Long-term capital growth: the good and bad properties of the Kelly and fractional Kelly capital growth criteria

Article *in* Quantitative Finance · April 2011
DOI: 10.1080/14697688.2010.506108 · Source: RePEc

Edward O. Thorp

William Ziemba
University of British Columbia - Vancouver

Good and bad properties of the Kelly criterion*

Leonard C. MacLean,
Herbert Lamb Chair (Emeritus),School of Business,
Dalhousie University, Halifax, NS

Edward O. Thorp,
E.O. Thorp and Associates, Newport Beach, CA
Professor Emeritus, University of California, Irvine

William T. Ziemba,
Professor Emeritus, University of British Columbia, Vancouver, BC
Visiting Professor, Mathematical Institute, Oxford University, UK
ICMA Centre, University of Reading, UK
University of Bergamo, Italy

January 1, 2010

Abstract

We summarize what we regard as the good and bad properties of the Kelly criterion and its variants. Additional properties are discussed as observations.

The main advantage of the Kelly criterion, which maximizes the expected value of the logarithm of wealth period by period, is that it maximizes the limiting exponential growth rate of wealth. The main disadvantage of the Kelly criterion is that its suggested wagers may be very large. Hence, the Kelly criterion can be very risky in the short term.

In the one asset two valued payoff case, the optimal Kelly wager is the edge (expected return) divided by the odds. Chopra and Ziemba (1993), reprinted in Section 2 of this volume, following earlier studies by Kallberg and Ziemba (1981, 1984) showed for any asset allocation problem that the mean is much more important than the variances and co-variances. Errors in means versus errors in variances were about 20:2:1 in importance as measured by the cash equivalent value of final wealth. Table 1 and Figure 1 show this and illustrate that the relative importance depends on the degree of risk aversion. The

*Special thanks go to Tom Cover and John Mulvey for helpful comments on an earlier draft of this paper.

lower is the Arrow-Pratt risk aversion, $R_A = -u\prime\prime(w)/u\prime(w)$, the higher are the relative errors from incorrect means. Chopra (1993) further shows that portfolio turnover is larger for errors in means than for variances and for co-variances but the degree of difference in the size of the errors is much less than the performance as shown in Figure 2.

Table 1: Average Ratio of Certainty Equivalent Loss for Errors in Means, Variances and Covariances. Source: Chopra and Ziemba (1993)

Risk Tolerance*	Errors in Means vs Covariances	Errors in Means vs Variances	Errors in Variances vs Covariances
25	5.38	3.22	1.67
50	22.50	10.98	2.05
75	56.84	21.42	2.68
	↓	↓	↓
	20	10	2
	Error Mean	Error Var	Error Covar
	20	2	1

*Risk tolerance=$R_T(w) = \frac{100}{\frac{1}{2}R_A(w)}$ where $R_A(w) = -\frac{u''(w)}{u\prime(w)}$

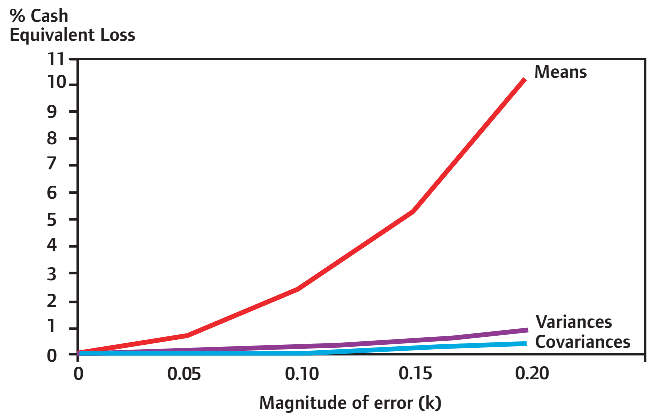

Figure 1: Mean Percentage Cash Equivalent Loss Due to Errors in Inputs.

