Operations Research Letters 39 (2011) 172–179

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

A note on constant proportion trading strategies☆

Martin B. Haugh

Department of Industrial Engineering and Operations Research, Columbia University, United States

ARTICLE INFO

Article history:
Received 18 February 2011
Accepted 10 March 2011
Available online 23 March 2011

Keywords:
Constant proportion trading strategy
Leveraged ETF

ABSTRACT

We consider constant proportion (CP) trading strategies when there are multiple underlying securities and use a recently derived expression for the terminal wealth of a CP strategy to address two issues. First, we characterize the performance of a CP strategy relative to the performance of the corresponding buy-and-hold strategy. We then explain the performance of leveraged ETFs which have been criticized for not performing as expected, particularly during the financial crisis of 2008.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

A *constant proportion* (CP) trading strategy is a strategy in which the fraction of wealth invested in each risky asset is constant and does not vary with time. These strategies are often referred to as *static* strategies in the literature but we will persist with referring to them as CP strategies in this paper. CP strategies require constant rebalancing and are therefore dynamic in nature. Moreover they are perhaps the most well known of all dynamic trading strategies. They appear as the optimal solution to the classic dynamic portfolio choice problem in which the investment opportunity set does not vary with time and the investor has *constant relative risk aversion*, i.e. a power or logarithmic utility function over terminal wealth and/or intermediate consumption. These problems were first studied and solved by Merton [24,25], Samuelson [27] and Hakansson [14]. Moreover, the optimality of these strategies is derived in just about every advanced financial economics textbook that discusses dynamic portfolio optimization. See Duffie [11], Merton [26] or Cvitanic and Zapatero [10], for example, or Karatzas and Shreve [17] for a more recent treatment of the optimality of constant proportion trading strategies. Browne [6] studies the *rate of return on investment* for CP strategies when the underlying securities follow geometric Brownian motions and lists several other problems for which CP strategies are optimal. They include, for example, the problem of minimizing the expected time to reach a given level of wealth and the problem of maximizing the expected discounted reward for reaching a given level of wealth.

CP strategies, of course, are also synonymous with the *Kelly criterion* [18] for optimizing the long-term growth rate of wealth. Other works related to the Kelly criterion include, for example, [4,28,13,12,8,9].

It is remarkable, however, that despite the ubiquity of CP strategies, until very recently we could not find an expression in the literature for the terminal wealth of a CP trading strategy in terms of the terminal security prices. In this note, we use the expression for the terminal wealth of a CP strategy that was recently derived by Haugh and Jain [15] for two purposes.

The first is to compare the performance of a CP strategy with the performance of the corresponding *buy-and-hold* strategy. Our main result here is that when no-short sales and no-borrowing constraints are imposed, the exposure of the CP strategy to realized variances and covariances can be interpreted as a (multiplicative) premium paid to the follower of the CP strategy for accepting a final wealth that is proportional to the geometric mean of the terminal security prices rather than the arithmetic mean. This result follows from a simple application of the geometric-mean inequality but we have not seen it elsewhere in the literature.

The second purpose is to explain the recent and controversial performance of leveraged ETFs (LETFs). Unlike regular ETFs which are passively managed, LETFs require active management. They have the stated goal of replicating some multiple of the *daily* performance of some underlying security or index. This multiple is greater than one for a positively leveraged ETF and less than zero for an *inverse* ETF. Typical leverage values are ± 2 and ± 3. Many investors who invested in these securities expected returns that would be very similar to the returns of a buy-and-hold investment in the same underlying security at the same leverage multiple. During the highly volatile period of the 2008 credit crisis this was not the case and so LETFs received much attention and criticism from the financial press.

We are certainly not the first to explain LETF performance. Indeed Avellaneda and Zhang [1] and Cheng and Madhavan [7] both derived (18) by arguing from first principles. (Cheng and Madhavan [7] derived (18) assuming geometric Brownian motion dynamics.) In this paper, we obtain (18) as a particular case of the more general expression derived by Haugh and Jain [15]. In fact, we argue that given an understanding of CP trading

☆ The first version of this paper was called "A Note on Constant Proportion Trading Strategies and Leveraged ETFs".

E-mail address: Martin.B.Haugh@gmail.com.

0167-6377/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2011.03.005

strategies, there should have been no surprise whatsoever when LETFs performed as they did during the financial crisis of 2008. Indeed the sensitivity of the performance of a CP trading strategy to realized variance does not appear to be widely known or at the very least, widely appreciated. This is remarkable, given the central role they played in the early development of dynamic portfolio optimization and their association with the Kelly criterion. This lack of appreciation is most likely explained by the fact that just about every treatment of CP strategies in the literature neglects to write the terminal wealth as a function of the terminal values of the underlying securities. Instead these treatments often end once they have demonstrated the optimality of a CP strategy and on some occasions, derived the optimal value function.

In this note, we write the terminal wealth of a CP trading strategy as a function of the terminal security prices and see that it immediately explains LETF performance when specialized to the case of just one underlying security. Moreover, we can use this expression to interpret the exposure of a CP strategy to realized variances and covariances as the cost or compensation for following a CP strategy as opposed to a buy-and-hold strategy. As stated above, when no short sales and no-borrowing constraints are imposed, we can interpret this exposure as a compensation. In the case of LETFs, it is a cost. We also propose a *constant proportion* ETF (CPETF) that requires neither borrowing nor short selling and argue that such a security would be more appealing to unsophisticated investors as well as having some positive systemic effects on market microstructure. In particular, the rebalancing requirements of a CPETF would require the manager to sell at the close after an up-day and to buy at the close after a down-day, thereby dampening volatility at the close.

The remainder of this note is organized as follows. In Section 2, we introduce our model for the security price dynamics and then derive the expression of Haugh and Jain [15] for the terminal wealth that results from following a CP strategy. After briefly mentioning some other applications of this expression we move to Section 3 where we compare the performance of a CP strategy with the performance of the corresponding buy-and-hold strategy. In Section 4, we use Eq. (18) to explain the performance of LETFs during the financial crisis of 2008. While not the first to derive (18), we do demonstrate that it follows immediately from the earlier work of Haugh and Jain [15].

