

ON THE OPTIMAL GROWTH RATE STRATEGY: LONG TERM OPPORTUNITIES AND SHORT TERM RISK

Abstract

This paper investigates in depth the optimal growth rate strategy as an optimal long-term investment strategy. We show that it is possible, for an arbitrary probability distribution of stock returns, to choose the optimal equity proportion of a long-term portfolio independently of the investor's preferences. We discuss the validity of the asymptotic results when the horizon is finite and show that the suggested criterion performs well. We then examine the risk associated with such a portfolio. We argue that traditional risk measures are ill-suited for long term investment problems and propose to use another measure, the drawdown from maximum, which quantifies the risk that an investment has to be ended prematurely while the market conditions are poor.

Résumé

Cet article traite des stratégies optimales de long terme, et en particulier de la stratégie de croissance optimale. Nous montrons qu'il est possible de choisir la proportion optimale d'un portefeuille de long terme à investir en actifs risqués de façon totalement indépendante des préférences des investisseurs et pour une distribution des rentabilités des actifs risqués quelconque. Nous discutons la validité de ces résultats asymptotiques lorsque l'horizon d'investissement est fini et nous montrons que le critère retenu est efficace. Puis, nous examinons le risque associé à ce portefeuille. Nous montrons que les mesures traditionnelles de risque sont peu adaptées aux problèmes d'investissement sur le long terme et proposons d'utiliser le *drawdown* par rapport au maximum, qui quantifie le risque portant sur une position de long terme qui doit être retournée prématurément.

ON THE OPTIMAL GROWTH RATE STRATEGY: LONG TERM OPPORTUNITIES AND SHORT TERM RISK

*Vivien Brunel*¹
*Jérôme Legras*²

Abstract

This paper investigates in depth the optimal growth rate strategy as an optimal long-term investment strategy. Following Kelly (1956), we show that it is possible, for an arbitrary probability distribution of stock returns, to choose the optimal equity proportion of a long-term portfolio independently of the investor's preferences. The resulting portfolio is the optimal growth rate portfolio. We discuss the validity of the asymptotic results when the horizon is finite and show that the suggested criterion performs well. We then examine the risk associated with such a portfolio. We argue that traditional risk measures are ill-suited for long term investment problems and propose to use another measure, the drawdown from maximum, which quantifies the risk that an investment has to be ended prematurely while the market conditions are poor. We then present some empirical investigations on the American stock markets.

¹ Financial engineer at the Department of Research and Innovation at CCF Group. Mailing address: CCF - DRI-DEFE, 75419 Paris Cedex 08, France. Phone: (+33) 1 40 70 79 78. Fax: (+33) 1 40 70 30 31. E-mail: vbrunel@ccf.fr.

² Financial engineer at the Department of Research and Innovation at CCF Group. Mailing address: CCF - DRI-DEFE, 75419 Paris Cedex 08, France. Phone: (+33) 1 40 70 34 89. Fax: (+33) 1 40 70 30 31. E-mail: jlegras@ccf.fr.

We thank François Longin and Nicolas Gaussel for stimulating discussions and helping to improve the manuscript. We are also grateful to Pr. Dumas and Pr. Ziemba as well as all the participants of the Fall 2000 Inquire seminar for their helpful comments. We thank the anonymous referee for his critical reading of the manuscript.

1. INTRODUCTION

Optimal long-term investment strategies have been investigated in two very different ways in the financial literature. Many authors have chosen to solve the optimal portfolio selection problem (with or without consumption) on an arbitrary finite horizon and to analyze the case of large horizons as a special case¹. This approach is grounded on the seminal work of Markowitz (1952) in the mono-periodic case, of Merton (1971) in the continuous time case and of Samuelson (1969) in the discrete time case and has culminated in the book by Merton (1990). Important contributions can also be found in Cox and Huang (1989), Bajeux-Besnainou and Portait (1998) or Karatzas (1989). The second approach consists in specifically analyzing long term investments, with no short term constraints, and in looking for criteria which are optimal for such investments, though they are clearly not optimal on arbitrary (possibly short) horizons². This way of tackling the problem was historically pioneered by Bernoulli³ and later, from a more financial point of view, by Kelly (1956) and Breiman (1961). These authors show that the optimal portfolio is the optimal growth rate portfolio and that it should be used for any long-term investment. Such optimal long-term strategies have been analyzed recently by Baviera et al. (1999) or Maslov and Zhang (1998a).

With the noticeable exception of Kelly (1956), whose approach is grounded on Shannon's information theory, all these works share the same general framework: specific dynamics of the assets are introduced on a finite horizon and the investor builds an optimal strategy that maximizes her expected utility. Some constraints, such as defined contributions for a pension fund (Taillard and Boulier (2000)), consumption problems (see Merton (1990) and numerous references therein), inflation hedging as in Campbell and Viceira (2000), are often added to the model but do not alter its core and the optimization process focuses on utility. Even the optimal growth rate portfolio can usually be recovered as the portfolio of an investor with logarithmic utility.

The major drawback of the utility-based approach is that the utility function is usually unknown. In incomplete markets there is no such thing as a representative agent and one has to consider the specific utility of each investor when building an optimal strategy. There is yet no consensus on the way that one can recover the utility function from empirical observations. In fact, since the work of Allais, the use of expected utility as a way of representing agent's preferences has been largely discussed by many authors.

The use of the expected utility in long term investment can also be questioned for other reasons. It is now well known that the maximization of some expected utility functions can be, on the long term, unreasonable and should not be used by any sensible investor. The first contribution of this paper is to show that indeed, on the long term, some utility functions should not be used. More precisely, we show that when the investment horizon goes to infinity, the optimal growth rate portfolio is almost surely superior to any other strategy, independently of the investor's preferences. The universal properties of this portfolio are thus described in the first section. Though we provide conclusive evidence that when the investment horizon tends to infinity, Kelly's prescriptions should be used, we believe it is important to assess the risk that stems from the obvious fact that all investments take place in finite time. We thus try to quantify what "long term" means and estimate the characteristic time after which it is safe to consider that the investment is a long term one. The second contribution of this paper is to give an indication of the horizon after which traditional utility theory can and probably should be replaced by Kelly's approach to long term investing.

If the investment horizon is long enough, as defined in the first section of the article, traditional risk measures such as the VaR on an arbitrary holding period or the variance of the terminal wealth are not good indicators of the risk of an investment strategy. This is why, in the second section, we deal with another important issue, namely the risk that, for liquidity

reasons or other constraints, the manager will have to terminate her investment before its expected maturity and with poor market conditions. We argue that this risk is best described by the distributions of drawdowns from historical maxima, also called VaR with no horizon. Following Maslov and Zhang (1998b), we show that it is possible to compute the asymptotic behavior of the distribution of drawdowns as a function of the investment strategy and that the distribution of drawdowns of an optimal growth rate portfolio exhibits universal features, independently of the stochastic process of the stock returns. The third contribution of this article is to provide empirical estimations to support this claim.

The third section of this article is devoted to empirical investigations. We provide some evidence on stock markets to show that Kelly's optimal strategy can be successfully implemented, at least with a degree of confidence equivalent to the one we have when using standard strategies. We discuss the practical uses of such strategies and conclude.

