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

Professor William T Ziemba
Alumni Professor of Financial Modeling and Stochastic Optimization
(Emeritus), UBC

Advanced Finance and Stochastics
June 2013
Moscow

The two trading routes available for cash, equity and equity Futures hedge funds

- The search for positive α and the generation of returns from β risk in smart index funds and active equity portfolio management

$$r_i = \alpha_i + \beta_i M$$

α_I=excess mean return

β_i=leveraging factor for long market exposure

- The search for absolute returns using research on market imperfections, security market biases, mispriced derivative securities, arbitrage, risk arbitrage, and superior investment criteria

$$r_i = \alpha_i$$

- The strategy is to win in all markets - up, down and even, and to achieve a smooth wealth path with few drawdowns.

To win consistently: you must

1. Get the "mean right" that is the direction of the market - adjusted for various types of hedging and positioning. This must consider your risk tolerance.
2. You must diversify so that regardless of the market path (scenario) the positions are not overbet.
3. You must bet (portfolio allocate) well.

I argue that getting the mean right is **the most** important ingredient in winning strategies.

This was especially crucial in the equity markets the past 10+ years: the mean was frequently negative.

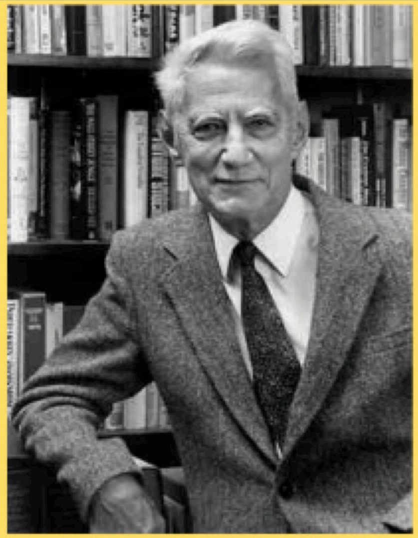
This talk focuses on #3 but I will try to cover #1 and #2 in various applications as we discuss the theory.

Log Utility: Bernoulli (1738)

- Daniel Bernoulli was a professor at the University of Basel
- Up to that time, the utility of wealth was simply the amount received, so utility was linear
- At the time the unit of exchange in Europe was the ducat, Michelangelo was paid in ducats for painting the ceiling of the Sistine Chapel
- Bernoulli's 1738 paper in Latin had two major contributions: the St Petersburg paradox and log utility
- Bernoulli postulated that marginal utility must decline with increasing wealth
  - ✧ Specifically: marginal utility declines not with wealth but with the reciproal of wealth
  - ✧ So $u'(w)=1/w$ or $u(w)=\log w$
- Bernoulli's paper was translated into english in 1954 and it is the first paper in my Kelly book with MacLean and Thorp (slide 20)

The great pioneers in capital growth investing

Daniel Bernoulli, 1700-1782



Claude Shannon, 1916-2001

John Kelly, 1923-1965

Buffett
Samuelson
Thorp

Brief Synopsis

- 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.
- But the elegant results on the Kelly (1956) criterion, as it is known in the gambling literature and the capital growth theory as it is known in the finance, economics and investments literature, see the survey by Hakansson and Ziemba (1995) and MacLean and Ziemba (2006), that were proved rigorously by Breiman (1961) and generalized by Algoet and Cover (1988) are long run asymptotic results.
- However, the Arrow-Pratt absolute risk aversion of the log utility criterion is essentially zero, where u is the utility function of wealth w,, and primes denote differentiation.
- The Arrow-Pratt risk aversion index.

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

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.
- Indeed I agrue it is the **most risky** utility function one should **ever consider**

Maximizing long run exponential growth is equivalent to maximizing the expected log of one period's final wealth (Kelly, 1956)

Each trial is win or lose 1 with probabilities p+q=1

Bernoulli Trails (p,q)

$X_N = (1+f)^M (1-f)^{N-M} X_0$

f = fraction bet

X_0 = initial wealth

win M of N trials

Exponential rate of growth

$$G = \lim_{N \to \infty} \log \left(\frac{X_N}{X_0} \right)^{1/N}$$

$$G = \lim_{N \to \infty} \left[\frac{M}{N} \log(1+f) + \frac{N-M}{N} \log(1-f) \right]$$

$$= p \log(1+f) + q \log(1-f) \qquad \text{by SLLN}$$

$$= E \log W$$

Maximizing long run growth

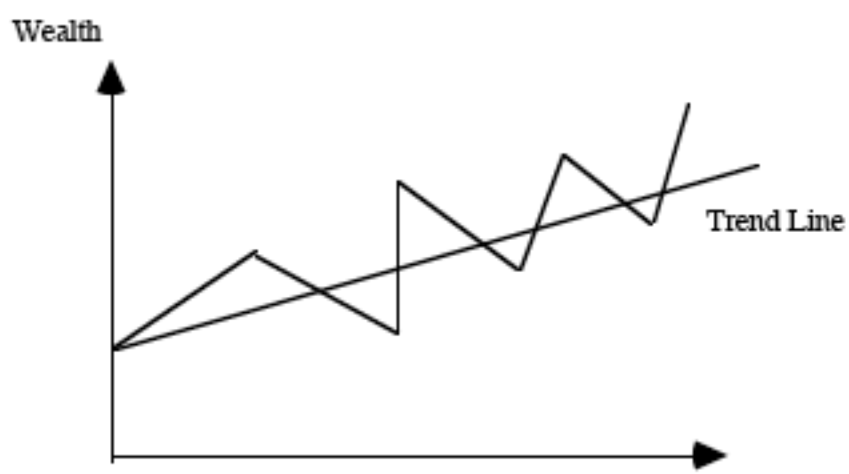
- Thus the criterion of maximizing the long run exponential rate of asset growth is equivalent to maximizing the one period expected logarithm of wealth.
- So an optimal policy is myopic - future and past do not affect current optimal decisions
- Max $G(f) = p \log (1+f) + q \log (1-f) \quad \rightarrow \quad f^* = p-q$
- The optimal fraction to bet is the edge p-q (the mean)
- So if the edge is large, the bet is larger
- $p = .99$, $q = .01 \rightarrow f^* = 98\%$ of wealth

The bets can be large: $\frac{u''}{u'} = \frac{1}{w} \approx 0$

p	.5	.51	.6	.99
q	.5	.49	.4	.01
f*	0	.02	.2	.98

$f^* = \frac{\text{edge}}{\text{odds}} = \frac{\text{edge}}{10}$ with 10-1 situation

The key to the size of the bet is not the edge, it is the risk



This formula is only correct for the one asset case; for n-assets it is a stochastic nonlinear programme.

Slew O’ Gold, 1984 Breeders Cup Classic

f*=64% for place/show; suggests fractional Kelly.

