Universal Portfolios

Thomas M. Cover*
Stanford University

October 23, 1996

Abstract

We exhibit an algorithm for portfolio selection that asymptotically outperforms the best stock in the market. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})^t$ denote the performance of the stock market on day i , where x_{ij} is the factor by which the j-th stock increases on day i . Let $\mathbf{b}_i = (\mathrm{b}_{i1}, \mathrm{b}_{i2}, \ldots, \mathrm{b}_{im})^t$, $\mathrm{b}_{ij} \geq 0, \sum_j \mathrm{b}_{ij} = 1$, denote the proportion b_{ij} of wealth invested in the j-th stock on day i . Then $S_n = \prod_{i=1}^n \mathbf{b}_i^t \mathbf{x}_i$ is the factor by which wealth is increased in n trading days.

Consider as a goal the wealth $S_n^* = \max_{\mathbf{b}} \prod_{i=1}^n \mathbf{b}^t \mathbf{x}_i$ that can be achieved by the best constant rebalanced portfolio chosen *after* the stock outcomes are revealed. It can be shown that S_n^* exceeds the best stock, the Dow Jones average, and the value line index at time n. In fact, S_n^* usually exceeds these quantities by an exponential factor.

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots,$ be an arbitrary sequence of market vectors. It will be shown that the nonanticipating sequence of portfolios $\hat{\mathbf{b}}_k = \int \mathbf{b} \prod_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i d\mathbf{b} / \int \prod_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i d\mathbf{b}$ yields wealth $\hat{S}_n = \prod_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k$ such that $(\frac{1}{n}) \ln(S_n^* / \hat{S}_n) \to 0$, for every bounded sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots,$ and, under mild conditions, achieves

$$\hat{S}_n \sim S_n^* (m-1)! (2\pi/n)^{(m-1)/2} / \mid J_n \mid^{1/2} \quad ,$$

where J_n is an $(m-1) \times (m-1)$ sensitivity matrix. Thus this portfolio strategy has the same exponential rate of growth as the apparently unachievable S_n^* .

*Departments of Statistics and Electrical Engineering, Durand Room 121, Stanford, CA, 94305. This work was partially supported by NSF Grant NCR 89-14538.

1 Introduction.

We consider a sequential portfolio selection procedure for investing in the stock market with the goal of performing as well as if we knew the empirical distribution of future market performance. Throughout the paper we are unwilling to make any statistical assumption about the behavior of the market. In particular, we allow for the possibility of market crashes such as those occurring in 1929 and 1987. We seek a robust procedure with respect to the arbitrary market sequences that occur in the real world.

We first investigate what a natural goal might be for the growth of wealth for arbitrary market sequences. For example a natural goal might be to outperform the best buy-and-hold strategy, thus beating an investor who is given a look at a newspaper n days in the future.

We propose a more ambitious goal. To motivate this goal let us consider all constant rebalanced portfolio strategies. Let $\mathbf{x} = (x_1, x_2, \ldots, x_m)^t \geq 0$ denote a *stock market vector* for one investment period, where x_i is the *price relative* for the ith stock, i.e., the ratio of closing to opening price for stock i . A *portfolio* $\mathbf{b} = (b_1, b_2, \ldots, b_m)^t$, $b_i \geq 0$, $\sum b_i = 1$, is the proportion of the current wealth invested in each of the m stocks. Thus $S = \mathbf{b}^t\mathbf{x} = \sum b_i x_i$, where $\mathbf{b}$ and $\mathbf{x}$ are considered to be column vectors, is the factor by which wealth increases in one investment period using portfolio $\mathbf{b}$.

Consider an arbitrary (nonrandom) sequence of stock vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \epsilon \mathbf{R}^m_+$. Here x_{ij} is the price relative of stock j on day i. A constant rebalanced portfolio strategy $\mathbf{b}$ achieves wealth

$$S_n(\mathbf{b}) = \prod_{i=1}^{n} \mathbf{b}^t\mathbf{x}_i \quad , \tag{1}$$

where the initial wealth $S_0(\mathbf{b}) = 1$ is normalized to one. Let

$$S_n^* = \max_{\mathbf{b}} S_n(\mathbf{b}) \tag{2}$$

denote the maximum wealth achievable on the given stock sequence maximized over all constant rebalanced portfolios. Our goal is to achieve S_n^*.

We will be able to show that there is a "universal" portfolio strategy $\hat{\mathbf{b}}_k$, where $\hat{\mathbf{b}}_k$ is based only on the past $\mathbf{x}_1, \mathbf{x}_2, \ldots, \hat{\mathbf{x}}_{k-1}$, that will perform asymptotically as well as the best constant rebalanced portfolio based on foreknowledge of the sequence of price relatives. At first it may seem surprizing that the portfolio $\hat{\mathbf{b}}_k$ should depend on the past, because the future has no relationship to the past. Indeed the stock sequence is arbitrary, and a malicious nature can structure future $\mathbf{x}_k$'s to take advantage of past beliefs as expressed in the portfolio $\hat{\mathbf{b}}_k$. Nonetheless the resulting wealth can be made to track S_n^*.

The proposed universal adaptive portfolio strategy is the performance weighted strategy specified by

$$\hat{\mathbf{b}}_1 = (\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}),$$

$$\hat{\mathbf{b}}_{k+1} = \int \mathbf{b} S_k(\mathbf{b}) d\mathbf{b} / \int S_k(\mathbf{b}) d\mathbf{b}, \tag{3}$$

where

$$S_k(\mathbf{b}) = \prod_{i=1}^{k} \mathbf{b}^t \mathbf{x}_i \ , \tag{4}$$

and the integration is over the set of $(m-1)$- dimensional portfolios

$$B = \{\mathbf{b} \epsilon R^m : b_i \geq 0, \sum_{i=1}^{m} b_i = 1 \}. \tag{5}$$

The wealth $\hat{S}_n$ resulting from the universal portfolio is given by

$$\hat{S}_n = \prod_{k=1}^{n} \hat{\mathbf{b}}_k^t \mathbf{x}_k \ . \tag{6}$$

Thus the initial universal portfolio $\hat{\mathbf{b}}_1$ is uniform over the stocks, and the portfolio $\hat{\mathbf{b}}_k$ at time k is the performance weighted average of all portfolios $\mathbf{b} \epsilon B$. An approximate computation will be given in Section 8, and a generalization of this algorithm will be given in Section 9.

We will show that

$$\frac{1}{n} \ln \hat{S}_n - \frac{1}{n} \ln S_n^* \to 0 \ , \tag{7}$$

for arbitrary bounded stock sequences $\mathbf{x}_1, \mathbf{x}_2, \ldots$. Thus $\hat{S}_n$ and S_n^* have the same exponent to first order. A more refined analysis shows

$$\hat{S}_n \sim \sqrt{\frac{2\pi}{nJ_n}} S_n^* \ , \tag{8}$$

in a sense that will be made precise. It is difficult to summarize the behavior of $\hat{S}_n$ relative to S_n^* because of the arbitrariness of the sequence and the fact that we cannot assume a limiting distribution. For example, even the limit of $\frac{1}{n} \ln S_n^*$ cannot be assumed to exist.