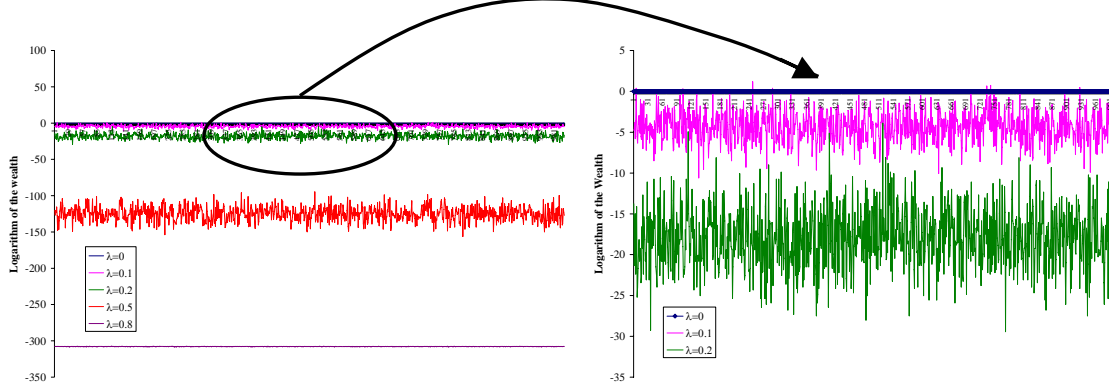
2. LONG TERM OPPORTUNITIES IN THE OPTIMAL GROWTH RATE STRATEGY

2.1 St Petersburg's paradox

It is interesting to notice that the problem discussed here is a very old one. It dates back to Bernoulli's (1730) presentation of St Petersburg's paradox. Suppose that a gambler has 100 rubbles that she can bet as she pleases on the toss of a fair coin. If she wins, her payoff is equal to her bet; if she loses, she forfeits her bet. She can play as many times as she likes provided she does not go bankrupt. Which proportion of her wealth should she bet on each toss of the coin? It is very simple to show that, in order to maximize her expected payoff, the gambler ends up betting all her money each time. So the probability that she is bankrupt tends quickly to 1: if she loses only once she will forfeit all her money, even though her expected gain is maximal. The reason for this is that the average wealth at maturity is artificially increased by random events with increasingly low probabilities. If you win the toss N times, your initial wealth is multiplied by 2^N , but the probability of this happening is extremely low (2^{-N}). In all the other cases, if you invested all your money each time you end up with nothing. Which rational investor would make such a choice in the long term?

One could argue that this is true only for a linear utility and that introducing risk aversion will change the whole picture, but it is not the case. Suppose that the investor chooses a strategy that consists in investing a fixed proportion λ of her wealth each time. This means that her utility function is isoelastic with constant relative risk aversion. In this framework, the possible utility functions or relative risk aversions can be mapped on the $[0, 1]$ interval of possible λ 's. We plot, in Figure Ia and Ib, 2000 observations of the final wealth after 2000 coin tossing for several values of λ . In this example, "long term" is reflected by the number of times the game is played. The results are given in the decimal logarithmic scale, since for some value of λ the wealth is so small that it is otherwise impossible to distinguish it from 0. These graphs clearly show that utility is not an issue here and that no rational investor should choose a long term strategy other than $\lambda=0$, which corresponds to the logarithmic utility function. This is shown graphically by the fact that the strategies do not "overlap", *i.e.* on the long term some strategies yield almost surely a smaller wealth than others.

Figure Ia, and Ib: Final wealth in St Petersburg's game



Note: The graphs show, in logarithmic scale, 2000 realizations of the final wealth after 2000 coin tossing. Figure Ib is a zoom of Figure Ia for small values of λ . It is clear that some strategies almost surely outperform others since the terminal wealths do not “overlap”.

These results are the essence of St Petersburg's paradox: in a multiplicative random walk, expected wealth is not a good indicator of a long term optimal strategy since it is essentially dominated by extremely unlikely and extremely favorable outcomes.

2.2 Kelly's theory of optimal long term investment

In this section we will consider the case of a discrete time market model. This choice stems from two important facts. Firstly, it is obvious that all investments and all price observations take place in discrete time since there is not such thing as “continuous trading”. The question of knowing whether there exists a more fundamental, somehow theoretical, price process which takes place in continuous time and which can only be observed in discrete time is a challenging one and we shall not discuss it here. Secondly, we believe it is easier to account for non-Gaussian distributions of returns in discrete time than in continuous time, especially if we wish to maintain the stationary random walk hypothesis⁵. However, all the results presented here can fairly easily be extended to all price processes such that an exponential growth rate exists, including the standard geometrical Brownian motion.

We will now be more precise and describe Kelly's criterion in a more general setting. Take the case of a market initially composed of one risky and one riskless asset with constant interest rate r . Kelly sets and solves the following problem: what is, on the long term, the optimal fraction of her wealth that an investor should invest in a risky asset? An agent who follows Kelly's strategy and principles will benefit from the optimal growth rate in the long run; in other words, her wealth will increase at a maximal rate. Conversely, an agent who seeks to maximize her expected utility will set great store by outcomes that are highly positive but increasingly improbable.

Mathematically, let W_t be the value of a portfolio and l the fraction of W_t invested in the risky asset at date t . We adopt a discrete time setting and suppose that the price process (S_t) of the risky asset follows a multiplicative random walk, i.e.:

$$S_{t+\tau} = S_t \cdot e^{X_t(\tau)} \quad (1)$$

The time increment τ can be taken arbitrary small but in our applications we will use $\tau=1$ day or one month. $X_t(\tau)$ is the random logarithmic return of the spot price between t and $t+\tau$. This model allows for non Gaussian returns as usually observed on most equity markets. We will assume that the returns are independent and identically distributed. This assumption

is very strong, and probably not very realistic, but it can be relaxed. We plan to allow for some predictability of the returns or varying volatility in future research. We also assume that there are no transaction costs and that the borrowing and lending rates are constant and equal to r .

The investor's strategy is to hold a constant fraction l of her wealth in the risky asset. The wealth balance for one period is thus the following:

$$W_{t+\tau} = (1-l)W_t \cdot e^{r\tau} + lW_t \cdot e^{X_t(\tau)} \quad (2)$$

We will now use the notation $U_t = e^{X_t(\tau)}$. Over a horizon of size $N \cdot \tau$ and by sending N to $+\infty$ we can obtain R the logarithmic rate of return of the portfolio that is equal to: $R(l) = E(\ln(1 + l(e^{X-r\tau} - 1))) + r\tau$. By the central limit theorem the typical - i.e the most likely value \bar{W}_T of the portfolio grows asymptotically at this exponential rate: $\bar{W}_T \approx W_0 e^{R(l)T}$.

