetitions was more relevant than the over $\$100$ million expected cumulative value –once they had viewed the resulting probability distribution of cumulative outcomes. There was little tendency to risk even a hypothetical $100, but those who *would* do that *after* viewing the distribution are not considered irrational by decision theorists.[12] The results for the separate subsamples are summarized in Table 2.

INSERT TABLE 2 HERE

We now turn to the remaining survey questions, which are concerned with the effects of diversification. Table 3 displays the data.

INSERT TABLE 3 HERE

Table 3 shows that $46/83 = 55\%$ of the pooled sample accepted diversification before seeing the distribution of cumulative wealth resulting from it, rising to $68/83 = 82\%$ after the subjects viewed the distribution. Note that the change in their willingness to accept the portfolio occurred despite having *already* seen the distribution of cumulative wealth resulting from the single investments used to construct it.[13] This pattern is also consistent with our Hypothesis H2' (stated in section 2), because the expected log gross return per repetition of the diversified portfolio is positive. But there was no significant difference in the before and after responses of the academicians, with $77\%$ accepting diversification before, rising just abit to $86\%$ after.[14]

While 19 of our 31 coffee shop customers rejected the portfolio before having access to its outcome distribution, 11 of those 19 (i.e. $58\%$ of them) changed their minds after viewing the cumulative wealth distribution.[15] While only 13 of our 30 undergraduate investment students rejected the portfolio before seeing the its distribution, 11 of those 13 (i.e. $85\%$ of them) changed their minds when they saw it. And while only 5 of our 22 academic respondents rejected the portfolio before having access to its outcome distribution, 3 of those 5 went on to change their minds, while the other 2 answered "No" to all five of the survey's questions, i.e. they did not want to invest under any circumstances.

### 3.2 An Alternative Experiment

To investigate the influence of volatility on the findings, we substituted a much less volatile asset with a much higher return/risk ratio, specifically, one that either increases by $40\%$ or decreases by $20\%$ with equal probability. The expected return per repetition is then $10\%$ with a standard deviation of $30\%$. This has a return/risk ratio of .33, which is $43\%$ higher than the return/risk ratio of our example with expected return and standard deviation of $15\%$ and $65\%$, respectively. As a consequence, the expected log gross return per repetition is high, i.e. $5.67\%$, resulting in median cumulative wealth after $T = 100$ repetitions of $e^{5.67} = \$28,900$ after starting with just $\$100$ (see Figure 3).[16] Unlike our earlier experiment, the extremely high expected cumulative return is accompanied by a prodigious median cumulative return. So we did not expect to see the statistically significant experimental results reported above. This was confirmed when we substituted this example into our survey and administered that to a different section of the same investment course used earlier. Only 11 out of 43 students rejected the investment before seeing the distribution of final outcomes, and while 7 of those 11 did accept the investment after seeing its (highly favorable!) distribution of cumulative wealth, this was not statistically significant. Not surprisingly, 31 of the 32 students who favored the investment before seeing the cumulative wealth distribution still favored it after seeing the distribution. Of course the rebalanced portfolio formed from two of these investments is even more desirable (see Figure 4). Hence only 3 of the 43 students rejected the diversified portfolio after seeing its results.

## 4. Conclusions

The beliefs and behavior of long-term investors is a subject of importance to both financial advisors and retirement fund managers. Research experiments that pose repeated fixed dollar gambles to subjects are not the best design for experimental modeling of long term investment behavior. Financial economists believe that a better simplified model for investment returns utilizes repeated *percentage* returns that cumulate multiplicatively rather than additively. This is standard textbook pedagogy, andis of more than academic importance because the cumulative wealth distributions from such textbook (as well as real-world) investments can be radically different than those arising from fixed dollar gambles. Specifically, the typically realized (e.g. median) cumulative wealth from a highly volatile investment will be much lower than the expected cumulative wealth that is only atypically realized. So we substituted a highly volatile, textbook-style investment with a high positive expected return for the highly volatile, high expected return fixed dollar gamble used in others' otherwise analogous experimental designs and theoretical investigations, and discovered two new findings of practical significance. First, subjects prone to making a highly volatile investment with a high expected return changed their minds after being shown the long-term distribution of cumulative wealth resulting from it. Second, after seeing that investment's long-term cumulative distribution, subjects often *at-first* rejected investment in a diversified portfolio of two such identical investments, but again changed their minds after being shown the future wealth distribution resulting from the portfolio. This happened despite the fact that the expected cumulative wealth from the single investment is the same as this portfolio's expected cumulative wealth. Hence many subjects acted as-if they did not intuit the stark qualitative difference between the respective long-term median and expected cumulative returns in a diversified portfolio, even *after* having witnessed effects of this disparity in the future wealth arising from investment in a single asset. Taken together, the effects can be reduced to a catchy phrase: *due to volatility, the expected cumulative investment return is not the cumulative return to be expected*!