Since log has $R_A(w) = 1/w$, which is close to zero, The Kelly bets may be exceedingly large and risky for favorable bets. In MacLean, Thorp, Zhao and Ziemba (2009) in this section of this volume, we present simulations of medium term Kelly, fractional Kelly and proportional betting strategies. The results show that with favorable investment opportunities, Kelly bettors attain large final wealth most of the time. But, because a long sequence of bad scenario outcomes is possible, any strategy can lose substantially even if there are many independent investment opportunities and the chance of losing at each investment

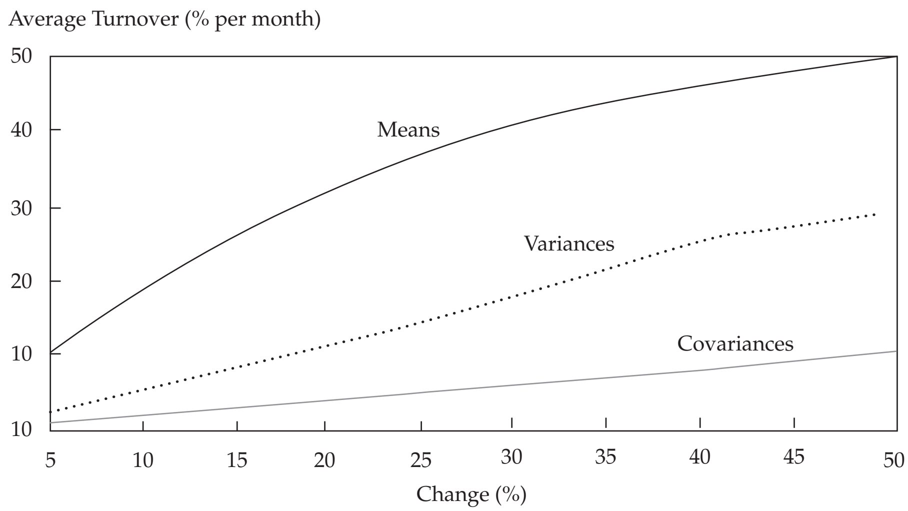

Figure 2: Average turnover for different percentage changes in means, variances and covariances. Source: Based on data from Chopra (1993)

decision point is small. The Kelly and fractional Kelly rules, like all other rules, are never a sure way of winning for a finite sequence.

In Section 6 of this volume, we describe the use of the Kelly criterion in many applications and by many great investors. Two of them, Keynes and Buffett, were long term investors whose wealth paths were quite rocky but with good long term outcomes. Our analyses suggest that Buffett seems to act similar to a fully Kelly bettor (subject to the constraint of no borrowing) and Keynes like a 80% Kelly bettor with a negative power utility function $-w^{-0.25}$, see Ziemba (2003). See the wealth graphs reprinted in section 6 from Ziemba (2005).

Graphs such as Figure 3 show that growth is traded off for security with the use of fractional Kelly strategies and negative power utility functions. Log maximizes the long run growth rate. Utility functions such as positive power that bet more than Kelly have more risk and lower growth. One of the properties shown below that is illustrated in the graph is that for processes which are well approximated by continuous time, the growth rate becomes zero plus the risk free rate when one bets exactly twice the Kelly wager.

Hence it never pays to bet more than the Kelly strategy because then risk increases (lower security) and growth decreases, so Kelly dominates all these strategies in geometric risk-return or mean-variance space. See Ziemba (2009) in this volume.

As you exceed the Kelly bets more and more, risk increases and long term growth falls, eventually becoming more and more negative. Long Term Capital is one of many real

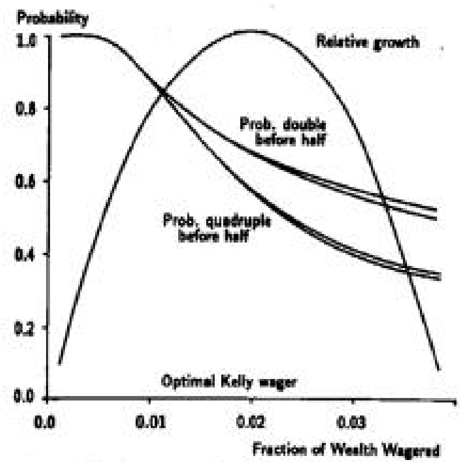

Figure 3: Probability of doubling and quadrupling before halving and relative growth rates versus fraction of wealth wagered for Blackjack (2% advantage, p=0.51 and q=0.49). Source: MacLean and Ziemba (1999)

world instances in which overbetting led to disaster. See Ziemba and Ziemba (2007) for additional examples.

