Expected Return or Growth Rate?

Choices in Repeated Gambles That Model Investments [1]

Michael Stutzer, Professor of Finance and Director,

Burridge Center for Securities Analysis and Valuation, Univ. of Colorado, Boulder, CO

michael.stutzer@colorado.edu

and Susan Jung Grant, Professor of Marketing,

Univ. of Colorado, Boulder, CO

Current Version: June 2010

Some experimental studies have posed repeated binary fixed dollar gambles, focused on short vs. long run behavior. Results of those experimental designs may be misleading when used to interpret short vs. long term investment behavior in analogous IID settings, where returns are more realistically modelled as possible percentages, rather than possible dollar amounts. Prior psychological, educational, and marketing research has documented computational errors by individuals, in situations where it is relevant to cumulate the effects of deterministic, sequential percentage changes. The random sequence of percentage changes that characterizes the cumulative return from an investment creates a possibility for additional errors. We substitute percentage gambles for fixed dollar gambles to determine whether or not previous experimental results, and hypotheses motivated by them, are robust to this substitution. Some are not. Did subjects act *as if* they understood the important differences between cumulative long-run outcomes of fixed dollar gambles and their (investment relevant) fixed percentage counterparts? Many did not – including some expert authors in this research area.

1 Introduction

Paul Samuelson (1963) [12] used simple binary fixed dollar gambles (i.e. in each gamble, one either wins a fixed amount or loses a fixed amount) to illustrate a counterintuitive normative result concerning the decision to accept or reject one gamble versus many such gambles. Benartzi and Thaler (1999) [2] experimentally studied this and related phenomena, as did Klos, et al. (2005) [8]. Researchers hope to interpret the experimental results in more general contexts of uncertain choice. For example, Benartzi and Thaler (op.cit.) conducted additional experiments concerning long-term asset allocation decisions characteristic of retirement fund choices. Klos, et al. (op.cit) inferred from their 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 binary fixed dollar gamble.

But there is an important *qualitative* difference between a repeated fixed dollar gamble and a realistic investment, even when the investment's returns/period are modelled to be binary and IID, like the aforementioned fixed dollar gambles. In the former, 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, most analysts follow Samuelson (1965) [13] in assuming that one stands to make or lose fixed *percentages* of money, so the outcome of investment over time (i.e. the cumulative return) is the *product* of the sequence of future gross returns.

There is good reason to believe that this distinction will result in important behavioral differences. Chen and Rao (2007) [4] thoroughly document, in both an experimental and a market setting, that 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, more subjects 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] Eisenstein and Hoch [6] documented the inability of many subjects to approximately forecast – without use of electronic calculators – the cumulative long-run return from a riskless asset growing at a constant percentage per year. Many subjects acted as-if invested wealth grew linearly over time, rather than exponentially, i.e. compounded at the fixed interest rate.

Moreover, ex-ante investment analysis embodies the additional complexity of *random* sequential percentage changes, which determine the probability distribution of cumulative investment returns. In the following section 2, we show that this creates significant differences, both quantitative and qualitative, between repeated fixed dollar gambles and their closest investment analog, i.e. the standard textbook binomial fixed percentage price tree for an investment's cumulative return. [3] In section 3, we substitute a binomial investment in otherwise identical experiments to determine whether or not previous experimental results and hypotheses based on fixed dollar gambles are robust to this change, and whether or not the additional possibilities for error induced by stochastic percentage returns occur. Survey responses indicated that many subjects answered questions *as if* they were surprised by the cumulative return distributions resulting from binomial investments, including some (otherwise anonymous) academics who have co-authored articles referenced in the experimental literature concerning fixed dollar gambles. The findings suggest a particular modification to Benartzi and Thaler's (op.cit.) Hypothesis 2 concerning fixed dollar gambles, which is more consistent with the survey responses we observed. In addition, two behavioral phenomena discussed (respectively) in Samuelson (1963) and Samuelson (1965) were not frequently observed in our subject pool. Section 4 concludes.