By concavity of the logarithm, if a l^* exists such that $R(l^*)$ is maximal, then it is unique and given by the following first order condition (p is the density of the random variable u):

$$\frac{dR}{dl}(l^*) = \int_0^{+\infty} \frac{ue^{-r\tau} - 1}{1 + l^*(ue^{-r\tau} - 1)} p(u) du = 0 \quad (3)$$

l^* is the optimal growth rate strategy and it is the one recommended by Kelly. Let us note that, in this setting, the actual results derived from Kelly's criterion are the same as the one we could obtain using a logarithmic utility function. However, this should not hide the fact that these two approaches have radically different interpretations.

At this point one could ask why is the typical wealth more universal than the average wealth or any utility function, when analyzing agent's preferences. The long-term hypothesis is the key to understanding why Kelly's criterion gives indeed the optimal strategy. Actually, the average rate of return between the initial date $t=0$ and time t , $z_t = t^{-1} \ln(W_t / W_0)$, is the only quantity that converges in probability to its expected value as time goes to infinity. The central limit theorem then gives the following Gaussian density at time t for the random variable z_t :

$$f(z_t) = \frac{1}{\sqrt{2\pi/t}\sigma(l)} e^{-\frac{(z_t - R(l))^2}{2\sigma^2(l)/t}} \quad (4)$$

The average of the random variable z_t is constant over time, whereas the dispersion decreases as the square root of time. Therefore, whatever the value of the dispersion, in the limit of an infinite horizon, the portfolio with the highest typical growth rate has to be preferred since the wealth will almost surely be greater. The dispersion $\sigma(l)$ can be computed numerically but in the asymptotic limit it has no impact. In practice, however, it will be important to assess its actual impact since all investments necessarily take place in finite time.

2.3 Universal properties of the optimal growth rate portfolio

To better grasp the universal nature of Kelly's theory, let us state the problem differently and consider the case of two agents with the same unspecified utility function V . The first one invests her money so as to maximize her expected utility using the ratio l_1 of risky asset, the other invests following Kelly's strategy l^* and maximizes her logarithmic rate of return. As the game repeats, the probability of the first agent's utility exceeding the other agent's utility tends to zero. And yet, the first agent has maximized her expected utility. Mathematically, we have, for $N \rightarrow \infty$:

$$\Pr[V(W_N(l_1)) \geq V(W_N(l^*))] \rightarrow 0 \text{ whereas } E[V(W_N(l_1))] > E[V(W_N(l^*))] \quad (5)$$

This equation is probably the key to understanding the normative properties of the optimal growth rate portfolio. To put it in plain English, if you maximize your expected utility – except, of course, log utility – instead of investing according to Kelly’s theory, you are almost certain, on the long term, to end up with a smaller utility. Expected utility is simply not a good measure of the terminal utility! This is why we argue that on the long term, utility is not a relevant measure of investor’s preferences. The reason for this is exactly the same as the reason for not betting all your money in St Petersburg’s paradox; in a multiplicative random walk, someone investing according to its expected utility (excluding logarithmic utility) sets exponentially increasing weights to outcomes with exponentially decreasing probability. We believe that on the long term, and we will try to quantify this later on, no rational investor should behave this way.

There is yet another way to visualize the optimality of Kelly’s criterion. Suppose that the portfolio has been rebalanced to maintain a fixed proportion l of risky asset during N trading periods and that the observed return at each date t is equal to X_t . Whatever the utility function of the investor, she seeks to earn more money, since her utility function is increasing with her wealth. Therefore, *ex-post*, knowing the actual returns, the optimal constant mix strategy is the one that maximizes W_T with respect to l . Then, independently of any stochastic model, maximizing the wealth leads to the following first order condition:

$$\frac{\partial \ln(W_T)}{\partial l}(l^*) = 0 \Leftrightarrow \sum_i \frac{e^{X_{it} - r\tau} - 1}{1 + l^*(e^{X_{it} - r\tau} - 1)} = 0 \quad (6)$$

This equation is exactly the equation (3) defining l^* except that we did not take an average following a particular distribution but the actual time average of the observed returns. When the time goes to infinity and if the process (X_t) is ergodic the law of large numbers tells us that these two definitions coincide and therefore that Kelly’s optimal l^* leads to the maximum wealth. This shows again, if necessary, that Kelly’s criterion has universal, independent of agent’s preferences, properties.

2.4 When is long term long ?

Kelly’s criterion is valid asymptotically. The problem is that we never invest on an infinite horizon. Let us call T the finite horizon of our investment. The question is to know whether T can be considered a long-term horizon or not. To answer it, we will compare Kelly’s criterion with another one, for example coming from utility theory. More specifically, we will calculate, for large T , the probability that Kelly’s strategy performs better than another one and study how this probability converges to 1. A utility-based criterion will lead to an investment strategy characterized by a given value of Kelly’s coefficient. This value can be time dependent, but we will only consider here time independent strategies, *i.e.* isoelastic utilities⁶. We calculate the following quantities that tell us when our long-term strategy is better than another strategy :

$$P[\ln W_T(l^*) > \ln W_T(l)] \quad (7)$$

This quantity goes to 1 when T goes to infinity. The convergence can be shown to be exponential in the case of i.i.d. returns. The characteristic time of this exponential is the limit after which T can be considered long term. In this section we calculate it explicitly in the binomial and geometrical Brownian motion models. In the more general case such quantities can be calculated in the asymptotic limit using equation (5), though the calculations can be quite challenging for arbitrary distributions of returns. However, one can also show that the convergence is exponential, with a correcting logarithmic term that accelerates convergence.

2.4.a. A simple binomial setting

We now suppose that the riskless interest rate is 0. In the binomial model, the stock price between date $t-1$ and date t can only take two values, u and $1/u$ with probability p and $1-p$ respectively. Let us call ε_t this random variable ; the expected growth rate \bar{R} of a portfolio invested into l risky assets and $1-l$ riskless assets is $\bar{R} = p \ln(1-l+lu) + (1-p) \ln(1-l+1/u)$.

The optimal portfolio is obtained by maximizing the expected growth rate ; we obtain :

$$l^* = \frac{p(1+u)-1}{u-1} \quad (8)$$

We can now compare this optimal portfolio to any other sub-optimal portfolio characterized by the proportion l ($l \neq l^*$) invested in the risky asset. Explicit calculations, available from the authors upon request, show that of the probability that Kelly's strategy is better than a sub-optimal strategy has a long-term approximation:

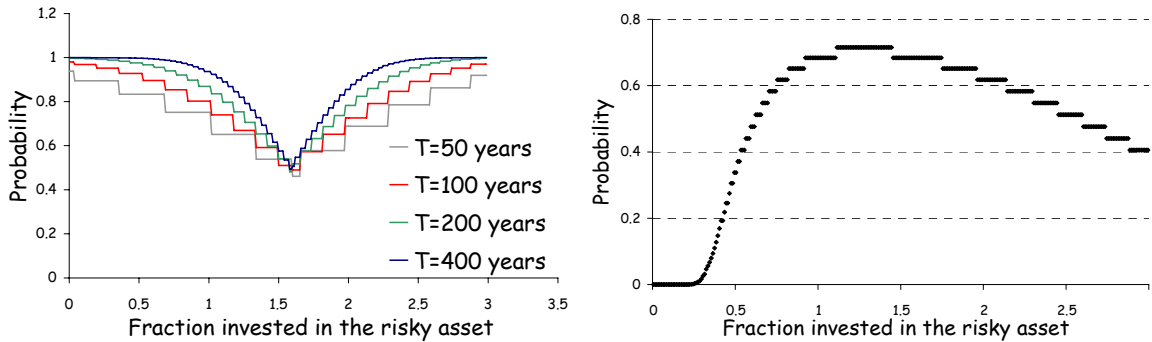
$$P[\ln W_T(l^*) > \ln W_T(l)] \approx 1 - k_0 \frac{e^{-T/\tau_0}}{\sqrt{T}}, \quad \tau_0 = \frac{4p(1-p)}{(\bar{R}^* - \bar{R})^2} \quad (9)$$