The goal of uniformly achieving $S_n^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$, as specified in (7), was partially achieved by Cover and Gluss (1986) for discrete valued stock markets by using the theory of compound sequential Bayes decision rules developed in Robbins (1951), Hannan and Robbins (1955), and the game-theoretic approachability-excludability theory of Blackwell (1956a, 1956b). Work on natural investment goals can be found in Samuelson (1967) and Arrow (1974). The vast theory of undominated portfolios in the mean-variance plane is exemplified in Markowitz (1952) and Sharpe (1963), while the theory of rebalanced portfolios for known underlying distributions is developed in Kelly (1956), Mossin (1968), Thorp (1971), Markowitz (1976), Hakansson (1979), Bell and Cover (1980, 1988), Cover and King (1978), Cover (1984), Barron and Cover (1988), and Algoet and Cover (1988). A spirited defense of utility theory and the incompatibility of utility theory with the asymptotic growth rate approach is made in Samuelson (1967, 1969, 1979) and Merton and Samuelson (1974).

We see the present work as a departure from the above model-based investment theories, whether they be based on utility theory or growth rate optimality. Here the goal $S_n^* = \max_{\mathbf{b}} \prod_{i=1}^n \mathbf{b}^t \mathbf{x}_i$ depends solely on the data and does not depend upon underlying statistical assumptions. Moreover, Theorem 1, for example, provides a finite sample lower bound for the performance $\hat{S}_n$ of the universal portfolio with respect to S_n^*. Therefore the case for success rests almost entirely on the acceptance of S_n^* as a natural investment goal.

The performance of the universal portfolio is exhibited in Section 8, where numerous examples are given of $S_n(\mathbf{b}), S_n^*$ and $\hat{S}_n$ for various pairs of stocks. In general, volatile uncorrelated stocks lead to great gains of S_n^* and $\hat{S}_n$ over the best buy-and-hold strategy. However, ponderous stocks like IBM and Coca Cola show only modest improvements.

2 Elementary Properties.

We wish to show that the wealth $\hat{S}_n$ generated by the universal portfolio strategy $\hat{\mathbf{b}}_k$ exceeds the value line index and that $\hat{S}_n$ is invariant under permutations of the stock sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. We will use the notation

$$W(\mathbf{b}, F) = \int \ln \mathbf{b}^t \mathbf{x} dF(\mathbf{x}) \tag{9}$$

$$W^*(F) = \max_{\mathbf{b}} W(\mathbf{b}, F) \ , \tag{10}$$

and we will denote by F_n the empirical distribution associated with $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, where F_n places mass $\frac{1}{n}$ at each $\mathbf{x}_i$. In particular we note that

$$S_n^* = \max_{\mathbf{b}} S_n(\mathbf{b}) = \max_{\mathbf{b}} \prod_{i=1}^{n} \mathbf{b}^t \mathbf{x}_i = e^{nW^*(F_n)} \ . \tag{11}$$

For purposes of comparison, we pay special attention to buy-and-hold strategies $\mathbf{b} = \mathbf{e}_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$, where e_j is the j-th basis vector. Note that

$$S_n(\mathbf{e}_j) = \Pi_{k=1}^{n} \mathbf{e}_j^t \mathbf{x}_k = \Pi_{k=1}^{n} x_{kj} \tag{12}$$

is the factor by which the j-th stock increases in n investment periods. Thus $S_n(\mathbf{e}_j)$ is the result of the buy-and-hold strategy associated with the j-th stock.

We now note some properties of the target wealth S_n^*:

Proposition 1 (*Target exceeds best stock*)**:**

$$S_n^* \geq \max_{j=1,2,\ldots,m} S_n(\mathbf{e}_j) \ . \tag{13}$$

Proof: S_n^* is a maximization of $S_n(\mathbf{b})$ over the simplex, while the right hand side is a maximization over the vertices of the simplex. □

Proposition 2 (*Target exceeds value line*)**:**

$$S_n^* \geq \left(\Pi_{j=1}^{m} S_n(\mathbf{e}_j) \right)^{1/m} \tag{14}$$

Proof: Each $S_n(\mathbf{e}_j)$ is $\leq S_n^*$. □

The next proposition shows that the target exceeds the DJIA.

Proposition 3 (*Target exceeds arithmetic mean*)**:** If $\alpha_j \geq 0$, $\sum \alpha_j = 1$, then

$$S_n^* \geq \sum_{j=1}^{m} \alpha_j S_n(\mathbf{e}_j) \tag{15}$$

Proof:

$$S_n(\mathbf{e}_j) \leq S_n^*, \ \ j = 1, 2, \ldots, m \ .\Box \tag{16}$$

Thus S_n^* exceeds the arithmetic mean, the geometric mean, and the maximum of the component stocks. Finally, it follows by inspection that S_n^* does not depend on the order in which $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ occur:

Proposition 4: $S_n^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ is invariant under permutations of the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

Now recall the proposed portfolio algorithm in (3) with the resulting wealth

$$\hat{S}_n = \Pi_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k \quad . \tag{17}$$

It will be useful to recharacterize $\hat{S}_n$ in the following way.

Lemma 1

$$\hat{S}_n = \prod_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k = \int S_n(\mathbf{b}) d\mathbf{b} / \int d\mathbf{b} \tag{18}$$

where

$$S_n(\mathbf{b}) = \prod_{i=1}^n \mathbf{b}^t \mathbf{x}_i \quad . \tag{19}$$

Thus the wealth $\hat{S}_n$ resulting from the universal portfolio is the average of $S_n(\mathbf{b})$ over the simplex.

Proof: Note from (3) and (4) that

$$\begin{aligned} \hat{\mathbf{b}}_k^t \mathbf{x}_k &= \frac{\int \mathbf{b}^t \mathbf{x}_k \prod_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i d\mathbf{b}}{\int \prod_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i d\mathbf{b}} & (20) \\ &= \int \prod_{i=1}^{k} \mathbf{b}^t \mathbf{x}_i d\mathbf{b} / \int \prod_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i d\mathbf{b} \quad . & (21) \end{aligned}$$

Thus the product in (17) telescopes into

$$\hat{S}_n = \prod_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k = \int \prod_{i=1}^n \mathbf{b}^t \mathbf{x}_i d\mathbf{b} / \int d\mathbf{b} = \int S_n(\mathbf{b}) d\mathbf{b} / \int d\mathbf{b} \quad .\Box \tag{22}$$

We observe two properties of the wealth $\hat{S}_n$ achieved by the universal portfolio.