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. After T repetitions:

$$\begin{aligned} W_T &= \sum_{t=1}^{T} X_t \\ &= (\frac{1}{T}\sum_{t=1}^{T} X_t)T \end{aligned} \tag{1}$$

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

$$\begin{aligned} E[W_T] &= \sum_{t=1}^{T} E[X_t] \\ &\overset{IID}{=} E[X_1]T \\ &= [\frac{1}{2}200 + \frac{1}{2}(-100)]T \\ &= 50T \end{aligned} \tag{2}$$

We see that this gamble has a positive expected value of $E[X_1] = 50$, and that its expected value after T repetitions is $50T$, which grows linearly to infinity as T (i.e. the number of repeated gambles) increases. Because the outcome (1) depends on (T multiplied by) an *average* whose distribution approaches normality (due to the Central Limit Theorem), the expectation (2) is close to the $Median[W_T]$ when T is suitably large. So (2) is a useful measure of the central tendency in repeated binary fixed dollar gambles. For example, with $T = 100$ repeated gambles, as 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) [10] 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.

For (1) and other repeated gambles with positive expected value, Benartzi and Thaler (op cit., p. 370) formulated and tested their "Hypothesis 2" ,i.e.

H2: Subjects will find repeated plays of a positive expected value gamble 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.

But Samuelson (1963) argued that while "at first glance" the calculations above seem to be a good reason to accept "a long sequence of favorable bets", he believed it really wasn't. He formulated this as a theorem: a decision maker who acts as if she maximized an expected utility that would lead her to decline the offer of a single one of these gambles *no matter how much wealth she has*, also should decline the offer of *any* number of repeated gambles.

Samuelson's result is true only under his maintained assumption that she would have rejected a single gamble at *any* possible level of wealth achievable by repetition. This assumption is an implicit but severe restriction on the form of the utility function. Ross [11, p.326] proves that "the only utility functions that reject the same gambles at all wealth levels are the risk-neutral function and the exponential, i.e., the constant coefficient of absolute risk aversion utility functions ". This raises the possibility that a different utility function could lead someone to eventually accept some number of repeated gambles, even if it leads to rejection of a single gamble. Indeed, Tversky and Bar-Hillel (1983) [15] constructed an expected utility counterexample that rationalizes rejecting one gamble while accepting many (they note that their example violates the aforementioned assumption maintained in Samuelson 1963). But they also devised a new, seemingly reasonable pair of normative axioms that *are* violated by rejecting one gamble while accepting many. Despite that normative claim, Chew and Epstein (1988) [5] argued that this behavior is not necessarily inconsistent with some seemingly reasonable *non*-expected utility decision criteria.

We do not attempt to shed additional light on this controversy, i.e. whether or not it is *rational in some sense* to reject one parlor gamble while accepting many independent, identical gambles. We restrict attention to whether or not it is *typical* behavior, formulated as hypothesis H1 below:

> **H1**: Subjects are more likely to accept repeated plays of a positive expected return gamble than a single play of it.

Benartzi and Thaler's (op.cit.) evidence did *not* support H1, although more of their subjects were willing to accept repeated plays after being shown the probability distribution of the outcomes of repeated plays, consistent with their Hypothesis 2 (listed as H2 above).

But to model an investment, Samuelson (1965) argued that random *percentage* gains and losses made more sense. When modeled in a particular way, this precludes the possibility of (nonsensical) negative asset prices. 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 fixed dollar gamble, after replacing its fixed dollar amounts with fixed percentage amounts.

Specifically, consider a risky investment that either increases 80% in value with probability $\frac{1}{2}$ or decreases in value by 50% with probability $\frac{1}{2}$, with those possibilities repeated in each subsequent period the investment is held. The expected return is $\frac{1}{2}80\% + \frac{1}{2}(-50\%) = +15\%$. For example, if \$100 is invested, the expected value after one period is \$115. Because this investment has a positive expected value, a researcher might expect to observe behavior in accord with the aforementioned Hypothesis 2 of Benartzi and Thaler (op.cit.).

We now turn to a mathematical analysis of this investment. The single-period gross return X_t from each dollar invested is either 1.80 with probability $\frac{1}{2}$ or is 0.5 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:

$$W_T = W_0 \prod_{t=1}^{T} X_t. \tag{3}$$

The mathematical expectation $E[\,]$ of wealth in the IID process (3) is

$$\begin{aligned} E[W_T] &\overset{I}{=} W_0 \prod_{t=1}^{T} E[X_t] \\ &\overset{ID}{=} W_0 E[X_1]^T \\ &= W_0[\frac{1}{2}1.8 + \frac{1}{2}(0.5)]^T \\ &= W_0[1.15]^T \qquad (4) \end{aligned}$$