Stark *qualitative* differences between expected and median cumulative returns do not arise in previously published papers positing experiments (both actual and theoretical thought experiments) based on repeated binary fixed dollar gambles (cited herein), nor do they arise in our repeated binary fixed percentage investments when the ratio of expected return to standard deviation (i.e. the return/risk ratio) is high. Hence in an alternative experiment, we did not find these statistically significant experimental results when we substituted an asset with a much higher return/risk ratio.

Because highly volatile assets often do have high expected returns, our first experimental finding suggests that some may overinvest in highly volatile assets with high expected returns but relatively low return/risk ratios. The second experimental finding suggests that some investors may underinvest in well-diversified portfolios of those same, highly volatile assets.

Finally, we note a limitation of our approach. It is possible that changes to our tabular displays of the repeated binary investment outcomes (i.e. the survey's Tables 1 and 2) could affect the responses of some surveyed. Table 1 was constructed so as to not obscure the very low median in the repeated, single asset investment, while Table 2 was constructed to use the same binning of possible outcomes as Table 1, so as not to introduce additional complexity to the visual display of results. It is possible that other representations of the outcome distribution could lead to different survey results, but it is unclear what the most neutral representation would be. In any event, the best visual representation of quantitative information is very important to effective disclosure of past mutual fund results, so studies focused on finding the most accurate yet understandable representation are good prospects for future research.

## Mathematical Appendix

This appendix contains the math to establish that the expected log gross return is the key determinant of the median long-run cumulative return, as asserted before equation (5). Take the logarithm of both sides of (3) and then re-exponentiate both sides to find:

$$W_T = W_0 e^{\sum_{t=1}^{T} \log X_t} \tag{7}$$

Because the exponential function is a monotone (increasing) function of $\sum_{t=1}^{T} \log X_t$ ,

$$Median\,[W_T] = W_0 e^{Median\left[\sum_{t=1}^{T} \log X_t\right]} \tag{8}$$

When $T$ is suitably large, Ethier (2004, p.1234) uses the following approximation:

$$Median\,[\sum_{t=1}^{T} \log X_t] \overset{IID}{\approx} E[\log X_1]T - \frac{E\left[(\log X_1 - E[\log X_1])^3\right]}{6Var\,[\log X_1]} \tag{9}$$

In practice, the second term in (9) is quite small compared to the first term. So substituting the first term of (9) into (8) yields :

$$Median\,[W_T] \approx W_0 e^{E[\log X_1]T} \tag{10}$$

which is the approximation used in formulae (5) and (6). Simulations used to compute the survey instrument's Tables 1 and 2 confirm the accuracy of the approximation for this purpose.

## REFERENCES

Armitage, P. and G. Berry, 1994, *Statistical Methods in Medical Research (3rd Ed.*, Oxford, Blackwell.

Benartzi, S. and R. Thaler, 1999, "Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments" *Management Science* 45(3),364-381.

Camerer, C., 1995, "Individual Decision Making". In J.H. Kagel and A.E. Roth, Ed., *Handbook of Experimental Economics*, Princeton, University of Princeton Press, 587-683.

Chen, H. and A.R. Rao, 2007, "When Two Plus Two is Not Equal to Four: Errors in Processing Multiple Percentage Changes" *Journal of Consumer Research* 34, 327-340.