After the characteristic time τ_0 we can consider that the horizon is long enough to beat another strategy. A corollary of this result is that we can compute the probability that Kelly's strategy will perform better than an investment with fixed return R_0 . The long-term approximation is given by:

$$P[\ln W_T(l) > R_0 T] \approx 1 - k_1 \frac{e^{-T/\tau_1}}{\sqrt{T}}, \quad \tau_1 = \frac{2p(1-p)}{(R_0 - \bar{R})^2} \quad (10)$$

After the characteristic time τ_1 we can consider that it is possible to secure a return equal at least to $(R_0 - \bar{R})$. Figure IIa and IIb show the exact values of these probabilities using realistic parameters of a binomial stock market.

Figure IIa and IIb: Probabilities that Kelly's strategy is superior to other strategies (left) or exceeds a fixed return (right)



Note: The left figure shows, for various horizons and fractions invested in the risky asset, the probability that the terminal wealth using Kelly's strategy is superior to other strategies. On the right figure, we plot, as a function of the fraction invested in the risky asset, the probability that the final return, after 500 months, is superior to a 5% yearly return. The risky asset follows a binomial random walk with the mean annual return set to 10% and the volatility equal to 25%. The asset manager rebalances her portfolio every month. The optimal fraction is $l^*=1.6$

The probability of exceeding a 5% yearly return is high at the Kelly optimum, almost 70%. Recall that these results were obtained using a riskless interest rate equal to zero. It is therefore possible to secure an almost riskless investment using Kelly's long term strategy, despite the lack of a riskless asset. We now turn to the case of continuous time modeling, namely Merton's (1971) model of optimal investing.

2.4.b. A continuous time setting

The standard approach to calculating the optimal fraction of a risky investment has been derived by Merton (1971) in the continuous time setting. We will now show that, even under the standard diffusion assumptions, any rational investor should use Kelly's criterion on the long term. Let us suppose that the risky asset follows a geometrical Brownian motion with drift μ and volatility σ and that the riskless asset has a constant return r . We also suppose that the investor has an isoelastic utility function⁷, i.e. $U(W) = \frac{W^{1-\alpha}}{1-\alpha}$ where α is the relative risk aversion of the investor. Merton (1971) shows that, whatever the time horizon considered, the optimal fraction of her wealth invested in the risky asset is equal to $l_M = \frac{1}{\alpha} \left(\frac{\mu - r}{\sigma^2} \right)$. The value $\alpha=1$ corresponds to the logarithmic utility function and is exactly equal to Kelly's strategy. However, if $\alpha \neq 1$ and if the time horizon is large enough, such a strategy is not optimal. The diffusion equation satisfied by the wealth W_t of the investor who uses a constant fraction l is the following (B_t is a standard Wiener process.):

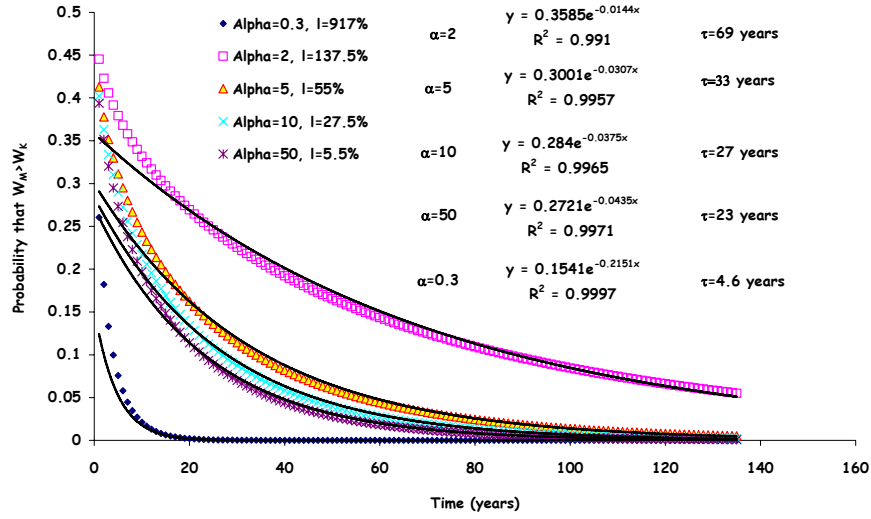
$$\frac{dW_t(l)}{W_t} = [r + l(\mu - r)]dt + \sigma \cdot l \cdot dB_t \quad (11)$$

This equation can be explicitly solved and we can compute the probability that one strategy outperforms the other. In the $\alpha > 1$ case, we have (B_1 has a standard Gaussian distribution):

$$\Pr(W_T(l_M) > W_T(l_K)) = \Pr\left(B_1 \leq -\sqrt{T}\lambda \left[1 - \frac{\alpha + 1}{2\alpha}\right]\right) \quad (12)$$

One should also be aware that extremely risky strategies, i.e. $\alpha < 1$, are in some sense stochastically dominated by the Kelly strategy since they provide a lower return with more risk⁸. From these equations we can easily compute the theoretical characteristic time, only a function of α and λ . In Figure III we plot this probability for various values of the risk aversion and maturities. The annualized parameters we used are $r=4\%$, $\mu=15\%$ and $\sigma=20\%$.

Figure III: Comparison of Merton's and Kelly's strategies



Note: For various risk aversions α , we plot, as a function of time, the probability that the terminal wealth is higher when using Merton's strategy instead of Kelly's. We give the exponential fit and the R^2 of this fit which are very close to 1. We also indicate the characteristic time after which Kelly's strategy outperforms Merton's strategy. The risky asset follows a geometric Brownian motion with $r=4\%$, $\mu=15\%$ and $\sigma=20\%$. We also indicate, for each risk aversion, the corresponding optimal fraction invested in the risky asset. The Kelly optimal fraction is $l^*=2.75$.

As seen from Figure III, the probability that Merton's strategy is superior to Kelly's tends quickly to 0 for long horizons, except for very small risk aversions (i.e. when $\alpha \rightarrow 1^+$) for which the characteristic time is rather long and the strategies close to the Kelly optimal strategy. For $\alpha < 1$, the investor is risk-seeker, and the characteristic time decreases with this coefficient. In this framework, it is important to notice that the decay of the probability is exponential only up to a correcting logarithm term. A more precise asymptotic expression is

given by $k \cdot e^{-\frac{T}{\tau} - \frac{1}{2} \ln(T)}$, as can be shown using the asymptotic expansion of the Gaussian distribution function. Therefore, the theoretical characteristic time given by τ is not exactly the characteristic time that we estimate empirically and it is larger.