Thus long term growth maximizing investors should bet Kelly or less. We call betting less than Kelly "fractional Kelly," which is simply a blend of Kelly and cash. Consider the negative power utility function $\delta\omega^{\delta}$ for $\delta < 0$. This utility function is concave and when $\delta \to 0$ it converges to log utility. As δ becomes more negative, the investor is less aggressive since his absolute Arrow-Pratt risk aversion index is also higher. For the case of a stationary lognormal process and a given δ for utility function δw^{δ} and $\alpha = 1/(1-\delta)$ between 0 and 1, they both will provide the same optimal portfolio when α is invested in the Kelly portfolio and $1-\alpha$ is invested in cash.

This handy formula relating the coefficient of the negative power utility function to the Kelly fraction is correct for lognormal investments and approximately correct for other distributed assets; see MacLean, Ziemba and Li (2005). For example, half Kelly is $\delta = -1$ and quarter Kelly is $\delta = -3$. So if you want a less aggressive path than Kelly pick an appropriate δ. This formula does not apply more generally. For example, for coin tossing, where $Pr(X=1) = p,\ Pr(X-1) = q,\ p+q = 1$,

$$f_{\delta}^{*} = \frac{p^{\frac{1}{1-\delta}} - q^{\frac{1}{1-\delta}}}{p^{\frac{1}{1-\delta}} + q^{\frac{1}{1-\delta}}} = \frac{p^{\alpha} - q^{\alpha}}{p^{\alpha} + q^{\alpha}}$$

which is not αf^{*}, where $f^{*} = p - q \geqslant 0$ is the Kelly bet.

We now list these and other important Kelly criterion properties, updated from MacLean, Ziemba and Blazenko (1992), MacLean and Ziemba(1999) and Ziemba and Ziemba (2007). See also Cover and Thomas (2006, chapter 16).

The Good Properties

Good Maximizing ElogX asymptotically maximizes the rate of asset growth. See Breiman (1961), Algoet and Cover (1988)

Good The expected time to reach a preassigned goal A is asymptotically least as A increases without limit with a strategy maximizing ElogX_N. See Breiman (1961), Algoet and Cover (1988), Browne (1997a)

Good Under fairly general conditions, maximizing ElogX also asymptotically maximizes median logX. See Ethier (1987, 2004, 2010)

Good The ElogX bettor never risks ruin. See Hakansson and Miller (1975)

Good The absolute amount bet is monotone increasing in wealth.

Good The ElogX bettor has an optimal myopic policy. He does not have to consider prior nor subsequent investment opportunities. This is a crucially important result for practical use. Hakansson (1971) proved that the myopic policy obtains for dependent investments with the log utility function. For independent investments and any power utility a myopic policy is optimal, see Mossin (1968). In fact past outcomes can be taken into account by maximizing the conditional expected logarithm given the past (Algoet and Cover, 1988)

Good Simulation studies show that the ElogX bettor's fortune pulls ahead of other "essentially different" strategies' wealth for most reasonable-sized samples. Essentially different has a limited meaning. For example $g^* \geqslant g$ but $g^* - g = \epsilon$ will not lead to rapid separation if ϵ is small enough The key again is risk. See Bicksler and Thorp (1973), Ziemba and Hausch (1986) and MacLean, Thorp, Zhao and Ziemba (2009) in this volume. General formulas are in Aucamp (1993).

Good If you wish to have higher security by trading it off for lower growth, then use a negative power utility function, δw^δ, or fractional Kelly strategy. See MacLean, Sanegre, Zhao and Ziemba (2004) reprinted in section 3, who show how to compute the coefficent to stay above a growth path at discrete points in time with given probability or to be above a given drawdown with a certain confidence limit. MacLean, Zhao and Ziemba (2009) add the feature that path violations are penalized with a convex cost function. See also Stutzer (2009) for a related but different model of such security.