Warren Buffett (Berkshire Hathaway) and George Soros (Quantum Fund) behave as if they were full Kelly bettors

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

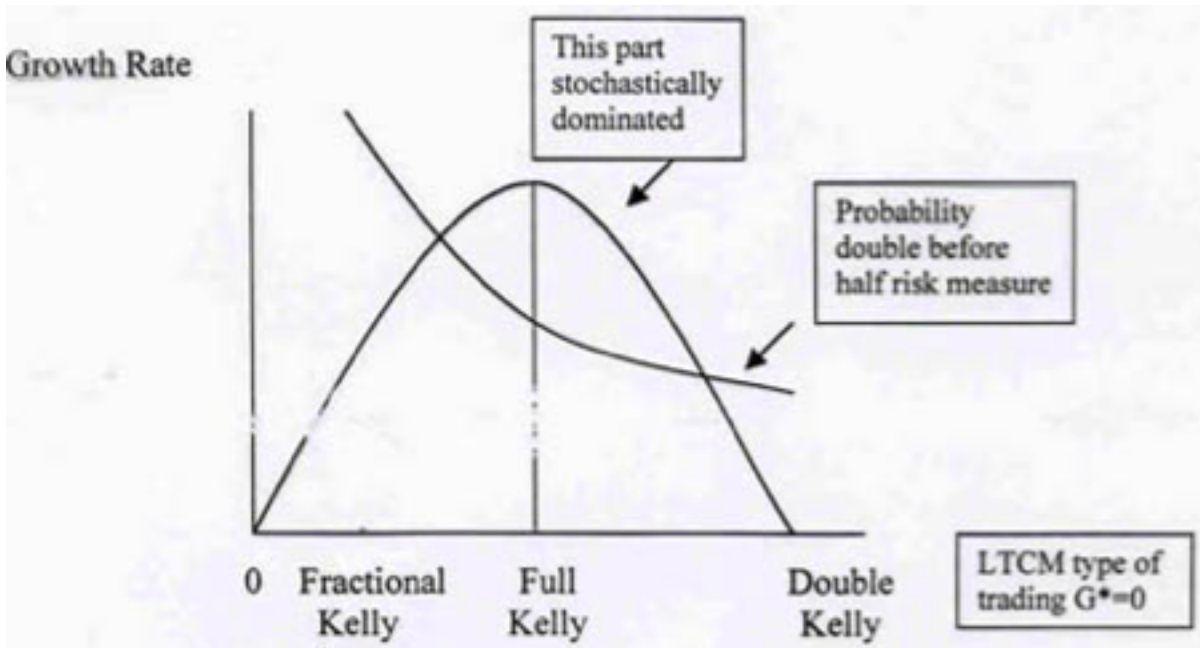
Berkshire Hathaway

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

The characteristics of Kelly bets

- Large positions in the few very best investments
- Not much diversification
- Typically a violent wealth path with many ups and downs
- Usually many monthly losses
- But usually but not always very high final wealth
- The possibility of low final wealth even without leveraging and bankruptcy with leveraging
- No matter how good the bets are and how long you play, you cannot be guaranteed to make large gains or even break even

What does the theory tell us about long term hedge fund trading and overbetting?



Kelly and fractional Kelly - explaining the overbetting that leads to hedge fund disasters: you cannot ever bet more than full Kelly and usually you should bet less

Fractional Kelly blends full Kelly with cash so it reduces the bet size but the growth rate is less

Kelly and half Kelly medium time simulations: Ziemba-Hausch (1986)

Simulation 700 investments, 1000 simulated runs $w_0 =$ \$1000

These were independent

Probability of Winning	Odds	Expected Return	Likelihood of Being Chosen in the Simulation	F*
0.57	1-1	1.14	0.1	.14
0.38	2-1	1.14	0.3	.07
0.285	3-1	1.14	0.2	.0475
0.228	4-1	1.14	0.2	.035
0.19	5-1	1.14	0.1	.028

The good, the bad and the ugly

	Final Wealth				Number of times the final wealth out of 1000 trials was				
Strategy	Min	Max	Mean	Median	>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

166 times out of 1000 the final wealth is more than 100 times the initial wealth with full Kelly but only once with half Kelly does the investor gain this much

But the probability of being ahead is higher with half Kelly, 87% vs 95.4%

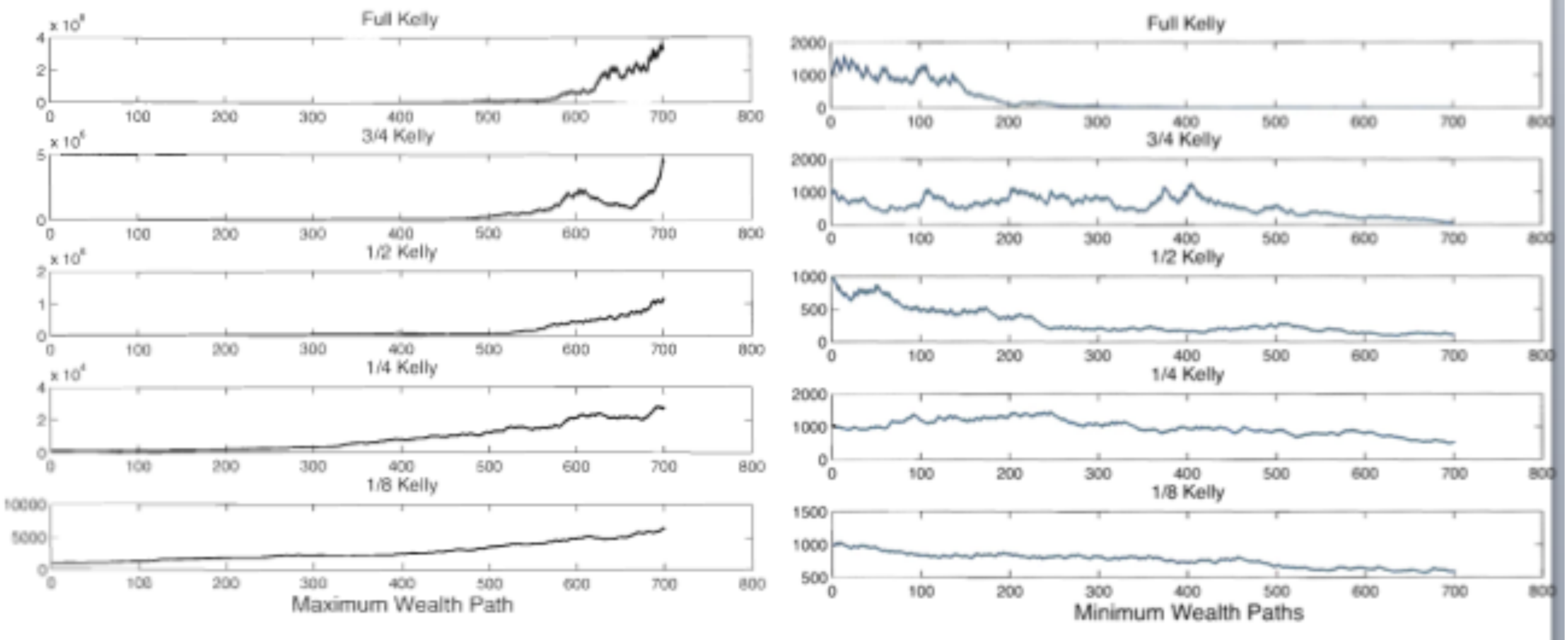
Min wealth is 18 and only 145 with half Kelly

700 bets all independent with a 14% edge: the result, you can still **lose** over 98% of your fortune with bad scenarios

With half Kelly, lose half of wealth only 1% of the time but is is 8.40% with full Kelly

So even after 700 plays, the strategy is still risky

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



Highest

Lowest

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

Mohnish Pabrai, investing in Stewart Enterprises - Thorp (2010) in our Kelly book

Hedge fund manager won bidding for 2008 lunch with Warren Buffett for \$600K+

Stewart Enterprises, Payoff ≤ 24 months

Prob	Net Return
0.8	>100%
0.19	zero
0.01	lose all investment

Pabrai bet 10% of his fund

What's the full Kelly bet?

f^* =0.975; half Kelly 0.3875; quarter Kelly=0.24375

Pabrai bet issues: should he have bet more?

Other opportunities: must compute against all options (nonlinear or stochastic optimization) for the available wealth

Risk tolerance: what fractional Kelly to use?

