Submitted to *Operations Research*
manuscript (Please, provide the mansucript number!)

Response to Paul A Samuelson letters and papers on the Kelly Capital Growth Investment Strategy

William T Ziemba
Alumni Professor of Financial Modeling and Stochastic Optimization (Emeritus),
University of British Columbia, Vancouver, BC,
Visiting Professor Sabanci University and the Korean Advanced Institute for Science & Technology

The Kelly Capital Growth Investment Strategy (KCGIS) is to maximize the expected utility of final wealth with a logarithmic utility function. This approach dates to Bernoulli's 1738 suggestion of log as the utility function arguing that marginal utility was proportional to the reciprocal of current wealth. In 1956 Kelly showed that static expected log maximization yields the maximum asymptotic long run growth. Later, others added more good properties such as minimizing the time to large asymptotic goals, maximizing the median, and being ahead on average for the first period. But there are bad properties as well such as extremely large bets for short term favorable investment situations because the Arrow-Pratt risk aversion index is near zero. Paul Samuelson was a critic of this approach and here we discuss his various points sent in letters to Ziemba and papers reprinted in MacLean, Thorp and Ziemba (2011). Samuelson's opposition has prevented many finance academics and professionals from using and suggesting Kelly strategies to students. For example, Ziemba was asked to explain this to Fidelity Investments, a major Boston investment firm close to and influenced by Samuelson at MIT. I agree that these points of Samuelson are theoretically correct and respond to them. I argue that they all make sense and caution users of this approach to be careful and understand the true characteristics of these investments including ways to lower the investment exposure. While Samuelson's objections help us understand the theory better, they do not detract from numerous valuable applications, some of which are discussed here.

Key words: capital growth, log utility, risk aversion, long run wealth maximization, Samuelson critique
History: This paper was first submitted on November 1, 2012.

1. Introduction to the Kelly capital growth criterion and Samuelson's objections to it

The Kelly capital growth criterion which maximizes the expected log of final wealth, provides the strategy that maximizes long run wealth growth asymptotically for repeated investments over time. But a shortcoming is its very risky short term behavior because of log's essentially zero risk aversion and consequently the large concentrated investments or bets that it suggests. The criterion is used by many investors, hedge funds, sports bettors and its seminal application is to a long sequence of favorable to investment situations.

Since I have a long history with Professor Paul A. Samuelson starting with his papers in my 1975 book Ziemba and Vickson, he wrote me from time to time on various topics. Paul was a critic of the theory concerned with the Kelly criterion and how that impacted its use in practice. Because of Paul's status, arguably the most important economist of the last century, people took note of the fact that he was objecting even though they did not actually know what these objections were. As a consequence, the Kelly strategy is not used much in most of the investment industry except for certain hedge funds looking for superior performance. Also it is not a standard topic in, for example, MBA finance courses.

My motivation for this paper comes from two sources. First, Paul wrote me three letters on this topic and his papers objecting to the Kelly criterion are reprinted in my recent book with Thorp and MacLean. So I wanted to respond to these letters and his papers. The second reason for this paper was a request from Fidelity Investments in August 2011 in Boston, a multi trillion dollar investment firm, to explain exactly what is Samuelson' objection to Kelly and should they be using Kelly strategies? After a five hour session on this and other topics, I think they were convinced to consider Kelly strategies and did understand the advantages and disadvantages of the strategy.

The first letter was the correspondence of November 16, 2005 to Professor Elwyn Berlekamp[1] and forwarded to me on December 13, 2006. Samuelson sent additional letters to me on 17 May 2007 and 12 May 2008. Sadly he died before I was able to finish this paper.The letters are downloadable from the web.[2]

Besides myself and my colleague Edward Thorp, who was the first one to employ this *Fortune's Formula* as he called it to the game of blackjack in his 1960 book *Beat the Dealer* that changed the way this game was played once he showed that there was a winning strategy, there are other notable investors who use such strategies in various forms. These include Jim Simons of the Renaissance Mediallion hedge fund, arguably the world's most successful hedge fund who I taught this approach to in 1992. Others who I believe behave as if they were full or close to full Kelly investors include George Soros, Warren Buffett and John Maynard Keynes. The reason we (Thorp and I) believe this is because of the portfolios they hold which are very concentrated in very few assets with huge positions in each asset, their monthly performance which include many losses but more gains and their very high long run growth of wealth.

The plan of this paper is a follows: Section 2 describes what is the Kelly criterion and what are its main properties. Section 3 describes Samuelson's objections one by one in general terms and my response to them aided by some research of Ed Thorp, David Luenberger and Harry Markowitz all of whom agree with me in this debate with the deceased giant thinker. Section 4 describes three investors Samuelson posed to me in letters with the addition of two tail investors.[3] Section 5 describes various applications and endorsements and information relevant to the actual use of

Kelly and fractional Kelly strategies in practice by me and others. Finally, Section 6 concludes where I argue that Samuelson's points are basically valid and sharpen our understanding of these strategies but the Kelly approach, if properly used, is extremely valuable in many applications.

2. What is the Kelly Strategy and what are its main properties?

Until Daniel Bernoulii's 1738 paper, linear utility of wealth was used so the value in ducats equalled the number of ducats one had. He postulated that the additional worth was less as wealth increased and was proportional to the reciprocal of wealth so

$$u'(w) = 1/w \text{ or } u(w) = \log(w).$$

Thus concave log utility was invented.

In the theory of optimal investment over time, it is not quadratic (the utility function behind the Sharpe ratio) but log that yields the most long term growth asymptotically. But the elegant results on the Kelly (1956) criterion, as it is known in the gambling literature and by capital growth criterion as its known in the finance economics literature, see the survey by Hakansson and Ziemba (1995), that were proved rigorously by Breiman (1961) and generalized by Algoet and Cover (1988) and Thorp (2011) are mostly long run asymptotic results. See also the papers in MacLean, Thorp and Ziemba (2011) and the discussion in Cover and Thomas (2006). These are discussed below.

The Arrow-Pratt risk aversion index for log w is:

$$R_A(w) = \frac{-u''}{u'} = \frac{1}{w}$$

which is essentially zero, where u is the utility function of wealth w, and primes denote differentiation. Hence, in the short run, log can be an exceedingly risky utility function with wide swings in wealth values.

In his 1738 paper, Bernoulli also formulated the St Petersburg paradox (see Summer, 1954 for an English translation of the original latin) which is discussed by Samuelson (1977), Aase (2001), Cover and Thomas (2006). See also MacLean, Thorp and Ziemba (2011) which reprints most of the papers cited here including those by Paul Samuelson.

John Kelly (1956) working at Bell Labs with information theorist Claude Shannon showed that for Bernoulli trials, that is win or lose 1 with probabilities p and q for $p+q=1$, the long run growth rate, namely

$$G = \lim_{t\to\infty} \log \left\{ \frac{w_t}{w_0} \right\}^{1/t}$$

where t is discrete time and w_t is the wealth at time t with W_0 the initial wealth is equivalent to $\max E \log w$.

Since $W_t = (1+f)^M(1-f)^{t-M}$ is the wealth after t discrete periods, f is the fraction of wealth bet in each period and M of the t trials are winners.

Then, substituting W_t into G gives

$$G = \lim_{t\to\infty} \left\{ \frac{M}{t} \log(1+f) + \frac{t-M}{t} \log(1-f) \right\} + p\log(1+f) + q\log(1-f)$$

and, by the strong law of large numbers

$$G = E \log w$$

Thus the criterion of maximizing the long run exponential rate of asset growth is equivalent to maximizing the one period expected logarithm of wealth. So an optimal policy is myopic. Since

$$\max G(f) = p\log(1+f) + q\log(1-f)$$

the optimal fraction to bet is the edge $f^* = p - q$. These bets can be large. For example, if $p = .99$ and $q = .01$, then $f^* = 0.98$, that is 98% of one's fortune. Some real examples of very large and very small bets appear later in the paper. If the payoff odds are +B for a win and -1 for a loss, then the edge is Bp-q and

$$f^* = \frac{Bp-1}{B} = \frac{edge}{odds}$$

So the size of the investments depend more on the odds, that is the probability of losing, rather than the mean advantage. Kelly bets are usually large and the more attractive the wager, the larger the bet. For example, in the trading on the January turn-of-the-year effect with a huge advantage, full Kelly bets approach 75% of initial wealth. Hence, Clark and Ziemba (1987) suggested a 25% fractional Kelly strategy as discussed in Section 5, for their trades.

Latané (1959) introduced log utility as an investment criterion to the finance world independent of Kelly's work. Focusing, like Kelly, on simple intuitive versions of the expected log criteria, he suggested that it had superior long run properties. Chopra and Ziemba (1993) have shown that in standard mean-variance investment models, accurate mean estimates are about twenty times more important than covariance estimates and ten times variances estimates in certainty equivalent value. But this is risk aversion dependent with the importance of the errors becoming larger for low risk aversion utility functions. Hence, for $\log w$ with minimal risk aversion, the impact of these errors is of the order 100:3:1. So $E \log$ bettors can easily over bet. Poundstone (2005) has popularized the term *Fortune's Formula* through his best selling trade book. This book is a good read but does not get some technical details right. As a consequence, investment newsletters such as *Morningstar* and *Motley Fool* show Kelly criterion calculations but only use the edge/odds formula which is only valid for single investments. The correct calculation as discussed in Section 5 requires the solution of an n-variable stochastic nonlinear program.

Leo Breiman (1961), following his earlier intuitive paper Breiman (1960), established the basic mathematical properties of the expected log criterion in a rigorous fashion. He proved three basic asymptotic results in a general discrete time setting with intertemporally independent assets.