Hence (4) shows that the *expected value* of wealth $E[W_T]$ increases exponentially to infinity as T increases, qualitatively similar but eventually outstripping the linear increase (2) of the fixed dollar gamble's expected cumulative value. But unlike (1), (3) 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. In fact, the distribution of W_T in (3) becomes extremely positively skewed for large T, so $E[W_T]$ in (4) is much higher than the *median value* 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, you can't lose more than your initial investment, unlike what occurs in the repeated binary gambles (like Samuelson's 1963 gamble) used in Benartzi and Thaler (op.cit.) and Klos, et.al. (op.cit.). When investment returns per period are IID (as our example is), the short mathematical appendix describes a good, large T approximation for the median invested wealth. Applying that approximation to our example yields:

$$
\begin{aligned}
Median[W_T] &\overset{IID}{\approx} W_0 e^{E[\log X_1]T} \\
&= W_0 e^{[\frac{1}{2}\log 1.8+\frac{1}{2}\log 0.5]T} \\
&= W_0 e^{-.0527T} \qquad (5)
\end{aligned}
$$

The approximation (5) shows that the median value depends on the expected *logarithm* of the gross return (i.e. -5.27%), rather than the 15% expected net return. The expected log gross return is the *expected growth rate* of invested wealth W. Comparing (4) and (5) shows that while the positive expected value $E[W_T]$ exponentially approaches infinity as T increases (due to the 15% expected return per period), the smaller positive $Median[W_T]$ exponentially decays toward zero, due to the negative (-5.27%) expected *log* gross return (growth rate). For example, starting from an initial investment of just $W_0 = \$100$, (4) shows that after $T = 100$ periods, the expected value of invested wealth is $\$100(1.15)^{100} = \$117,431,345$, while (5) shows that the median wealth will be only about $\$100e^{-5.27} \approx \$0.50$! Results from 200 simulations, reported in the survey instrument's Table 1 (see the appendix), are consistent with this; in 49% of the simulations, one's initial investment of \$100 deteriorated to a mere fifty cents or less.[4] This probability of almost-complete ruin grows asymptotically toward 100% as $T \to \infty$.

Yet decision theorists do *not* all agree that it would be irrational to accept an investment opportunity like this (which after all, has a 15% expected return per period, leading to the gigantic expected value of invested wealth computed above), despite the virtual certainty of almost-complete ruin for suitably large T.[5] 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. [6]

Simulations summarized in the survey instrument's Table 1 (see the appendix) indicate that there is indeed a small chance (6.5%) of winding up with more than \$2500 (i.e. a multiple of 25 times the initial investment) after $T = 100$ periods, which might "console" subjects enough to accept the opportunity, despite the near 90% probability of losing at least some of the initial \$100 invested, and the near 50-50 chance of winding up with less than fifty cents.

Hence we also experimentally test the following hypothesis suggested by the aforementioned quote from Samuelson (1965):

> **H3**: Some subjects will still be willing to accept repeated plays of a negative expected *log* gross return (fixed percentage) gamble after they are presented with the distribution of final outcomes.

The final area investigated is the important, possibly counterintuitive, value of investment diversification. The most common multi-period 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. Benartzi and Thaler (op cit.) did a separate experiment, asking each subject to choose her/his preferred allocation weight on a stock-like index, in a two-asset portfolio of the stock-like index with a bond-like index. But these questions were not direct extensions of their repeated binary gamble questions, nor did they present the outcome distribution of the *portfolio* selected by a subject; instead, they presented the outcome distributions of the stock-like index and the bond-like index. Of course, financial theory holds that the portfolio's outcome distribution is the relevant distribution, and that is a subject's weighted mixture of those two distributions. Benartzi and Thaler (op.cit.) did not show that to their subjects.

Hence we added diversification questions to our survey, which were the simplest and most transparent extensions of our prior questions concerning a single investment. We wished to see whether or not subjects would choose to put half of their \$100 initial stake in one asset, and the other half in an *identical* second asset whose return in each period 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 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 portfolio's outcome distribution.

One might conjecture that no one should agree to split their money between two identical investments that, when considered separately as above, tend to lose money over time – 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 per period is $\frac{1}{2}15\% + \frac{1}{2}15\% = 15\%$, i.e. the same expected value as each of the separate investments. As a result, the portfolio will have the same enormous expected cumulative return as the single investment, and one might guess that the portfolio's outcome distribution will also look similar.*However, this heuristic reasoning frames the issue too narrowly; a quantitative calculation of a typical (i.e. median) outcome results in a radically different conclusion than occurs* *for sole investment in either one of the investments.* To see this, do the following simple math:

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

In contrast to the typical almost-complete ruin stemming from investment in a single asset for large-T periods (see (5)), (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 value of wealth will be around \$7750 after 100 periods, in contrast to the \$0.50 median value of wealth that occurs when investing everything in (either) single asset used to form it. This is why the survey instrument's Table 2 (see the appendix) shows that in 200 simulation runs for the rebalanced portfolio, 60% of the simulation runs ended with more than \$2500, while only 20.5% of the runs lost some of the initial \$100 (and of course, none lost more than the \$100 initial investment). The mathematical reason for this obviously does not lie in the portfolio's expected return per period (which is the same 15% that characterizes sole investment in a single asset), 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 the middle line in (6)). The stark qualitative difference between (5) (for a single losing investment) and (6) (for a rebalanced portfolio of two identical losing investments) appears to be the simplest example of what has been dubbed a *Parrondo Paradox* in more complex and artificial situations.[7]

Because of the paradoxically strong and positive effects of diversification on the central tendency of the portfolio value distribution, we formulate the following hypothesis:

> **H4**: When $E[\log W] > 0$ due to diversification, subjects are more likely to accept diversification *after* they are presented with the distribution of final outcomes.

Subjects who reject the equally-weighted, rebalanced portfolio *before* examining Table 2, but then accept the portfolio *after* examining Table 2, act in accord with H4 above. Their behavior would confirm the conjecture of Klos, et. al. (op.cit., p. 1788) 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." Because of diversification, the narrowly framed heuristic that 'half of two identical losers is the same as one loser' is extremely misleading. The behavioral literature has identified many other examples of unfortunate behavior resulting from such narrow framing (e.g. see Kahneman and Lovallo 1993 [7]), including a non-investment "diversification" bias (Simonson and Winer 1992 [14]).

3 The Survey

We substituted the fixed percentage investment (analyzed in the previous section) for Samuelson's (1963) fixed dollar gamble that was used in the Benartzi and Thaler (op cit.) experimental protocol. Question 1 of the survey instrument (see the appendix) models the decision to invest $100 in a single-period investment. Question 2 models the decision to buy-and-hold that investment for $T = 100$ "periods". Care is taken to point out that a computer would determine the outcome immediately – one would not have to wait an appreciable length of time for the outcome of $T = 100$ "periods". The use of a $100 bet and 100 "periods" matches the initial bet and number of

IID repetitions in the Samuelson (1963) binary fixed dollar experiment conducted by Benartzi and Thaler (op.cit.). To help ensure that the subjects' answers reflect their respective internal decision processes in stochastic settings, rather than the influence of normative opinions from finance professors, investment advisors and financial journalists, the phrases "investment" and "buy and hold" were not used in the wording of the survey questions. Question 3 poses the same question, *after* the subjects are shown the probability distribution of the cumulative outcome after 100 "periods" (Table 1 of the survey). 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 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 poses the same question, *after* the subjects are shown the probability distribution of the cumulative outcome after 100 "periods" (Table 2).

Following Benartzi and Thaler (op.cit.), our subjects included both coffee shop customers and students (University of Colorado). The students were nearing the end of an undergraduate investment course. Hence they had been exposed to typical undergraduate investment coursework which surely includes use of percentage returns, expected returns, etc. *Unlike Benartzi and Thaler (op.cit.) and Klos, et.al. (op.cit.), we also 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 Klos, et.al. (op.cit.). 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*, we conjectured that subjects in this subsample would be unlikely to change their responses upon observing the cumulative outcome distributions, because they would anticipate the relevant features of the outcome distribution before seeing our presentation of it and hence have no need to change their prior responses.

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 either the median outcome or 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) [3] 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

Hypothesis **H1** examines the frequency of behavior that bothered Samuelson (1963): were subjects more likely to accept the result of $T = 100$ than $T = 1$, *before* seeing the distribution of outcomes? The contingency table of the results from survey Questions 1 and 2 in our pooled sample are presented in the following Figure 1.

Figure 1

Hypothesis H1: When $E[W_1] > 0$, subjects are more likely to accept W_{100} then W_1

Pooled Samples	Accept $T = 100$	Reject $T = 100$	Row Total
Accept $T = 1$	33	11	44
Reject $T = 1$	16	23	39
Column Total	49	34	83