Chew, S.H. and L.G. Epstein, 1988, "The Law of Large Numbers and the Attractiveness of Compound Gambles" *Journal of Risk and Uncertainty* 1, 124-132.

Ethier, S.N., 2004, "The Kelly System Maximizes Median Fortune" *Journal of Applied Probability* 41, 1230-1236.

Friedman, M., 1970, "Methodology of Positive Economics". In M. Friedman, Ed., *Essays in Positive Economics*, Chicago, Univ. of Chicago Press.

Klos, A., E.U. Weber, and M. Weber, 2005, "Investment Decisions and Time Horizon: Risk Perception and Risk Behavior in Repeated Gambles" *Management Science* 51(12), 1777-1790.

Rabin, M. and R. Thaler, 2001, "Risk Aversion" *Journal of Economic Perspectives* 15(1), 219-232.

Ross, S.A., 1999, "Adding Risks: Samuelson's Fallacy of Large Numbers Revisited" *Journal of Financial and Quantitative Analysis* 34(3), 323-339.

Samuelson, P.A., 1963, "Risk and Uncertainty: A Fallacy of Large Numbers" *Scientia* 98, 108-113.

Samuelson, P.A., 1965, "Rational Theory of Warrant Prices" *Journal of Industrial Management Review* 6, 13-39.

Sundali, J.A. and F. Guerrero, 2009, "Managing a 401k Account: An Experiment in Asset Allocation" *Journal of Behavioral Finance* 10(2), 108-124.

Tverski, A. and M. Bar-Hillel, 1983, "The Long and the Short" *Journal of Experimental Psychology: Human Learning, Memory, and Cognition* 9, 713-717.

**Table 1**
**Many subjects rejected the volatile long-term investment with high expected return, but only after they were shown its cumulative wealth distribution.**

| *Pooled Samples* | Accept $T = 100$ After | Reject $T = 100$ After | Row Total |
|---|---|---|---|
| Accept $T = 100$ Before | 18 | 31 | 49 |
| Reject $T = 100$ Before | 3 | 31 | 34 |
| Column Total | 21 | 62 | 83 |

**Table 2**
**The sampled groups' responses were qualitatively similar.**

| *Subsample* | Accept $T = 100$ After | Reject $T = 100$ After |
|---|---|---|
| Coffee Shop | 5 | 26 |
| Students | 7 | 23 |
| Academics | 9 | 13 |
| Total | 21 | 62 |

**Table 3**
**Subjects were more likely to accept repetition of the rebalanced portfolio of two identical binary investments after they were presented with the distribution of its cumulative outcomes, despite having already seen the distribution of cumulative outcomes from the single binary investment.**

| *Pooled Samples* | Accept $T = 100$ After | Reject $T = 100$ After | Row Total |
|---|---|---|---|
| Accept $T = 100$ Before | 43 | 3 | 46 |
| Reject $T = 100$ Before | 25 | 12 | 37 |
| Column Total | 68 | 15 | 83 |

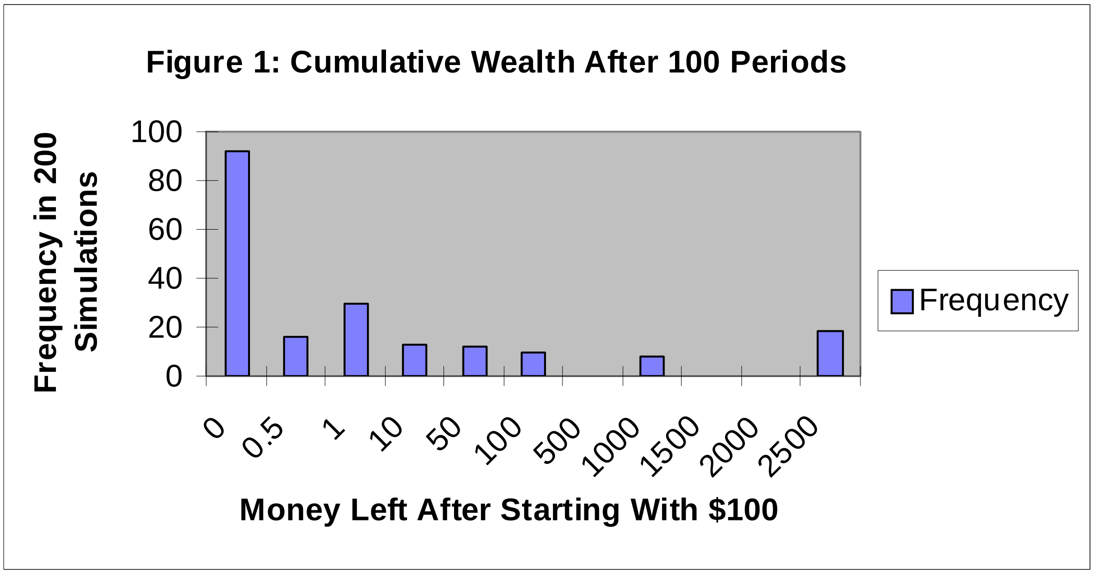
Figure 1: Cumulative Wealth After 100 Periods
Frequency in 200 Simulations
100
80
60
40
20
0
0
0.5
1
10
50
100
500
1000
1500
2000
2500
Money Left After Starting With $100
Frequency