Proposition 5 (*Universal portfolio exceeds value line index***):**

$$\hat{S}_n \geq \left(\Pi_{j=1}^m S_n(\mathbf{e}_j) \right)^{1/m} \quad . \tag{23}$$

Proof: Let F_n be the empirical cumulative distribution function induced by $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. By two applications of Jensen's inequality, and writing

$$\int S_n(\mathbf{b}) d\mathbf{b} / \int d\mathbf{b} = E_{\mathbf{b}} S_n(\mathbf{b}), \tag{24}$$

we have

$$
\begin{aligned}
\hat{S}_n &= E_{\mathbf{b}} S_n(\mathbf{b}) \\
&= E_{\mathbf{b}} e^{nW(\mathbf{b},F_n)} \\
&\geq e^{nE_b W(\mathbf{b},F_n)} \\
&= e^{nE_b \int \ln \mathbf{b}^t \mathbf{x} dF_n(\mathbf{x})} \\
&= e^{nE_b \int \ln(\sum_{j=1}^m \mathrm{b}_j \mathbf{e}_j^t \mathbf{x}) dF_n(\mathbf{x})} \\
&\geq e^{nE_b \sum_{j=1}^m \mathrm{b}_j \int \ln(\mathbf{e}_j^t \mathbf{x}) dF_n(\mathbf{x})} \\
&= e^{n(\frac{1}{m}\sum \int \ln \mathbf{e}_j^t \mathbf{x} dF_n(\mathbf{x}))} \\
&= \left(\Pi_{j=1}^m S_n(\mathbf{e}_j)\right)^{\frac{1}{m}} . \square \qquad (25)
\end{aligned}
$$

Thus the wealth induced by the proposed portfolio dominates the value line index for any stock sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, for all n.

Next, we observe that although $\hat{\mathbf{b}}_k$ depends on the order of the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, the resulting wealth $\hat{S}_n = \Pi \hat{\mathbf{b}}_k^t \mathbf{x}_k$ does not.

Proposition 6: $\hat{S}_n$ is invariant under permutations of the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

Proof: Since the integrand in

$$
\hat{S}_n = \Pi_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k = \int_B S_n(\mathbf{b}) d\mathbf{b} \Big/ \int_B d\mathbf{b} = \int_B \Pi_{i=1}^n \mathbf{b}^t \mathbf{x}_i d\mathbf{b} \Big/ \int_B d\mathbf{b}, \qquad (26)
$$

is invariant under permutations, so is $\hat{S}_n$. $\square$

This observation guarantees that the crash of 1929 will have no worse consequences for wealth $\hat{S}_n$ than if the bad days of that time had been sprinkled out among the good.

3 The Reason the Portfolio Works.

The main idea of the portfolio algorithm is quite simple. The idea is to give an amount $d\mathbf{b}/\int_B d\mathbf{b}$ to each portfolio manager indexed by rebalancing strategy $\mathbf{b}$, let him make $S_n(\mathbf{b}) = e^{nW(\mathbf{b},F_n)} d\mathbf{b}/\int d\mathbf{b}$ at exponential rate $W(\mathbf{b}, F_n)$ and pool the wealth at the end.

Of course, all dividing and repooling is done "on paper" at time k , resulting in $\hat{\mathbf{b}}_k$. Since the average of exponentials has, under suitable smoothness conditions, the same asymptotic exponential growth rate as the maximum, one achieves almost as much as the wealth S_n^* achieved by the best constant rebalanced portfolio. The trap to be avoided is to put a mass distribution on the market distributions $F(\mathbf{x})$. It seems that this cannot be done in a satisfactory way.

4 Preliminaries.

We now introduce definitions and conditions that will allow characterization of the behavior of $\hat{S}_n/S_n^*$.

Let $F_n(\mathbf{x})$ denote the empirical probability mass function putting mass $\frac{1}{n}$ on each of the points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbf{R}_+^m$. Let the portfolio $\mathbf{b}^* = \mathbf{b}^*(F_n)$ achieve the maximum of $S_n(\mathbf{b}) = \Pi_{i=1}^n \mathbf{b}^t\mathbf{x}_i$. Equivalently, since $S_n(\mathbf{b}) = e^{nW(\mathbf{b},F_n)}$, the portfolio $\mathbf{b}^*(F_n)$ achieves the maximum of $W(\mathbf{b}, F_n)$. Thus

$$S_n^* = \max_{\mathbf{b}\in B} S_n(\mathbf{b}) = e^{nW^*(F_n)} \quad . \tag{27}$$

Definition: We shall say *all stocks are active* (at time n) if $(\mathbf{b}^*(F_n))_i > 0$, $i = 1, 2, \ldots, m$, for some $\mathbf{b}^*$ achieving $W^*(F_n)$. All stocks are *strictly active* if inequality is strict for all i and all $\mathbf{b}^*$ achieving $W^*(F_n)$.

Definition: We shall say $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbf{R}^m$ are of *full rank* if $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ spans $\mathbf{R}^m$.

The condition of full rank is usually true for observed stock market sequences if n is somewhat larger than m, but the condition that all stocks be active often fails when certain stocks are dominated. The next definition measures the curvature of $S_n(\mathbf{b})$ about its maximum and accounts for the second order behavior of $\hat{S}_n$ with respect to S_n^*.

Definition: The *sensitivity matrix function* $J(\mathbf{b})$ of a market with respect to distribution $F(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}_+^m$ is the $(m-1)\times(m-1)$ matrix defined by

$$J_{ij}(\mathbf{b}) = \int \frac{(x_i - x_m)(x_j - x_m)}{(\mathbf{b}^t\mathbf{x})^2} dF(\mathbf{x}) \ , \quad 1 \le i \, , j \le m-1 \quad . \tag{28}$$

The *sensitivity matrix* J^* is $J(\mathbf{b}^*)$, where $\mathbf{b}^* = \mathbf{b}^*(F)$ maximizes $W(\mathbf{b}, F)$.

We note that

$$J^*_{ij} = -\frac{\partial^2 W((b^*_1, b^*_2, \ldots, b^*_{m-1}, 1 - \sum_{i=1}^{m-1} b^*_i), F)}{\partial b_i \partial b_j} \quad . \tag{29}$$

Lemma 2 *J^* is nonnegative definite, and is positive definite if all stocks are strictly active.*

5 Analysis for Two Assets.

We now wish to show that $\hat{S}_n/S^*_n \sim \sqrt{2\pi/nJ_n}$ where J_n is the curvature or volatility index. We show in detail that $\sqrt{2\pi/nJ_n}$ is an asymptotic lower bound on $\hat{S}_n/S^*_n$ and indeed develop explicit lower bounds on $\hat{S}_n/S^*_n$ for all n and any market sequence $\mathbf{x}_1, \ldots, \mathbf{x}_n$. We develop an upper bound by invoking strong conditions on the market sequence. Section 6 outlines the proof for m assets.