While the fraction accepting $T = 100$ (i.e. 49/83) was higher than the fraction accepting $T = 1$ (i.e. 44/83), this could be due to sampling error. To formally test the null hypothesis that the probabilities of accepting (rejecting) $T = 1$ and $T = 100$ are the same, we employ Liddell's Exact Test [1, p.127]. Under the null, one would expect each of the corresponding row and column totals to be equal, but for sampling error. If the corresponding row and column totals are equal, then simple algebra implies that the number of subjects who accepted $T = 1$ while rejecting $T = 100$ (i.e. 11) will equal the number who rejected $T = 1$ while accepting $T = 100$ (i.e. 16), but for sampling error. But are they far enough apart to reject the null of equality in population? The Liddell test statistic is the ratio $11/16 = 0.69$, and its 95% confidence interval is computed from Liddell's formula to be $[.29, 1.58]$. Because the null hypothesis is that the population value of the ratio equals one, which in this case is within the confidence interval of the test statistic, Liddell's Exact Test applied to the pooled samples fails to reject the null of equal probabilities of accepting the $T = 1$ and $T = 100$ investments, at the 5% level of significance. As a result, Hypothesis **H1** was *not* confirmed in the pooled sample.

Separate tests using the separate subsamples of coffee shop customers and students also failed to reject the null of equality at the 5% level. Benartzi and Thaler (op.cit.) also failed to find typical coffee shop customers and students behaving like Samuelson's(1963) friend. So our results show that those particular results of theirs are robust to our substitution of a binary fixed percentage investment for the binary fixed dollar gamble used by them and Samuelson (1963). But Benartzi and Thaler (op.cit.) did not survey an expert panel of academic authors of closely related literature.

Applied to our academic subsample, we found that the Liddell test rejected the null of equality at the 10% level, i.e. was marginally significant. The test statistic was 1/7, with a 90% confidence interval $[.006, .089]$ below the null hypothesized population value of 1. The contingency table is shown in the following Figure 1A.

Figure 1A

Hypothesis H1: When $E[W_1] > 0$, subjects are more likely to accept W_{100} then W_1

Academic Subsample	Accept $T = 100$	Reject $T = 100$	Row Total
Accept $T = 1$	9	1	10
Reject $T = 1$	7	5	12
Column Total	16	6	22

Note that the fraction of academicians accepting $T = 100$ ($16/22 = 73\%$) was (marginally significantly) higher than the fraction accepting $T = 1$ ($10/22 = 45\%$), consistent with Samuelson's friend's behavior in Hypothesis **H1**.

All the aforementioned results occurred *before* subjects were shown the probability distribution of W_{100} for the $T = 100$ investment. So Hypothesis **H2** examines the robustness of Benartzi and Thaler's (op.cit.) Hypothesis 2, i.e. did our subjects find the $T = 100$ investment (with its gigantic expected cumulative return) 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 the cumulative "buy-and hold" investment more attractive after seeing the distribution of outcomes. Liddell's test statistic for the pooled sample was 10.3 with a 95% confidence interval of $[3.2, 52.8]$, so the null of equality *was* rejected at the 5% level. But the following Figure 2 shows while $49/83 = 59\%$ of the pooled sample accepted the $T = 100$ investment before being shown the distribution of W_{100}, only $21/83 = 25\%$ did so *after* being shown that distribution. Hypothesis **H2** is that the opposite would happen! Many subjects acted *as if* they did not anticipate the negative expected growth rate and its associated decay of the median cumulative return toward zero, which they did not like – despite the huge expected cumulative return.

Figure 2

Hypothesis H2: When $E[W_1] > 0$, subjects are more likely to accept W_{100} after they are presented with the distribution of final outcomes.

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

The subsamples all exhibited the same pattern of *less* frequent acceptance of $T = 100$ after seeing the distribution of outcomes. 26% of coffee shop customers were willing to accept $T = 100$ before seeing the outcome distribution, while only 16% were willing to after seeing it. 83% of the students were willing to accept $T = 100$ before seeing the outcome distribution, while only 23% were willing to after seeing it. Our academic panel was no different; 73% were willing to accept before, while only 41% were willing to do so after – a statistically significant difference at the 5% level. *The academic authors of closely related literature did not act as if they had anticipated the relevant features of the cumulative outcome distribution before seeing it.*

Hence our data show that Benartzi and Thaler's Hypothesis 2 was *not* robust to the substitution of a binary fixed percentage investment for their binary fixed dollar gamble, despite the fact that both have very large expected cumulative values. Because (5) shows that the median value of wealth is determined by the expected *logarithm* of the gross return per period – which is negative in this case – we propose the following simple modification of Benartzi and Thaler's Hypothesis 2:

H2$'$: Subjects will find a long-term investment with a positive expected *log* gross return (i.e. a positive expected growth rate) more attractive if they are presented with the distribution of final outcomes.