3. SHORT TERM DRAWDOWN RISK IN THE OPTIMUM GROWTH RATE PORTFOLIO

3.1 Drawdowns risk and VaR with no horizon

We have seen that, on the long term, Kelly's long term criterion is clearly relevant. However it is important to assess the risk taken by an investor whose horizon is not as long as she expected. Imagine for example the case of an asset manager who chooses to invest her portfolio on a long-term strategy and who faces sudden liquidity constraints. The risk she is facing is the risk that her wealth at that specific point in time will be under an historical high and below her long-term expectations. Though the analysis of VaR or standard deviations can bring a light on such questions, we believe they are not good indicators of risk for a long-term investment. The main reason for rejecting volatility measures lies in the strongly non-Gaussian character of stock returns. We do not use VaR either because it is impossible to know the horizon over which VaR should be calculated since the risk is precisely the risk of having to terminate the investment at a future unknown date. Therefore, we need a measure of risk that

is not explicitly linked to a time horizon. Moreover, VaR is above all a solvency criterion and as such is not very relevant to a long-term investor since solvency should not be an issue and short-term fluctuations should not constrain too much his long-term goals. Another well known drawback of the VaR risk measure is that it is not sub-additive as pointed out by Artzner and al. (1997) ; this drawback is relevant for portfolio issues.

Another possibility is to use the VaR over the entire investment horizon. However, the problem is that, on the long term, the law of large number will lead to a smaller VaR for the optimal growth rate strategy, whatever the chosen probability threshold. It is only if the probability threshold is a function of the investment horizon that such a measure is meaningful. More generally, if the risk measure is defined only on the wealth at the final date of the investment, the optimality of Kelly's strategy will lead to the - false - idea that it is a riskless investment, at least if the investment is carried over a long enough period of time. The real risk stems from the fact that there is no guaranty that this will actually be the case.

We believe that the best way to study long term risk is to analyze the drawdowns distribution. The definition of a drawdown is precisely the following : at time t , the drawdown from maximum is :

$$D(t) = \frac{\max_{s \in [0, t]} W_s}{W_t} \quad (13)$$

where W_t is the value of the portfolio at date t . Please note that this risk is more an opportunity risk than the risk of loosing money in the investment, since the drawdown is measured from the previous maximum and not from the initial wealth. $D(t)$ is also called VaR with no horizon since it describes losses on no prespecified horizon. It also provides a measure of the frustration of the investor, forced to liquidate a position that was worth more in the past. However, it is not a coherent measure of risk as described by Artzner and al. but this is out of the scope of this article.

The ability to control these drawdowns might turn to be crucial if one has no guarantee that the investment can be maintained until maturity. Maslov and Zhang (1998b) have shown that, in the multiplicative random walk framework, these drawdowns have a stationary unconditional distribution which is asymptotically equivalent to a power law if the drift of the random walk is positive (*i.e.* X has positive mean). Therefore, controlling the drawdown risk is equivalent to controlling the decay factor of this power law, at least for large drawdowns which are certainly the most relevant for investors. More precisely, these authors show that if the behavior of the wealth obeys the dynamics of equations (2)-(3) and if we denote by π the distribution function of the log returns of the wealth, then for large values of the drawdown D , the distribution of D behaves as:

$$P(D > x) = P(x) \approx \frac{1}{x^\Gamma} \quad \text{with} \quad \int \pi(x) e^{-\Gamma x} dx = 1 \quad (14)$$

In the Gaussian case with drift equal to μ and variance equal to σ^2 we have $\Gamma = 2\mu / \sigma^2$.

3.2 Drawdown risk in the optimal growth rate portfolio

How can the investor control these drawdowns ? Let us assume, for the sake of simplicity, that the interest rate is 0. If the risky asset follows a multiplicative random walk, so does the wealth if a constant fraction l of the wealth is invested at each date. The new logarithmic return is $Y = \ln(1 - l + le^X)$. At the Kelly optimal ratio, equation (14) with $\Gamma = 1$ written for the random variable Y is equivalent to the first order condition defining l^* and implies that the drawdown exponent is equal to 1 for any probability law for the variable X . At the Kelly optimum, whatever the probability distribution, the drawdowns have a well-known and universal asymptotic behavior, though a risky one.

The investor using the Kelly optimal strategy has an advantage in doing so, apart from the optimal growth rate: she knows exactly the asymptotic behavior of the probability distribution of the drawdowns since it decays like a power law with fixed exponent. The Kelly strategy appears to be a very aggressive one, since the drawdowns have theoretically an infinite average at the Kelly optimum. It is therefore of crucial importance that this strategy be limited to cases where the manager is absolutely positive that no exterior constraints will force him to liquidate his position before the prespecified horizon. This strategy is probably well suited for investors who are not likely to face sudden liquidity constraints and who really have the time to capitalize their investments. In the next section we perform empirical investigations on stock markets to show that both the long-term opportunities and the short-term risk of the optimal growth rate strategy that we just describe are supported by empirical data.

4. EMPIRICAL INVESTIGATIONS

In this section we perform empirical investigations to support the claims made in the rest of the article. We first analyze the stability of the empirical estimation of l^* , then we show that the long term opportunities of the optimal growth rate strategy are genuine and finally provide conclusive evidence to show the validity of the theoretical results of section 3 regarding drawdown risk.

4.1. Stability of l^* and choice of sample

We first give two ways of estimating l^* . Consider the modified returns:

$$Y_t = \frac{U_t e^{-r\tau} - 1}{1 + l(U_t e^{-r\tau} - 1)} \quad (15)$$

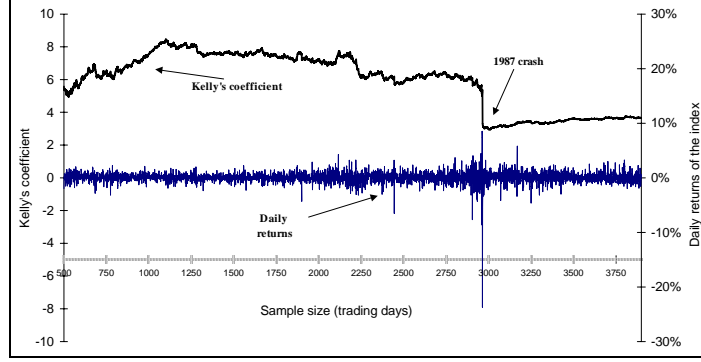
Using the first order condition defining l^* one shows that these returns have a zero average for $l=l^*$ and therefore a standard optimization algorithm can give l^* . Another approach is to compute the cumulants of the standard returns and to use the following formula (see Legras and de Monts de Savasse (2000)) :

$$l^* = \frac{\mu}{\sigma^2} + \frac{\chi}{\sigma^3} \mu^2 + \frac{2\chi^2 - (\kappa + 3)}{\sigma^4} \mu^3 + o((\mu/\sigma)^3) \quad (16)$$

The parameter μ is the average, σ the standard deviation, χ the skewness and κ the kurtosis of the returns. Similar formulas are available for the standard deviation of l^* .