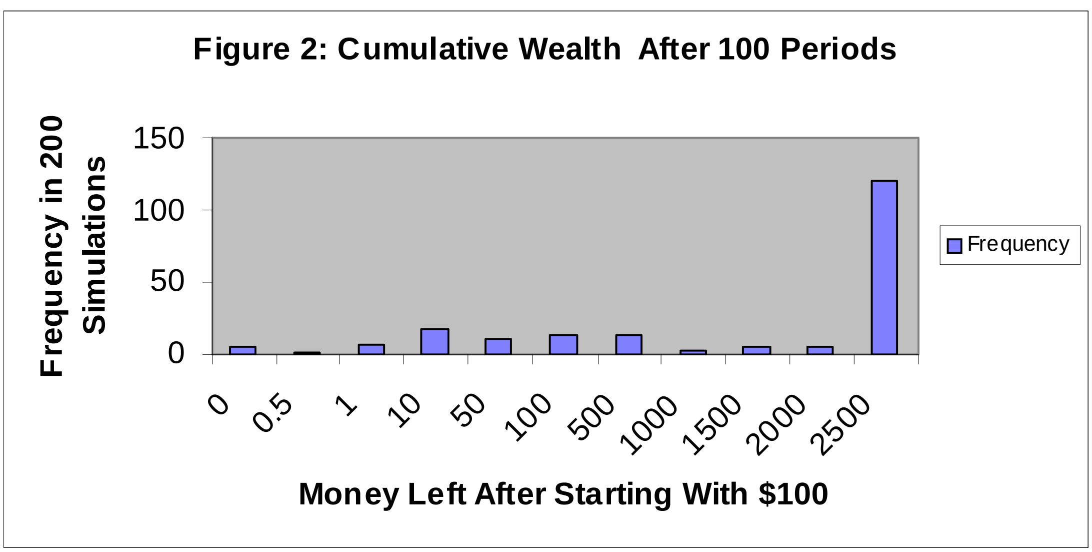
Figure 2: Cumulative Wealth After 100 Periods
Frequency in 200 Simulations
150
100
50
0
0
0.5
1
10
50
100
500
1000
1500
2000
2500
Money Left After Starting With $100
Frequency