Suppose in each period, N, there are K investment opportunities with returns per unit invested $X_{N_1}, \ldots, X_{N_K}$. Let $\Lambda = (\Lambda_1, \ldots, \Lambda_K)$ be the fraction of wealth invested in each asset. The wealth at the end of period N is

$$w_N = \left(\sum_{i=1}^{K} \Lambda_i X_{Ni} \right) w_{N-1}.$$

Property 1 In each time period, two portfolio managers have the same family of investment opportunities, X, and one uses a Λ^* which maximizes $E \log w_N$

whereas the other uses an *essentially different* strategy, Λ, so they differ infinitely often, that is,

$$E \log w_N \Lambda^* - E \log w_N(\Lambda) \to \infty.$$

Then

$$\lim_{N\to\infty} \frac{w_N(\Lambda^*)}{w_N(\Lambda)} \to \infty.$$

So the wealth exceeds that with any other strategy by more and more as the horizon becomes more distant.

This generalizes the Kelly Bernoulli trial setting to intertemporally independent and stationary returns.

Property 2 The expected time to reach a preassigned goal A is, asymptotically least as A increases with a strategy maximizing $E \log w_N$.

Property 3 Assuming a fixed opportunity set, there is a fixed fraction strategy that maximizes $E \log w_N$, which is independent of N.

Consider the example described in Table 1. There are five possible investments and if we bet on any of them, we always have a 14% advantage. The difference between them is that some have a higher chance of winning and, for some, this chance is smaller. For the latter, we receive higher odds if we win than for the former. But we always receive 1.14 for each 1 bet on average. Hence we have a favorable game. The optimal expected log utility bet with one asset (here we either win or lose the bet) equals the edge divided by the odds.[4] So for the 1-1 odds bet, the wager is 14% of ones fortune and at 5-1 its only 2.8%. We bet more when the chance that we will lose our bet is smaller. Also we bet more when the edge is higher. The bet is linear in the edge so doubling the edge doubles the optimal bet. However, the bet is non-linear in the chance of losing our money, which is reinvested so the size of the wager depends more on the chance of losing and less on the edge.

Table 1 The Investments

Probability of Winning	Odds	Probability of Being Chosen in the Simulation at at Each Decision Point	Optimal Kelly Bets Fraction of Current Wealth
0.57	1-1	0.1	0.14
0.38	2-1	0.3	0.07
0.285	3-1	0.3	0.047
0.228	4-1	0.2	0.035
0.19	5-1	0.1	0.028

Source: Ziemba and Hausch (1986)

The simulation results shown in Table 2 assume that the investor's initial wealth is 1000 and that there are 700 investment decision points. The simulation was repeated 1000 times. The numbers in Table 2 are the number of times out of the possible 1000 that each particular goal was reached. The first line is with log or Kelly betting, The second line is half Kelly betting. That is you compute the optimal Kelly wager but then blend it 50-50 with cash. For lognormal investments α-fractional Kelly wagers are equivalent to the optimal bet obtained from using the concave risk averse, negative power utility function, $-w^{-\beta}$, where $\alpha = \frac{1}{1-\beta}$. For non lognormal assets this is an approximation (see MacLean, Ziemba and Li, 2005 and Thorp, 2010, 2011). For half Kelly ($\alpha = 1/2$), $\beta = -1$ and the utility function is $w^{-1} = \frac{1}{w}$. Here the marginal increase in wealth drops off as w^2, which is more conservative than log's w. Log utility is the case $\beta \to -\infty, \quad \alpha = 1$ and cash is $\beta \to -\infty, \quad \alpha = 0$.

Table 2 Statistics of the Simulation

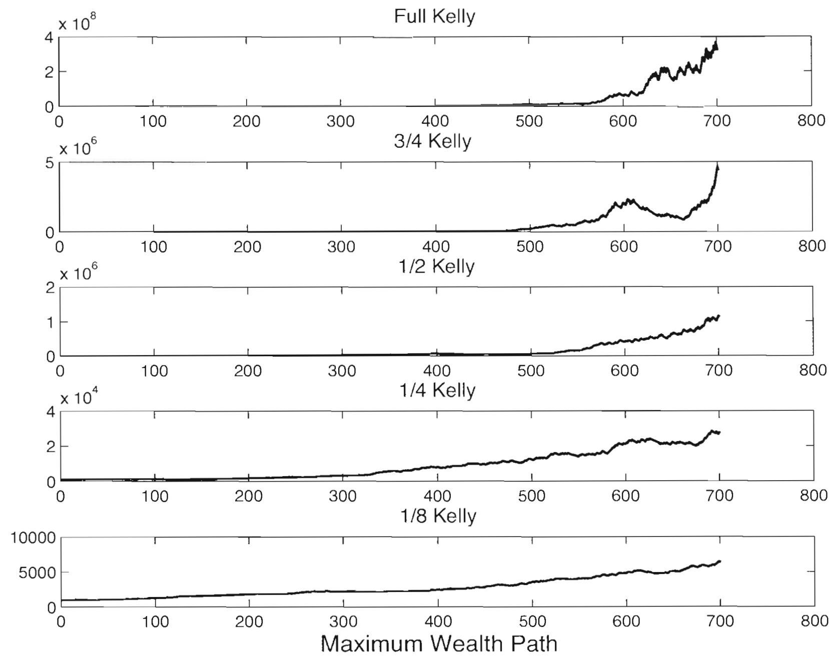
Final Wealth Strategy	Min	Max	Mean	Median	Number of times the final wealth out of 1000 trials was >500	>1000	>10,000	>50,000	>100,000
Kelly	18	483,883	48,135	17,269	916	870	598	302	166
Half Kelly	145	111,770	13,069	8,043	990	954	480	30	1

Source: Ziemba and Hausch (1986)

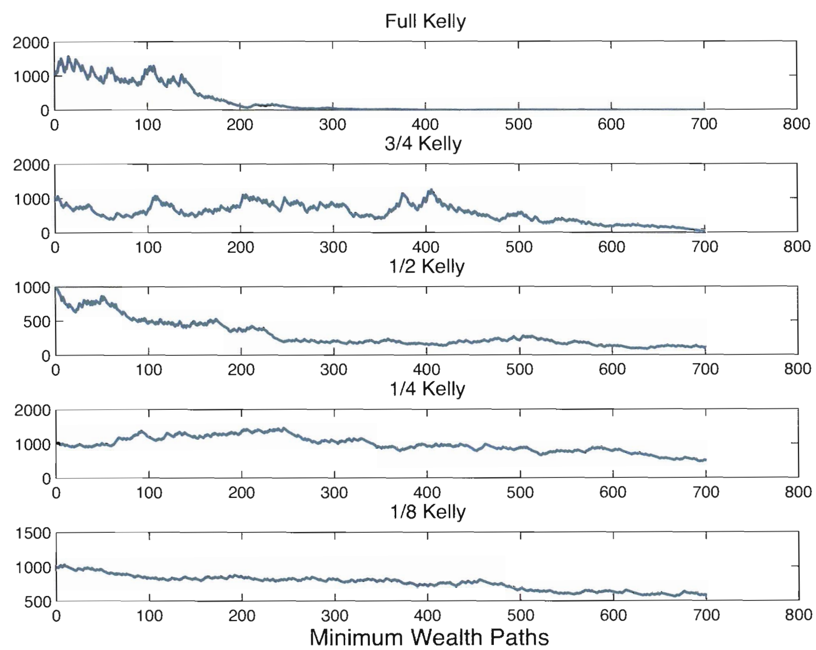
A major advantage of full Kelly log utility betting is the 166 in the last column. In fully 16.6% of the 1000 cases in the simulation, the final wealth is more than 100 times as much as the initial wealth. Also in 302 cases, the final wealth is more than 50 times the initial wealth. This huge growth in final wealth for log is not shared by the half Kelly strategies, which have only 1 and 30, respectively, for their 50 and 100 times growth levels. Indeed, log provides an enormous growth rate but at a price, namely a very high volatility of wealth levels. That is, the final wealth is very likely to be higher than with other strategies, but the wealth path will be very very bumpy. The maximum, mean, and median statistics in Table 2 illustrate the enormous gains that log utility strategies usually provide.

Let us now focus on bad outcomes. The first column provides the following remarkable fact: one can make 700 independent bets of which the chance of winning each one is at least 19% and usually is much more, having a 14% advantage on each bet and still turn 1000 into 18, a loss of more than 98%. Even with half Kelly, the minimum return over the 1000 simulations was 145, a loss of 85.5%. Half Kelly has a 99% chance of not losing more than half the wealth versus only 91.6% for Kelly. The chance of not being ahead is almost three times as large for full versus half Kelly. Hence to protect ourselves from bad scenario outcomes, we need to lower our bets and diversify across many independent investments.

Figures 1(a) and 1(b) show the highest and lowest final wealth trajectories for full, $\frac{3}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ and $\frac{1}{8}$ Kelly strategies for this example. Most of the gain is in the final 100 of the 700 decision points. Even with these maximum graphs, there is much volatility in the final wealth with the amount of volatility generally higher with higher Kelly fractions. Indeed with $\frac{3}{4}$ Kelly, there were losses from about decision points 610 to 670.



(a) Highest



(b) Lowest

Figure 1 Final Wealth Trajectories: Ziemba-Hausch (1986) Model. Source: MacLean, Thorp, Zhao and Ziemba (2011)

The final wealth levels are much higher on average, the higher the Kelly fraction. As you approach full Kelly, the typical final wealth escalates dramatically. This is shown also in the maximum wealth levels in Table 3.

Table 3 Final Wealth Statistics by Kelly Fraction: Ziemba-Hausch (1986) Model

	Kelly Fraction				
Statistic	1.0k	0.75k	0.50k	0.25k	0.125k
Max	318854673	4370619	1117424	27067	6330
Mean	524195	70991	19005	4339	2072
Min	4	56	111	513	587
St. Dev.	8033178	242313	41289	2951	650
Skewness	35	11	13	2	1
Kurtosis	1299	155	278	9	2
$> 5 \times 10$	1981	2000	2000	2000	2000
10^2	1965	1996	2000	2000	2000
$> 5 \times 10^2$	1854	1936	1985	2000	2000
$> 10^3$	1752	1855	1930	1957	1978
$> 10^4$	1175	1185	912	104	0
$> 10^5$	479	284	50	0	0
$> 10^6$	111	17	1	0	0

There is a chance of loss (final wealth is less than the initial $1000) in all cases, even with 700 independent bets each with an edge of 14%. The size of the losses can be large as shown in the > 50, > 100, and > 500 and columns of Table 3. Figure 1(b) shows these minimum wealth paths.