A question remains to be answered : what is the sample that should be used when estimating l^* ? The dilemma can be expressed as follows. If one uses a very long sample it is less likely that the returns will be identically distributed. However, if one uses short data sets, the statistical estimations are not very significant. Since the essence of Kelly's optimal leverage is to be a long-term optimum we chose long term estimations. It seems important to use long data sets especially if some extreme events, such as the October 1987 crash, occurred some time ago. Not taking these values into account could lead to overestimating l^* thus leading to possible bankruptcy⁹. In Figure IV we illustrate our point by plotting the estimation of l^* for various sample sizes on the S&P 100 dataset.

Figure IV: Estimation of l^* for various sample sizes, S&P 100 Index



Note: we plot the estimated value of l^* on the S&P 100 Index using the non-parametric approach for various sample sizes. The data set comes from Datastream and gives the daily closing prices of the S&P 100 index for the period going from March the 3rd 1984 to December the 31st 1998. We also indicate the daily returns. The impact of the 1987 crash is blatant. These estimations were performed with $r=0$.

Without the 1987 crash the estimated value of l^* is almost 50% above the one obtained using the whole sample. The optimal leverage estimated with 2000 trading days (around 7) leads to bankruptcy if it used before the 1987 crash. In fact, any value above $4.74=1/21\%$ leads to bankruptcy. This shows the importance of having upper limits to the optimal leverage and stress the importance of the boundedness of the distribution. Estimations on other indexes, available from the authors upon request, yield similar results.

4.2 Empirical stability of l^* and long term opportunities

We have seen that the statistical estimation of l^* is not very precise. However, what is really important is to know whether this estimation can be used to implement a good investment strategy. What are the results of a long-term strategy designed using a long-term estimation of l^* ? It is clear from the simple statement of this question that we will need very long data sets to answer it. The first option is to use intraday data, for which millions of prices are sometimes available. In this article we focus on the asset manager's point of view, for which systematic intraday trading is not very realistic.

We first studied five indexes of the NYSE for which 33 years of daily data are available, from January the 3rd 1966 to December the 31st 1998. These indexes are the NYSE Composite, Industrial, Transports, Utility and Finance indexes. For each index we estimated the optimal value of l^* using the whole sample and the optimal value of l^* using two sub-samples of the same size, the separating date being November the 25th 1982. The stability of l^* was investigated when switching from the first subsample to the second one. This stability is not a theoretical issue only and is crucial to the investor since, as we have seen in the first section, the estimated value of l^* for a sample is exactly the value of l^* that guarantees the maximum wealth at the end of the sample.

There is a major difference between our theoretical framework and empirical data, namely the fact that interest rates change. If we want to take this effect into account we may either assume that the daily returns minus the daily interest rate follow a random walk or assume that the daily returns themselves follow a random walk. We can therefore use one of these two models :

$$\begin{aligned} S_{t+1} &= S_t \cdot e^{X_t} \quad (\text{I}) \\ S_{t+1} e^{-r_{t+1}} &= S_t e^{-r_t} e^{X_t} \quad (\text{II}) \end{aligned} \tag{17}$$

If one model is correct, the corresponding X_t variable is i.i.d. and therefore l^* is stable. In order to check which model is the best we thus looked at the stability of l^* under both assumptions. The interest rates we used are the daily yield to maturity of a 3 months T-Bill provided by the Federal Reserve Bank of Atlanta¹⁰. The results are given in Table I.

Table I: Stability of l^* for NYSE Indexes

			COMPOSITE	INDUSTRIAL	TRANSPORT	UTILITY	FINANCE
Model I Constant interest rate (variable risk premium)	Whole Sample	l^*	3.46	3.36	3.26	2.70	2.69
		Std dev	0.99	0.93	0.67	0.93	0.95
	Sample 1	l^*	2.05	2.20	1.20	-0.55	2.24
		Std dev	1.86	1.72	1.40	2.51	1.76
	Sample 2	l^*	3.94	3.76	3.97	3.10	2.87
		Std dev	0.98	0.94	0.84	0.78	1.09
Model II Variable interest rate (constant risk premium)	Whole Sample	l^*	0.70	0.96	2.64	0.29	0.50
		Std dev	1.22	1.13	0.39	0.94	1.04
	Sample 1	l^*	-1.95	-1.22	-1.08	-7.60	-1.37
		Std dev	1.84	1.71	1.39	2.39	1.75
	Sample 2	l^*	2.50	2.48	3.30	2.14	1.49
		Std dev	1.41	1.32	0.77	1.18	1.25

Note: We estimate the values of l^* as well as standard deviations using the non-parametric approach for the three different samples. In model I and II the estimated values of l^* are very different simply because in model I we do not take interest rates into account (we assume $r=0$), whereas we do in model II.

The estimated values for l^* are not very stable. The standard deviations are high and values of l^* can switch from 0.29 to -7.6 in the worst case! However model II is clearly not the most realistic. This justifies the estimations performed in the previous sections with $r=0$. Under this hypothesis the values for the whole sample and for sample 2 are quite close. The values for sample 1 are not as close – except for the Finance index – but the standard deviations are high.

This means that, in order to implement Kelly's optimal strategy in practice, the actual value of l^* one should use has to be computed with constant interest rates and then needs daily adjustment as the short term (one period) interest rate moves. Similar results, again available from the authors upon request, can be obtained on individual French stocks.

This short empirical study has shown that though some errors remain, there is some stability in the estimation of the optimal strategy. However, this stability is obtained if the dynamics of the interest rates are not taken into account, *i.e.* if the estimations of l^* are performed with a constant interest rate. This suggests that the risk premium on the stock market is not constant and that a long-term investor should adapt her strategy to the daily values of short-term interest rates.

We have seen that the optimal value of l^* is correct but far from excellent. These results should be compared to the one obtained using more standard approaches, such as Merton's (1971) optimal investment strategy. It is well known that the relative risk aversion of an agent is difficult to estimate and this imprecision has a direct impact on the optimal strategy since the relative error made when estimating the risk aversion is exactly equal to the relative error made on the optimal leverage. Therefore, a simple application of standard portfolio optimization can also lead to substantial errors in the calculation of the optimal leverage even if the parameters of the diffusion are perfectly known.

Moreover, in Merton's (1971) case the errors can also stem from improper estimation of the parameters of the diffusion. This issue has been discussed in depth by many authors. For instance, Chopra (1993) showed that tiny errors in the inputs could drastically modify the optimal portfolios in the mean variance framework. As regards the utility function, Kallberg and Ziemba (1984) found that utility functions with similar risk aversion yield roughly the same optimal portfolios whereas different risk aversions lead to substantially different portfolios. A simple way to put it is that basically if you cannot predict the future average

return you cannot find the optimal portfolio. This is true both for the mean variance approach and in Merton's (1971) model and it is not surprising that it still holds for Kelly's strategy.

In Table II we give the estimated optimal leverage with different relative risk aversions for NYSE data using the same sample and sub-samples as before.