This data also sheds light on Hypothesis **H3**. As we just concluded, most subjects acted *as if* the near-zero median value of cumulative wealth after $T = 100$ was more relevant than the over \$117 million expected value of cumulative wealth – once they had access to the resulting probability distribution of cumulative wealth. There was little tendency to risk \$100, even when the risk was hypothetical as it was here, for a small chance to earn what Samuelson (1965, p.17) would have dubbed a gain that is "prodigiously large – in order to balance the complete ruin of the many losers." Still Hypothesis H3 is confirmed, in the sense that "some" subjects were willing to risk \$100 on this. *But that behavior was not typical, despite Samuelson's previously quoted opinion that a very small chance to win a huge sum of money could compensate for the high probability of "almost-complete ruin".* The results are summarized in the following Figure 3.

Figure 3

Hypothesis H3: When $E[\log W_1] < 0$, some subjects will accept W_{100} even after seeing the distribution of final outcomes

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

Hypothesis **H4** examines whether subjects will be more likely to accept (the favorable effects of) diversification after seeing the distribution of outcomes resulting from it, despite its having no effect on the expected cumulative return. *Our data show that this behavior was common.* Liddell's test statistic is .12 with the 95% confidence interval $[.02, .39]$, well outside the null of equality's hypothesized population value of one. The data are depicted in the following Figure 4.

Figure 4

When $E[\log W] > 0$ due to diversification, subjects are more likely to accept diversification after they are presented with the distribution of final outcomes.

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 4 shows $46/83 = 55\%$ of the pooled sample accepted diversification before seeing the distribution of outcomes resulting from it, rising to 82% after the subjects viewed the distribution of outcomes in the survey instrument's Table 2 (see the appendix). Note that the change in their willingness to accept the portfolio occurred despite having *already* seen the distribution of outcomes from the investment used to construct it (i.e they had seen the survey instrument's Table 1) when they answered the survey's Question 3. This pattern is also consistent with our proposed Hypothesis **H2**', because the expected log gross return of the diversified portfolio is positive. But the effect was concentrated among the coffee shop customers and students. There was no significant difference in the before and after responses of the academicians, with 77% accepting diversification before, rising just a bit to 86% after. The academics may have inferred the underlying math from their study of the single investment's outcome distribution (i.e. the survey instrument's Table 1 used in the survey's Question 3). Alternatively, because the academicians previously *were* surprised by the distribution of outcomes from the single investment, it is possible that they suspected another surprise could be lurking with the diversified investment, and 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. Some if not most of them had experience constructing and/or reading about closely related surveys.

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 receiving the outcome distribution with its positive median value (and hence positive expected growth rate).[8] While only 13 of our 30 undergraduate business students rejected the portfolio before seeing the outcome 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. Moreover, only one subject in each of our three sampled groups who had *accepted* the portfolio *before* having access to the outcome distribution, decided to reject the portfolio *after* having access to the outcome distribution.

Those many subjects who changed their minds acted *as if* they were surprised by the high positive median cumulative return, associated with the positive expected log gross return (i.e. growth rate) created by diversification among two assets that separately had near-zero median cumulative returns.

4 Future Research

The survey results could vary with the win-loss parameters. The experiment used sizable percentage gains and losses per "play", characteristic of highly speculative investments. The results might also be dependent on the particular wording of the survey questions. Also, more data is needed to test our Hypothesis **H2**' that hypothesizes the primacy of the expected log gross return (rather than the expected return itself) as an important determinant of long-run investment behavior after the outcome distribution is known and understood. Are there interesting gender-specific differences in the responses? Does it matter if subjects had prior investment experience? In future research, we plan to analyze data from larger,Qualtrics survey panels to help address these issues.

5 Conclusions

Experiments that pose both one-shot and repeated binary, fixed dollar gambles are not the best design for understanding aspects of short vs. long term investment behavior. Financial economists believe that a better simplified model for investment returns utilizes repeated binary, fixed *percentage* gambles that cumulate multiplicatively rather than additively. This is standard textbook pedagogy dubbed the *binomial* model. The central tendency of the cumulative wealth distribution resulting from the fixed *dollar* gambles used in previous studies is largely determined by the expected return per gamble, but with the fixed *percentage* gambles of the binomial investment model, it is the expected *log* gross return per gamble that determines the central tendency, i.e. the median cumulative return. Our experiment confirmed (the finding of previous studies) that behavior concerning long-run prospects often changes once the subjects are given access to the probability distribution of long-run outcomes. Surprisingly, this also characterized subjects drawn from the population of researchers engaged in closely related research, who presumably had much better ability to grasp the relevant features of the outcome distribution before seeing it. But our experiments did not generally confirm (the finding of previous studies) that a positive expected return per gamble – even when unusually high – made subjects more likely to accept repeated percentage gambles once they saw the outcome distribution. In light of our results, we alternatively hypothesized that a positive expected *log* gross return will do that. Subjects acted *as if* they were surprised by the radically different performance of investments with extremely high expected cumulative returns when the investments also had negative expected log gross returns. This held for our relatively uninformed subsample of coffee shop customers, our somewhat better informed subsample of college investment students, and (perhaps unique in this particular literature) for our presumably best informed subsample comprised of authors of closely related papers.