If capital is infinitely divisible and there is no leveraging then the Kelly bettor cannot go bankrupt since one never bets everything (unless the probability of losing anything at all is zero and the probability of winning is positive). If capital is discrete, then presumably Kelly bets are rounded down to avoid overbetting, in which case, at least one unit is never bet. Hence, the worst case with Kelly is to be reduced to one unit, at which point betting stops. Since fractional Kelly bets less, the result follows for all such strategies. For levered wagers, that is, betting more than one's wealth with borrowed money, the investor can lose much more than their initial wealth and become bankrupt. See MacLean, Thorp, Zhao and Ziemba (2011).

3. The objections of Professor Paul A. Samuelson to Kelly capital growth investing and a response

The great economist Paul A. Samuelson was a long time critic of the Kelly strategy which maximizes the expected logarithm of final wealth, see, for example, Samuelson (1963, 1969, 1979) and Merton and Samuelson (1974). His objections to Elog investing were:

Objection 1. It does not maximize expected utility for utility functions other than log.

That is correct and I agree, Elog maximizes only log utility. Indeed no utility function can maximize expected utility for other utility functions. Thorp and Whitley (1972) show that different concave utility functions do indeed produce different optimal decisions.

Samuelson seemed to imply that Kelly proponents thought that the Kelly strategy maximizes for other utility functions but this was neither argued nor implied. It is true that the expected value of wealth is higher with the Kelly strategy but bad outcomes are very possible.

Mike Stutzer pointed out to me

"I think Samuelson was referring to a claim or conjecture (likely a footnote) in the Latané article. So the real problem is that some of Samuelson's critiques were misused by later readers to falsely tar other non-problematic growth optimal results in that and other papers."

In his correspondence with me (private correspondence, 2006, 2007, 2008), Samuelson seemed to imply that half Kelly (assuming lognormal asset distributions) or $u(w) = -\frac{1}{w}$ explains the data better. I agree that in practice, half Kelly is a toned down version of full Kelly that provides a lot more security to compensate for its loss in long term growth.

Objection 2 Despite the fact that in the long run, E log investors asymptotically dominate all essentially different utility functions it does not follow that an Elog investor will have good performance. Indeed, no matter how long the investment sequence is and how favorable the investment situations are, it is possible to lose a lot of money.

In his letters to me, he formulated this as

Theorem (Samuelson): In no run, however long, does Kelly's Rule effectuate a *dominating* retirement next egg.

I agree completely and illustrate this with the simple simulated example above in which with a 14% advantage in each period and many independent wagers over 700 periods it is possible with no leveraging to lose 98% of one's initial wealth. Other examples are in MacLean, Thorp, Zhao and Ziemba (2011).

There are at least three approaches for dynamic investment that one could consider as stated, for example, by Luenberger (1993):

1. $E \log w$
2. $\max Eu(w)$ for u concave for $u \neq \log$
3. $\max E \sum_{t=1}^{T} \beta^t u(c_t)$, where c_t is consumption drawn out of wealth in period t and $0 < \beta < 1$ is a discount factor (Samuelson, 1969)

Many great investors use full Kelly $E \log w$ and fractional Kelly $\alpha w^\alpha, \ \ \alpha < 0$ successfully; some examples follow in Section 5. But this does not mean that they will have optimal policies for the 2nd and 3rd approaches or will always win in finite time. Indeed, $E \log$ betting can yield substantial losses even without leveraging and with leveraging the losses can be many times the initial wealth and leads to bankruptcy.

The Kelly strategy maximizes the asymptotic long run growth of the investor's wealth, and I agree that this is a Breiman (1961) property.

Objection 3. The Kelly strategy always leads to more wealth than any essentially different strategy. I know from the simulations that this is not true since it is possible to have a large number of very good investments and still lose most of one's fortune even without leveraging. So this could not be claimed by anyone and Samuelson's theorem above demonstrates this.

Luenberger (1993) investigated investors who are only interested in tail returns in the iid case. In response to Samuelson regarding the long run Kelly behavior, Luenberger shows when $E\log W_t$ is optimal (simple utility functions) and when a $(E\log W_t), var\log w_t)$ tradeoff is optimal (compound utility functions)

Previously Samuelson (1970) showed that there was an accurate log mean-log variance approximation to concave terminal utility if uncertainty is small and the distributions are compact; see also Ohlson (1975) on this power expansion approximation. This is because a two term power expansion to E log will be accurate with compact distributions.

Objection 4. A long run technical criticism of Samuelson articulated in Merton and Samuelson (1974) while pointing out math errors in Hakansson (1971a) is that $\lim_{t\to\infty} E(u(w_t))$ is not an expected utility. So log mean criteria and log mean-log variance criteria are not consistent with expected utility.

To respond to point 4, I refer to Luenberger (1993).

Luenberger uses compound utility functions that subtract $m = E\log x_1$ and deal with the $m = 0$ case, and $m < 0, m > 0$ are dealt with using simple utility functions.

$$u(w) = \overline{\lim_{t\to\infty}}\psi(\log w_t - tm, m, t)$$

a.s. m=0

An investor with $m = 0$ prefers increased variance (up gains for the investor). The compound utility function is equivalent to a function of the expected logarithm and variance of the logarithm of wealth. A tail utility function involving the limits of total return must be equivalent to a log mean-variance criterion. Thus there is an efficient frontier and the investor chooses a point on this frontier

Luenberger establishes preferences on infinite sequences of wealth rather than wealth at a fixed (but later taken to the limit) terminal time.

Simple $\to$ tail utility function: $u(w) = w(\overline{w})$ if w, and $\overline{w}$ differ in at most a finite number of elements.

$$u(w) = \overline{\lim_{t\to\infty}} \overline{\rho}(w_t, t)$$

$$\overline{\rho}(w_t, t) = \rho(\log w_t, t)$$

$$u(w) = \overline{\lim_{t\to\infty}} \rho(\log w_t, t)$$

$\overline{\rho}$ is continuous and increasing in w_t for each t.

Tail events have probability of either zero or one. The criterion is not expected value $E\log w_1$ but is actually an "almost sure criterion". I accept this and conclude that

- $E\log w$ is one approach
- $\max Eu(w)$ for u concave is another approach, for $u \neq \log$
- $\max E\sum_{t=1}^{T}\beta^t u(c_t)$, where c_t is consumption drawn out of wealth in period t and $0 < \beta < 1$ is a discount factor (Samuelson, 1969) is yet another approach

This is, of course, a theoretical discussion of the long run properties of Elog investing which while interesting has little to do with the various applications in long but finite time discussed in Section 5. But there is a bit more. Markowitz (1976, 2006) adds to this in the simple utility function case. Assuming iid investments in discrete time, as Luenberger did, he shows that

> "with probability one, there comes a time such that forever after the wealth of the investor who rebalances to portfolio P exceeds that of the investor who rebalances to portfolio Q, surely one can say P does better than Q in *the long run*"

where P maximizes $E\log(1 + r_t^p)$, r_t^p is the return on the portfolio during time $t-1$ and t, and Q is another iid portfolio, possibly correlated with P, where $\mu_p = E\log(1 + r_t^P) > E\log(1 + r_t^Q)$.

This, of course, as discussed above

> "does not necessarily imply that any particular investor with a finite life and imminent consumption needs, should invest in P rather than Q. But it seems an unobjectionable use of language to summarize relationship 17.10 by saying that portfolio P does better than portfolio Q in the long run (Markowitz (2006, p 256)"

where 17.10 says that with probability 1, there is a time T_0 such that w_T^P exceeds w_T^Q ever after, that is

$$\exists T_0 \forall T > T_0, \quad w_P^T > w_Q^T.$$

Markowitz (1976) relaxes the iid assumption. See also Algoet and Cover (1988) and Thorp (2011).

So what do we conclude on this Samuelson objection number 4? We can just dismiss it based on the Luenberger and Markowitz results as its maximizing the wrong quantity? Or we can say, yes, he is right but does that matter as we have other limiting results supporting the E log case?

The essence of one of Samuelson's objections to the MaxE log rule,as articulated by Markowitz (2006) is that: if the investor seeks to maximize the expected value of a certain kind of function of final wealth, for a long game of fixed length, then maximizing E log is not the optimal strategy.

What Samuelson has in mind here is $u(w) = \alpha w^{\alpha}$, namely,the negative and power utility functions of which log, namely $\alpha \to 0$ is the limiting member. Of course, we know from above that $\alpha > 0$, positive power is definitely over betting so let's assume that $\alpha < 0$, namely the utility function not dominated by having less growth and more risk.

This argument rests on the Samuelson (1969) and Mossin (1968) results for power utility that show myopic behavior assuming independent period by period assets where the investor rebalances to the same portfolio in each period. So the optimal strategy is this portfolio which is not the E log portfolio. When $u(w) = \log u$, the $\alpha \to 0$ case, then there is a myopic policy even for dependent assets, see Hakansson (1971b).

> "the wealth of the investor who rebalances to portfolio P exceeds that of the investor who rebalances to portfolio Q, surely one can say that P does better than Q in the *long run*."

Then, as Markowitz (2006, p 260) concludes:

> "Indeed, if we let the length of the game increase, the utility supplied by the max E log strategy does not even approach that supplied by the optimal strategy. This assumes that utility of final wealth remains the same as game length varies. On the other hand, if we assume that it is the utility of rate-of-growth-achieved, rather than utility of final wealth, that remains the same as length of game varies, then the E log rule is asymptotically optimal."

And Markowitz has a nice way of reminding us that betting more than full Kelly is dominated, as Figure 9 shows graphically.