We investigate the behavior of $\hat{S}_n$ for $m = 2$ stocks. Consider the arbitrary stock vector sequence

$$\mathbf{x}_i = (x_{i1}, x_{i2}) \epsilon \mathbf{R}^2_+, \; i = 1, 2, \ldots \quad . \tag{30}$$

We now proceed to recast this 2-variable problem in terms of a single variable. Since the portfolio choice requires the specification of one parameter, we write

$$\mathbf{b} = (\mathrm{b}, 1 - \mathrm{b}) \; , \; 0 \leq \mathrm{b} \leq 1 \; , \tag{31}$$

and rewrite $S_n(\mathbf{b})$ as

$$S_n(\mathrm{b}) = \prod_{i=1}^{n} (\mathrm{b}x_{i1} + (1 - \mathrm{b})x_{i2}) \; , 0 \leq b \leq 1 \quad . \tag{32}$$

Let

$$S^*_n = \max_{0 \leq \mathrm{b} \leq 1} S_n(\mathrm{b}) \; , \tag{33}$$

and let b^*_n denote the value of b achieving this maximum. Section 8 contains examples of these graphs.

The universal portfolio

$$\hat{\mathbf{b}}_k = (\hat{\mathrm{b}}_k \ , 1 - \hat{\mathrm{b}}_k) \tag{34}$$

is defined by

$$\hat{\mathrm{b}}_k = \int_0^1 \mathrm{b} S_k(\mathrm{b}) d\mathrm{b} / \int_0^1 S_k(\mathrm{b}) d\mathrm{b} \ , \tag{35}$$

and achieves wealth

$$\hat{S}_n = \prod_{i=1}^{n} (\hat{\mathrm{b}}_i x_{i1} + (1 - \hat{\mathrm{b}}_i) x_{i2}) \ . \tag{36}$$

Let

$$\begin{aligned} W_n(b) &= \frac{1}{n} \ln S_n(b) && (37) \\ &= \frac{1}{n} \sum_{i=1}^{n} \ln(b x_{i1} + (1-b) x_{i2}) && (38) \\ &= \int \ln(b x_1 + (1-b) x_2) dF_n(\mathbf{x}) \ , && (39) \end{aligned}$$

where $F_n(\mathbf{x})$ is the empirical cdf of $\{\mathbf{x}_i\}_{i=1}^n$. By Lemma 1, the wealth $\hat{S}_n$ achieved by the universal portfolio $\hat{\mathbf{b}}_k$ is given by

$$\hat{S}_n = \int_0^1 \mathrm{e}^{nW_n(\mathrm{b})} d\mathrm{b} \ . \tag{40}$$

In order to characterize the behavior of $\hat{S}_n$ we define the following functions of the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Define the *relative range* τ_n of the sequence $\mathbf{x}_1, \mathbf{x}_2 \ \ldots, \mathbf{x}_n$ to be

$$\tau_n = 2^{1/3} (\frac{\max\{x_{ij}\}}{\min\{x_{ij}\}} - 1) \ , \tag{41}$$

where the minimum and maximum are taken over $i = 1, 2, \ldots, n$; $j = 1, 2$. Let

$$J_n = \frac{1}{n} \sum_{i=1}^{n} \frac{(x_{i1} - x_{i2})^2}{(\mathrm{b}_n^* x_{i1} + (1 - \mathrm{b}_n^*) x_{i2})^2} \ , \tag{42}$$

where b_n^* maximizes $W_n(\mathrm{b})$. Let

$$W_n^* = \max_{0 \le \mathrm{b} \le 1} W_n(\mathrm{b}) = W_n(\mathrm{b}_n^*) \ . \tag{43}$$

Thus τ_n corresponds to the relative range of the price relatives and J_n denotes the curvature of $\ln S_n(\mathrm{b})$ at the maximum.

Theorem 1 *Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be an arbitrary sequence of stock vectors in $\mathbf{R}_+^2$ and let $a_n = \min\{b_n^*, 1-b_n^*, 3J_n/\tau_n^3\}$. Then for any $0<\epsilon<1$, and for any n,*

$$\frac{\hat{S}_n}{S_n^*} \geq \sqrt{\frac{2\pi}{nJ_n(1+\epsilon)}} - \frac{2}{\epsilon(1+\epsilon)a_nJ_nn}\mathrm{e}^{-\epsilon^2(1+\epsilon)a_nJ_nn/2} \tag{44}$$

Remarks: This theorem says roughly that $\hat{S}_n/S_n^* \geq \sqrt{2\pi/nJ_n}$. So the universal wealth is within a factor of $C/\sqrt{n}$ of the (presumably) exponentially large S_n^*. It will turn out that every additional stock in the universal portfolio costs an additional factor of $1/\sqrt{n}$. But these factors become negligible to first order in exponent. It is important to mention that this theorem is a bound for each n. The bound holds for any stock sequence with bound τ_n and volatility J_n.

Proof: We wish to bound $\hat{S}_n = \int_0^1 \mathrm{e}^{nW_n(b)}db$. We first expand $W_n(\mathrm{b})$ about the maximizing portfolio b_n^* , noting that $W_n(\mathrm{b})$ has different local properties for each n and indeed a different maximizing b_n^* . We have

$$W_n(\mathrm{b}) = W_n(\mathrm{b}_n^*) + (\mathrm{b}-\mathrm{b}_n^*)W_n^{'}(\mathrm{b}_n^*) + \frac{(\mathrm{b}-\mathrm{b}_n^*)^2}{2}W_n^{''}(\mathrm{b}_n^*) + \frac{(\mathrm{b}-\mathrm{b}_n^*)^3}{3!}W_n^{'''}(\tilde{\mathrm{b}}_n)\ , \tag{45}$$

where $\tilde{\mathrm{b}}_n$ lies between b and b_n^* .

We examine the terms:

(i) The first term is

$$W_n(\mathrm{b}_n^*) = W^*(F_n) = \frac{1}{n}\log S_n^*\ , \tag{46}$$

where S_n^* is the target wealth at time n .

(ii) The second term is

$$\begin{aligned} W_n^{'}(\mathrm{b}_n^*) &= \int \frac{x_1 - x_2}{\mathbf{b}_n^{*t}\mathbf{x}}dF_n(\mathbf{x}) \\ &= 0\ ,\ \text{if } 0 < \mathrm{b}_n^* < 1\ , \end{aligned} \tag{47}$$

by the optimality of b_n^* ;

(iii) The third term is

$$W_n^{''}(\mathrm{b}_n^*) = -\int \frac{(x_1-x_2)^2}{(\mathbf{b}_n^{*t}\mathbf{x})^2} dF_n(\mathbf{x}) = -J_n^* \quad . \tag{48}$$

Thus $W_n^{''}(\mathrm{b}_n^*) \geq 0$, with strict inequality if $0 < \mathrm{b}_n^* < 1$ and $x_{i1} \neq x_{i2}$ for some time i . This term provides the constant in the second order behavior of $\hat{S}_n$.