2. Security price and wealth dynamics

We assume that there are n risky assets and a single risk-free asset available in the economy. The time t vector of risky asset prices is denoted by $P_t = (P_t^{(1)} \ \dots \ P_t^{(n)})^\top$ and the time t continuously compounded risk-free rate of return is denoted by r_t. We assume the price dynamics of the risky assets satisfy

$$\frac{\mathrm{d}P_t}{P_t} = \mu_t \mathrm{d}t + \Sigma_t \mathrm{d}B_t \tag{1}$$

where $\mathrm{d}P_t/P_t$ should be interpreted component-wise, B_t is an m-dimensional standard Brownian motion, μ_t is an n-dimensional adapted process and Σ_t is an $n \times m$ adapted matrix process. Note that while the dynamics in (1) do not allow for jumps, they are otherwise quite general. We could, for example, include additional dynamics for some state variables, X_t say, that drive μ_t and Σ_t. Such an assumption would not change any of our analysis and so we do not bother to explicitly specify any state variable dynamics.

Consider now an investor who follows a CP trading strategy, $\theta = (\theta_1 \ \dots \ \theta_n)^\top$, so that at any time t, the fraction of wealth invested in the ith risky asset is constant and equal to θ_i. The fraction invested in the cash account is then given by $1 - \sum_{i=1}^n \theta_i$. The value of the portfolio, W_t, then has the following dynamics

$$\frac{\mathrm{d}W_t}{W_t} = [(1 - \theta^T \mathbf{1})r + \theta^T \mu_t]\mathrm{d}t + \theta^T \Sigma_t \mathrm{d}B_t. \tag{2}$$

2.1. The terminal wealth of a constant proportion trading strategy

We have the following proposition which solves for the terminal wealth corresponding to any CP trading strategy when the price dynamics satisfy (1). In particular, we solve the SDE in (2). This result was originally obtained by Haugh and Jain [15] who used it to study the dual approach for portfolio evaluation that was proposed by Haugh et al. [16]. To be precise, Haugh and Jain [15] assumed that the volatility matrix in (1) was constant but it was clear that their derivation also worked for an adapted volatility matrix process.

Proposition 1. *Suppose price dynamics satisfy (1) and that a static trading strategy is employed so that at each time $t \in [0, T]$ a proportion, θ_i, of time t wealth is invested in the ith risky security for $i = 1, \dots, n$ with $1 - \theta^\top \mathbf{1}$ invested in the risk-free asset. Then the terminal wealth, W_T, resulting from this strategy satisfies*

$$W_T = W_0 \exp\left(\int_0^T \left[(1 - \theta^\top \mathbf{1})r_t + \frac{1}{2}\theta^\top(\mathrm{diag}(\Sigma_t \Sigma_t^\top) - \Sigma_t \Sigma_t^\top \theta)\right] \mathrm{d}t\right) \prod_{i=1}^n \left(\frac{P_T^{(i)}}{P_0^{(i)}}\right)^{\theta_i}. \tag{3}$$

Proof. Using (1) and applying Itô's lemma to $\ln P_T$ we obtain

$$\ln P_T = \ln P_0 + \int_0^T \left(\mu_t - \frac{1}{2}\mathrm{diag}(\Sigma_t \Sigma_t^\top)\right) \mathrm{d}t + \int_0^T \Sigma_t \mathrm{d}B_t. \tag{4}$$

As the wealth dynamics satisfy (2) another simple application of Itô's lemma to $\ln W_T$ then implies

$$W_T = W_0 \exp\left(\int_0^T \left[(1 - \theta^\top \mathbf{1})r_t + \theta^\top \mu_t - \frac{1}{2}\theta^\top \Sigma_t \Sigma_t^\top \theta\right] \mathrm{d}t + \theta^\top \int_0^T \Sigma_t \mathrm{d}B_t\right). \tag{5}$$

Substituting (4) into (5) we then obtain (3) as desired. □

It is worth mentioning again that Proposition 1 is no longer valid if we allow jumps in the security price dynamics. Nonetheless, it is straightforward to show that the proposition would hold approximately in the presence of jumps and we will return to this issue in Section 4.1. Before discussing our main applications of Proposition 1 in Sections 3 and 4, we will briefly discuss some other applications of the proposition.

2.2. Merton's problem

Consider the classic dynamic portfolio optimization problem that was originally considered by Merton [24]. The drift vector, volatility matrix and interest rate are now all assumed to be constant. For an investor with a constant relative risk aversion the optimal solution is to adopt a CP strategy. In this case (3) reduces to

$$W_T = W_0 \exp\left((1 - \theta^\top \mathbf{1})rT + \frac{1}{2}\theta^\top(\mathrm{diag}(\Sigma\Sigma^\top) - \Sigma\Sigma^\top\theta)T\right) \times \prod_{i=1}^n \left(\frac{P_T^{(i)}}{P_0^{(i)}}\right)^{\theta_i}. \tag{6}$$

Despite all the attention that has been paid to this problem in the literature, we have only seen the expression for W_T in (6) in [15].

2.3. Studying return predictability

The CP strategy is often used as a base case when studying the value of *predictability* in security prices. Predictability is often (see, for example, [23]) induced by setting $\Sigma_t = \Sigma$, a constant, and allowing the drift term, μ_t, to be a function of some state vector process, X_t say. In that case the expression in (6) still applies and terminal wealth is a function of *only* the terminal security prices. Moreover, the expected utility of any CP strategy can often be determined in closed form when the distribution of P_T is also known. For example, if $\log(P_t)$ is a (vector) Gaussian process, then W_T is log-normally distributed. Moreover if we assume CRRA utility then the value function for the CP strategy,

$$V_t := \mathrm{E}_t[W_T^{1-\gamma}/(1-\gamma)], \tag{7}$$

can be computed analytically. Using (7) it is also possible to compute the optimal CP strategy and compare it to the CP strategy that an investor would employ if he ignored the predictability of returns and assumed a constant investment opportunity set. (The static strategy obtained from maximizing (7) is the CP strategy that an investor would employ if he was *forced* to select a CP strategy and knew the true price dynamics. In contrast, an investor who was ignorant of the true price dynamics and believed the investment opportunity set was not time varying would *willingly* select a CP strategy. However, the two CP strategies would not coincide.)