> "Perhaps this is a sufficient caveat to attach to the observation that the cautious investor should not select a mean-variance efficient portfolio higher on the frontier than the point which approximately maximizes expected log(1+return): for a point higher on the frontier subjects the investor to greater volatility in the short run and, almost surely, no greater rate-of-growth in the long run."

4. The Samuelson investors

Samuelson postulated three people with concave risk averse utility functions. They are

- Tom with $u(w) = w^{1/2}$, a positive power maximizer.
- Dick with $\log w$, a geometric mean Kelly criterion optimizer and
- Harriet with $u(w) = -1/w$, who is a half Kelly optimizer (exactly if assets are lognormal and approximately otherwise).

Harriet has a limited degree of risk aversion and according to Samuelson fits well with lots of empirical Wall Street equity premium data. I believe that Tom is over betting, see Figure 9, and will eventually go bankrupt. I add two more investors to complete the spectrum. They include one very conservative investor and one very risky investor.

- Ida who is approaching infinitely risk averse has $u(w) = -\frac{1}{Nw}$, with N$\rightarrow\infty$
- Victor is on the other extreme, infinitely risky with linear utility

Ida resembles the famous Ida May Fuller of Ludlow, Vermont who was the first US social security recipient receiving check number 00-000-001 on January 31, 1940. Ida paid \$24.75 into the social security fund, then lived to be 100 and collected nearly 1000 times her investment, namely \$22,889, before she died at age 100; see Schwartz and Ziemba (2007) who discuss social security.

On the other extreme is Victor, who is inspired by hedge fund trader Victor Niederhoffer. who historically has alternated between huge returns and disasters with a much greater than full Kelly betting strategy. This is known, see MacLean et al (1992), to be over betting and dominated in a mean risk sense where risk is the probability of not achieving a high goal before falling to a low wealth level, see Figure 9 for example. An account of some of his trading up to mid 2007 is in Ziemba and Ziemba (2013). In the ensuing years, more ups and downs have occurred. For our Victor, I assume that the trader is at the absolute limit of 0 absolute and relative risk aversion. Table 4 describes the absolute and relative risk aversion properties of these five investors, all of which had constant relative risk aversion (CRRA).

Table 4 The five investors in the Samuelson Experiment

The Investors

	Victor	Tom	Dick	Harriet	Ida
	w linear	$w^{\frac{1}{2}}$ positive power	$\log w$ geometric mean optimizer	$-\frac{1}{w}$ half Kelly	$-\frac{1}{NW}$, $N \rightarrow \infty$ infinitely risk averse
Absolute $R_A - \frac{u'}{u'(w)}$	0	$\frac{1}{2w}$	$\frac{1}{2}$	$\frac{2}{w}$	∞
Relative $R_A - \frac{wu''(w)}{u'(w)}$	0	$\frac{1}{2}$	1	2	∞

Consider the investment where cash returns zero and stock returns with equal 1/2 probability either \$4 or \$0.25 for each \$1 bet in each period.

Test 1: If you must put 100% of your nest egg in only one option, which do you pick?

$$\max_{x=0,1}\left[\frac{1}{2}xu(4) + (1-x)u\left(\frac{1}{4}\right)\right]$$

Tom and Victor choose all stock, $x^* = 1$

Dick is indifferent

Harriet and Ida choose all cash x^*=0 stock

Given a horizon of $n > 1$ periods until the final date of your retirement. All three say *no change.*

Test 2: The blending portfolio optimization case. Using

$$\max_x \left[\frac{1}{2} u(4x + 1 - x) + \frac{1}{2} u \left(\frac{1}{4} \right) x + 1 - x \right]$$

gives

Harriet	$x^* = \frac{2}{9}$ stock, $1 - x^* = \frac{7}{9}$ cash
Tom	$x^* = 1$ stock, cash=0
Dick	$x^* = \frac{1}{2}$ stock, $1 - x^* = \frac{1}{2}$ cash

Certainty Equivalents

$$u(CE) = \frac{1}{2} u(4) + \frac{1}{2} u \left(\frac{1}{4} \right)$$

$$CE = u^{-1} \left[\frac{1}{2} u(4) + \frac{1}{2} u \left(\frac{1}{4} \right) \right] \equiv E(4, 1)$$

The certainty equivalents are:

Tom	$CE = KM = \left[\frac{1}{2} \sqrt{4} + \frac{1}{2} \sqrt{1/4} \right]^2 = 1 + \frac{9}{16}$
Dick	$CE = GM = \sqrt{4 * \frac{1}{4}} = 1$
Harriet	$CE = HM = \left[\frac{1}{2} \left(\frac{1}{4} \right) + \frac{1}{2} (4) \right]^{-1} = 1 - \frac{9}{17}$
Victor	bets 2 by borrowing 1 at zero interest so $CE = AM = 2 + \frac{1}{8}$.
Ida	bets zero so $CE = IM = 0$.

Here, IM, HM, GM, KM and AM are infinitely risk averse mean, harmonic mean, geometric mean, root-squared mean, and the arithmetic mean, respectively.

The $\sqrt{w}$ formula of 1728 Kramer defines for **Tom**'s CE a "root-squared mean" KM.

$$0 = IM < HM < GM < KM \leqslant AM = \frac{1}{2}(4) + \frac{1}{2} \left(\frac{1}{4} \right) = 2\frac{1}{8}$$

For the double full Kelly betting case, we know that the growth rate is zero plus the risk free rate which is assumed to be zero (see MacLean, Ziemba and Blazenko, 1992, Thorp, 2011, and Ziemba, 2003). Dick who bet $x^* = \frac{1}{2}$ stock maximizing Elog now bets $x^* = 1$, all stock. Victor, with his linear utility $u(w) = w$ bets $x^* = 1$, all stock, the same as Tom, but in more complex multi-asset cases he may bet differently.

- The 4^{th} investor **Victor**, with a linear utility w, bets $x^* = 1$ all stock, the same as **Tom** but in more complex multi-asset cases, **Victor** will bet even more than **Tom** and have a negative growth rate and go bankrupt faster than **Tom**

- The 5^{th} investor, who is infinitely risk averse, **Harriet**'s sister **Ida**, bets nothing and lives off her cash until she dies

Dick is a pushy guy. What if he persuades Harriet to replace her $x^* = \frac{2}{9}$ with his $x^* = \frac{1}{2}$? If she agrees to shoot herself in her own foot, the loss in her CE dollars below her best CE* dollars is equivalent to her having agreed to throw away a definable percentage of her initial wealth. What is left,invested her *proper* way, will fall short of what she could have received by "being true to herself" by measurable deadweight loss. Persuasive Dick could also do a measurable monetary harm to Tom if Tom gives up $x^* = 1$ and goes along with Dick's $x^* = \frac{1}{2}$

Can these one-period harms erode away after Tom and Harriet come to shoot themselves in their respective feet two times, three times ... N= 100^{10} times? No. No such Limit Theorem is valid. For N large, N$\gg 1, x^* = \frac{2}{9}$ and $x^* = \frac{1}{2}$ and $x^* = 1$ each produce *on retirement date* three different wide-spread Log Normal limit distributions. Tom's Log Normal has the largest absolute arithmetic mean dollars. Harriet's has the least absolute arithmetic mean dollars. However, at Harriet's request we calculate the three H.M's. Hers is the largest!

In a vanity duel between any two neighbors, where what's to be maximized is **A**'s probability of being ahead of **B** when they both retire at the same time and start to invest at the same time, Dick types will beat out both Harriet types and Tom types. And for Methusala-ish neighbors, Dick's probability edge will go to 1 (*almost*) as N$\to\infty$ These are then the Breiman Theorems, discussed in Section 2, in the limit.

In Samuelson's words

"the MacLean and Ziemba (see MacLean, Ziemba and Blazenko, 1992) probability, $t_{w_{large}} < t_{w_{small}}$ versus the growth rate do this, but in finite, calculable time, as shown in Figure 9."

My response is that this is correct and I have no disagreement with Professor Samuelson. As the simulation in Section 2 shows, the full Kelly strategy gives very high final wealth most of the time. But it is possible to have low final wealth with no leveraging (-98%) and many times w_0 lost with leveraging as examples in MacLean, Thorp, Zhao and Ziemba (2011) show.

For two outcome stocks, we can solve for x^* as the root of the equation

$$\frac{d}{dx}\left[\frac{1}{2}U(3x+1)+\frac{1}{2}U\left(1-\frac{3}{4}x\right)\right]=0.$$

x^* can be found for these three neighbors by solving *linear* equation. That is only because all three utilities have Constant *Relative* Risk Aversion,

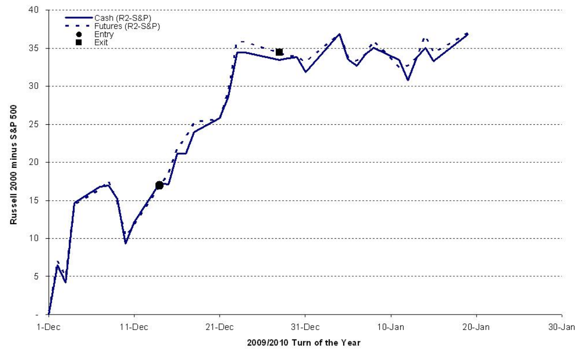
In general, to obtain the max E log portfolio, one must solve a constrained non-linear one-period stochastic programming model to calculate the optimal portfolio weights like the model in Section 5 where the effect of our bets on the odds is in the model.