Good Competitive optimality . Kelly gambling yields wealth X^* such that $E(\frac{X}{X^*}) \leqslant 1$, for all other strategies X. This follows from the Kuhn Tucker conditions. Thus by Markov's inequality, $Pr\left[X \geqslant tX^*\right] \leqslant \frac{1}{t}$, for $t \geqslant 1$, and for all other induced wealths x. Thus an opponent cannot outperform X^* by a factor t with probability greater than $\frac{1}{t}$. This inequality can be improved when $t = 1$ by allowing fair randomization U. Let U be drawn according to a uniform distribution over the interval [0,2], and let U be independent of X^*. Then the result improves to $Pr\left[X \geqslant UX^*\right] \leqslant \frac{1}{2}$ for all portfolios X. Thus fairly randomizing one's initial wealth and then investing it according to the Kelly criterion, one obtains a wealth UX^* that can only be beaten half the time. Since a competing investor can use the same strategy, probability $\frac{1}{2}$ is the best competitive performance one can expect. We see that Kelly gambling is the heart of the solution of this two-person zero sum game of who ends up with the most money. So we see that X^* (actually UX^*) is competitively optimal in a single investment period (Bell and Cover 1980, 1988).

Good If X^* is the wealth induced by the log optimal (Kelly) portfolio, then the expected wealth ratio is no greater than one, i.e., $E(\frac{X}{X^*}) \leqslant 1$, for the wealth X induced by any other portfolio (Bell and Cover, 1980, 1988).

Good Super St Petersburg. Any cost c for the St Petersburg random variable X, $Pr\left[X = 2^k\right] = 2^{-k}$, is acceptable. But the larger the cost c, the less wealth one should invest. The growth rate G^* of wealth resulting from repeated such investments is

$$G^* = \max_{0 \leqslant f \leqslant 1} E \ln(1 - f + \frac{f}{c} X),$$

where f is the fraction of wealth invested. The maximizing f^* is the Kelly proportion (Cover and Bell, 1980). The Kelly fraction f^* can be computed even for a super St Petersburg random variable $Pr\left[Y = 2^{2^k}\right] = 2^{-k},\ k = 1, 2, \ldots$, where $E \ln Y = \infty$, by maximizing the relative growth rate

$$\max_{0 \leqslant f \leqslant 1} E \ln \frac{1 - f + \frac{f}{c} Y}{\frac{1}{2} + \frac{1}{2c} Y}.$$

This is bounded for all f in [0,1].

Now, although the exponential growth rate of wealth is infinite for all proportions f and it seems that all $f \in [0, 1]$ are equally good, the maximizing f^* in the previous equation guarantees that the f^* portfolio will asymptotically exponentially outperform any other portfolio $f \in [0, 1]$. Both investors' wealth have super exponential growth, but the f^* investor will exponentially outperform any other essentially different investor.

The Bad Properties

Bad The bets may be a large fraction of current wealth when the wager is favorable and the risk of loss is very small. For one such example, see Ziemba and Hausch (1986; 159-160). There, in the inaugural 1984 Breeders Cup Classic $3 million race, the optimal fractional wager, using the Dr Z place and show system using the win odds as the probability of winning, on the 3-5 shot Slew of Gold was 64%. (See also the 74% future bet on the January effect in MacLean, Ziemba and Blazenko (1992) reprinted in this volume). Thorp and Ziemba actually made this place and show bet and won with a low fractional Kelly wager. Slew finished third but the second place horse Gate Dancer was disqualified and placed third. Wild Again won this race; the first great victory by the masterful jockey Pat Day.

Bad For coin tossing, any fixed fraction strategy has the property that if the number of wins equals the number of losses then the bettor is behind. For n wins and n losses and initial wealth W_0 we have $W_{2n} = W_0(1 - f^2)^n$.

Bad The unweighted average rate of return converges to half the arithmetic rate of return. Thus you may regularly win less than you expect. This is a consequence of weighting equally rather than by size of the wager. See Ethier and Tavaré (1983) and Griffin (1985).

Some Observations

- For an i.i.d. process and a myopic policy, which results from maximizing expected utility in case the utility function is log or a negative power, the result is fixed fraction betting, hence fractional Kelly includes all these policies.
- A betting strategy is "essentially different" from Kelly if $S_n \equiv \sum_{i=1}^{n} Elog(1+f_i^* X_i) - \sum_{i=1}^{n} Elog(1 + f_i X_i)$ tends to infinity as n increases. The sequence $\{f_i^*\}$ denotes the Kelly betting fractions and the sequence $\{f_i\}$ denotes the corresponding betting fractions for the essentially different strategy.
- The Kelly portfolio does not necessarily lie on the efficient frontier in a mean-variance model (Thorp, 1971).
- Despite its superior long-run growth properties, Kelly, like any other strategy, can have a poor return outcome. For example, making 700 wagers all of which have a 14% advantage, the least of which has a 19% chance of winning can turn $1000 into $18. But with full Kelly 16.6% of the time $1000 turns into at least $100,000, see Ziemba and Hausch (1996). Half Kelly does not help much as $1000 can become $145 and the growth is much lower with only $100,000 plus final wealth 0.1% of the time. For

more such calculations, see Bicksler and Thorp (1973) and MacLean, Thorp, Zhao and Ziemba (2009) in this volume.