Table II: Stability of the optimal strategy with Merton's approach

			COMPOSITE	INDUSTRIAL	TRANSPORT	UTILITY	FINANCE
RRA = 2	Whole Sample	I^*	1.42	1.36	0.27	0.90	0.90
	Sample 1	I^*	0.52	0.60	0.10	-0.78	0.62
	Sample 2	I^*	2.08	1.94	0.33	1.29	1.05
RRA = 4	Whole Sample	I^*	0.71	0.68	0.14	0.45	0.45
	Sample 1	I^*	0.26	0.30	0.05	-0.39	0.31
	Sample 2	I^*	1.04	0.97	0.17	0.65	0.53
RRA = 1.5	Whole Sample	I^*	1.89	1.81	0.37	1.20	1.21
	Sample 1	I^*	0.70	0.80	0.13	-1.04	0.82
	Sample 2	I^*	2.78	2.59	0.45	1.72	1.41

Note: We estimate the optimal strategies for the three different samples with various risk aversions. Following our previous findings, we did not take interest rates into account.

The stability of the optimal leverage is as good (or as poor) as the one obtained using Kelly's theory. In fact, this instability is not intrinsic to one approach or the other. It comes from the more fundamental instability of the average and standard deviations of daily returns. This is why we believe that Kelly's criterion should be used in the same way as asset managers use Merton's (1971) theory or mean variance allocation. The actual parameters of the multiplicative random walk used in the optimization should be representative of the manager's anticipations regarding the behavior of the risky asset and not be the result of past statistical estimations. Once the manager has formulated her anticipations for the average return, standard deviation and higher moments of these returns, she can use formula (16) to determine an optimal strategy coherent with her expectations. Of course, this means that the manager is able to anticipate properly not only the drifts and volatilities but also skewness or kurtosis. This is the price to pay in a non-Gaussian world!

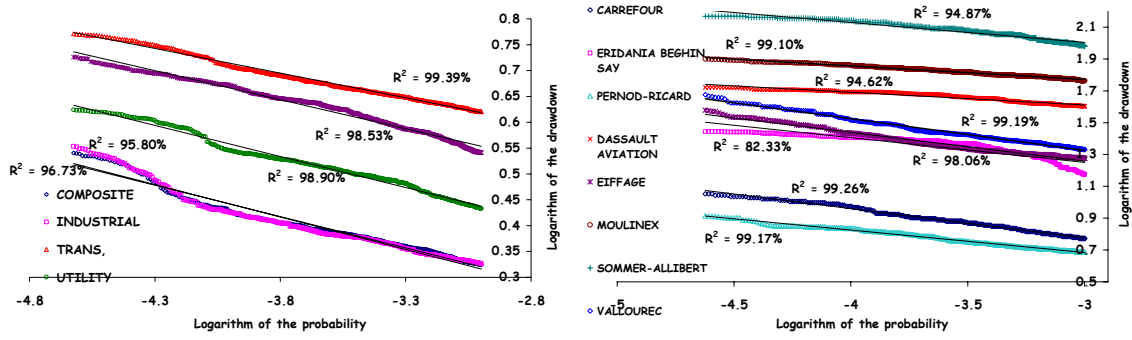
4.3 The asymptotic drawdown distribution

In this subsection we show that the theoretical presentation of section 3 is indeed relevant. More specifically, we show that the unconditional asymptotic distribution of drawdowns is well described by a power law and that at the Kelly optimum the exponent of this power law is close to 1.

As always in statistics we are confronted with a sample problem. The theoretical result is an asymptotic one, therefore in the analysis we need to use only a small fraction of the observed drawdowns and discard the small drawdowns. Moreover, finite size effects will inevitably appear at the other spectrum of the drawdowns, *i.e.* for the largest drawdowns. In order to estimate the best sample to be used, we first performed a Monte Carlo simulation in a case where the exact asymptotic behavior of the drawdowns and the value of Γ are known. The results, not indicated for the sake of brevity, show that a good estimation of Γ can be obtained if we do not use the most extreme part of the distribution, *i.e.* if we discard the worst 1% of the drawdowns.

We then checked the asymptotic property of the distribution of drawdowns on several individual stocks as well as stock indexes. We studied the five indexes of the NYSE described above for the same time period. For each data set, we performed a linear regression of the log-drawdowns on their log-probabilities, estimated empirically, on the range going from the worst 5% to the worst 1% of the drawdowns. The results are presented Figure Va and Vb.

Figure Va and Vb: Asymptotic distribution of drawdowns

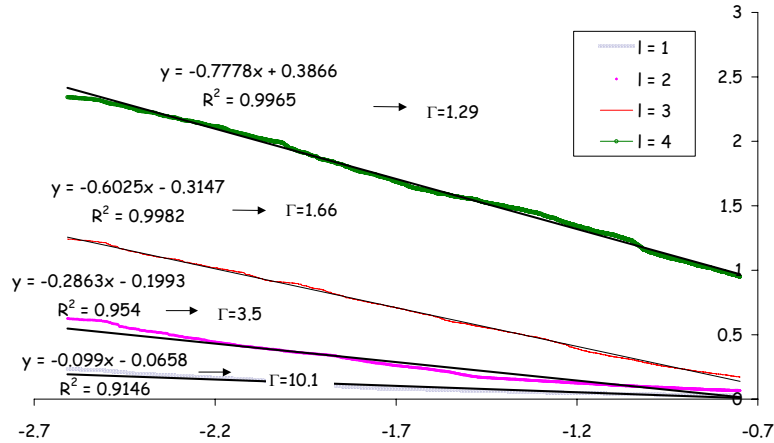


Note: For five sectorial indexes and seven individual stocks, we plot, in logarithmic scale, the drawdowns as a function of their cumulative probability of occurrence estimated by their rank divided by the number of data. The data set, coming from Datastream and the NYSE, gives the daily closing of the stocks and indexes. The good linear fit shows the validity of Equation (20). The individual stocks are quoted on the Paris Stock Exchange for the period going from 1974 to 1999.

The results are good and the R^2 range from 82.33% to 99.8%. There is only one case where the R^2 is under 94%. This shows that the asymptotic expression given in equation (16) is very robust both for individual stocks and for indexes.

We now want to test the asymptotic property of the drawdowns of the optimal growth rate portfolio. We tested these theoretical results on the S&P 100 index for the period going from March the 3rd 1984 to December the 31st 1998. This sample, taken from Datastream, has 3870 daily values and includes the dramatic crash of October 1987. Figure VI illustrates the distribution of drawdowns in logarithmic scales for a portfolio invested in the S&P 100 index with several values of the leverage l . Values of $l > 1$ are authorized. The Kelly optimal value of l can be estimated at 4.02 on this market with a zero interest rate. Therefore, we estimated the distribution of drawdowns for $l=1$, $l=2$, $l=3$ and $l \approx l^*=4$.

Figure VI: Distributions of drawdowns for strategies in the S&P Index



Note: We first plot, for various constant mix strategies, in logarithmic scale, the drawdowns as a function of their cumulative probability of occurrence estimated by their rank divided by the number of data. The data set comes from Datastream and gives the daily closing prices of the S&P 100 index for the period going from March the 3rd 1984 to December the 31st 1998. The good linear fits show the validity of Equation (16) for all the strategies.