5. Selected applications and endorsements

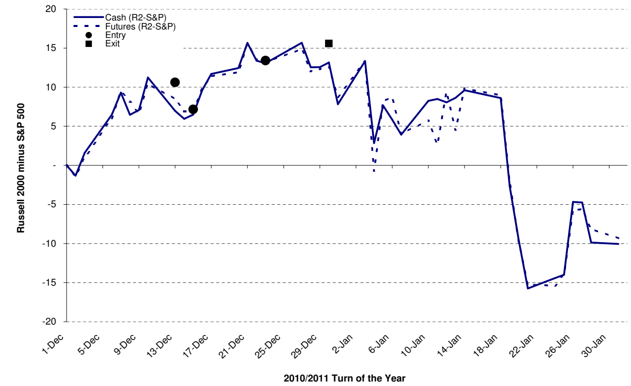
The purpose of the applications section of *Operations Research* is to focus on novel applications that were actually used and use operations research techniques in an important way such as non-linear programming for the applications in this paper. I start with two applications of mine. The first is trading the turn-of-the-year effect using futures in the stock market. The first paper on that was published in *Operations Research* by Clark and Ziemba (1987) and because of the huge advantage at the time suggested a large full Kelly wager approaching 75% of initial wealth. But there are risks, transaction costs, margin requirements, and other uncertainties which suggested a lower wager of 25% Kelly. They traded successfully for the years 1982/83 to 1986/87 - the first four years of futures in the TOY. Futures in the S&P500 having just begun. I then continued this trade of long small cap minus short large cap measured by the Value Line small cap index and the large cap S&P500 index for ten more years with gains each year. The plots and tables describing these trades for all 14 years from 1982/83 to 1995/ are in Ziemba (2012a). At the end, I was 7% of the Value Line futures market volume and I stopped trading this for two reasons. First, the Value Line contracts volume was becoming less and less which suggested trouble exiting positions. Secondly, I taught the trade as a consultant to Peter Muller's group at Morgan Stanley - a group that later became famous gaining $5 billion in profits, the most of any group at that firm. I knew that competing with them was too dangerous for me since their wealth was so much greater than mine. I continued to write papers on this but did not trade it until the 2009/10 TOY. Consulting to noted futures trader Blair Hull got me back in this market using the Russell 2000 as the small cap futures index. This year and the next three years going into 2013 produced profits.

Figure 2 shows updates of graphs betting with the author's money successfully in December 2009, 2010 and 2011, where the dots are the entries and the squares are the exits. These trades were made in in my investment account, see Figure 3. The size of the position is 15% fractional Kelly. The profit on these trades can be seen in the three December periods in the graph. The January effect still exists in the futures markets but now is totally in December contrary to the statements in most finance books such as Malkiel (2011). The 2012/13 TOY was also successful and one of the greatest ever. That graph is available from me. The fractional Kelly wager suggested in the much more dangerous market situation now is lower than in the Clark and Ziemba. Programmed trading, high frequency trading and other factors add to the complexity so risk must be lowered as one sees in the 2011/12 graph. In 2010/11, the exit was intraday at prices better than those at the close on the black line.

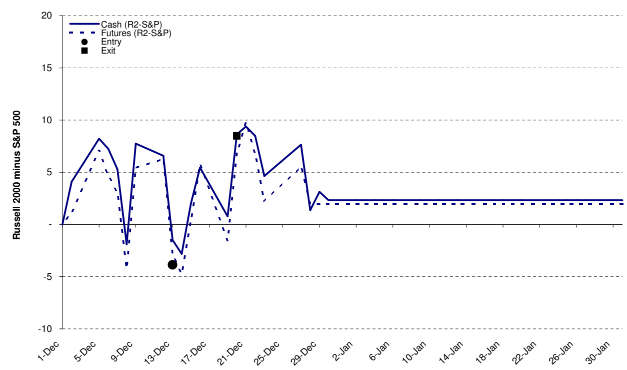
These turn of the year bets are large, however, the Kelly wagers can be very small even with a large edge if the probability of winning is low. An example is betting on unpopular numbers in



(a) 2009-2010



(b) 2010-2011



(c) 2011-2012

Figure 2 Russell2000 - S&P500 spread with our entries (dots) and exits (squares). The cash market spread is the black line and the dotted line is the futures spread, the one actually traded

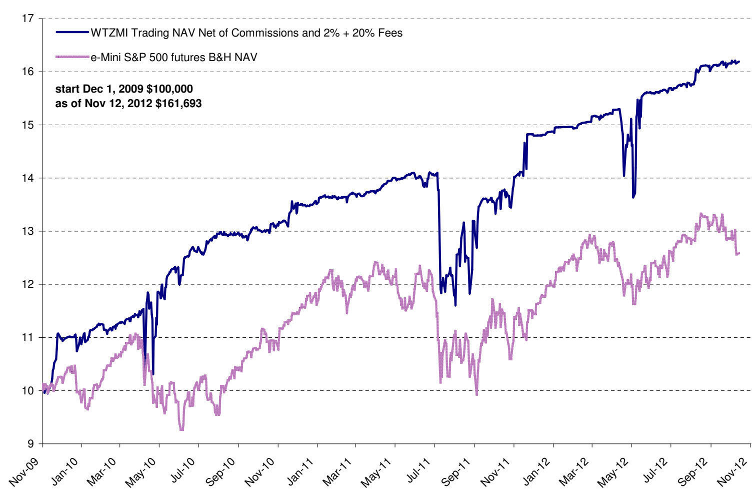


Figure 3 Private Futures Account at Interactive Brokers in Canada, December 1, 2009 to November 12, 2012

lotto games. MacLean, Ziemba and Blazenko (1992) show that with an 82.7% edge, the full Kelly wager is only 65 $1 tickets per $10 million of ones fortune. This is because most of the edge is in very low probability of winning the Jackpot and second prize. While there is a substantial edge, the chance of winning a substantial amount is small and indeed to have a high probability of a large gain requires a very long time, in the millions of years.

The second application of mine is in horse racing. First in a simulation of betting on the Kentucky Derby from 1934 to 2005, the full Kelly log bettor has the most total wealth at the horizon but has the most bumpy ride: \$2500 becomes \$16,861; see Figure 4. The half Kelly bettor ends up with much less, \$6945 but has a much smoother ride. The system is based on the place and show betting system described in Hausch, Ziemba and Rubinstein (1981), using a dosage breeding filter to eliminate horses that do not have the stamina to win the $1\frac{1}{4}$ mile derby on the first Saturday in May of their three year old career. A comparison with random betting proxied by betting on the favorite in the race, shows how difficult it is to win at horseracing with the 16% track take plus breakage (rounding payoffs down to the nearest 20 cents per \$2 bet) at Churchill Downs. Betting on the favorite turns \$2500 into \$480. Random betting has even lower final wealth at the horizon since favorites are underbet.

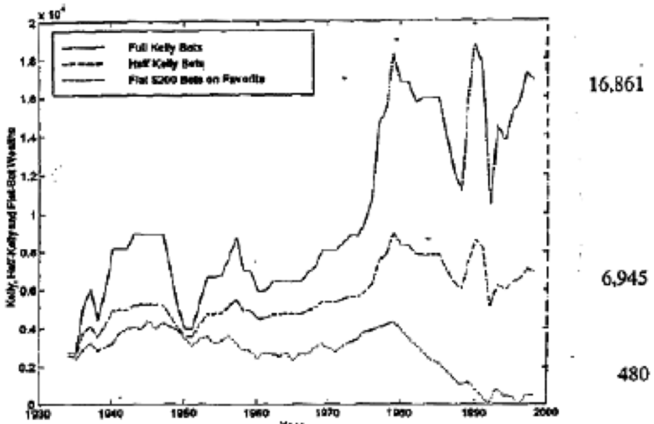


Figure 4 Wealth history of some Kentucky Derby bets, 1934-2005, Kelly, half Kelly and betting on the favorite, using a dosage filter.

Professional racetrack betting in the modern era is different from standard at the track betting through the mutuel pools. The bets are made in an office with computer and TV screens. Rebate is given to large bettors so their effective track take is less. The actual track take is shared by the track which takes less than usual, the rebator and the bettor. Bettors outside the US can also make short as well as long bets on Betfair in London for example. There are about eight worldwide syndicates or teams making successful racetrack bets all who are Kelly type wagering to size their bets. I have consulted for two of them plus do my own betting.

In general, Kelly bets are large and risky short term and proceed by computing an optimization like the following, that considers **exact transaction costs**, that is the effect of the bets on the prices (odds).

The model to maximize the expected utility capital growth model for racetrack place and show bets is

$$\max_{p_i, s_i} \sum_{i=1}^{n} \sum_{\substack{j=i \\ j\neq i}}^{n} \sum_{\substack{k=i \\ k\neq i,j}}^{n} \frac{q_i q_j q_k}{(1-q_i)(1-q_i-q_j)} \log \left[\begin{array}{c} \frac{Q\left(P+\sum_{l=1}^{n} p_l\right)-(p_i+p_j+P_{ij})}{2} \\ \times \left[\frac{p_i}{p_i+P_i}+\frac{p_j}{p_j+P_j}\right] \\ +\frac{Q\left(S+\sum_{l=1}^{n} s_l\right)-(s_i+s_j+s_k+S_{ijk})}{3} \\ \times \left[\frac{s_i}{s_i+S_i}+\frac{s_j}{s_j+S_j}+\frac{s_k}{s_k+S_k}\right] \\ +w_0 - \sum_{\substack{l=i \\ l\neq i,j,k}}^{n} s_l - \sum_{\substack{l=i \\ l\neq i,j}}^{n} p_l \end{array} \right]$$

$$\text{s.t.} \sum_{l=1}^{n} (p_l + s_l) \leqslant w_0, \quad p_l \geqslant 0, \quad s_l \geqslant 0, \quad l = 1, \ldots, n,$$

If rebate is available it is then added to final wealth inside the large brackets by adding the rebate rate times all the bets,winners and losers, namely

$$r(\sum_{l=1}^{n} (p_l + s_l)$$

where

- The effect of transactions costs which is called slippage in commodity trading is illustrated with the above place/show horseracing optimization formulation; see Hausch, Ziemba and Rubinstein (1981).
- Here q_i is the probability that i wins, and the Harville probability of an ij finish is $\frac{q_i q_j}{1-q_i}$, etc.
- That is $q_j/1-q_j$ is the probability that j wins a race that does not contain i, that is, comes second to i.
- Q, the track payback, is about 0.82 (but is about 0.90 with professional rebates).
- The players' bets are to place p_j and show s_k for each of the about ten horses in the race out of the players' wealth w_0. The bets by the crowd are P_i with $\sum_{i=1}^{n} P_i = P$ and S_k with $\sum_{k=1}^{n} S_k = S$.
- The payoffs are computed so that for place, the first two finishers, say i and j, in either order share the net pool profits once each P_i and p_i bets cost of say \$1 is returned.
- The show payoffs are computed similarly.