(iv) And

$$W_n^{'''}(\tilde{\mathrm{b}}_n) = 2\int \frac{(x_1-x_2)^3}{(\tilde{\mathrm{b}}x_1 + (1-\tilde{\mathrm{b}})x_2)^3} dF_n(\mathbf{x}) \quad . \tag{49}$$

We have the bound

$$\begin{aligned} \mid W_n^{'''}(\tilde{\mathrm{b}}_n) \mid \quad &= \quad \left| \int \frac{2(x_1-x_2)^3}{(\tilde{\mathrm{b}}_n x_1 + (1-\tilde{\mathrm{b}}_n)x_2)^3} dF_n(\mathbf{x}) \right| & (50) \\ &\leq \quad \tau_n^3 \ , \text{ for all } \tilde{\mathrm{b}}_n \ \epsilon \ [0,1] \quad . & (51) \end{aligned}$$

Thus

$$S_n(\mathrm{b}) \geq \exp\left(nW_n^* - \frac{n}{2}\,(\mathrm{b}-\mathrm{b}_n^*)^2 J_n - \frac{n \mid \mathrm{b}-\mathrm{b}_n^* \mid^3 \tau_n^3}{6} \right) \quad , \tag{52}$$

for $0 \leq \mathrm{b} \leq 1$,

where

$$J_n = \int \frac{(x_1-x_2)^2}{(\mathrm{b}_n^* x_1 + (1-\mathrm{b}_n^*)x_2)^2} dF_n(\mathbf{x}) \quad . \tag{53}$$

We now make the change of variable

$$u = \sqrt{n}\,(\mathrm{b}-\mathrm{b}_n^*) \ , \tag{54}$$

where the new range of integration is

$$-\sqrt{n} \ \ \mathrm{b}_n^* \leq u \leq \sqrt{n} \ \ (1-\mathrm{b}_n^*) \quad . \tag{55}$$

Then, noting $\mathrm{e}^{nW_n^*} = S_n^*$, we have

$$\hat{S}_n = \int_0^1 S_n(\mathrm{b})d\mathrm{b} \tag{56}$$

$$\geq \frac{S_n^*}{\sqrt{n}} \int_{-\sqrt{n}\mathrm{b}_n^*}^{\sqrt{n}(1-\mathrm{b}_n^*)} \mathrm{e}^{-\frac{1}{2}u^2 J_n - \frac{1}{6\sqrt{n}} |u^3|\tau_n^3} du \ . \tag{57}$$

We wish to approximate this by the normal integral. To do so let $0 < \epsilon \leq 1$ and note that

$$-\frac{1}{2}u^2 J_n - \frac{|\ u\ |^3\ \tau_n^3}{6\sqrt{n}} \geq -\frac{1}{2}\ u^2 J_n(1+\epsilon) \tag{58}$$

for

$$u \leq 3\epsilon\sqrt{n}J_n/\tau_n^3 \ . \tag{59}$$

Let Φ denote the cdf of the standard normal,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-u^2/2}\ du\ , \tag{60}$$

and let

$$a_n = \min\{\mathrm{b}_n^*\ , 1 - \mathrm{b}_n^*\ , 3J_n/\tau_n^3\} \ . \tag{61}$$

Thus a_n is a measure of the degree to which $S_n(b)$ has a maximum of reasonable curvature within the unit interval. Then from (57), for any $0 < \epsilon \leq 1$,

$$\frac{\sqrt{n}\ \hat{S}_n}{S_n^*} \geq \int_{-\sqrt{n}\ b_n^*}^{\sqrt{n}\ (1-b_n^*)} \mathrm{e}^{-\frac{1}{2}u^2 J_n - \frac{1}{6\ \sqrt{n}}|u|^3\tau_n^3}\ du \tag{62}$$

$$\geq \int_{-\sqrt{n}\ a_n\epsilon}^{\sqrt{n}\ a_n\epsilon} \mathrm{e}^{-\frac{1}{2}u^2 J_n(1+\epsilon)}\ du \tag{63}$$

$$= \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}u^2 J_n(1+\epsilon)}\ du - 2\int_{-\infty}^{-\sqrt{n}\ a_n\epsilon} \mathrm{e}^{-\frac{1}{2}u^2 J_n(1+\epsilon)} du \tag{64}$$

$$= \sqrt{\frac{2\pi}{J_n(1+\epsilon)}} - \sqrt{\frac{8\pi}{J_n(1+\epsilon)}}\Phi\left(-\epsilon a_n\sqrt{nJ_n(1+\epsilon)}\right) \ . \tag{65}$$

We use the inequality

$$\frac{1}{\sqrt{2\pi x^2}}e^{-x^2/2}(1-\frac{1}{x^2}) < \Phi(-x) < \frac{1}{\sqrt{2\pi x^2}}e^{-x^2/2} \ , \tag{66}$$

for $x > 0$, to obtain the bound

$$\Phi\left(-\epsilon a_n\sqrt{nJ_n(1+\epsilon)}\right) < \frac{1}{\sqrt{2\pi\epsilon^2 a_n^2 nJ_n(1+\epsilon)}}e^{-\epsilon^2 a_n^2 nJ_n(1+\epsilon)/2} \ . \tag{67}$$

Hence

$$\frac{\sqrt{n}\,\hat{S}_n}{S_n^*} \geq \sqrt{\frac{2\pi}{J_n(1+\epsilon)}} - \frac{2}{\epsilon a_n J_n(1+\epsilon)\sqrt{n}}e^{-\epsilon^2 a_n^2 nJ_n(1+\epsilon)/2} \ . \tag{68}$$

for any $0 < \epsilon \leq 1$, for all n , and all $\mathbf{x}_1, \mathbf{x}_2, \ldots$, which proves the theorem. $\Box$

The explicit bounds in Theorem 1 may be useful in practice, but a cleaner summary of performance is given in the following weaker theorem.

Theorem 2 *Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ be a sequence of stock vectors in $\mathbf{R}_+^2$ and suppose $\delta \leq b_n^* \leq 1-\delta$, $\tau_n \leq \tau < \infty$, and $J_n \geq J > 0$, for a subsequence of times $n_1, n_2, \ldots$. Then*

$$\liminf_{n\to\infty} \frac{\hat{S}_n/S_n^*}{\sqrt{2\pi/nJ_n}} \geq 1 \ , \tag{69}$$

along this subsequence.