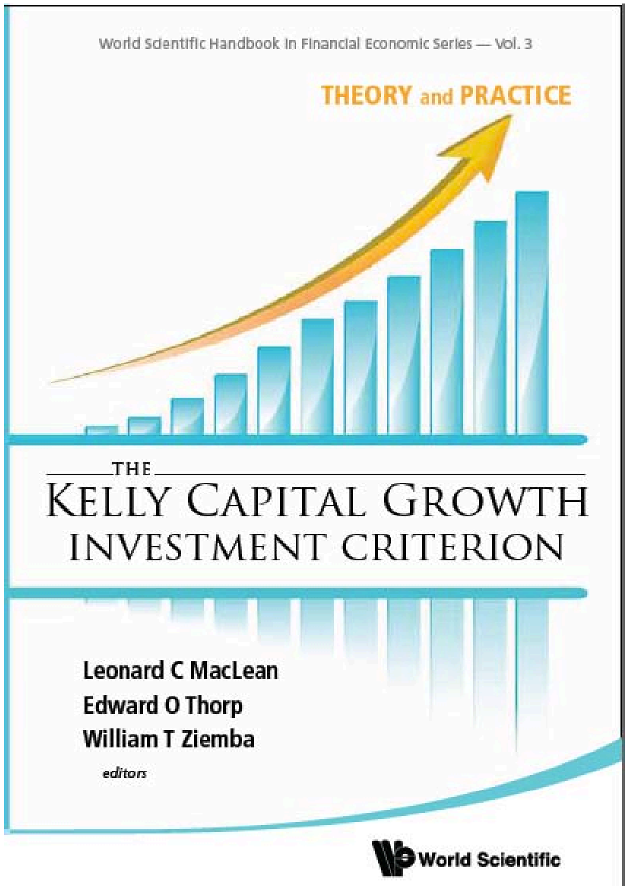
Black Swans: we call them bad scenarios

Long vs short run planning

Bill Gross, the world's largest bond trader: 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 Company's (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 it's $200 or $1 trillion … Professional blackjack is being played in this trading room from the standpoint of risk management and that's a big part of our success.".



Full treatment in Kelly book with Ed Thorp has major literature and new papers

Also discussion in Simulations *Journal of Portfolio Management*, Summer 2011 and Good and Bad, *Quantitative Finance*, September 2011

Easier book to read is Ziemba and Ziemba (2013) *Investing for the Modern Age,* World Scientific (with color and paperback)

The Classic Breiman Results

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.

In continuous time

$$f^* = \frac{\mu - r}{\sigma^2} \approx \frac{\text{edge}}{\text{risk (odds)}}$$

$$g^* = \frac{1}{2}\left(\frac{(\mu - r)^2}{\sigma^2}\right) + r$$

$$= \frac{1}{2}(\text{Sharpe Ratio})^2 + \text{risk free asset}$$

The plan of this talk is a follows:

Section 2 described 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 focuses on one particular type of application, namely, professional racetrack betting with some applications and endorsements on other topics

The Samuelson paper has other things as well including applications and endorsements and information relevant to the actual use of Kelly and fractional Kelly strategies in practice.

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).

We will look at his four major objections.

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 $\alpha w^{\alpha}, \quad \alpha < 0$ successfully.

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

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

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

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

$\overline{\rho}$ 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

Many great investors use $E \log w$ and $E[w^{\alpha}]$, $\alpha < 0$, successfully

But this does not mean that they will have optimal policies for the 2^{nd} and 3^{rd} approaches or will win always in finite time

Indeed $E \log$ betting can yield substantial losses

Markowitz Response to Samuelson

Markowitz (1976, 2006) adds to Luenberger's analysis in the simple utility function case.

Assuming iid investments in discrete time, as Luenberger did, he shows

> 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,

> 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 **A** by saying that portfolio P does better than portfolio Q in the long run (Markowitz (2006, p 256)

where **A** 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 (2010).

So what do we conclude on this Samuelson objection? We can just dismiss it based on the Luenberger and Markowitz results as its maximizing the wrong quantity?

Or we can say, yes,he's 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 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 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 the 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

> 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.

Tracing elements of syndicate betting procedures in professional racetrack betting

- My articles Hausch, Ziemba and Rubinstein (1981) and Hausch and Ziemba (1985) in *Management Science* presented a system with positive expectation for place and show betting at racetracks using behavioral biases and weak market inefficiency
- That provided an edge which could then be bet on using Kelly capital growth ideas
- The essence is to use probabilities from a simple market win in more complex markets to place and show and make repeated bets when there was a weak market inefficiency in these place and show markets
- This was a perfect application of Kelly betting techniques that maximize the expected log of final wealth as it is the approximation to an infinite sequence of bets
- The optimization proceeds by solving a sequence of one period nonlinear stochastic programs that take into account the effect of our bet on the odds

The books and calculator

- The trade books *Beat the Racetrack (BTR)* (1984, revised 1987) and *Betting at the Racetrack (BATR)* (1986) made this available to the general public,
- There is some game theory to grab the edge
- In BTR, following the 1985 paper we compute how many people can be playing the system optimally before the advantage is gone
- BATR was written to respond to the request

  write me a book for the person who does not know the front end from the back end of a horse but wants to win
- This system changed the way the racetrack was considered treating it as a financial market
- Only prices are involved to estimate everything
- The public's odds to win are used to generate win probabilities
- A calculator computes the mean return and the Kelly bet through regression approximations
- This system is still used today and we have a program searching 80 North American racetracks for bets

The market in 2013

- It is more difficult because of
  1. Intense competition from other syndicates near the finish of betting. These syndicates rely heavily on the book Hausch, Lo and Ziemba (1994, 2008) which has key papers with ideas for betting;
  2. cross track betting (as high as 87%) where bets from one racetrack are on races at another racetrack and all the moneys are in one pool so about 50% of the bets enter the pools after the race is being run and thus the odds/probabilities change and must be forecast;
  3. rebates are available and bets return some of the track take so the effective track take is about 10-15% rather than 13-30%
  4. betting exchanges like betfair where you can short as well as go long bets influence the odds but are separate pools so arbitrage is possible
- The second area is other more complex wagers and the use of factor models and other ways to estimate means … that is probabilities and how to get P_{ijkl} for a superfecta from P_i, etc
- This estimation is by far the most difficult issue
- Some bets like place and show are high probability low payoff and some are low probability high payoff
- One bet in Hong Kong has 48 million possible winners
- Some of these bets use a tree based modified Kelly approach with multiple equal valued bets where you bet more as in Kelly on the higher probability outcomes

Racetrack betting is simply an example of portfolio analysis

The investment horizon is short.

There are many types of bets

• High probability low payoff	win
• Low probability high payoff	pick 3-6
• One horse is involved	win
• Two horses are involved	place, exacta
• Three horses are involved	show, triactor
• Four horses are involved	superfecta
• Two races are involved	double
• Three races are involved	pick 3
• Four races are involved	pick 4
• Five races are involved	pick 5
• Six races are involved	pick 6
• N races are involved	place pick all

In all cases the strategy to win is the same:

- Get the mean right - probabilities here
- Bet right - Kelly betting useful
- All can be modeled as stochastic programs or network trees to approximate SPs with equal bets