In practice, given limited time to make bets, one uses regression approximations to the expected value and optimal wager that are functions of only four numbers, namely, the totals to win and place for the horse in question and the totals bet.

The expected value approximations are (using 1000s) of sample calculations of the nonlinear programming model:

$$\text{Ex Place}_i = 0.319 + 0.559 \left(\frac{w_i/w}{p_i/p} \right)$$
$$\text{Ex Show}_i = 0.543 + 0.369 \left(\frac{w_i/w}{s_i/s} \right).$$

Then, this system gets implemented easily looking at only these four numbers for place or show bets. There are corresponding Kelly bet approximations that can be used in a hand held calculator. Other Kelly bets on different racetrack wagers with low probability high payoff or high probability low payoff bets are discussed in Ziemba (2012b).

An application of real money bet with this system in 2004 is shown in Figure 5. The initial wealth was w_0=US$5,000. At each wager opportunity there is either no bet or a full Kelly bet using the model above with rebate collected on winning and losing bets. Then $w(t)$ became $w(t+1)$ after each wager winning or losing. The system was programmed by John Swetye to search for bets at 80 racetracks in North America. They system lost about 7% largely because the racetrack market combines bets made at many other racetracks and betting sites into one pool at the last minute. So betting like ours is not recorded into the pools until after the race is running. About half the money is entered then and that alters the odds we used at the end of betting. Our calculation takes into account the bets by other people and the effect of our bets on the odds. The rebate averaged 9% so we had a net gain of about 2% or $26,500.

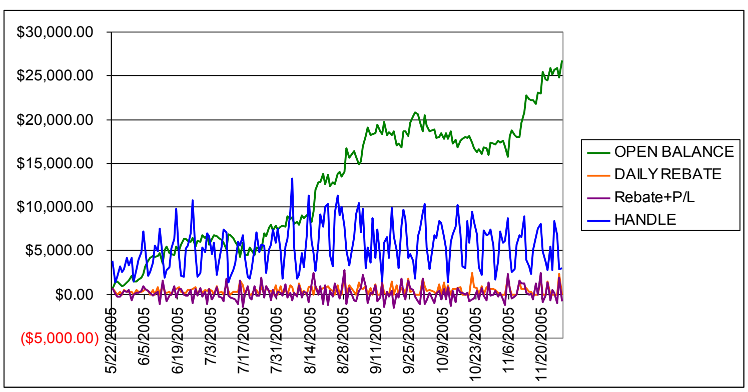


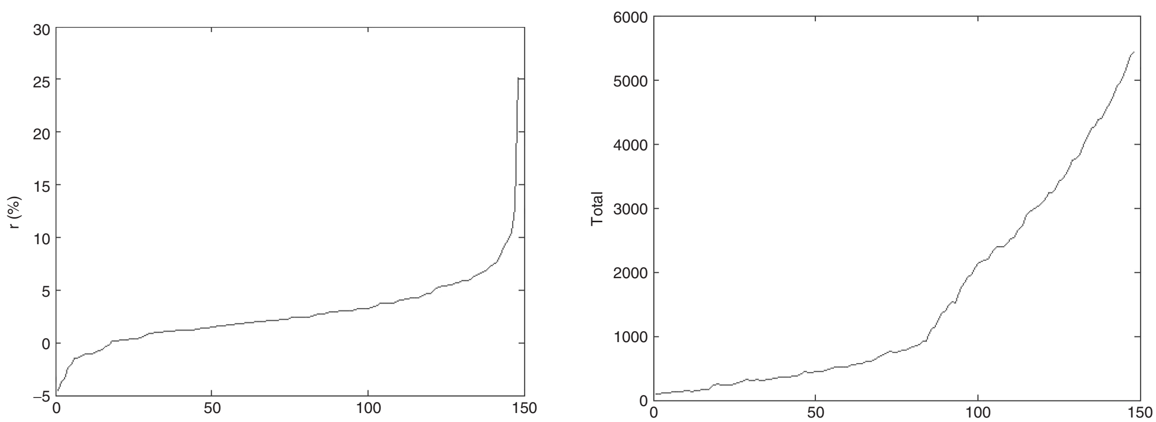
Figure 5 Racetrace betting record of the place and show system

The remainder of the section has examples and endorsements by other people. There is much secrecy in many management and most will not provide information even if I know they are Kelly bettors.

Kelly investing has several characteristics. It is not diversified but instead places large bets on the few very best assets. Hence, given the large bets, the portfolio can have a considerable monthly losses. But the long run growth of wealth is usually high.

The seminal application of the Kelly strategy is to a large sequence of similar investments. A good example of this is the Renaissance Medallion hedge fund which has thousands of 3-8 second trades. I taught Jim Simons, head of the Medallion Fund, about the advantages of Kelly betting in 1992. The Kelly strategy provides good wagers where the size of the bets depend upon the characteristics of the situation. Despite very high fees of 5% for management plus 44% of the net new profits, the gains have been outstanding as shown in Figure 6. There are very few monthly losses and a smooth wealth graph. The data available was monthly from January 1993 to April 2005, see the discussion in Gergaud and Ziemba (2012). Subsequent monthly results are not available but the yearly net returns up to the end of 2009 according to Insider Monkey (2010) were

2006 44.3%
2007 73.0%
2008 80.0%
2009 39.0%.



a) Monthly rates of return in increasing order b) Wealth over time

Figure 6 Renaissance Medallion Fund, January 1993 to April 2005

Ed Thorp and I believe that George Soros and Warren Buffett, two of the world's most successful investors, bet as if they were fully Kelly investors. We do not have direct proof of this but there is a lot of circumstantial evidence for our belief. Each of them has many investments but as Table 5 shows, these portfolios are very concentrated in very few investments have characteristics of full Kelly portfolios..

Both Soros and Buffett go for long term growth with many monthly losses but large final wealth, another characteristic of full Kelly betting. Figure 7 shows their results from December 1985 to April 2000.

Table 5 Top ten equity holdings of Soros Fund Management and Berkshire Hathaway, September 30, 2008. Source: SEC Filings

Soros Fund Management Company	Current Value x 1000	Shares	% Portfolio
Petroleo Brasileiro SA	$1,673,048	43,854,474	50.53
Potash Corp Sask Inc	378,020	3,341, 027	11.58
Wal Mart Stores Inc	195,320	3,791,890	5.95
Hess Corp	115,001	2,085,988	4.49
Conoco Phillips	96,855	1,707,900	3.28
Research in Motion Ltd	85,840	1,610,810	2.88
Arch Coal Inc	75,851	2,877,486	2.48
iShares TR	67,236	1,300,000	2.11
Powershares QQQ Trust	93,100	2,000,000	2.04
Schlumberger Ltd	33,801	545,000	1.12

Berkshire Hathaway Company	Current Value x 1000	Shares	% Portfolio
ConocoPhillips	$4,413,390	77,955,80	8.17
Procter & Gamble Co	4,789,440	80,252,000	8.00
Kraft Fods Inc	3,633,985	120,012,700	5.62
Wells Fargo & Co	1,819,970	66,132,620	3.55
Wesco Finl Corp	1,927,643	5,703,087	2.91
US Bancorp	1,1366,385	49,461,826	2.55
Johnson & Johnson	1,468,689	24,588,800	2.44
Moody's	1,121,760	48,000,000	2.34
Wal Mart Stores, Inc	1,026,334	19,944,300	1.71
Anheuser Busch Cos, Inc	725,201	13,845,000	1.29

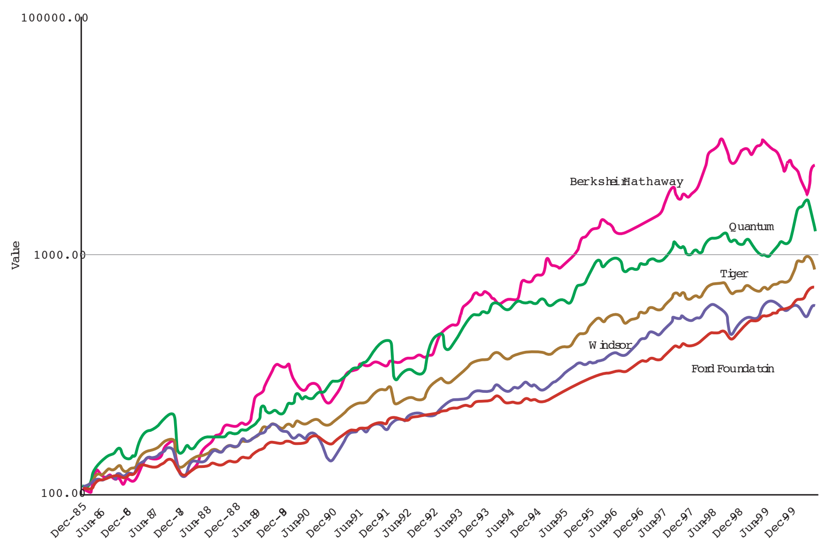


Figure 7 Growth of assets, log scale, various high performing funds, 1985-2000. Source: Ziemba (2005) using data from Siegel, Kroner and Clifford (2001)

Figure 8 shows that Berkshire has had the most large monthly gains and the most large monthly losses for the funds in the sample that Ziemba (2005) got from Larry Siegel of the Ford Foundation. So both the Sharpe ratio and my downside symmetric Shapre ratio (DSSR) are not high compared to other great traders such as Thorp or Simons. But as shown by Frazzani, Kabiller and Petersen (2020), Berkshire has a higher Sharpe ratio than any US stock or mutual fund with a history of more than 30 years. The secret to Buffett's success seems to be leveraging (about 1.6 to 1 using low cost and stable sources of financing much from his insurance businesses) plus a focus on cheap, safe, quality, low beta stocks.