Eqs. (6) and (7) can also be used to determine the *myopic* strategy where at each time t, the investor solves his portfolio optimization problem by assuming that the instantaneous moments of asset returns are fixed at their current values for the remainder of the investment horizon. The myopic strategy ignores the hedging component of the optimal trading strategy and has also been studied extensively in the literature. See, for example, [19,21,5,3]. Again (3) and (7) will often simplify the numerical calculations when studying the performance of myopic policies.

2.4. The dual approach to portfolio evaluation

Haugh and Jain [15] used the preceding observations to compute duality-based upper bounds on the value function of the optimal dynamic trading strategy when return predictability was induced via the drift process, μ_t. In addition to improving the efficiency of their numerical algorithms, the closed form expression for the value function in (7) allowed them to construct (in the case of CP strategies) upper bounds on the optimal value function that were superior and theoretically more satisfying than those calculated originally by Haugh et al. [16].

2.5. The Kelly criterion

The Kelly criterion is a particular constant proportion strategy that is only consistent with the goal of maximizing expected utility when the investor has log utility and when there is a constant investment opportunity set. When the investment opportunity set is constant it will outperform any other strategy in the "long run". The related *fractional* Kelly criterion is also a CP strategy and is reputedly commonly employed by some of the most well known and respected investors in the world. In a continuous-time framework, our expression in (3) for W_T lends itself to the analysis of both Kelly and fractional Kelly. It would be quite straightforward using (3), for example, to study the errors made by Kelly when the investment opportunity set is not constant.

3. The CP strategy versus the buy-and-hold portfolio

In this section, we will use (3) to compare the performance of a given CP strategy with the corresponding buy-and-hold strategy. We will now refer to the risk-free asset as the 0th security and use θ_0 to denote the fraction of wealth invested in this security. Unless otherwise stated, we will assume that $0 \le \theta_i \le 1$ for each $i = 0, \ldots, n$ and that $\sum_{i=0}^n \theta_i = 1$ so that borrowing and short sales are forbidden. Consider now a buy-and-hold strategy where at time $t = 0$ we invest a constant proportion, θ_i, of our time $t = 0$ wealth in the ith security for $i = 0, \ldots, n$. Assuming that we started with an initial wealth of W_0, then the gross return at date T is given by

$$\frac{W_T}{W_0} = \sum_{i=0}^n \theta_i \frac{P_T^{(i)}}{P_0^{(i)}}. \tag{8}$$

Similarly (and noting that we now use θ to denote $(\theta_0, \ldots, \theta_n)^\top$) we can rewrite (3) as

$$\frac{W_T}{W_0} = \exp\left(\frac{1}{2}\theta^\top \int_0^T (\mathrm{diag}(\Sigma_t^a \Sigma_t^{a\top}) - \Sigma_t^a \Sigma_t^{a\top}\theta)\mathrm{d}t\right) \times \prod_{i=0}^n \left(\frac{P_T^{(i)}}{P_0^{(i)}}\right)^{\theta_i} \tag{9}$$

$$= \exp\left(\frac{1}{2}\int_0^T \left(\sum_{i=0}^n \theta_i \mathrm{Var}(R_t^{(i)}) - \mathrm{Var}(\theta^\top \mathbf{R}_t)\right)\mathrm{d}t\right) \times \prod_{i=0}^n \left(\frac{P_T^{(i)}}{P_0^{(i)}}\right)^{\theta_i} \tag{10}$$

where Σ_t^a is the instantaneous variance–covariance matrix of the $n+1$ securities and $\mathbf{R}_t = (R_t^{(0)}, \ldots, R_t^{(n)})$ is the time t vector of their instantaneous returns. That is, $R_t^{(i)} = \mathrm{d}P_t^{(i)}/P_t^{(i)}$. Also note that the first row and first column of Σ_t^a contain only zeros and the sub-matrix beginning at the $(2,2)$th element is identical to Σ_t. We now have the following lemma which immediately establishes that the exponential term in (10) is strictly greater than 1.

Lemma 3.1. *If $0 \le \theta_i \le 1$ for each $i = 0, \ldots, n$ and $\sum_{i=0}^n \theta_i = 1$ then*

$$\sum_{i=0}^n \theta_i \mathrm{Var}(R_t^{(i)}) - \mathrm{Var}(\theta^\top \mathbf{R}_t) \ge 0 \quad \text{for all } t \in [0, T]. \tag{11}$$

Proof. Let $\sigma_i^2 := \mathrm{Var}(R_t^{(i)})$ and $\sigma_{ij} := \mathrm{Cov}(R_t^{(i)}, R_t^{(j)})$. First note that since $\sigma_{ij} \le \sigma_i\sigma_j$ for all i, j it follows that

$$\sum_{i<j} \theta_i\theta_j\sigma_{i,j} \le \sum_{i<j} \theta_i\theta_j\sigma_i\sigma_j$$

for all $\theta_i, \theta_j \ge 0$ and so we immediately obtain that $\mathrm{Var}(\theta^\top \mathbf{R}_t) \le (\sum_{i=0}^n \theta_i\sigma_i)^2$. It therefore follows that

$$\sum_{i=0}^n \theta_i \mathrm{Var}(R_t^{(i)}) - \mathrm{Var}(\theta^\top \mathbf{R}_t) \ge \sum_{i=0}^n \theta_i\sigma_i^2 - \left(\sum_{i=0}^n \theta_i\sigma_i\right)^2 = \sum_{i=0}^n \theta_i\sigma_i x_i$$

where $x_i := \sigma_i - \sum_{j=0}^n \theta_j\sigma_j$. But $\sum_{i=0}^n \theta_i x_i = 0$ since $\sum_{i=0}^n \theta_i = 1$ and so it follows that $\sum_{i=0}^n \theta_i\sigma_i x_i \ge 0$ and the result follows. □