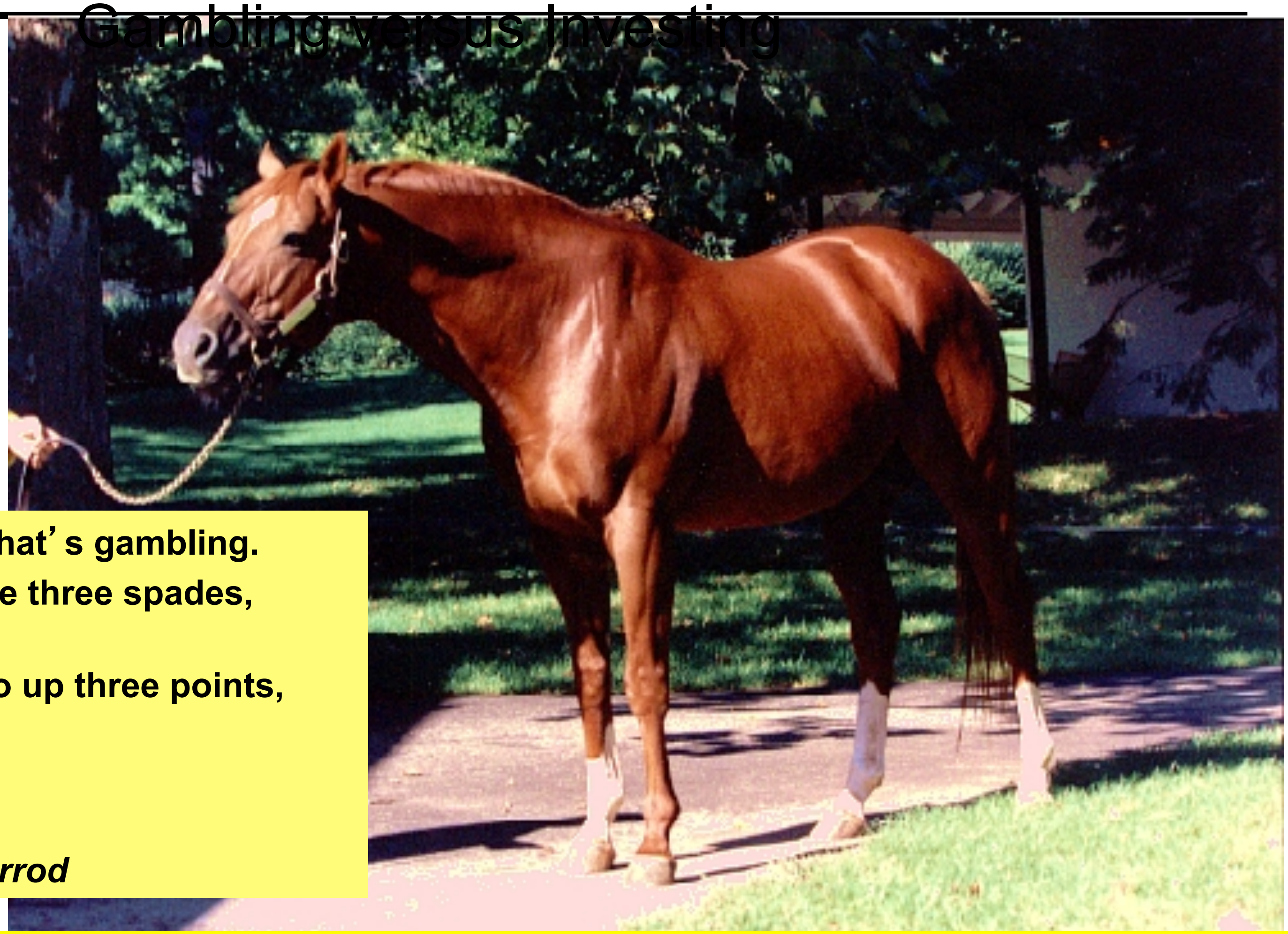
Racetrack betting is simply an application of portfolio theory

- The racetrack offers many bets that involve the results of one to about ten horses.
- Each race is a special financial market with betting then a race that takes one or a few minutes.
- Unlike the financial markets, one cannot stop the race when one is ahead, except possibly in Betfair.
- There is a well-defined end point.
- In the US internet gambling is illegal but racetrack betting has an exemption.
- In other countries, it is legal and active in other wagers.
- Professional syndicates or teams have been successful as small hedge funds with gains approaching one billion for the most successful.
  - Their models are **better** than Goldman Sachs in the financial markets!
  - Sophisticated factor models and good optimization
- I have been involved in this research since the late 1970s with five books and a number of articles, consulting, betting, etc.
- I will try to relate the theory, computations and examples of real races and experiences for various bets such as win, place and show, exactas, triactors, superfectas, super hi five, place pick all, double, pick 3, 4, 5 and 6 (see chapter 9 of Gassman and Ziemba, 2012).
- Various important races are on www.chef-de-race.com a wonderful website that my colleague Dr Steve Roman maintains

Investing in traditional financial markets has many parallels with racetrack & lottery betting

- Behavioral anomalies such as the favorite-longshot bias are pervasive and also exist and are exploitable in political elections, sports betting, etc
- In the S&P500 and FTSE100 futures and equity puts and calls options markets biases favor buying high probability favorites and selling low probability longshots.
  - Not in this lecture but I use this in hedge funds and personal trading
- In complex exotic wagers such as the Pick 6, the bias is to overbet the favorite so one must include other value wagers in the betting program.
- Fundamental information such as breeding is important and is especially useful for the Kentucky Derby and Belmont Stakes.

Gambling versus Investing



If you bet on a horse, that's gambling.

If you bet you can make three spades, that's entertainment.

If you bet cotton will go up three points, that's business.

See the difference?

Blackie Sherrod

Secretariat at Claiborne Farm, Paris, Kentucky

1973 Belmont -consiered possibly the greatest race ever is on www.chel-de-race.com

Storm Cat at Overbrook : 166 breeds @ $500k each = $58 million/yr

The 5 most expensive yearlings in the Keeneland fall 2005 sale were all Storm Cat's: $9.7M, $6.3M, $5M, $4M, $3.5M - all bought by Shiek Mohammed of Dubai

Storm Cat: Secretariat never reproduced himself but became the top broodmare sire such as the dam of Storm Cat

Ghostzapper, sired by Awesome Again, the fastest horse since Secretariat, winner of the Breeders' Cup Classic 2004 and provider of the superfecta for WTZ

4 $5 tickets = $20 bet:

$5000 payoff

ABCD

BACD

ABDC

BADC

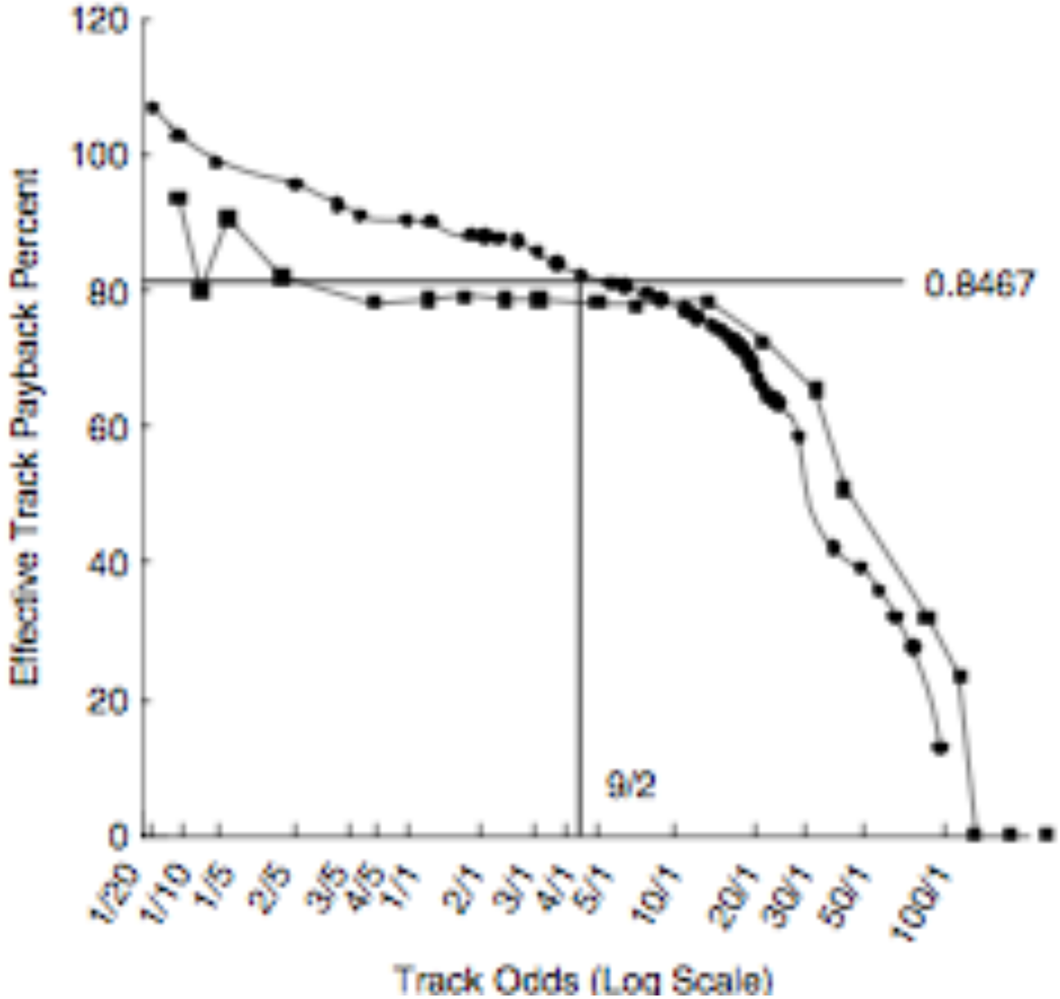
Favorite-Longshot bias at racetracks and in other gambling events