- Fallacy: If maximizing $ElogX_N$ almost certainly leads to a better outcome then the expected utility of its outcome exceeds that of any other rule provided N is sufficiently large. *Counterexample:* $u(x) = x, 1/2 < p < 1$, Bernoulli trials $f = 1$ maximizes $EU(x)$ but $f = 2p - 1 < 1$ maximizes $ElogX_N$. See Samuelson (1971) and Thorp (1971, 2006).
- It can take a long time for any strategy, including Kelly, to dominate an essentially different strategy. For instance, in continuous time with a geometric Wiener process, suppose $\mu_\alpha = 20\%$, $\mu_\beta = 10\%$, $\sigma_\alpha = \sigma_\beta = 10\%$. Then in five years A is ahead of B with 95% confidence. But if $\sigma_\alpha = 20\%, \sigma_\beta = 10\%$ with the same means, it takes 157 years for A to beat B with 95% confidence. As another example, in coin tossing suppose game A has an edge of 1.0% and game B 1.1%. It takes two million trials to have an 84% chance that game A dominates game B, see Thorp (2006).

The theory and practical application of the Kelly criterion is straightforward when the underlying probability distributions are fairly accurately known. However, in investment applications this is usually not the case. Realized future equity returns may be very different from what one would expect using estimates based on historical returns. Consequently practitioners who wish to protect capital above all, sharply reduce risk as their drawdown increases.

Prospective users of the Kelly Criterion can check our list of good properties, bad properties and observations to test whether Kelly is well suited to their intended application. Given the extreme sensitivity of $E\log$ calculations to errors in mean estimates, these estimates must be accurate and to be on the safe side,the size of the wagers should be reduced.

For long term compounders, the good properties dominate the bad properties of the Kelly criterion. But the bad properties may dampen the enthusiasm of naive prospective users of the Kelly criterion. The Kelly and fractional Kelly strategies are very useful if applied carefully with good data input and proper financial engineering risk control.

Appendix

In continuous time, with a geometric Wiener process, betting exactly double the Kelly criterion amount leads to a growth rate equal to the risk free rate. This result is due to Thorp (1997), Stutzer (1998) and Janacek (1998) and possibly others. The following simple proof, under the further assumption of the Capital Asset Pricing Model, is due to Harry Markowitz and appears in Ziemba (2003).

In continuous time

$$g_p = E_p - \frac{1}{2}V_p$$

E_p, V_p, g_p are the portfolio expected return, variance and expected log, respectively. In the CAPM

$$E_p = r_o + (E_M - r_0)X$$
$$V_p = \sigma_M^2 X^2$$

where X is the portfolio weight and r_0 is the risk free rate. Collecting terms and setting the derivative of g_p to zero yields

$$X = (E_M - r_0)/\sigma_M^2$$

which is the optimal Kelly bet with optimal growth rate

$$g^* = r_0 + (E_M - r_0)^2 - \frac{1}{2}[(E_M r_0)/\sigma_M^2]^2\sigma_M^2$$

$$= r_0 + (E_M - r_0)^2/\sigma_M^2 - \frac{1}{2}(E_M - r_0)^2/\sigma_M^2$$

$$= r_0 + \frac{1}{2}[(E - M - r))/\sigma_M]^2.$$

Substituting double Kelly, namely $Y = 2X$ for X above into

$$g_p = r_0 + (E_M - r_0)Y - \frac{1}{2}\sigma_M^2 Y^2$$

and simplifying yields

$$g_0 - r_0 = 2(E_M - r_0)^2/\sigma_M^2 - \frac{4}{2}(E_M - r_0)^2/\sigma_M^2 = 0.$$

Hence $g_0 = r_0$ when $Y = 2S$.

The CAPM assumption is not needed. For a more general proof and illustration, see Thorp (2006).