Proof: The conditions of the theorem, together with Theorem 1, imply

$$\frac{\hat{S}_n/S_n^*}{\sqrt{2\pi/nJ_n}} \geq \sqrt{\frac{1}{(1+\epsilon_n)}} - \frac{2\sqrt{J_n}}{\epsilon_n\sqrt{2\pi nJ}\min\{\delta, 3J/\tau^3\}} \ , \tag{70}$$

where τ is the bound ratio, and where we are free to choose $\epsilon_n \epsilon [0,1]$ at each n. Noting that $J_n \leq \tau^2 < \infty$ and letting $\epsilon_n = n^{-1/4}$ proves the theorem. $\Box$

We have just shown that $\hat{S}_n/S_n^*$ is as good as $\sqrt{2\pi/nJ_n}$. We now show that it is no better. For this we consider a subsequence of times such that $W_n(\mathrm{b})$ is approximately equal to some function $W(\mathrm{b})$ and argue that upper bounds on $\int_0^1 e^{nW(\mathrm{b})}d\mathrm{b}$ suffice to limit the performance of the wealth $\hat{S}_n$. Toward that end, let us consider functions W such that

(i) $W(\mathrm{b})$ is strictly concave on [0,1] ,

(ii) $W'''(\mathrm{b})$ is bounded on [0,1] ,

(iii) $W(\mathrm{b})$ achieves its maximum at $\mathrm{b}^* \epsilon (0,1)$. (71)

We plan to pick out a subsequence of times such that $W_n(\mathrm{b}) = \frac{1}{n}\sum_{i=1}^n \ln \mathbf{b}^t \mathbf{x}_i$ approaches $W(\mathrm{b})$. We can expect such limit points from Arzelà's Theorem on the compactness of equicontinuous functions on compact sets. Let b_n^* maximize $W_n(\mathrm{b})$. Let $\{n_i\}$ be a subsequence of times such that for $n = n_1, n_2, \ldots$,

(i) $W_n(\mathrm{b}) \leq W(\mathrm{b})$, $0 \leq \mathrm{b} \leq 1$,

(ii) $W_n''(\mathrm{b}_n^*) \to W''(\mathrm{b}^*)$. (72)

Recall the notation $J_n = -W_n''(\mathrm{b}_n^*)$. The following theorem establishes the tightness of the lower bound in Theorem 2.

Theorem 3 *For any $\mathbf{x}_1, \mathbf{x}_2, \ldots \epsilon \mathbf{R}_+^2$ and for any subsequence of times $n_1, n_2, \ldots$ such that $W_n(b)$ satisfies the conditions (72) for $W(b)$ satisfying (71),*

$$\frac{\hat{S}_n}{S_n^*} \sim \sqrt{\frac{2\pi}{nJ_n}} , \tag{73}$$

along the subsequence.

Proof: The lower bound follows from Theorem 2.

From Laplace's method of integration we have

$$\int_0^1 \mathrm{e}^{ng(u)}\, du \sim \mathrm{e}^{ng(u^*)} \sqrt{\frac{2\pi}{n|g''(u^*)|}} \tag{74}$$

if g is three times differentiable with bounded third derivative, strictly concave, and the u^* maximizing $g(\cdot)$ is in the open interval $(0,1)$. Consequently,

$$\hat{S}_n = \int_0^1 \mathrm{e}^{nW_n(\mathrm{b})}\, d\mathrm{b} \leq \int_0^1 \mathrm{e}^{nW(\mathrm{b})}\, d\mathrm{b}$$

$$
\begin{aligned}
&\sim \mathrm{e}^{nW(\mathrm{b}^*)}\sqrt{\frac{2\pi}{n|W''(\mathrm{b}^*)|}} \\
&= S_n^*\sqrt{\frac{2\pi}{n|W''(\mathrm{b}^*)|}} \sim S_n^*\sqrt{\frac{2\pi}{nJ_n}} \qquad (75)
\end{aligned}
$$

and the theorem is proved. □

6 Main Theorem.

Here we prove the result for m assets under the assumption that all stocks are active and of full rank and $\mathbf{b}_n^*(F_n) \to \mathbf{b}^* \in \mathrm{int}(B)$. We discuss removing the conditions in Section 9. For example, lack of full rank reduces the dimension from m to m', as does the existence of inactive stocks. Finally, $\mathbf{b}_n^*(F_n)$ need not have a limit, in which case we can describe the behavior of $\hat{S}_n$ for convergent subsquences of $\mathbf{b}_n^*(F_n)$, as well as develop explicit bounds for all n.

From Lemma 1, we have

$$
\hat{S}_n = \int_B S_n(\mathbf{b})d\mathbf{b} / \int_B d\mathbf{b} \ .
$$

where

$$
\begin{aligned}
S_k(\mathbf{b}) &= \Pi_{i=1}^k \mathbf{b}^t \mathbf{x}_i \ , \\
\hat{\mathbf{b}}_{k+1} &= \frac{\int \mathbf{b} S_k(\mathbf{b})d\mathbf{b}}{\int S_k(\mathbf{b})d\mathbf{b}} \ , \\
\hat{S}_n &= \Pi_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k \ .
\end{aligned}
$$

A summary of the performance of $\hat{\mathbf{b}}_k$ is given by the following theorem.

Theorem 4 *Suppose* $\mathbf{x}_1, \mathbf{x}_2, \ldots \epsilon [a,c]^m, 0 < a \le c < \infty$, *and at a subsequence of times* $n_1, n_2, \ldots$, $W_n(\mathbf{b}) \nearrow W(\mathbf{b})$ *for* $\mathbf{b} \epsilon B$, $J_n^* \to J^*$, $\mathbf{b}_n^* \to \mathbf{b}^*$, *where* $W(\mathbf{b})$ *is strictly concave, the third partial derivatives of* W *are bounded on* B, *and* $W(\mathbf{b})$ *achieves its maximum at* $\mathbf{b}^*$ *in the interior of* B. *Then*

$$
\frac{\hat{S}_n}{S_n^*} \sim \left(\sqrt{\frac{2\pi}{n}}\right)^{m-1} \frac{(m-1)!}{|J^*|^{1/2}} \qquad (76)
$$

in the sense that the ratio of the right and left hand sides converges to 1 along the subsequence.