- Behavioral finance
  Kahneman-Tversky (1979)
  low probability events are overestimated
  high probability events are underestimated
- More bragging rights from picking longshots than from favorites
  50-1 *wow, was I smart*
  2-5 *easy pick*
- Transactions costs
  bet $50 to win $10 it's hardly worth the effort
  - a 1-10 horse having more than a 90% chance of winning has an expected value of about $1.03 (for every $1 bet)
  - a 100-1 horse has only has an expected value of about 14 cents per dollar invested. The fair odds are about 700-1 not 100-1.
- Early literature goes back to at least 1949
  US, British, Asian, etc classic papers reprinted in Hausch, Lo and Ziemba (1994), *Efficiency of racetrack betting markets* (Academic Press) 2nd editon (2008), World Scientific.
- This book became a cult item, the bible for new racetrack syndicates and sold for as high as US $12,000
- Some wanted to buy all copies and destroy all copies not held by them to eliminate all competition
- New papers in Hausch and Ziemba (2008) *Handbook of Sports and Lotto Investments,* North Holland

Update to era with rebates and betting exchanges

Effective track payback less breakage for various odds levels in California.

Source: Ziemba and Hausch (1986) and Ziemba (2008). Right side similar to past but short odds horses have lost their advantage



A typical use of the favorite-longshot bias is to correct probabilities for first, second and third.

For first

$$q_i = \frac{Q + \Delta Q(O_i)}{O_i}$$

where
q_i is the win fraction bet on i, namely W_i/W the estimate that i wins the race from the win pool,
Q is the track payback and
$\Delta Q(O_i)$ is the favorite-longshot bias correlation at US odds O_i to 1.

Harville

The Harville formulas

$$q_i = \frac{q_i q_j}{(1-q_i)} \quad \text{and} \quad q_{ijk} = \frac{q_i q_j q_k}{(1-q_i)(1-q_i-q_j)}$$

are the probabilities of ij and ijk finishes assuming that $q_j/(1-q_i)$ is the probability that j wins a race that does not contain i.

Similarly,

$$\frac{q_k}{(1-q_i-q_j)}$$

is the probability that k wins a race that does not contain i or j.

These formulas are handy but are biased because favorites do not come in second and third as much as these formulas indicate.

The correction

A correction can be used to decrease the second and third probabilities for the favorites and increase it for the longer priced horses. This is called discounted Harville. One decays q_i by a coefficient α about 0.81. So

$$r_i = \frac{q_i^\alpha}{\sum q_i^\alpha}$$

are used instead of q_i for second and $\alpha = (0.81)^2 \cong 0.64$ is used for third.

So

$$q_{ij} = \frac{q_i r_j}{(1 - r_j)} \quad \text{and} \quad q_{ijk} = \frac{q_i r_j s_k}{(1 - r_i)(1 - s_i - s_j)}$$

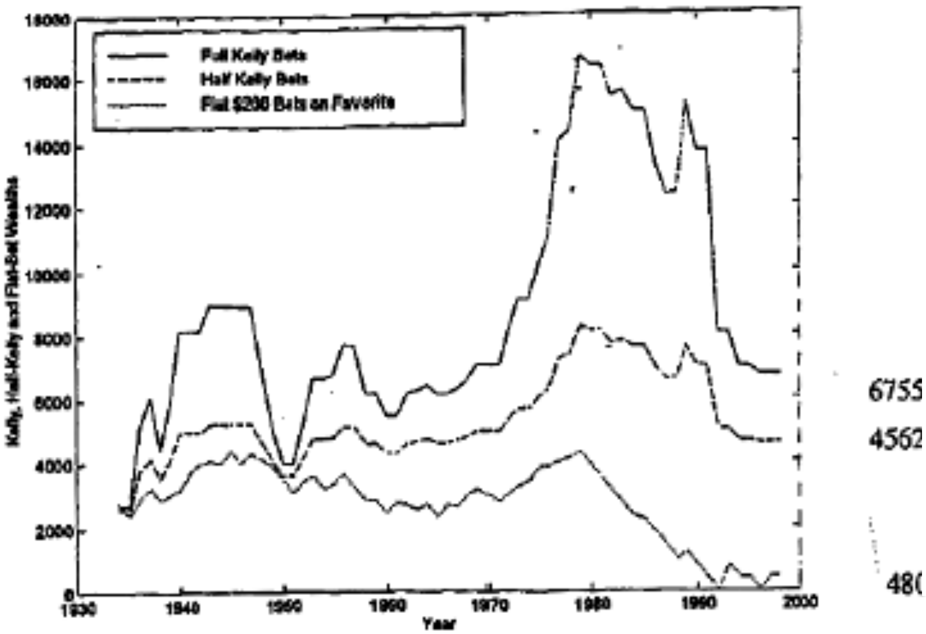
where the

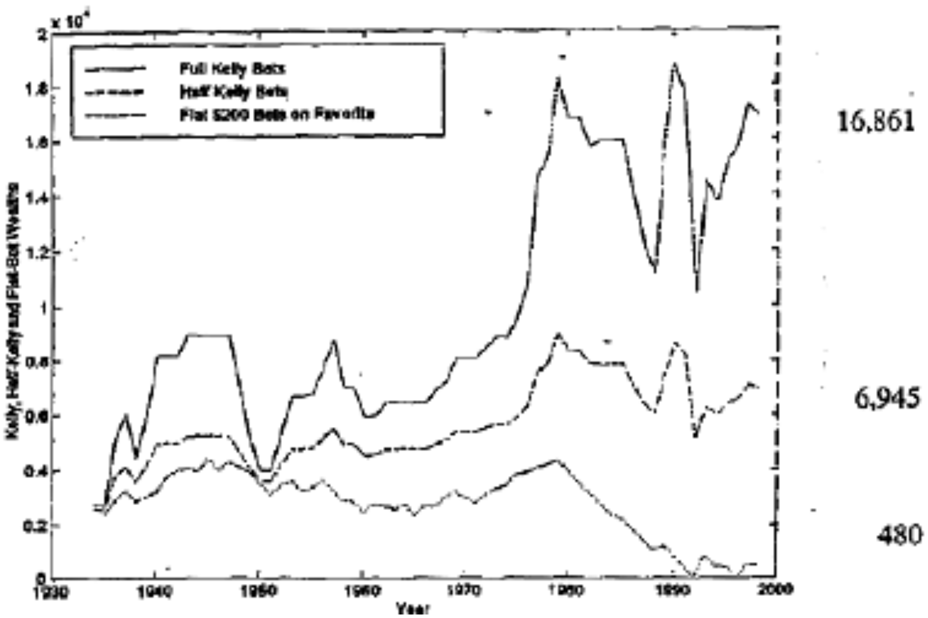
$$s_i = \frac{q_i^{0.64}}{\sum q_i^{0.64}}.$$

Then these formulas given good ijk probabilities for various bets involving multiple horses.