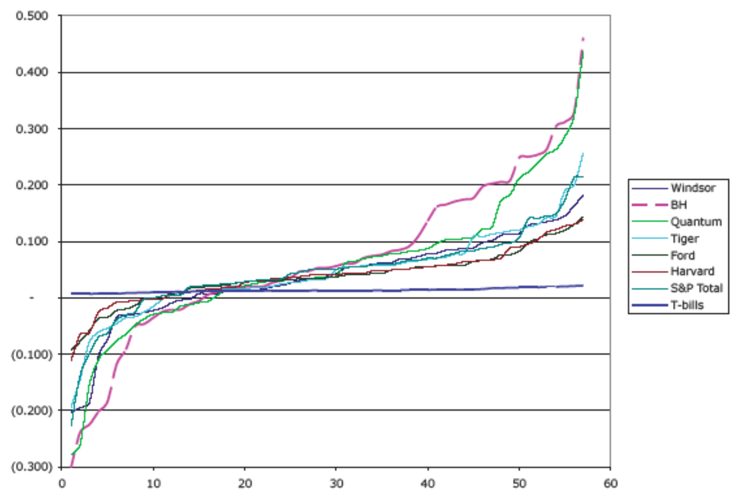


Figure 8 Return distributions of all the funds, quarterly returns distribution, December 1985 to March 2000.

Other famed investors such as John Maynard Keynes, running the King's College Cambridge endowment from 1927-1945; Bill Benter, the famed Hong Kong racing guru; and Ed Thorp, running the Princeton Newport hedge fund from 1968-88, all had excellent records and used Kelly and fractional Kelly strategies. Graphs of these and more discussion is in Ziemba and Ziemba (2013)

How much should you bet?

A real example of this by Mohnish Pabrai (2007), who won the bidding for the 2008 lunch with Warren Buffett paying more than $600,000, had the following investment in Stewart Enterprises as discussed by Edward Thorp (2010). Over a 24-month period, with probability 0.80 the investment at least doubles, with 0.19 probability the investment breaks even and with 0.01 probability all the investment is lost.

The optimal Kelly bet is 97.5% of wealth and half Kelly is 38.75%. Pabrai invested 10%. While this seems rather low, other investment opportunities, miscalculation of probabilities, risk tolerance, possible short run losses, bad scenario *Black Swan* events, price pressures, buying in and exiting suggest that a bet a lot lower than 97.5% is appropriate.

The original application of Kelly betting strategies was in Ed Thorp's book *Beat the Dealer* (1960) where card counting was introduced to get better mean gain estimates. The edge for a successful card counter varies from about -5% to +10% depending upon the favorability of the deck. By wagering more in favorable situations and less or nothing when the deck is unfavorable, an average weighted edge is about 1-2%. Thorp's book changed the casino business dramatically and expanded the play by professionals and advanced amateurs. Throughout the years, millions of people bought this book and attempted to use the card counting betting system and the Kelly sizing procedure. An approximation to provide insight into the long-run behavior of a player's fortune is to assume that the game is a Bernoulli trial with a probability of success p = 0.51 and probability of loss q=1-p= 0.49.

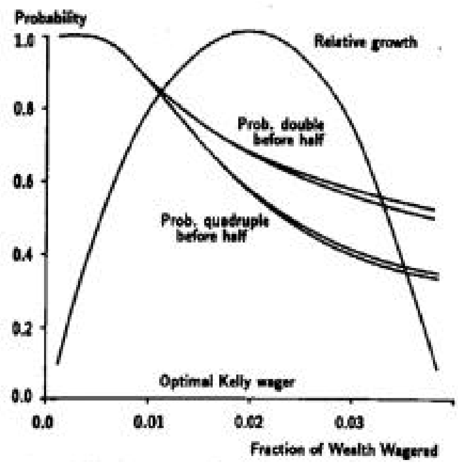


Figure 9 Probability of doubling and quadrupling before halving and relative growth rates versus fraction of wealth wagered for Blackjack (2% advantage, p=0.51 and q=0.49). Source: McLean, Ziemba and Blazenko (1992).

Figure 9 shows the relative growth rate $\pi ln(1+p)_{(}1-\pi)ln(1-\pi)$ versus the fraction of the investor's wealth wagered, π. The security curves show the bounds on the true probability of doubling or quadrupling before halving. This is maximized by the Kelly log bet $\pi^* = p - q = 0.02$. The growth rate is lower for smaller and for larger bets than the Kelly bet. Superimposed on this graph is also the probability that the investor doubles or quadruples the initial wealth before losing half of this initial wealth. Since the growth rate and the security are both decreasing for $\pi > \pi^*$,, it follows that it is never advisable to wager more than π^*. Also it can be shown[5] that the growth rate of a bet that is exactly twice the Kelly bet, namely $2\pi^* = 0.04$, is zero plus the risk-free rate of interest. Figure 9 illustrates this. Hence log betting is the most aggressive investing that

Table 6 Growth Rates Versus Probability of Doubling Before Halving for Blackjack. Source: MacLean and Ziemba (1999)

	0.1		0.999		0.19
	0.2		0.998		0.36
Range	0.3		0.98		0.51
for	0.4	Safer	0.94	Less Growth	0.64
Blackjack	0.5		0.89		0.75
Teams	0.6	Riskier	0.83	More Growth	0.84
	0.7		0.78		0.91
	0.8		0.74		0.96
	0.9		0.70		0.99
	1.0	Kelly	0.67		1.00
	1.5		0.56		0.75
Overkill →	2.0		0.50		0.00
Too Risky					

one should ever consider. The root of hedge fund disasters is frequently caused by bets above π^* when they are highly levered as well. Table 6 shows typical fractional Kelly strategies that have been used by blackjack teams. See Ziemba and Ziemba (2013) for discussions of several hedge fund disaster examples including Long Term Capital Management, Amarath, Niederhoffer and Societié Generale.

Bill Gross, the world's largest bond trader. uses Kelly betting at PIMCO. During an interview in the Wall Street Journal (March 22-23, 2008) Bill Gross and Ed Thorp discussed turbulence in the markets, hedge funds and risk management. Bill considered the question of risk management after he read Ed Thorp's *Beat the Dealer* in 1966. That summer he was off to Las Vegas to beat blackjack. Just as Ed did some years earlier, he sized his bets in proportion to his advantage, following the Kelly Criterion as described in Beat the Dealer, and ran his \$200 bankroll up to \$10,000 over the summer. Bill has gone from managing risk for his tiny bankroll to managing risk for Pacific Investment Management Companys (PIMCO) investment pool of almost \$1 trillion. He still applies lessons he learned from the Kelly Criterion. As Bill said, "Here at PIMCO it doesn't matter how much you have, whether its \$200 or \$1 trillion Professional blackjack is being played in this trading room from the standpoint of risk management and thats a big part of our success". In a cover quote for the book MacLean, Thorp and Ziemba (2011), Gross added that

> "Ed Thorp and the Kelly criterion have been a lighthouse for risk management for me and PIMCO for over 45 years. First at the blackjack tables and then in portfolio management, the Kelly system has helped to minimize risk and maximize return for thousands of PIMCO clients"

Thorp in an email to me on September 1, 2010 added

> "The background here, briefly, is that he first read Beat the Dealer, went to Las Vegas for the summer after he graduated, ran \$200 into \$10,000, and went into the service. Two years

later he read Beat the Market, wrote a thesis on convertible bonds at UCI, and as a result was hired by Pacific Mutual, leading to his founding PIMCO, etc."

Most applications choose the Kelly fraction in an ad hoc way. One approach to determining discrete time *optimal* Kelly fractions was proposed by MacLean, Sanegre, Zhao and Ziemba (2004). Their model has a pre-determined ex ante wealth path through time. Then the fractional Kelly wagers that maximize the growth rate are determined subject to the constraint that the portfolio stays above the path a high percentage of the time. In MacLean, Zhao and Ziemba (2009, 2012), they extend the analysis to have the additional feature that if the path is violated, then the violations are penalized with a convex function. So the larger the violation, the larger the penalty. This model then tends to force the decisions to be such that the path is not violated. One cannot have too aggressive wealth paths. But for example, if the wealth path is constant at $w(0)$ you have a form of portfolio insurance aiming for close to zero losses but with reasonably high growth.

6. Conclusions

The Kelly capital growth strategy has been used successfully by many investors and speculators including me during the past fifty years. In this paper I describe its main advantages, namely its superiority in producing long run maximum wealth from a sequence of favorable investments. The seminal application is to an investment situation that has many repeated similar bets over a long time horizon. In all cases one must have a winning system that is one with a positive expectation. Then the Kelly and fractional Kelly strategies (those with less long run growth but more security) provide a superior bet sizing strategy. The mathematical properties prove maximum asymptotic long run growth. But short term there can be high volatility. Examples include the initial application by Thorp in 1960 to blackjack and my applications to horseracing, the turn-of-the-year effect, and other financial market anomalies. It is known that important successful hedge funds such as Renaissance Medallion, Bill Gross at PIMCO, Bill Benter's Hong Kong racing syndicate and Harry McPike's trend following syndicate have made extra millions from the use of the Kelly strategies. Other examples of investors who I believe behave as if they were full or close to full Kelly investors are Warren Buffett, George Soros and John Maynard Keynes who all have portfolios with the following Kelly characteristics.:

> "the portfolios are highly concentrated, not diversified, with huge positions in the few very best investments with much monthly variation and many monthly losses but very high final wealth most of the time."