Proof: (Outline) We define

$$C = \{(c_1, c_2, \ldots, c_{m-1}) \ : \ c_i \geq 0, \ \Sigma c_i \leq 1\} \tag{77}$$

and

$$S_n(\mathbf{c}) = \Pi_{i=1}^n \mathbf{b}^t(\mathbf{c})\mathbf{x}_i \ , \quad \mathbf{c} \in C \ , \tag{78}$$

where

$$\mathbf{b}(\mathbf{c}) = (c_1, c_2, \ldots, c_{m-1}, 1 - \sum_{i=1}^{m-1} c_i) \ . \tag{79}$$

Note that

$$\text{Vol}(C) = \int_C d\mathbf{c} = \frac{1}{(m-1)!} \ . \tag{80}$$

We shall prove only the lower bound associated with (76). From Lemma 1, the universal portfolio algorithm yields

$$\hat{S}_n = \int S_n(\mathbf{b})d\mathbf{b} / \int d\mathbf{b} = E_{\mathbf{b}} S_n(\mathbf{b}) \ , \tag{81}$$

where $\mathbf{b}$ is uniformly distributed over the simplex B. Since uniform over B induces uniform over C, we have

$$\hat{S}_n = (m-1)! \int_C S_n(\mathbf{c})d\mathbf{c} \ . \tag{82}$$

We now expand $S_n(\mathbf{c})$ in a Taylor series about $\mathbf{c}^* = (\mathrm{b}_1^*, \ldots, \mathrm{b}_{m-1}^*)$ where $\mathbf{b}^*$ maximizes $W(\mathbf{b}, F_n)$. We drop the dependence of $\mathbf{b}^*$ on n for notational convenience. By assumption, $\mathrm{b}_i^* > 0$, for all i. We have

$$S_n(\mathbf{c}) = e^{nW_n(\mathbf{c})} \ , \tag{83}$$

where

$$\begin{aligned} W_n(\mathbf{c}) &= \frac{1}{n} \sum_{i=1}^n \ln \mathbf{b}^t \mathbf{x}_i \\ &= \int \ln \mathbf{b}^t \mathbf{x} dF_n(\mathbf{x}) \\ &\triangleq E_{F_n} \ln \mathbf{b}^t \mathbf{X} \end{aligned} \tag{84}$$

and

$$\mathbf{b} = (\mathbf{c}, 1 - \Sigma c_i) \ . \tag{85}$$

Expanding $W_n(\mathbf{c})$, we have

$$\begin{aligned} W_n(\mathbf{c}) &= W_n(\mathbf{c}^*) + (\mathbf{c} - \mathbf{c}^*)^t \bigtriangledown W_n(\mathbf{c}^*) - \frac{1}{2}(\mathbf{c} - \mathbf{c}^*)^t J_n^* (\mathbf{c} - \mathbf{c}^*) \\ &\quad + \frac{1}{6} \sum_{i,j,k} (c_i - c_i^*)(c_j - c_j^*)(c_k - c_k^*) E_{F_n} \frac{2(X_i - X_m)(X_j - X_m)(X_k - X_m)}{S^3(\tilde{\mathbf{c}})} \ , \end{aligned} \tag{86}$$

where $\tilde{\mathbf{c}} = \lambda \mathbf{c}^* + (1-\lambda)\mathbf{c}$, for some $0 \le \lambda \le 1$, where λ may depend on $\mathbf{c}$, and

$$S(\mathbf{c}) = \sum_{i=1}^{m-1} c_i X_i + \left(1 - \sum_{i=1}^{m-1} c_i\right) X_m \ . \tag{87}$$

Here

$$\begin{aligned} S_n(\mathbf{c}) &= \Pi_{i=1}^n \mathbf{b}^t(\mathbf{c})\mathbf{x}_i \ , &(88)\\ \mathbf{b}(\mathbf{c}) &= \left(c_1, c_2, \ldots, 1 - \sum_{i=1}^{m-1} c_i\right) \ , \\ W(\mathbf{c}) &= \int \ln\left(\sum_{i=1}^{m-1} c_i x_i + (1 - \sum_{1}^{m-1} c_i) x_k\right) dF_n(\mathbf{x}) \ , \\ \frac{\partial W_n}{\partial c_i} &= \int \frac{(x_i - x_m)}{S(\mathbf{c})} dF_n(\mathbf{x}) \ , \\ \frac{\partial^2 W_n}{\partial c_i \partial c_j} &= -\int \frac{(x_i - x_m)(x_j - x_m)}{S^2(\mathbf{c})} dF_n(\mathbf{x}) \ , \\ J_n^* &= -\left[\frac{\partial^2 W_n(\mathbf{c}^*)}{\partial c_i \partial c_j}\right] . &(89) \end{aligned}$$

The condition that all stocks be strictly active implies by Lemma 2 that $|J_n^*| > 0$, where $|\cdot|$ denotes determinant. We treat the terms one by one:

(i) By definition of $\mathbf{b}^*$,

$$W(\mathbf{c}^*) = W(\mathbf{b}^*, F_n) = W^*(F_n) \ \ . \tag{90}$$

(ii) The second term is 0, because $\mathbf{b}^*$ is in the interior of $B, W_n(\mathbf{b})$ is differentiable, and $\mathbf{b}^*$ maximizes W_n. Thus,

$$\begin{aligned} \frac{\partial W_n(\mathbf{c}^*)}{\partial c_i} &= E_{F_n} \frac{X_i - X_m}{\mathbf{b}^{*t}\mathbf{X}} &(91)\\ &= 0, \ i = 1, 2, \ldots, m-1 \ . \end{aligned}$$

(iii) The third term is a positive definite quadratic form, where $J_n^* = J^*(\mathbf{b}^*(F_n))$.

(iv) For the fourth term

$$\frac{1}{6} \sum_{i,j,k=1}^{m-1} (c_i - c_i^*)(c_j - c_j^*)(c_k - c_k^*) E_{F_n} \frac{2(X_i - X_m)(X_j - X_m)(X_k - X_m)}{S^3(\tilde{\mathbf{c}})} \ , \tag{92}$$

we examine

$$E_{F_n}\frac{(X_i - X_m)(X_j - X_m)(X_k - X_m)}{S^3(\tilde{\mathbf{c}})} \ . \tag{93}$$

We note

$$\begin{aligned} S^3(\tilde{\mathbf{c}}) &= (\tilde{\mathbf{b}}^t X)^3 \\ &\geq (\Sigma \tilde{b}_i a)^3 \\ &\geq a^3 \ , \end{aligned} \tag{94}$$

since $X_i \geq a$, for all i. Also since $X_i - X_m \leq 2b$, we have

$$-\frac{8b^3}{a^3} \leq E\frac{(X_i - X_m)(X_j - X_m)(X_k - X_m)}{S^3(\tilde{\mathbf{c}})} \leq \frac{8b^3}{a^3} \ . \tag{95}$$

We now make the change of variable $\mathbf{u} = \sqrt{n}(\mathbf{c} - \mathbf{c}^*)$, where we note the new range of integration $\mathbf{u} \in U = \sqrt{n}(C - \mathbf{c}^*)$. Thus

$$\begin{aligned} S_n(\mathbf{c}) &= e^{nW_n(\mathbf{c}^*) - \frac{n}{2}(\mathbf{c}-\mathbf{c}^*)^t J_n^*(\mathbf{c}-\mathbf{c}^*) + \frac{n}{3}\Sigma_3} \\ &= e^{nW_n^* - \frac{1}{2}\mathbf{u}^t J_n^* \mathbf{u} + \frac{1}{3\sqrt{n}}\Sigma_3}, \end{aligned} \tag{96}$$

where

$$\Sigma_3 = \sum_{i,j,k=1}^{m-1} u_i u_j u_k \ E_{F_n} \ \left(\frac{(X_i - X_m)(X_j - X_m)(X_k - X_m)}{S^3(\tilde{\mathbf{c}})}\right) \ . \tag{97}$$