The aforementioned difference between the role of expected return and expected log gross return in determining the central tendency of the distribution of long-term cumulative outcomes can be so stark that an investment with a high positive expected return can have a near-zero central tendency (i.e. almost complete-ruin). In a different paper, Samuelson (1965) argued that it wouldn't be irrational to consider making a long-term investment like that. Some of our experimental subjects did indeed do that, even after seeing the long-term outcome distribution, but this behavior was atypical.

The role of diversification was also explored. Using survey questions that made no use of the (now normative) term "diversified portfolio", we posed an opportunity that modeled a long-term, rebalanced portfolio of two uncorrelated investments, each of which was identical to the investment used in the earlier survey questions (unlike previous studies cited herein). Taken separately, each investment would have resulted in a near-zero median long-term cumulative value. But an equally value-weighted portfolio of the two had a healthy positive median long-term cumulative value. Subjects were still more likely to invest in this portfolio after having been shown its outcome distribution, despite having previously seen the distribution of outcomes from the investments that comprise it.

Mathematical Appendix

This appendix contains the math to establish that the expected log gross return (a.k.a. the expected growth rate) 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.

Survey

1. Would you agree to spend $100 for a 50-50 (i.e. equal chances) opportunity to either earn 80% (i.e. wind up with $180) or lose 50% (i.e. wind up with $50)?
The outcome would be determined immediately, by a computer-simulated coin toss.

Would you spend $100 for this opportunity?

__________________ YES, I WOULD SPEND   __________________ NO, I WOULD NOT SPEND

2. Instead, suppose you had to leave the money in this opportunity for 100 “periods”. The outcome would still be determined immediately, by 100 successive computer-simulated toin tosses, so a “period” is just a tiny fraction of second. In the first “period”, your $100 would either grow to $180 or shrink to $50, depending on the outcome of the first coin toss. In the second period, that amount (either $180 or $50, depending on the outcome of the first coin toss) would again either grow by 80% or decline by 50%, dependent on the outcome of the second coin toss. This process would continue until the 100th coin is tossed. It would take less than a second to find out how much money you wound up with.

Would you spend $100 for this opportunity?

__________________ YES, I WOULD SPEND   __________________ NO, I WOULD NOT SPEND

3. Table 1 on page 3 tabulates the results of many computer simulations of the opportunity just described in Question 2. After each simulation of 100 successive coin tosses, the computer calculated how much money you would have wound up with. Table 1 shows the percentage of simulation runs in which you wound up with various amounts of money.

After studying Table 1, would you spend $100 for this opportunity (as described in Question 2)?

__________________ YES, I WOULD SPEND   __________________ NO, I WOULD NOT SPEND

4. Now suppose another university offers you another opportunity as described in Question 2, to be used in conjunction with our university's opportunity. Our university computer's coin tosses will determine what happens here, while their university computer's coin tosses will determine what happens there. As you would guess, their computer tosses coins completely independently of our computer.

To make matters concrete, suppose the two computers automatically keep your money split equally between the two places. Here is how that would work. To start, you would give $50 to our computer and $50 to their computer. Then, our computer and their computer will each toss a fair coin. Your $50 here will either grow by 80% (to $90) or decline by 50% (to $25). Your $50 there will also either grow to $90 or decline to $25.

In any event, the amounts in the two places will then be added together, half of which will be reallocated to each place. For example, if after each computer tosses its coin, the total is $115, $57.50 will be reallocated to each computer. Each computer will then toss its own coin again, again determining either an 80% gain or 50% loss of the money within it. The amounts in the two computers will again be totaled and split equally between the two computers. The process will continue like this until 100 coins are tossed in both places. Due to the amazing speed of computers, it will again take less than a second to tell you how much money you wound up with.

Would you spend $100 on this two-computer opportunity?