3.1. Compensation for earning the geometric mean

We are now in a position to compare (8) and (10). In particular, we can apply the general arithmetic–geometric mean inequality to conclude that

$$\sum_{i=0}^n \theta_i \frac{P_T^{(i)}}{P_0^{(i)}} \ge \prod_{i=0}^n \left(\frac{P_T^{(i)}}{P_0^{(i)}}\right)^{\theta_i} \tag{12}$$

since, by assumption, the θ_i's are all non-negative. In light of Lemma 3.1 and the inequality in (12) we can therefore interpret the exponential term in (10) as the (multiplicative) compensation that an investor receives for accepting the geometric mean of a CP strategy instead of the arithmetic mean of the corresponding buy-and-hold strategy. This compensation is similar to holding a regular option in that the CP strategy is long *gamma*: it therefore profits from the act of rebalancing by selling high and buying low and, conditional on the terminal security prices, it is long volatility.

Moreover an investor in a CP trading strategy benefits from knowing that the geometric mean of the underlying security returns will constitute a lower bound on his overall portfolio return. The degree to which the realized return outperforms this lower bound will depend on the realized variances and covariances of the securities. This long volatility feature of CP strategies has been (at least informally) identified by others. For example, Luenberger [22] demonstrates how an investor can benefit from volatility by rebalancing his portfolio in each period and he refers to this phenomenon as *volatility pumping*. More generally, the large literature on the Kelly criterion (see the references listed earlier in Section 1) and proportional betting has long been aware of this fact. But as stated earlier, the expression in (9) does not seem to be known in the literature and though it is simple to derive, the link between the geometric and arithmetic means also seems to be new.

While these observations apply whenever $0 \leq \theta_i \leq 1$ and $\sum_{i=0}^{n} \theta_i = 1$, i.e. when there are no short selling and no-borrowing constraints on the CP strategy, they can also apply for certain θ vectors that do not satisfy these constraints. This is demonstrated in the following result.

Lemma 3.2. *Suppose $\Sigma_t = \Sigma$ is a constant matrix for all t and that $\sigma_i^2 := \mathrm{Var}(R_t^{(i)})$. Let $((1-\sum_{i=1}^{n}\theta_i^*), \theta_1^*, \ldots, \theta_n^*)$ denote the vector of portfolio weights that maximizes the exponential term in (10). Then the vector of optimal weights on the risky securities, $\theta^* = (\theta_1^*, \ldots, \theta_n^*)$, satisfies*

$$\theta^* := \frac{1}{2}(\Sigma\Sigma^\top)^{-1}[\sigma_1^2 \ \cdots \ \sigma_n^2]^\top \tag{13}$$

in which case the exponential term in (10) reduces to

$$\exp\left(\frac{T}{4}[\sigma_1^2 \ \cdots \ \sigma_n^2](\Sigma\Sigma^\top)^{-1}[\sigma_1^2 \ \cdots \ \sigma_n^2]^\top\right) \tag{14}$$

which is always greater than or equal to one.

Proof. Let $f(\theta_0, \ldots, \theta_n) := \sum_{i=0}^{n} \theta_i \mathrm{Var}(R_t^{(i)}) - \mathrm{Var}(\theta^\top \mathbf{R}_t)$ which is the integrand in the exponential term in (10). Since the first security is assumed to be risk-free we may write

$$f(\theta_0, \ldots, \theta_n) = \sum_{i=1}^{n} \theta_i(1-\theta_i)\sigma_i^2 - 2\sum_{1\leq i<j\leq n} \theta_i\theta_j\sigma_{ij}. \tag{15}$$

The first-order optimality conditions for maximizing f are given by

$$\Sigma\Sigma^\top\theta^* = [\sigma_1^2 \ \cdots \ \sigma_n^2]^\top \tag{16}$$

which has (13) as its solution. It is also easy to check that the second-order conditions for a maximum are satisfied. Substituting (13) into (15), integrating from 0 to T and applying the exponential function we immediately obtain (14). Moreover, since the inverse of a positive-definite covariance matrix is also positive-definite it follows that the left-hand side of (14) is greater than or equal to one. Finally the position in the risk-free asset is given by $(1 - \sum_{i=1}^{n} \theta_i^*)$. □

Referring to Lemma 3.2, note that there is no reason why some components of θ^* cannot be negative or exceed one. As a result, it is possible that the CP strategy with the greatest, i.e. most positive, exposure to realized variances and covariances requires short selling and leveraged positions in the underlying securities. Of course, while the arithmetic–geometric mean inequality will generally no longer apply, we can still interpret the exponential term in (10) as the (multiplicative) premium that you earn for following the CP strategy instead of the corresponding buy-and-hold strategy. Also note that the assumption of a constant matrix, Σ, in Lemma 3.2 is not strictly necessary to draw these conclusions.

These observations are perhaps surprising, given our results on leveraged ETFs in Section 4 where $n = 1$ and $|\theta_1| > 1$. In this case, we will see that the exposure to realized variance is always negative, the arithmetic–geometric mean inequality does not apply and that the exponential term in (10) can be interpreted as the (multiplicative) premium that you must *pay* for following the CP strategy, i.e. for purchasing the leveraged ETF.