References

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

Aucamp, D. (1993). On the extensive number of plays to achieve superior performance with the geometric mean strategy. *Management Science 39*, 1163–1172.

Bell, R. M. and T. M. Cover (1980). Competitive optimality of logarithmic investment. *Math of Operations Research 5*, 161–166.

Bell, R. M. and T. M. Cover (1988). Game-theoretic optimal portfolios. *Management Science 34*(6), 724–733.

Bicksler, J. L. and E. O. Thorp (1973). The capital growth model: an empirical investigation. *Journal of Financial and Quantitative Analysis 8*(2), 273–287.

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

Browne, S. (1997). Survival and growth with a fixed liability: optimal portfolios in continuous time. *Math of Operations Research 22*, 468–493.

Chopra, V. K. (1993). Improving optimization. *Journal of Investing 2*(3), 51–59.

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.

Cover, T. and Thomas (2006). *Elements of Information Theory* (2 ed.).

Ethier, S. (1987). The proportional bettor's fortune. Proceedings 7th International Conference on Gambling and Risk Taking, Department of Economics, University of Nevada, Reno.

Ethier, S. (2004). The Kelly system maximizes median fortune. *Journal of Applied Probability 41*, 1230–1236.

Ethier, S. (2010). *The Doctrine of Chances: Probabilistic Aspects of Gambling.* Springer.

Ethier, S. and S. Tavaré (1983). The proportional bettor's return on investment. *Journal of Applied Probability 20*, 563–573.

Finkelstein, M. and R. Whitley (1981). Optimal strategies for repeated games. *Advanced Applied Probability 13*, 415–428.

Griffin, P. (1985). Different measures of win rates for optimal proportional betting. *Management Science 30*, 1540–1547.

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

Hakansson, N. H. and B. Miller (1975). Compound-return mean-variance efficient portfolios never risk ruin. *Management Science 22*, 391–400.

Janacek, K. (1998). Optimal growth in gambling and investing. M.Sc. Thesis, Charles University, Prague.

Kallberg, J. G. and W. T. Ziemba (1981). Remarks on optimal portfolio selection. In G. Bamberg and . Opitz (Eds.), *Methods of Operations Research, Oelgeschlager*, Volume 44, pp. 507–520. Gunn and Hain.

Kallberg, J. G. and W. T. Ziemba (1984). Mis-specifications in portfolio selection problems. In G. Bamberg and K. Spremann (Eds.), *Risk and Capital*, pp. 74–87. Springer Verlag, New York.

MacLean, L., R. Sanegre, Y. Zhao, and W. T. g (2004). Capital growth with security. *Journal of Economic Dynamics and Control reprinted in this volume*.

MacLean, L., E. O. Thorp, Y. Zhao, and W. T. Ziemba (2009). Medium term simulations of Kelly and fractional Kelly strategies. this volume.

MacLean, L. and W. T. Ziemba (1999). Growth versus security tradeoffs in dynamic investment analysis. *Annals of Operations Research 85*, 193–227.

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

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

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

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

Samuelson, P. A. (1971). The fallacy of maximizing the geometric mean in long sequences of investing or gambling. *Proceedings National Academy of Science 68*, 2493–2496.

Stutzer, M. (2009). On growth optimality versus security against underperformance. this volume, 19 pages.

Thorp, E. O. (1971). Portfolio choice and the Kelly criterion. *Proceedings of the Business and Economics Section of the American Statistical Association*, 215–224.

Thorp, E. O. (2006). The Kelly criterion in blackjack, sports betting and the stock market. In S. A. Zenios and W. T. Ziemba (Eds.), *Handbook of asset and liability management*, Handbooks in Finance, pp. 385–428. North Holland.

Ziemba, R. E. S. and W. T. Ziemba (2007). *Scenarios for Risk Management and Global Investment Strategies*. Wiley.

Ziemba, W. T. (2003). *The stochastic programming approach to asset liability and wealth management*. AIMR, Charlottesville, VA.

Ziemba, W. T. (2005). The symmetric downside risk Sharpe ratio and the evaluation of great investors and speculators. *Journal of Portfolio Management Fall*, 108–122.

Ziemba, W. T. (2009). Utility theory for growth versus security. Working Paper.

Ziemba, W. T. and D. B. Hausch (1986). *Betting at the Racetrack*. Dr. Z. Investments, San Luis Obispo, CA.