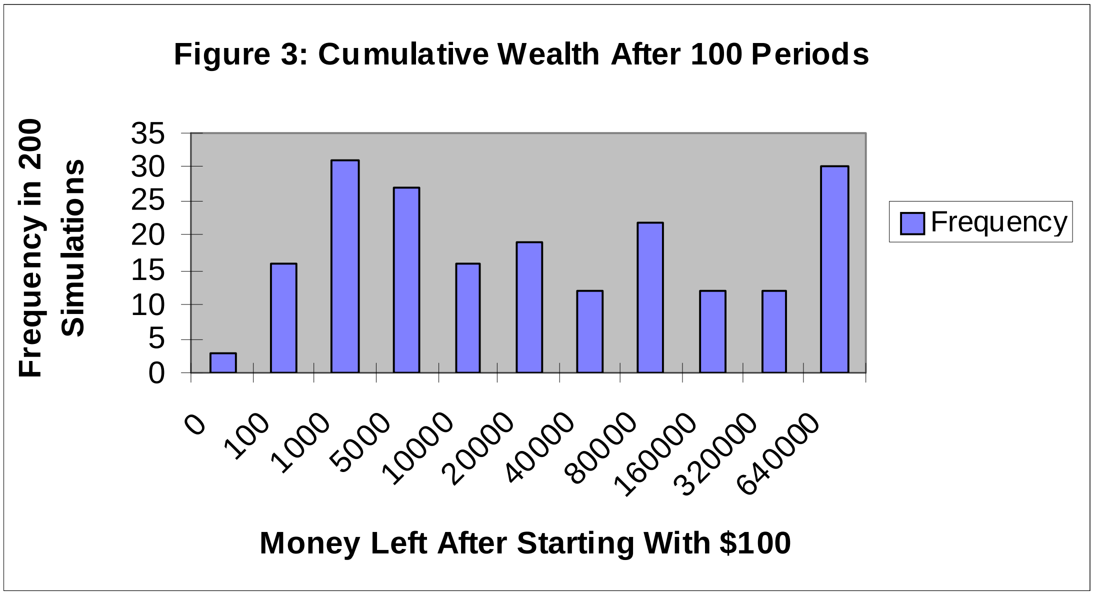
Figure 3: Cumulative Wealth After 100 Periods
Frequency in 200 Simulations
35
30
25
20
15
10
5
0
0
100
1000
5000
10000
20000
40000
80000
160000
320000
640000
Money Left After Starting With $100
Frequency

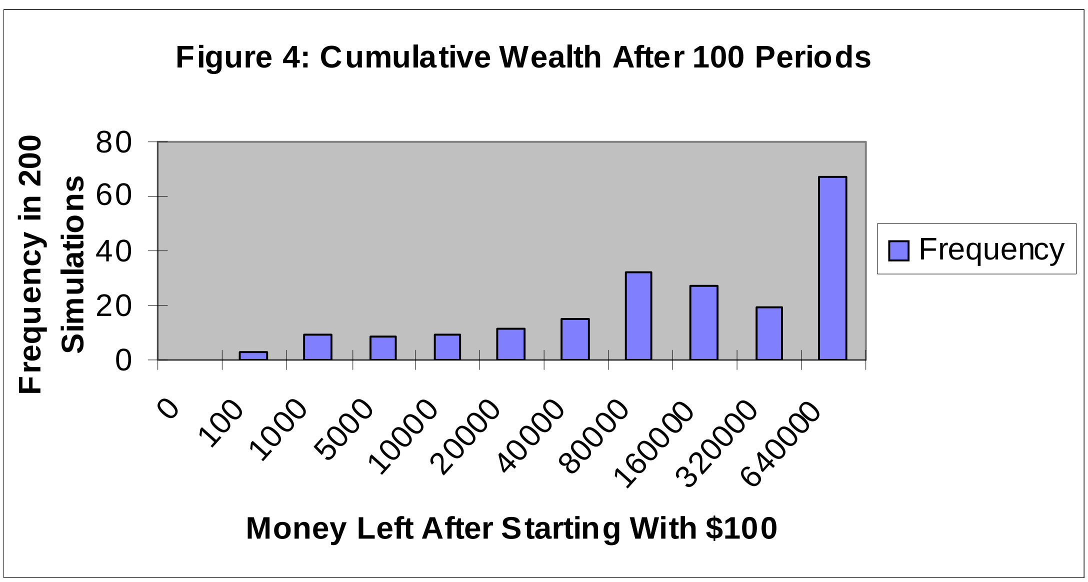
Figure 4: Cumulative Wealth After 100 Periods
Frequency in 200 Simulations
80
60
40
20
0
0
100
1000
5000
10000
20000
40000
80000
160000
320000
640000
Money Left After Starting With $100
Frequency

[1]The authors acknowledge the value of comments from the editors, from Profs. Ravi Jagannathan, Chris Yung and Peter McGraw, as well as comments from participants at the SPXI conference in Vienna.