4. Leveraged ETFs

A leveraged ETF is an exchange-traded derivative security with just a single underlying security or index. It promises to track θ times the *daily* performance of the underlying index and usually achieves this through the use of total return swaps. As in [1], we can approximate the value of the LETF with the following stochastic differential equation

$$\frac{dL_t}{L_t} = \theta\frac{dP_t}{P_t} + (1-\theta)r\,dt - f\,dt \tag{17}$$

where L_t is the time t value of the LETF and f is the constant *expense ratio* of the LETF. The $(1-\theta)r\,dt$ term in (17) reflects the cost of funding the leveraged position (when $\theta > 1$) or the risk-free income from an inverse ETF (when $\theta < 0$). Avellaneda and Zhang [1] and (in the case where P_t follows a geometric Brownian motion) Cheng and Madhavan [7] solved (17) to obtain

$$\frac{L_T}{L_0} = \left(\frac{P_T}{P_0}\right)^\theta \exp\left((1-\theta)rT - fT + \frac{1}{2}\theta(1-\theta)\int_0^T \sigma_t^2 dt\right) \tag{18}$$

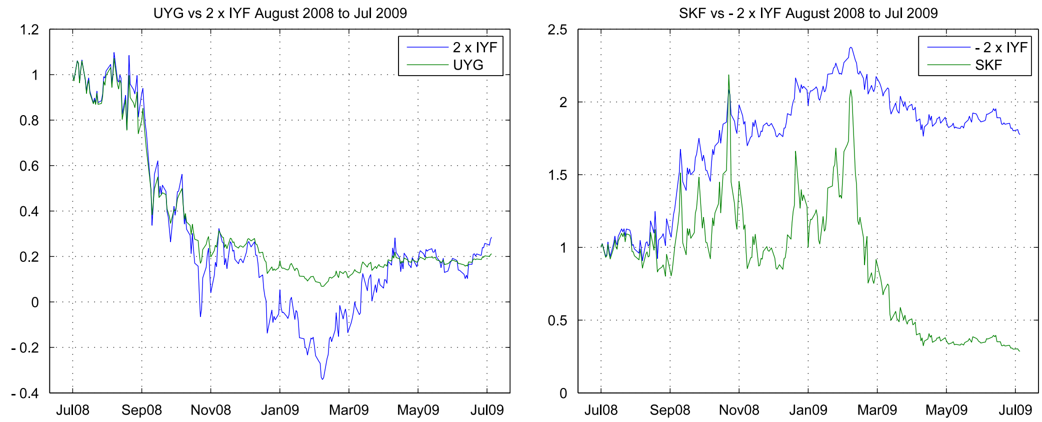
and used this expression to explain the empirical performance of LETFs.

One of the principal motivations for this note is to argue that this performance should have been anticipated in the market, given the ubiquity of CP trading strategies in the literature. Indeed if we ignore the expense ratio, then it is clear from (17) that the dynamics of L_t are simply those of a CP trading strategy and indeed (3) reduces to (18) in the single risky asset case where we now write σ_t for Σ_t. The expense ratio, being deterministic, simply results in the time T value of the LETF being reduced by a factor of $\exp(-fT)$.

It is worthwhile contrasting (18) with the time T value of a *static* position of θ times the underlying index that was initiated at time $t = 0$. If we denote the time T value of such a position by S_T, then it is clear that

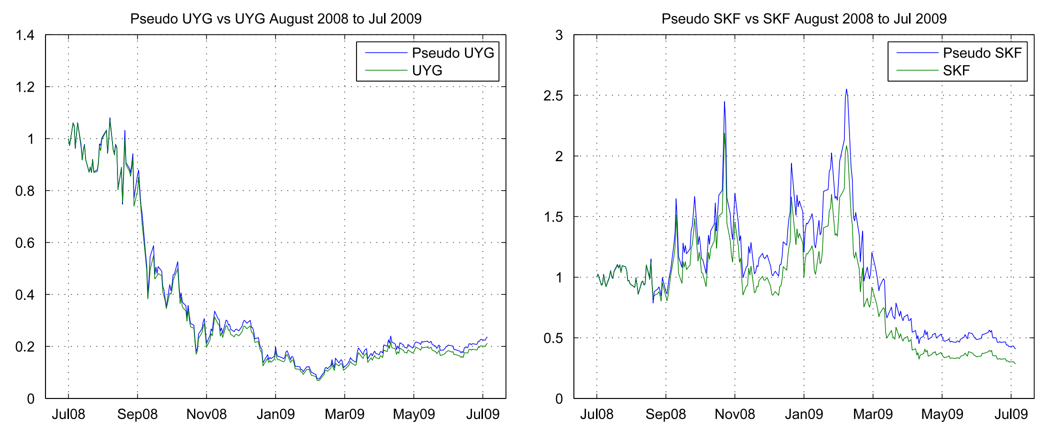
$$\frac{S_T}{S_0} = \frac{\theta P_T - (\theta - 1)\exp(rT)P_0}{P_0}. \tag{19}$$

This of course is just (8) in the case of a single risky security together with the risk-less cash account. Many of the original investors in LETFs believed that their returns would resemble the returns in (19) once they had adjusted for the expense ratio, f. And while they would have been justified in this belief in times of low volatility and short investment horizons, the difference between (18) and (19) can be quite remarkable when realized volatility is high.



(a) ProShares ultra financial. (b) ProShares ultra short financial.

Fig. 1. Performance of LETFs versus leveraged buy-and-hold positions in underlying index.



(a) ProShares ultra financial. (b) ProShares ultra short financial.

Fig. 2. Actual performance of leveraged ETFs versus performance predicted by (18).

Consider, for example Fig. 1(a) where we have plotted the performance of the ProShares Ultra Financial ETF (ticker *UYG*) against a static position of 2 $\times$ the I-Shares Financial Sector ETF (ticker *IYF*) between August 2008 and 2009. We assume that we entered into both positions at the end of July 2008. The ProShares ETF is a leveraged ETF that is designed to track two times the *daily* performance of the Dow Jones Financial Index (DJFI) while the I-Shares ETF is designed to simply track the DJFI. The discrepancy between the two performances is dramatic and is explained by the very high level of realized variance in that period. Note that an investor in a leveraged ETF with $\theta = 2$ is short realized variance as suggested by (18).