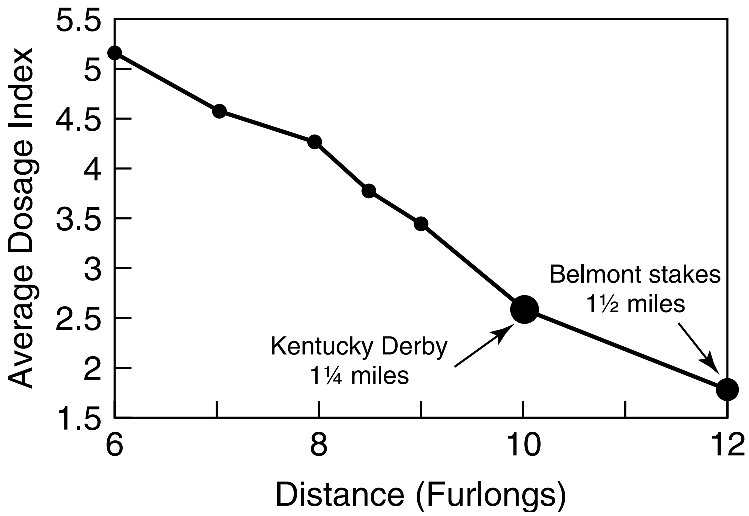
More on these biases and their correction is in Hausch, Lo, and Ziemba (2008).

Growth versus security





- Security rises if the bet sizes are decreased with fractional Kelly (negative power) but growth decreases as well. Half Kelly is $-w^{-1}$.
- Fractional Kelly $f=1/1-\alpha$, $\alpha<0$, in αw^{α} is exact for lognormal, approx otherwise. See MacLean-Ziemba, Time to Wealth (2005), MacLean, Thorp, Ziemba (2010) Kelly book.
- Half Kelly is half the Kelly wager plus cash.
- Right graph has dosage filter; do not bet on horses that cannot run 1 1/4 miles on the first Saturday in May.

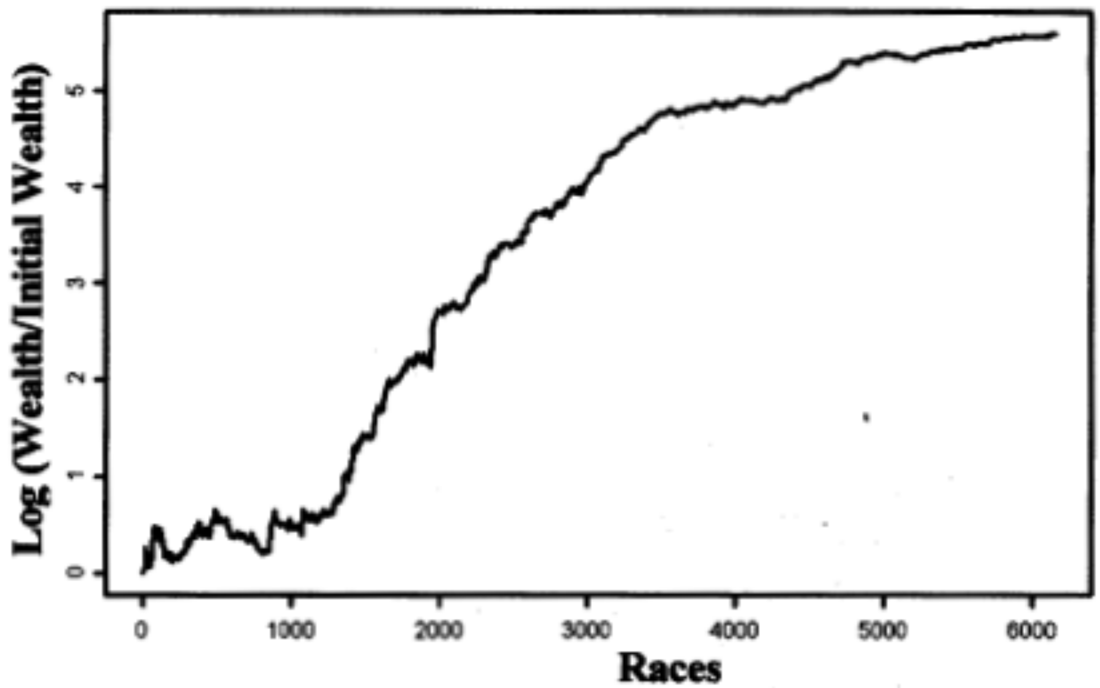


This figure shows the average dosage index or speed over stamina for the average winner of the $1\frac{1}{4}$ mile Kentucky Derby and $1\frac{1}{2}$ mile Belmont stakes as well as a large number of high quality races at different distances.

It was compiled years ago from race data of Steve Roman.

Observe that the winners of longer races have lower dosages which means more stamina and less speed.

Does it work? Update on a terrific, smooth record of Bill Benter.



Hong Kong racing syndicate to 2001.

Optimal capital growth model, assumes our bet influences the odds and we can bet on multiple horses. Exact transition costs.

$$\underset{\{p_l\}\{s_l\}}{\text{Maximize}} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} \frac{q_i q_j q_k}{(1-q_i)(1-q_i-q_j)} \log \left[\begin{array}{l} \dfrac{Q(P + \sum_{l=1}^{n} p_l) - \quad\quad + P_{ij})}{2} \\ \quad \times \left[\dfrac{p_i}{p_i + P_i} + \dfrac{p_j}{p_j + P_j} \right] \\ + \dfrac{Q(S + \sum_{l=1}^{n} s_l) - (s_i + s_j + s_k + S_{ijk})}{3} \\ \quad \times \left[\dfrac{s_i}{s_i + S_i} + \dfrac{s_j}{s_j + S_j} + \dfrac{s_k}{s_k + S_K} \right] \\ + w_0 - \sum_{\substack{l=1 \\ l \neq i,j,k}}^{n} s_l - \sum_{\substack{l=1 \\ l \neq i,j}}^{n} p_1 \end{array} \right]$$

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

Non concave program but it seems to converge.

Expected value approximation equations needed as 15 seconds to bet

$$\text{Ex Place}_{\text{i}} = 0.319 + 0.559\left(\frac{w_i / w}{p_i / p}\right)$$

$$\text{Ex Show}_{\text{i}} = 0.543 + 0.369\left(\frac{w_i / w}{s_i / s}\right)$$

- Notice expected value (and optimal wager) are functions of only four numbers - the totals and the horse in question.
- These equations approximate the full optimized optimal growth model.
- Solving the complex NLP: too much work and too much data for most people.
- This is used in the calculators.
- Track take etc in equations, Dr Z's *Beat the Racetrack*, 1987.

1983 Kentucky Derby

	Totals	#8 Sunny's Halo	Expected Value Per Dollar Bet	Optimal Bet (W_0=1000)
Odds		5-2		
Win	3,143,669	745,524		
Show	1,099,990	179,758	1.14	52

Sunny's Halo won the race

Win	Place	Show
7.00	4.80	4.00

$\Pi = \$52$

15 second bet!