It appears that the actual value of Γ is not far from the theoretical value for $l=l^*$ since we have $\Gamma=1.29$. The difference with our theoretical framework could come from serial dependence in the data. One could imagine, for example, that when the market is extremely low, compared to historical maxima, investors feel that it is not likely to go down any further and therefore tend to be a little more bullish, pushing the market up. A mean reverting

mechanism or psychological barriers could thus be responsible for such a change in I . Following that interpretation, the distribution of drawdowns should have thinner tails as compared with the random walk case and $I > 1$.

6. CONCLUSION

In this article, we have investigated the relevance of the optimal growth rate portfolio for long term fund management. We have shown that on the long term the expected utility can be a poor investment criterion. We then have been able to quantify the time horizon over which Kelly's criterion is relevant and over which it is possible to claim that a strategy is superior to another. This provides quantitative comparisons between different criteria on a long-term horizon. We also analyzed the risk associated with such strategies and showed that it is best described by the distribution of drawdowns from maxima, or VaR with no horizon. We studied the asymptotic distribution of these drawdowns and performed empirical estimations that confirmed the theoretical power law behavior. We argued the optimal growth rate strategy has a universal behavior, and is a rather risky one ! This confirms that the long term opportunities of the optimal growth rate portfolio are somehow balanced by an increased risk if the investment has to be terminated earlier than expected.

In practice, however, the real distribution of returns is not known and the optimal strategy is only estimated. We have focused on the empirical estimation of Kelly's parameter with various techniques and the stability of the estimations has been investigated. We have shown that this stability is as good – or as poor – as the one obtained using other investment criteria such as expected utility. We argued that as long as you cannot predict the future mean return, all investment criteria are somehow doomed to fail. A more pragmatic approach is therefore to use the investor's expectations regarding future fluctuations of the market and to choose the optimal strategy accordingly. We provided a parametric expression giving the optimal growth rate portfolio as a function of the first four cumulants of the stock returns.

We believe that future research should address several important questions. The first interesting problem is the calculation of optimal growth rate portfolio with conditional distributions of returns (for example time varying mean and volatility). How should an investor adapt her investment policy to stochastic returns or volatility when these variations are observable ? Another interesting question is to estimate the cost of using unconditional strategies when there is some predictability on the returns or volatility, as some works suggest (see Barberis (1999) and references therein).

REFERENCES

- Artzner P., Delbaen F., Eber JM., and Heath D. (1999)**, *Coherent measures of risk*, Mathematical Finance, 9, n°3, p. 203-228.
- Bajeux-Besnainou, I. and Portait, R., (1998)**, *Dynamic asset allocation in a mean-variance framework*, Management Science, November 1998
- Barberis N.C., (1999)**, *Investing for the long run when returns are predictable*, Forthcoming Journal of Finance
- Baviera R., Pasquini M., Serva M., and Vulpiani A. (1999)**, *Optimal strategies for prudent investors*, Working Paper
- Boulier, J.-F., Huang S., and Taillard, G. (1999)**, *Optimal management under stochastic interest rates : the case of a protected pension fund*, Working Paper, Crédit Commercial de France
- Breiman, L. (1961)**, *Optimal gambling system for favorable games*, Proceedings of the 4th Berkeley symposium on mathematics, statistics and probability 1, p. 63-68.
- Campbell J. Y. and Viceira, L., (2000)**, *Why should one buy long term bonds*, Working paper, Harvard University
- Chopra, V. (1991)**, *Mean-variance revisited : near optimal portfolios and sensitivity to input variations*, Russel Research Commentary
- Cox, J. and Huang, C. (1989)**, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, Journal of Economic Theory, 19, 33-83
- Kallberg, J. and Ziemba, W. (1984)**: *Mispecifications in portfolio selection problems*, in Bamberg G and Spremann K. eds, Risk and Capital: Lecture Notes in Economics and Mathematical Systems, New York Springer Verlag, 1984
- Karatzas, I. (1989)**, *Optimization problems in the theory of continuous trading*, SIAM Journal of Control and Optimization, 27, 1221-1259
- Kelly, J. L. Jr. (1956)**, *A new interpretation of the information rate*, Bell System Tech. Journal, 35, 917-926
- Legras, J., de Monts de Savasse, P.H. (2000)**, *Marchés incomplets, distributions non Gaussiennes : quel prix pour les options ?*, Banque et Marchés, March-April 2000 (in French)
- Markowitz, H. (1959)**, *Portfolio Selection: Efficient Diversification of an Investment*, Journal of Finance, 7, 77-91
- Maslov S. and Zhang Y. (1998a)**, *Optimal investment strategy for risky assets*, International Journal of theoretical and applied finance, vol 1 n°3
- Maslov S. and Zhang Y. (1998b)**, *Probability of Drawdowns in risky investments*, Working Paper
- Merton, R.C.(1971)**, *Optimum consumption and portfolio rules in a continuous time model*, Journal of Economic Theory, 3, 373-413
- Merton, R.C. (1990)**, Continuous-time finance, Blackwell Publishers Inc.
- Samuelson, P. A., (1969)**, *Lifetime portfolio selection by dynamic stochastic programming*, Review of economics and statistics, 37, 537-542

¹ In many cases, however, the results are independent of the selected horizon

² Please note that the point is not to know if the **portfolio** is optimal or not but to know if the **criterion** is optimal

³ It can be argued, however, that Bernoulli also pioneered the first approach!

⁴ In this simple game we assume no short selling or leverage is possible.

⁵ We can extend our results to markovian dynamics but the case of long memory is more challenging.

⁶ At the asymptotic limit, if the process is a simple random walk, it can be argued that time-varying strategies are meaningless since the horizon is unchanged.

⁷ Other utility functions do not lead to a constant value for l^* and thus cannot benefit from the optimal growth rate either.

⁸ We are grateful to Pr. Ziemba for pointing this out.

⁹ In discrete time, the question of the boundedness of the distribution is actually an important one. The actual definition of l^* gives precise bounds to its value since the non-bankruptcy condition implies an upper bound to l^* . Assuming $r = 0$ (this simplifies equations without changing any of the results) if the probability distribution of X is not bounded then we necessarily have $l^* \in [-1, 1]$. Any other value of l^* means implicitly that the distribution is bounded. Roughly, the maximum possible value of l^* is $1/X_{max}$ where X_{max} is the maximum possible drop of the risky asset's price. In many cases however, with unbounded distributions of X such as the Gaussian distribution, the actual solution of equation (3) which defines l^* is greater than 1. This does not mean that the optimal value l^* is greater than 1 but that the first order condition does not hold and the optimum is to be found at the frontier of the admissible values, i.e. $l^*=1$. The problem with this reasoning is that it is impossible to distinguish empirically a bounded distribution from an unbounded one. For Gaussian returns, the data set required to observe a 99% drawdown in one time period with a volatility of 1% is unbelievably large! This means that it is impossible to tell if the estimated value of l^* is relevant or not. In practice, we assume that the distribution is bounded by an arbitrary market move.

¹⁰ This was used as a proxy for an overnight, or spot/next, interest rate for which no historical data was available.