Similarly in Fig. 1(b), we have plotted the performance of the ProShares Ultra Short Financial ETF (ticker *SKF*) against a static position of $-2 \times$ the I-Shares Financial Sector ETF, again between August 2008 and 2009. Note that the ProShares Ultra Short Financial ETF is a leveraged ETF that is designed to track *minus* two times the *daily* performance of the DJFI. The discrepancy between the two is again dramatic and is of course explained by the very high level of realized variance in that period. Note that an investor in a leveraged ETF with $\theta = -2$ is once again short realized variance.

In fact, any value of $\theta < 0$ or $\theta > 1$ results in a negative exposure to realized variance for a fixed value of the terminal price of the underlying index, P_T. As demonstrated in Lemma 3.2, this does not hold in the more general case when following a CP strategy with multiple risky securities. Moreover the effect is asymmetric in that an LETF with a leverage of $\theta > 1$ is not as short variance as an inverse ETF with a leverage of $-\theta$. This short position in realized variance is best explained by noting that for values of $\theta \notin [0, 1]$ the act of daily rebalancing will require the manager of the LETF to "sell low" and "buy high". The greater the realized variance, the greater the magnitude of rebalancing and so the greater the losses on the LETF. Thus the daily rebalancing is similar to delta-hedging a short position in a vanilla European option where one is also short realized variance. In contrast, an ETF corresponding to a value of $\theta \in (0, 1)$ is long realized variance and benefits from high levels of realized variance, again conditional on P_T. We will return to this issue again in Section 4.2.

Our discussion on the performance of leveraged ETFs has implicitly assumed that their returns are well approximated by (18). As demonstrated by Avellaneda and Zhang [2,1], this is indeed the case, even when markets are highly volatile. They analyze the tracking error when the actual performance of LETFs is approximated by (18) and conclude that even in very volatile markets, the error is small. For example, in Fig. 2(a) we graph the performance of the ProShares Ultra Financial LETF ($\theta = 2$) against the performance implied in (18). We assumed in the latter case that $r = f = 1\%$. (Over this period the 1-month risk-free rate moved from approximately 1.6% to.1% but we find using a constant rate of 1% makes no discernable difference to the results.) Note that the two graphs are in extremely close agreement with one another despite the very high levels of realized volatility during that period.

Similarly in Fig. 2(b), we graph the performance of the ProShares Ultra Short Financial LETF ($\theta = -2$) against the performance implied in (18). We again assumed in the latter case that

$r = f = 1\%$. While the two graphs are very similar there is nonetheless a clear discrepancy between the two which Avellaneda and Zhang put down to the difficulty in shorting financial stocks during this period. Of the 56 LETFs considered by Avellaneda and Zhang, this was atypical and the tracking error was generally closer to that of Fig. 2(a). These observation help to justify our earlier claim that (3) should hold approximately even in the presence of jumps.

Further empirical evidence concerning the ability of (18) to explain LETF performance can be found in [2]. (Indeed we have intentionally examined in this section some of the same datasets considered by Avellaneda and Zhang [2].)

Before concluding this section, it is worth mentioning the effects that the presence of LETFs can have on market microstructure. Because LETFs need to buy at the close when the market is up and sell at the close when the market is down, they have been blamed (see, for example, [20]) for increasing volatility at the close. Furthermore, because the direction of the daily rebalancing trades are widely known in the market, it is suspected that many proprietary trading desks have regularly front-run these trades. They are therefore suspected of adding to market volatility at the close as well as negatively impacting LETF performance. Cheng and Madhavan [7] provide an account of these microstructure effects and estimate the aggregate daily hedging demand of LETFs in the market. (It is perhaps also worth mentioning that variance-swaps are somewhat similar in that the hedging of variance-swaps also requires daily rebalancing at the close.)

4.1. How good is the approximation?

As mentioned above, the approximation in (3) appears to work very well even when markets are volatile. This of course was the case during the financial crisis of 2008, a period included in Figs. 1 and 2 above. It is worth mentioning, however, that the largest daily log-return in the I-Shares Financial Sector ETF (ticker *IYF*) between August 1st 2008 and July 31st 2009 was 14.6% whereas the largest negative return was −17.1%. While such daily moves are indeed large, it would of course be possible to obtain even more extreme returns if the security underlying the LETF was a stock say, rather than an index which is typically less volatile.

In this section, we therefore consider the performance of the approximation in (3) (or equivalently in (18)) when more extreme moves are possible. In particular, we consider the case where the true price process is a jump-diffusion process and analyze the performance of the approximation under these dynamics. Before doing so, however, it is worth emphasizing that when sampling discretely from the underlying price process (as is the case when an LETF is rebalanced daily rather than continuously) it is not possible to infer whether the process is a diffusion process or a discontinuous process. As a result, the question of how well (18) performs for very volatile markets might just as well be answered by considering diffusions with very large instantaneous volatilities as opposed to explicitly considering jump-diffusions. Nonetheless we will consider the jump-diffusion case as it is easier to isolate the approximation errors that occur in this case. We therefore assume that the price dynamics now satisfy

$$\frac{dP_t}{P_t} = \mu_t dt + \sigma_t dB_t + dJ_t \tag{20}$$

where J_t is a pure-jump process. We can solve (20) to obtain

$$P_T = P_0 \exp\left(\int_0^T \left(\mu_t - \frac{\sigma_t^2}{2}\right) dt + \int_0^T \sigma_t dB_t\right) \times \prod_{0<s\leq T} (1+\eta_s) \tag{21}$$

where the product in (21) is taken over the jump times in $[0, T]$ and $\eta_s \in [-1, \infty)$ is the relative jump size when a jump occurs at time s.