Watch board in lineup
while everyone is at the TV

Watch this and other races on www.chef-de-race.com

1991 Breeders' Cup Race 5

Number	Entry	Name	CRC Pools Win	CRC Pools Place	CRC Pools Show	CRC Win Odds	CD Win Odds	CRC EXPL	CRC EXSH	CD EXPL	CD EXSH
1		Agincourt	2188	652	437	36.9	62.3	0.87	0.83	0.65	0.71
2		Pine Bluff	4510	1894	1182	17.4	27.5	0.70	0.75	0.56	0.67
3		Tri to Watch	21944	6135	3247	2.7	5.0	0.91	0.95	0.69	0.79
4		Bertrando	22483	5522	2967	2.6	2.5	0.99	1.00	1.04	1.04
5		Showbrook	4608	1904	1401	17.0	25.0	0.71	0.72	0.59	0.66
6		Star Recruit	2372	1165	937	34.0	36.5	0.64	0.67	0.63	0.6[illegible]
7		Snappy Landing	1792	870	780	45.3	60.7	0.65	0.66	0.57	0.62
8		Bag	1462	677	392	55.8	47.5	0.66	0.75	0.74	0.80
9		Big Sur	1105	502	581	74.2	69.0	0.67	0.63	0.71	0.64
10		Dance Floor	17241	6068	2829	3.8	3.7	0.78	0.90	0.80	0.92
11		Arazi	18674	3674	1657	3.4	2.1	1.16	1.24	1.55	1.57
12	e	Field	3113	1676	1373	25.7	18.5	0.59	0.62	0.73	0.69
			101,492	30,739	17,783						

ARAZI had about
1/6 of WIN Pool at CALDER
≈ 1/8 of place pool
≈ 1/11 of show pool

Ise win pools at CHURCHILL DOWNS TV PRICE
PLACE + SHOW ON ARAZI at CALDER

CRC Payoffs

	W	P	S
11	8.80	6.20	5.60
4		4.80	4.00
7			9.80

$2 Exacta: $34 40

CD Payoffs

W	P	S
6.20	4.80	4.80
	4 40	4.20
		14.40

$2 Exacta $22.60

Note CD EXPL and CD EXSH are based on CD win odds and CRC place and show pools

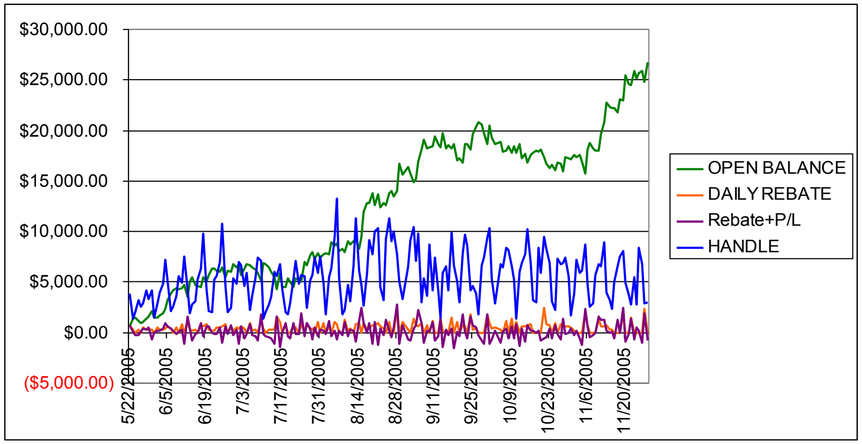
Courtesy Bill Stein Ron Dalhero

Separate pools no longer exist.

An application of real money bet with this system in 2004:

Racetrace betting record - place & show system:

- Initial wealth=US$5,000, total bet=US$1.5 million
- At each wager opportunity there is either no bet or a full Kelly bet using the model 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 race-track market combines bets made at many other racetracks and betting sites into one pool at the last minute.
- 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.



Low probability high payoff bets

- Favorites are usually underbet to win; this is useful in situations like the place and show systems. These are high probability low payoff situations.
- However, the situation is reversed in low probability high payoff situations.
- Then favorites are overbet because in exotics there are usually multiple races or multiple horses or both so that many tickets are needed to have a good chance of winning.
- Most bettors pick the favorites and overbet them so their tickets are not too expensive.
- Then they are most likely not to win and if they win, they get very little because they have picked the favorite so the payoff is very low.
- Need 4-7,000 in $2 tickets to have a good chance of winning.
- sls code to compute probabilities of winning, evaluate tickets, etc

An example: Consider the Pick 6
Santa Anita, March 6, 2002 with $202,790 carryover from Sunday's wagers

Race	3	4	5	6	7	8
Win Payoff	3.60	4.60	12.00	9.20	18.00	7.20

1-5 shot out of the money

$$\text{Advantage} \equiv \prod_{i=1}^{6}(1+\text{edge}_i)$$

Pick 6 $86,347.60 – one large transaction cost

Parlay $15,452.40 - six races, six smaller transactions costs

Bet = 1922	1 Pick 6		Score 6	$4
Rebate = 154	11 Pick 5/6	692.00	7	$58
Net bet = 1768	Gross =	93,961.80	8	$376
			9	$1484
				1922

Score 9: you win if the score is 9 or less, here 3 I's, 2 II's, total 9 so we won

Race	3	4	5	6	7	8
I	1, 3	3	7	3	2	2
II	7	1,5	1,2,5,6	4,7,8,10	1,4	5.9
III	5,6	2,4,8		1,2,5		1,6,7,8

8 was scratched

This $1922 ticket was actually 66 separate tickets.
The winning ticket will have one 6/6 winner and multiple 5/6 winners

			Morning Line
Race 7	1	Madame Pietra	3-1
	2	Love at Noon	3-5
	3	Harvest Girl	8-1
	4	Filigree	4-1
	5	Farah Love	12-1

- This was the behavioral key.
- Filigree, the third choice in the morning line went off at 8-1.
- The 3rd and 5th races back, he ran faster than the favorite Love at Noon ran in his last two races. So he had a chance to win and he did.
- Love at Noon went off at 1-5 and had most of the P5 money

Top number is final speed number, other numbers are pace within the race.

EQUIFORM ®

'TH SAX MARCH 6 - 6 FURLONGS DIRT

MADAME PIETRA (118) 40

70	72¾	67¾	70½	64¾	w76½	w71¾	64½
	/////	/////					
80¾	71¾	73¾	79	77¾	78	77	67
SA6½	SA-6½	SA8	DMR6	SA6	SA6	DMR6	HOL6
70¾	71	67¾					
82	70¾		81	83¾	81	63	68½
JAN24	10/17	09/27	SEP02	03/16	JAN31	AUG05	05/28

LOVE AT NOON (118) 44

72¾	73¾	74½	75½	w78¾	w^70½
71¾	82½	75¾	74¾	85	70¾
SA8½	SA7	HOL8½	CD8½	CD6½	SA6
69¾	74½	72	72¾	76¾	
	84			85¾	77¾
01/20	DEC29	06/16	05/26	05/05	04/07

HARVEST GIRL (118) 19

w71½	71½	68½	--	70¾	71	66¾	66¾	68¾	w71¾	68	69¾	w68½	69½	w67½	61½
	/////		/////					/////		/////			/////		
80½	77	80½	71½	83¾	81½	82½	79¾	81	78½	75¾	79½	77¾	82	79½	69¾
SA6	SA8	HOL6½	BM8½	FPX6½	SR6	SOL6	GG6	SA-6½	FPX6½	BMF8	SOL6	GG6	HOL5½	BM5½	BMF5½
	68¾	69	65¾	71½				69½	71½	68					
87		83		98	81½	82¾	80½	84½	83¾		78¾	77¾	75½	86½	80
02/14	01/05	11/21	10/13	09/14	08/06	JUL14	11/25	10/28	09/19	08/12	07/22	05/31	MAY18	09/12	08/15