In this paper, I describe how one makes the investment bets using static stochastic non-linear programming that takes the effect of our wagers on the prices into account. The paper also responds to the critique of Professor Paul A Samuelson in letters to me and papers reprinted in the Kelly book by MacLean, Thorp and myself (2011). The basic criticisms are well explained and largely are concerned with over betting, the major culprit of hedge fund disasters and theoretical long run properties of the strategy.

My response is that if properly used the Kelly strategy provides the best long term wealth maximizing technique and the examples in the paper show its use in practice. The main conclusions of the simulation studies are:

1. the great superiority of full Kelly and close to full Kelly strategies over longer horizons with very large gains a large fraction of the time;
2. that the short term performance of Kelly and high fractional Kelly strategies is very risky;
3. that there is a consistent tradeoff of growth versus security as a function of the bet size determined by the various strategies; and
4. that no matter how favorable the investment opportunities are or how long the finite horizon is, a sequence of bad scenarios can lead to very poor final wealth outcomes, with a loss of most of the investor's initial capital.

Endnotes

1. Berlekamp was a main intellectual force in the Renaissance Medallion hedge fund, arguably the world's most successful hedge fund, see Gergaud and Ziemba (2012). Later he was a professor of mathematics at the University of California, Berkeley.
2. insert website
3. Paul used considerable colorful jargon. At the suggestion of a referee on an earlier version of this paper, I have toned down the rhetoric so the argument is more clear.
4. For one or two assets with fixed odds, take derivatives and solve for the optimal wagers; for multi-asset bets under constraints; and when portfolio choices affect returns (odds), one must solve a stochastic nonlinear program which, possibly, is non-convex.
5. See Harry Markowitz's proof in Ziemba (2003) and the more general proof of Edward Thorp (2011) and the graphs in MacLean, Ziemba and Blazenko (1992).

Acknowledgments

Without implicating them, thanks go to Harry M. Markowitz and Edward O. Thorp for helpful discussions on this topic and the referees for useful comments on an earlier draft of this paper.

References

Aase, K. K. (2001). On the St. Petersburg Paradox. *Scandinavian Actuarial Journal 3*(1), 69–78.

Algoet, P. H. and T. M. Cover (1988). Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *Annals of Probability 16*(2), 876–898.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk (translated from the latin by Louise Sommer). *Econometrica 22*, 23–36.

Breiman, L. (1960). Investment policies for expanding businesses optimal in a long run sense. *Naval Research Logistics Querterly 4*(4), 647–651.

Breiman, L. (1961). Optimal gambling system for favorable games. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability 1*, 63–8.

Chopra, V. K. and W. T. Ziemba (1993). The effect of errors in mean, variance and co-variance estimates on optimal portfolio choice. *Journal of Portfolio Management 19*, 6–11.

Clark, R. and W. T. Ziemba (1988). Playing the turn-of-the-year effect with index futures. *Operations Research XXXV*, 799–813 (1988.

Cover, T. M. and J. Thomas (2006). *Elements of Information Theory, 2nd Edition.* Wiley & Sons, New York.

Frazzini, A., D. Kabiller, and L. H. Peterson (2012). Buffett's alpha. NYU technical report, August 30.

Gergaud, O. and W. T. Ziemba (2012). Great investors: their methods, results and evaluation. *Journal of Portfolio Management 28*(4), 28–147.

Hakansson, N. (1970). Optimal investment and consumption strategies under risk for a class of utility functions. *Econometrica 38*, 587–607.

Hakansson, N. H. (1971a). Capital growth and the mean-variance approach to portfolio selection. *Journal of Financial Quantitative Analysis 6*, 517–557.

Hakansson, N. H. (1971b). On optimal myopic portfolio policies with and without serial correlation. *Journal of Business 44*, 324–334.

Hakansson, N. H. and W. T. Ziemba (1995). Capital growth theory. In R. A. Jarrow, V. Maksimovic, and W. T. Ziemba (Eds.), *Finance, Handbooks in OR & MS*, pp. 65–86. North Holland.

Hausch, D. G., W. T. Ziemba, and M. E. Rubinstein (1981). Efficiency of the market for racetrack betting. *Management Science XXVII*, 1435–1452.

Insider Monkey (2010). Seeking alpha: best hedge funds, Jim Simons Medallion Fund. December 31.

Kelly, Jr., J. R. (1956). A new interpretation of the information rate. *Bell System Technical Journal 35*, 917–926.

Latané, H. (1959). Criteria for choice among risky ventures. *Journal of Political Economy 67*, 144–155.

Latané, H. (1978). The geometric-mean principle revisited – a reply. *Journal of Banking and Finance 2*(4), 395–398.

Luenberger, D. G. (1993). A preference foundation for log mean-variance criteria in portfolio choice problems. *Journal of Economic Dynamics and Control 17*, 887–906.

MacLean, L. C., R. Sanegre, Y. Zhao, and W. T. Ziemba (2004). Capital growth with security. *Journal of Economic Dynamics and Control 28*(4), 937–954.

MacLean, L. C., E. O. Thorp, and W. T. Ziemba (2011). *The Kelly Capital Growth Criterion: Theory and Practice*. World Scientific, Singapore, Hardback; 2012 revised paperback.

MacLean, L. C., E. O. Thorp, Y. Zhao and W. T. Ziemba (2011). How does the Fortunes Formula-Kelly capital growth model perform? *Journal of Portfolio Management 37*(4), 96-111.

MacLean, L. C., Y. Zhao, and W. T. Ziemba (2009). Optimal capital growth with convex loss penalties. Working paper, Dalhousie University.

MacLean, L. C., Y. Zhao, and W. T. Ziemba (2012). Optimal capital growth with shortfall penalties. Working paper, Dalhousie University.

MacLean, L. C. and W. T. Ziemba (1999). Growth versus security tradeoffs in dynamic investment analysis. In R. J.-B. Wets and W. T. Ziemba (Eds.), *Stochastic Programming: State of the Art, Annals of Operations Research, 1998*, pp. 193–226.

MacLean, L. C., W. T. Ziemba, and G. Blazenko (1992). Growth versus security in dynamic investment analysis. *Management Science 38*, 1562–85.

MacLean, L. C., W. T. Ziemba, and Y. Li (2005). Time to wealth goals in capital accumulation and the optimal trade-off of growth versus security. *Quantitative Finance 5*(4), 343–357.

Malkiel, B. (2011). *A random walk down Wall Street* (10th ed.). Norton.

Markowitz, H. M. (1976). Investment for the long run: New evidence for an old rule. *Journal of Finance 31*(5), 1273–1286.

Markowitz, H. M. (2006). Samuelson and investment for the long run. In M. Szenberg, L. Ramrattan, and A. A. Gottesman (Eds.), *Samuelsonian economics and the twenty-first century*, pp. Chapter 17, 252–261. Oxford University Press.

Merton, R. C. and P. A. Samuelson (1974). Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods. *Journal of Financial Economics 1*, 67–94.

Mossin, J. (1968). Optimal multi period portfolio policies. *Journal of Business 41*(2), 215–229.

Ohlson, J. (1975). The asymptotic validity of quadratic utility with trading interval approaches zero. In W. T. Ziemba and R. G. Vickson (Eds.), *Stochastic optimization models in finance*, pp. 221–234. Academic Press.

Pabrai, M. (2007). *The Dhandho investor: the low-risk value method to high returns*. Wiley.

Poundstone, W. (2005). *Fortunes Formula: The Untold Story of the Scientific System that Beat the Casinos and Wall Street*. Hill and Wang, New York.

Roll, R. (1973). Evidence on the growth optimum model. *The Journal of Finance 28*(3), 551–566.

Rubinstein, M. (1976). The strong case for the generalized logarithmic utility model as the premier model of financial markets. *Journal of Finance 31*(2), 551–571.

Samuelson, P. A. (1963). Risk and uncertainty: a fallacy of large numbers. *Scientia* (6th Series, 57th year, April-May), 153–158.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics 51*, 239–246.

Samuelson, P. A. (1970). The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Review of Economic Studies 37*(4), 537–542.

Samuelson, P. A. (1977). St. Petersburg paradoxes: Defanged, dissected and historically described. *Journal of Economic Literature 15*(1), 24–55.

Samuelson, P. A. (1979). Why we should not make mean log of wealth big though years to act are long. *Journal of Banking and Finance 3*, 305–307.

Samuelson, P. A. (1991). Long-run risk tolerance when equity returns are mean regressing: Pseudoparadoxes and vindication of businessmen's risk. In W. C. Brainard, W. D. Nordhaus, and H. W. Watts (Eds.), *Money, Macroeconomics and Economic Policy*, pp. 181–200. MIT Press.

Samuelson, P. A. (various). Letters to W. T. Ziemba, December 13, 2006, May 17, 2007, and May 12, 2008.

Schwartz, S. L. and W. T. Ziemba (2007). ALM in social security. In S. A. Zenios and W. T. Ziemba (Eds.), *Handbook of Asset and Liability Modeling, Volume 2: Applications and Case Studies*, pp. 1069–1117. Handbooks in Finance Series, North Holland.

Thorp, E. O. . (2010). Understanding the Kelly criterion. *Wilmott*.

Thorp, E. O. . (2011). The Kelly criterion in blackjack, sports betting and the stock market. In L. C. MacLean, E. O. Thorp, and W. T. Ziemba (Eds.), *The Kelly capital growth investment criterion*, pp. 789–832. World Scientific, Singapore.

Thorp, E. O. and Whitley (1972). Concave utilities are distinguished by their optimal strategies. Colloquia Mathematica Societatis Janos Bolyai. pp. 813-830.

Ziemba, W. T. (2012a). *Calendar anomalies and arbitrage*. World Scientific.

Ziemba, W. T. (2012b). Stochastic programming and optimization in horserace betting. In H. I. Gassman and W. T. Ziemba (Eds.), *Stochastic Programming Applications in Finance, Energy and Production*. World Scientific.

Ziemba, W. T. and R. E. S. Ziemba (2013). *Investing in the Modern Age*. World Scientific.