Consider now a constant proportion trading strategy where θ denotes the target fraction of wealth invested in the risky security. (We say "target" since the fraction of wealth invested in the security will no longer be θ immediately after a jump in the price of the underlying security.) The terminal wealth of the CP strategy is then easily seen to satisfy

$$\begin{aligned} W_T &= W_0 \exp\left(\int_0^T \left[(1-\theta)r_t + \theta\mu_t - \frac{\theta^2\sigma_t^2}{2}\right] dt + \theta \int_0^T \sigma_t dB_t\right) \prod_{0<s\leq T} (1+\theta\eta_s) \\ &= W_0 \left(\frac{P_T}{P_0}\right)^{\theta} \exp\left((1-\theta)\int_0^T r_t dt + \frac{1}{2}\theta(1-\theta)\int_0^T \sigma_t^2 dt\right) \prod_{0<s\leq T} \frac{(1+\theta\eta_s)}{(1+\eta_s)^{\theta}}. \end{aligned} \tag{22}$$

Comparing (22) with (3) (in the case of a single underlying security) we see that the only *apparent* difference between the two expressions is

$$\prod_{0<s\leq T} \frac{(1+\theta\eta_s)}{(1+\eta_s)^{\theta}} \tag{23}$$

which clearly converges to 1 as the permitted relative jump sizes decrease in magnitude to 0. However, this difference is somewhat misleading as the relative error of (23) would be partially offset in the approximation (3). To see this more clearly, suppose that the true price process is a pure-jump process which therefore has no diffusive component. This implies that the term

$$\exp\left(\frac{1}{2}\theta(1-\theta)\int_0^T \sigma_t^2 dt\right)$$

in (22) vanishes. A user of the approximation in (3), however, will still calculate such a term. In particular they will include the term

$$\exp\left(\frac{1}{2}\theta(1-\theta)\sum_{0<s\leq T} \log(1+\eta_s)^2\right) \tag{24}$$

when they approximate (3) and this follows since $\sum_{0<s\leq T} \log(P_s/P_{s-})^2 = \sum_{0<s\leq T} \log(1+\eta_s)^2$. (The expression in (24) is consistent with our approximation of $\int_0^T \sigma_t^2 dt$ with $\sum \log(P_{i+1}/P_i)^2$ which we have used to generate Figs. 1 and 2. We are therefore implicitly ignoring the drift terms when approximating $\int_0^T \sigma_t^2 dt$. This is a standard practice (see the variance-swaps market, for example) and is justified because the drift terms are negligible at the frequency of daily rebalancing.)

Therefore the relative error that results from jumps, or equivalently daily rebalancing rather than continuous rebalancing, is more accurately expressed as

$$\exp\left(-\frac{1}{2}\theta(1-\theta)\sum_{0<s\leq T} \log(1+\eta_s)^2\right) \prod_{0<s\leq T} \frac{(1+\theta\eta_s)}{(1+\eta_s)^{\theta}}. \tag{25}$$

Note in particular the inclusion of the minus sign in the exponential term in (25). Note also that our argument implicitly assumes that there is at most one jump per day but this assumption was only made to simplify the exposition and is easily relaxed.

In Table 1, we calculate both (23) and (25) for the case where exactly one jump occurs in $[0, T]$. We consider relative jump sizes ranging from −50% to +100% and leverage values equal to −2, −1, 2 and 3. First note that, we have used 0^* to denote those cases where (23) or (25) is *negative*. This occurs when W_T in

Table 1
Approximation errors when using (18) to calculate LETF performance.

Relative jump size (%)	Approximation error $= (1+\theta\eta_s)/(1+\eta_s)^\theta$				Corrected approximation error $= e^{(-\frac{1}{2}\theta(1-\theta)\log(1+\eta_s)^2)}(1+\theta\eta_s)/(1+\eta_s)^\theta$			
	$\theta=-2$	$\theta=-1$	$\theta=2$	$\theta=3$	$\theta=-2$	$\theta=-1$	$\theta=2$	$\theta=3$
−50	0.500	0.75	0	0*	2.1132	1.213	0	0*
−40	0.648	0.84	0.556	0*	1.4176	1.090	0.721	0*
−30	0.784	0.91	0.816	0.292	1.1483	1.034	0.927	0.427
−20	0.896	0.96	0.938	0.781	**1.0404**	1.009	**0.985**	0.907
−10	0.972	0.99	0.988	0.960	**1.005**	1.001	**0.999**	0.993
10	0.968	0.99	0.992	0.977	**0.995**	0.999	**1.001**	1.004
20	0.8640	0.96	0.972	0.926	**0.955**	0.992	**1.005**	1.023
30	0.676	0.91	0.947	0.865	0.8311	0.975	1.014	1.063
40	0.392	0.84	0.918	0.802	0.551	0.941	1.029	1.126
50	0	0.75	0.889	0.741	0	0.884	1.048	1.213
60	0*	0.64	0.859	0.684	0*	0.798	1.072	1.326
70	0*	0.51	0.830	0.631	0*	0.676	1.101	1.469
80	0*	0.36	0.803	0.583	0*	0.509	1.134	1.644
90	0*	0.19	0.776	0.539	0*	0.287	1.171	1.857
100	0*	0	0.750	0.500	0*	0	1.213	2.113

(22) is negative due to the combination of relative jump size and the leverage, θ. Due to the limited liability of an LETF position, however, we have replaced these terms with 0. As expected, the corrected approximation error is closer to 1 than the uncorrected approximation error. Moreover, we see that these approximation errors become more significant at the higher levels (positive or negative) of leverage. This of course is to be expected.

We have also highlighted in bold those parts of Table 1 that are most relevant to Figs. 1 and 2, which have leverage levels of $\theta = 2$ and $\theta = -2$, respectively. As mentioned earlier the largest daily log-return in the underlying security, the I-Shares Financial Sector ETF, was +14.6% in the relevant time period, whereas the largest negative daily log-return was −17.1%. It is clear from Table 1 that the corrected approximation error should be very small for jumps of these magnitudes. Indeed it is easy to check that the estimated relative errors for these realized returns (−17.1% and +14.6%) are closer to the ±10% errors rather than the ±20% errors. The performance of the approximation in (3) in Figs. 1 and 2 is therefore not surprising.