FILIGREE (118) 40

71¾	67¾	74	71½	w74½	w71½	71½	71¾	70½	70	70	70½	w73½	w70¾	w70	68¾
77¾	72	82½	78½	75¾	74½	75¾	73¾	70½	81½	61½	77	74	77¾	78	73½
SA6½	SA5½	DMR6	DMR6½	HOL6½	HOL6	SA7	SA8½	SA6	SA5½	SA9	FPX6½	DMR6½	DMR6½	HOL6½	SA6
71½			70¾	72½		67¾	70			61½	68½	68	70	70	
77½	75½	85½	73¾	73½	74¾	75		67½	75½		74	68¾	77½	79	75¾
01/24	JAN02	09/05	07/23	06/27	05/28	04/16	03/03	01/31	JAN01	10/28	09/30	09/03	JUL31	05/26	04/16

FARAH LOVE (120) 26

73	w70¾	w70½	67	66½	65	63¾	67	62¾	w66	69½	w66
/////						/////	/////				
78¾	74½	79	72	72½	72	76¾	76¾	71½	76¾	74½	67
SA-6½	SA6	SA6½	HOL6	SA6	SA6	SA-6½	HOL5½	HOL8½	SA6½	SA6½	HOL6
71¾		69½				62½		65¾	65¾	69	
81¾	77½	82½	72	68½	74½	80	66½		84	70	70½
02/07	01/17	01/05	DEC09	02/16	01/31	12/30	12/20	11/23	11/01	OCT11	07/23

Approximate Kelly with more money on higher probability wagers.

$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 5 ,6 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 2 ,4 ,8 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 1 ,2 ,5 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 7 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 5 ,9
$2.00 ;P6; (3) 5 ,6 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 5 ,6 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 5 ,6 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 5 ,6 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 5 ,6 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 2 ,4 ,8 /(5) 7 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 2 ,4 ,8 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 2 ,4 ,8 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 2 ,4 ,8 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 2 ,4 ,8 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 1 ,2 ,5 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 1 ,2 ,5 /(7) 2 /(8) 2

$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 1 ,2 ,5 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 1 ,2 ,5 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 1 ,2 ,5 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 3 /(7) 2 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 1 ,6 ,7 ,8
$2.00 ;P6; (3) 7 /(4) 1 ,5 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 1 ,5 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 7 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 7 /(4) 3 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 2 ,3 ,5 ,6 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 1 ,5 /(5) 7 /(6) 3 /(7) 1 ,4 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 4 ,7 ,8 ,10 /(7) 1 ,4 /(8) 2
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 4 ,7 ,8 ,10 /(7) 2 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 2 ,3 ,5 ,6 /(6) 3 /(7) 1 ,4 /(8) 5 ,9
$2.00 ;P6; (3) 1 ,3 /(4) 3 /(5) 7 /(6) 4 ,7 ,8 ,10 /(7) 1 ,4 /(8) 5 ,9

		Amount bet
Score 6	$2	$4
Score 7	$2	$58
Score 8	$2	$376
Score 9	$2	$1,484
Score 10	$2	$0
Score 11	$2	$0
Total 0.54%		$1,922

Bernoulli's Second Contribution

Bernoulli's second idea in his 1738 paper was his contribution to to the St Petersburg paradox. This problem actually originates from Daniel Bernoulli's cousin, Nicolas Bernoulli, professor at the University of Basel. In 1708, he submitted five important problems to Professor Pierre Montmort, one of which was the St Petersburg paradox.

The idea is to determine the expected value and what you would pay for the following gamble:

> A fair coin with $\frac{1}{2}$ probability of heads is repeatedly tossed until heads occurs, ending the game. The investor pays c dollars and receives in return 2^{k-1} with probability 2^{-k} for $k = 1, 2, \ldots$ should a head occur. Thus, after each succeeding loss, assuming a head does not appear, the bet is doubled to 2, 4, 8, . . . etc. Clearly the expected value is $\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots$ or infinity with linear utility.

Bell and Cover (1980) argue that the St Petersburg gamble is attractive at any price c, but the investor wants less of it as $c \to \infty$.

The proportion of the investor's wealth invested in the St Petersburg gamble is always positive but decreases with the cost c as c increases. The rest of the wealth is in cash.

Bernoulli offers two solutions since he feels that this gamble is worth a lot less than infinity.

In the first solution, he arbitrarily sets a limit to the utility of very large payoffs. Specifically, any amount over 10 million is assumed to be equal to 2^{24}.

Under that assumption, the expected value is

$$\frac{1}{2}(1)+\frac{1}{4}(2)+\frac{1}{8}(4)+\ldots+\left(\frac{1}{2}\right)^{24}(2^{24})+\left(\frac{1}{2}\right)^{25}(2^{24})+\left(\frac{1}{2}\right)^{26}(2^{24})+\ldots = 12+\text{ the original 1 } = 13.$$

If utility is $\sqrt{w}$, the expected value is

$$\frac{1}{2}\sqrt{1}+\frac{1}{4}\sqrt{2}+\frac{1}{8}\sqrt{4}+\ldots=\frac{1}{2-\sqrt{2}}\cong 2.9.$$

When utility is log, as Bernoulli proposed, the expected value is

$$\frac{1}{2}\log 1+\frac{1}{4}\log 2++\frac{1}{8}\log 4+\ldots=\log 2=0.69315.$$

Use of a concave utility function does not eliminate the paradox.

For example, the utility function $U(x) = x/\log(x + A)$, where $A > 2$ is a constant, is strictly concave, strictly increasing, and infinitely differentiable yet the expected value for the St Petersburg gamble is $+\infty$.

As Menger (1934) pointed out, the log, the square root and many other, but not all, concave utility functions eliminate the original St Petersburg paradox but it does not solve one where the payoffs grow faster than 2^n.

So if log is the utility function, one creates a new paradox by having the payoffs increase at least as fast as log reduces them so one still has an infinite sum for the expected utility.

With exponentially growing payoffs one has

$$\frac{1}{2}\log(e^1) + \frac{1}{4}\log(e^2) + \ldots = \infty$$

The super St Petersburg paradox, in which even $ElogX = \infty$ is examined in Cover and Thomas (2006: p181, 182) where a satisfactory resolution is reached by looking at relative growth rates of wealth.

Another solution to such paradoxes is to have bounded utility.

For example, as Bernoulli suggested above 10 million.

To solve the St. Petersburg paradox with exponentially growing payoffs, or any other growth rate, a second solution, in addition to that of bounding the utility function above, is simply to choose a utility function which, though unbounded, grows ”sufficiently more” slowly than the inverse of the payoff function, e.g. like the log of the inverse function to the payoff function.

The key is whether the valuation using a utility function is finite or not; if finite, the specific value does not matter since utilities are equivalent to within a positive linear transformation $(V = aU + b, a > 0)$.

So for any utility giving a finite result there is an equivalent one that will give you any specified finite value as a result.

Only the behavior of $U(x)$ as $x \to \infty$ matters and strict monotonicity is necessary for a paradox.

For example, $U(x) = x,\ x \leqslant A$ will not produce a paradox.

But the continuous concave utility function

$$U(x) = \frac{x}{2} + \frac{A}{2},\ x > A$$

will have a paradox. Samuelson (1977) provides an extensive survey of the paradox. See also Menger (1967) and Aase (2001).