__________________ ____________________

YES, I WOULD SPEND NO, I WOULD NOT SPEND

5. Table 2 on page 3 tabulates the results of many computer simulations of the two-computer opportunity just described in Question 4. After each simulation of 100 successive coin tosses by each computer, both computers calculated how much money you would have wound up with, to make sure no errors in calculation were made. Table 2 shows the percentage of simulation runs in which you wound up with various amounts of money.

After studying Table 2, would you spend $100 in the two-computer opportunity described in Question 4?

__________________ ____________________

YES, I WOULD SPEND NO, I WOULD NOT SPEND

TABLE 1 (for Question 3)

AFTER STARTING WITH $100:

Left After 100 Periods	**Percentage of Simulations**
0 - $0.50	49%
$0.50 - $1.00	10.5%
$1 - $10	15.5%
$10 - $50	4%
$50 - $100	6.5%
$100 -$500	4%
$500 - $1000	0%
$1000 - $1500	4%
$1500 - $2000	0%
$2000 - $2500	0%
$2500 -	6.5%
TOTAL	100%

TABLE 2 (for Question 5)

AFTER STARTING WITH $100:

Left After 100 Periods	**Percentage of Simulations**
0 - $0.50	2.5%
$0.50 - $1.00	1%
$1 - $10	3.5%
$10 - $50	8.5%
$50 - $100	5%
$100 -$500	6.5%
$500 - $1000	6%
$1000 - $1500	1.5%
$1500 - $2000	2.5%
$2000 - $2500	3%
$2500 -	60%
TOTAL	100%

Notes

[1]The authors acknowledge the value of comments 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 binomial tree 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]The (finite) simulation was conducted on a spreadsheet.

[5]Unlike our example, Ross (op.cit.,sec.4) only analyzed investments that had *positive* expected log gross returns, and hence a median cumulative return growing to infinity as T does. He proved that anyone who had constant *relative* risk aversion less than one (i.e. less risk-averse than log utility!) will accept those investments for some T.

[6]Thanks to Prof. Moshe Milevsky for providing this reference.

[7]See Parrondo, et al. (2000) [9].

[8]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.

References

[1] P. Armitage and G. Berry. *Statistical Methods in Medical Research*. Blackwell, 3 edition, 1994.

[2] Shlomo Benartzi and Richard Thaler. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 110:73–92, 1995.

[3] Colin Camerer. Individual decision making. In J.H. Kagel and A.E. Roth, editors, *Handbook of Experimental Economics*, pages 587–683. Princeton University Press, 1995.

[4] Haipeng Chen and Akshay R. Rao. When two plus two is not equal to four: Errors in processing multiple percentage changes. *Journal of Consumer Research*, 34:327–340, 2007.

[5] S.H. Chew and Larry G. Epstein. The Law of Large Numbers and the attractiveness of compound gambles. *Journal of Risk and Uncertainty*, 1:124–132, 1988.

[6] Eric Eisenstein and Stephen Hoch. Intuitive compounding: Framing, temporal perspective, and expertise. Working Paper, Johnson School, Cornell University, 2007.

[7] Daniel Kahneman and Dan Lovallo. A cognitive perspective on risk taking. *Management Science*, 39(1):17–31, 1993.

[8] Alexander Klos, Elke U. Weber, and Martin Weber. Investment decisions and time horizon: Risk perception and risk behavior in repeated gambles. *Management Science*, 51(12):1777–1790, 2005.

[9] Juan M.R. Parrondo, Gregory P. Harmer, and Derek Abbott. New paradoxical games based on Brownian ratchets. *Physics Review Letters*, 85:5226–5229, 2000.

[10] Matthew Rabin and Richard Thaler. Risk aversion. *Journal of Economic Perspectives*, 15(1):219–232, 2001.

[11] Stephen A. Ross. Adding risks: Samuelson's fallacy of large numbers revisited. *Journal of Financial and Quantitative Analysis*, 34(3):323–339, 1999.

[12] Paul A. Samuelson. Risk and uncertainty: A fallacy of large numbers. *Scientia*, 98:108–113, 1963.

[13] Paul A. Samuelson. Rational theory of warrant prices. *Industrial Management Review*, 6:13–39, 1965.

[14] Itamar Simonson and Russell Winer. The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research*, 19:133–138, 1992.

[15] A. Tversky and M. Bar-Hillel. Risk: The long and the short. *Journal of Experimental Psychology: Human Learning, Memory, and Cognition*, 9:713–717, 1983.