[2]Thanks are extended to Haipeng ("Allen") Chen for providing this data. Readers interested in the extensive literature documenting computational errors made by individuals faced with deterministic, sequential percentage changes should see the thorough literature review in Chen and Rao (op.cit.).

[3]The model is exposited in most college investment or derivatives textbooks, and can even be found in the internet Wikipedia, under its entry for the binomial option pricing model.

[4]At best, they may have seen summary statistics of historical returns (e.g. the average and standard deviation of annual historical returns) or a model-based projection of the future cumulative return distribution estimated from historical returns.

[5]The rationality of this pattern was questioned by Samuelson (1963), and re-analyzed in subsequent papers by eminent scholars, e.g. Tversky and Bar-Hillel (op.cit.), Chew and Epstein (op.cit.), and Ross (op.cit.).

[6]Highly leveraged investments do present high expected returns accompanied by high volatility, and played a role in fomenting the Financial Crisis of 2008.

[7]This is quite unlike what can happen with repeated binary fixed dollar gambles, where the worst case potential loss increases steadily with the number of repetitions $T$.

[8]Benartzi and Thaler (op. cit.) did a separate experiment, asking each subject to choose her/his preferred allocation weight for a stock-like index, in a two-asset portfolio of the stock-like index with a bond-like index. But these questionswere not direct extensions of their repeated binary gamble questions.

[9]The reason for this does not lie in the portfolio's expected return (which is the same 15% per repetition that characterizes either of the single binary investments), but in the reduction of return volatility. Instead of receiving either an 80% gain or 50% loss each period with probability $\frac{1}{2}$, the diversified investor reduces the probability of each of those extreme outcomes to $\frac{1}{4}$, in order to receive an intermediate gain of 15% with probability $\frac{1}{2}$ (see (6)).

[10]Do you know the probability that your TIAA-CREF investments will go up or down this year, much less the amounts by which it might go up or down? You may gather some historical annualreturn statistics or seek some "expert advice", but nonetheless tomorrow you will either continue to hold your investments or you will change them.

[11]But the differences in these row and column totals (shown in Table 1) might be attributable to sampling error. So to formally test for that, we test the null hypothesis that the before and after

results are the same, using Liddell's Exact Test (see Armitage and Berry (1994, p. 127)). Under the null, one would expect each of the twocorresponding row and column totals to be equal. If the corresponding row and column totals are equal, then simple algebra implies that the number of subjects who accepted before viewing the outcome distribution but rejected after viewing it (i.e. 31) will equal the number of subjects who rejected before viewing but accepted after viewing (i.e. 3). The Liddell test statistic is the ratio $31/3 = 10.3$ with a 95% confidence interval of $[3.2, 52.8]$, so the null of equality (i.e. that the test statistic would equal 1, but for sampling error) was *rejected* at the 5% level.

[12]No less an authority than Samuelson (1965, p.17) argued that " This virtual certainty of almost-complete ruin bothers many writers. They forget, or are not consoled by, the fact that the gains of those (increasingly few) people who are not ruined grow prodigiously large – in order to balance the complete ruin of the many losers."(Thanks to Prof. Moshe Milevsky for providing this reference.)

[13]Liddell's test statistic for equality between the before and after results is $3/25 = .12$ with a confidence interval of $[.02, .39]$, well outside the null of equality's hypothesized value of 1. Hence the null of equality is again rejected at the 5% level.

[14]The academics may have inferred the underlying math from their study of the single investment's cumulative wealth distribution. But recall that some if not most of the academics surveyed had experience constructing and/or reading about closely related surveys. Because the academicians previously *were* surprised by the distribution of outcomes from the single investment, it is possible that their experience led them to suspect that another surprise could be lurking in the diversified portfolio, and hence hesitated to employ the reasoning that led them astray earlier. Moreover, the inherent complexity of the description of de-facto rebalanced diversification may have contributed to a feeling that something counterintuitive was at work.

[15]While 8 of those customers did not change their minds, all but one of them answered "No" to all five questions, thus refusing to invest under any circumstances presented.

[16]With the current market malaise, let's hope we all can locate a retirement vehicle with such favorable long-run returns.