What is interesting about Table 1 is how significant the approximation errors can be when we consider extreme moves that are greater in magnitude than 30%, say. It is clear then that the approximation becomes considerably less accurate and should not be relied upon as a guide to LETF *pathwise* performance. This then suggests that (3) would be a poor approximation for LETF performance when the underlying securities are volatile single stocks.

4.2. A constant proportion ETF?

Given the inability of most investors to time the market, constant proportion trading strategies should, at least in the absence of market frictions, be reasonably close to optimal for investors with power or logarithmic utility. The costs associated with daily rebalancing, however, would be prohibitively expensive and time consuming for individual investors. It might be possible, however, for an actively managed ETF to employ such a strategy. It would be similar to a regular LETF only instead of one underlying risky security, there could be n underlying risky securities. For example, an investor wishing to invest in global equity markets might be interested in an ETF that tracks the daily returns of the S&P 500, the Eurostoxx 50 and the Nikkei 225. In this case, we would have $n = 3$. Moreover, if $0 \le \theta_i$ for $i = 1, 2, 3$ and $\sum_{i=1}^{3}\theta_i \le 1$, then we know such a product would have a long exposure to market volatility, in contrast to LETFs. Such a product, a *constant proportion* ETF (CPETF) say, could be suitable for unsophisticated investors.

In addition, the manager of a CPETF would necessarily sell at the close after an up-day and buy at the close after a down-day and would therefore tend to dampen market volatility at the close. If rebalancing costs were too expensive, then the CPETF could be allowed to rebalance less frequently, say once a week or once a month. Or alternatively, it might be required to be balanced at the close only one day a month. This would make it difficult for proprietary trading desks to front-run the CPETF manager. Of course, such a CPETF would only be permitted as long as it satisfied various regulatory and transparency requirements. While less frequent rebalancing would render (9) a less useful approximation, the insights from (9) should still apply.

Acknowledgments

The author would like to thank an anonymous referee for valuable comments.

References

[1] M. Avellaneda, S.J. Zhang, Path-dependence of leveraged ETF returns, SIAM Journal of Financial Mathematics 1 (2010) 586–603.
[2] M. Avellaneda, S.J. Zhang, Leveraged ETFs: all you wanted to know but were afraid to ask, Risk Professional (February) (2010) 54–60.
[3] S. Basak, G. Chabakauri, Dynamic hedging in incomplete markets: a simple solution, Working paper, London Business School, 2008.
[4] L. Breiman, Optimal gambling systems for favorable games, in: Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1961, pp. 65–78.
[5] C. Brooks, O.T. Henry, G. Persand, The effect of asymmetries on optimal hedge ratios, Journal of Business 75 (2002) 333–352.
[6] S. Browne, The return on investment from proportional portfolio strategies, Advances in Applied Probability 30 (1998) 216–238.
[7] M. Cheng, A. Madhavan, The dynamics of leveraged and inverse exchange traded funds, 2009. Available at SSRN: http://ssrn.com/abstract=1539120.
[8] T.M. Cover, Universal portfolios, Mathematical Finance 1 (1) (1991) 1–29.
[9] T.M. Cover, J. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[10] J. Cvitanic, F. Zapatero, Introduction to the Economics and Mathematics of Financial Markets, MIT Press, Cambridge, Massachusetts, 2004.
[11] D. Duffie, Dynamic Asset Pricing Theory, Princeton University Press, Princeton, New Jersey, 2004.
[12] S.N. Ethier, The proportional bettor's fortune, in: Proceedings of the 5th International Conference on Gambling and Risk Taking, vol. 4, 1988, pp. 375–383.
[13] S.N. Ethier, S. Tavare, The proportional bettor's return on investment, Journal of Applied Probability 20 (1983) 563–573.
[14] N.H. Hakansson, Optimal investment and consumption strategies under risk for a class of utility functions, Econometrica 38 (1970) 587–607.
[15] M.B. Haugh, A. Jain, The dual approach to portfolio evaluation: a comparison of the static, myopic and generalized buy-and-hold strategies, Quantitative Finance 11 (1) (2011) 81–99.
[16] M.B. Haugh, L. Kogan, J. Wang, Evaluating portfolio strategies: a duality approach, Operations Research 54 (3) (2006) 405–418.
[17] I. Karatzas, S.E. Shreve, Methods of Mathematical Finance, Springer-Verlag, New York, New York, 1997.
[18] J. Kelly, A new interpretation of information rate, Bell System Technical Journal 35 (1956) 917–926.
[19] K.F. Kroner, J. Sultan, Time-varying distributions and dynamic hedging with foreign currency futures, Journal of Financial and Quantitative Analysis 28 (1993) 535–551.

[20] T. Lauricella, S. Pulliam, D. Gullapalli, Are etfs driving late-day turns? leveraged vehicles seen magnifying other bets; last-hour volume surge, Wall Street Journal article, December 15, 2008.
[21] A. Lioui, P. Poncet, The minimum variance hedge ratio under stochastic interest rates, Management Science 46 (2000) 658–668.
[22] D.G. Luenberger, Investment Science, Oxford University Press, New York, New York, 1998.
[23] A. Lynch, Portfolio choice and equity characteristics: characterizing the hedging demands induced by return predictability, Journal of Financial Economics 62 (2001) 67–130.
[24] R.C. Merton, Lifetime portfolio selection under uncertainty: the continuous-time case, Review of Economics and Statistics 51 (1969) 247–257.
[25] R.C. Merton, Optimum consumption and portfolio rules in a continuous-time model, Journal of Economic Theory 3 (1971) 373–413.
[26] R.C. Merton, Continuous-Time Finance, Basil Blackwell, New York, 1990.
[27] P. Samuelson, Lifetime portfolio selection by dynamic stochastic programming, Review of Economics and Statistics 51 (1969) 239–246.
[28] E.O. Thorp, Portfolio choice and the Kelly criterion, in: Proceedings of the 1971 Business and Economics Section of the American Statistical Association, 1972, pp. 215–224.