Note that

$$|\Sigma_3| \leq \left(\sum_{i=1}^{m-1} |u_i|\right)^3 \frac{8b^3}{a^3} \ . \tag{98}$$

Observing

$$\Sigma|u_i| \leq (\Sigma u_i^2)^{1/2}\sqrt{m} \tag{99}$$

yields

$$S_n(\mathbf{c}) \geq e^{nW_n^* - \frac{1}{2}\mathbf{u}^t J_n^* \mathbf{u} - \frac{m^{3/2}}{3\sqrt{n}}\|\mathbf{u}\|^3 8\mathrm{b}^3/a^3} \ . \tag{100}$$

The lower bound on $\hat{S}_n$ becomes

$$\begin{aligned} \hat{S}_n &= (m-1)! \int_{\mathbf{c}\in C} S_n(\mathbf{c}) d\mathbf{c} \\ &\geq (m-1)! S_n^* \int_{\mathbf{u}\in U} e^{(-\frac{1}{2}\mathbf{u}^t J_n^* \mathbf{u} - \frac{m^{3/2}}{3\sqrt{n}}\|\mathbf{u}\|^3 8\mathrm{b}^3/a^3)} (\frac{1}{\sqrt{n}})^{m-1} d\mathbf{u} \ , \end{aligned} \tag{101}$$

which can now be bounded using the techniques in the 2-stock proof. The upper bound follows from Laplace's method of integration as in Theorem 3, from which the theorem follows. □

7 Stochastic Markets.

Another way to see the naturalness of the goal $S_n^* = e^{nW(\mathbf{b}^*(F_n),F_n)}$ is to consider random investment opportunities. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be independent identically distributed random vectors drawn according to $F(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^m$, where F is some known distribution function. Let $S_n(\mathbf{b}) = \Pi_{i=1}^n \mathbf{b}^t\mathbf{X}_i$ denote the wealth at time n resulting from an initial wealth $S_0 = 1$, and a reinvestment of assets according to portfolio $\mathbf{b}$ at each investment opportunity. Then

$$\begin{aligned} S_n(\mathbf{b}) = \Pi_{i=1}^n \mathbf{b}^t\mathbf{X}_i &= e^{\Sigma_{i=1}^n \ln \mathbf{b}^t\mathbf{X}_i} \\ &= e^{n(E \ln \mathbf{b}^t\mathbf{X} + o_p(1))} \\ &= e^{n(W(\mathbf{b},F)+o_p(1))} \quad , \end{aligned} \tag{102}$$

by the strong law of large numbers, where the random variable $o_P(1) \to 0$, a.e. We observe from the above that, to first order in the exponent, the growth rate of wealth $S_n(\mathbf{b})$ is determined by the expected log wealth

$$W(\mathbf{b}, F) = \int \ln \mathbf{b}^t\mathbf{x} dF(\mathbf{x}) \tag{103}$$

for portfolio $\mathbf{b}$ and stock distribution $F(\mathbf{x})$.

It follows for $\mathbf{X}_1, \mathbf{X}_2, \ldots,$ i.i.d. $\sim F$ that $\mathbf{b}^*(F)$ achieves an exponential growth rate of wealth with exponent $W^*(F)$. Moreover Breiman (1961) establishes for i.i.d. stock vectors for any nonanticipating time-varying portfolio strategy with associated wealth sequence S_n that

$$\overline{\lim} \frac{1}{n} \ln S_n \leq W^*(F), \text{ a.e.} \tag{104}$$

Finally, it follows from Breiman (1961), Finkelstein and Whitley (1981), Barron and Cover (1988), and Algoet and Cover (1988), in increasing levels of generality on the stochastic process, that $\lim_{n\to\infty} \frac{1}{n} \ln \frac{S_n}{S_n^*} \leq 0$, a.e., for every sequential portfolio. Thus $\mathbf{b}^*(F)$ is asymptotically optimal in this sense, and $W^*(F)$ is the highest possible exponent for the growth rate of wealth.

We omit the proof of the following.

Theorem 5 *Let* $\mathbf{X}_i$ *be i.i.d.* $\sim F(\mathbf{x})$. *Let* $\mathbf{b}^*(F)$ *be unique and lie in the interior of* B. *Then the universal portfolio* $\hat{\mathbf{b}}_k$ *yields a wealth sequence* $\hat{S}_n$ *satisfying*

$$\frac{1}{n}\ln \hat{S}_n \rightarrow W^*(F) \text{ , a.e.} \tag{105}$$

Thus, in the special case where the stocks are independent and identically distributed according to some unknown distibution F , the universal portfolio essentially learns F in the sense that the associated growth rate of wealth is equal to that achievable when F is known.

8 Examples.

We now test the portfolio algorithm on real data. Consider, for example, Iroquois Brands Ltd. and Kin Ark Corp., two stocks chosen for their volatility listed on the New York Stock Exchange. During the 22 year period ending in 1985, Iroquois Brands Ltd. increased in price (adjusted in the usual manner for dividends) by a factor of 8.9151, while Kin Ark increased in price by a factor of 4.1276, as shown in Fig. 1.

Prior knowledge (in 1963) of this information would have enabled an investor to buy and hold the best stock (Iroquois) and earn a 791% profit. However, a closer look at the time series reveals some cause for regret. Table 1 lists the performance of the constant rebalanced portfolios $\mathbf{b} = (\mathrm{b}, 1-\mathrm{b})$, for $\mathrm{b} = 1, .95, \ldots, .05, 0.0$. The graph of $S_n(b)$ is given in Figure 2. For example, reinvesting current wealth in the proportions $\mathbf{b} = (.8, .2)$ at the start of each trading day would have resulted in an increase by a factor of 37.5 . In fact, the best rebalanced portfolio for this 22 year period is $\mathbf{b}^* = (.55, .45)$, yielding a factor $S_n^* = 73.619$. Here S_n^* is the target wealth (with respect to the coarse quantization of $B = [0,1]$ we have chosen). The universal portfolio $\hat{\mathbf{b}}_k$ achieves a factor of $\hat{S}_n = 38.6727$. While $\hat{S}_n$ is short of the target, as it must be, $\hat{S}_n$ dominates the 8.9 and 4.1 factors of the constituent stocks. The daily performance of both stocks, the universal portfolio, and the target wealth are exhibited in Fig. 3. The portfolio choice $\hat{\mathrm{b}}_k$ as a function of time k is given in Figure 4.

To be explicit in the above analysis, we have quantized all integrals, resulting in the replacements of

$$S_n^* = \max_{\mathrm{b}} S_n(\mathrm{b}) \quad \text{by} \quad S_n^* = \max_{i=0,1,\ldots,20} S_n(i/20) \tag{106}$$

and

$$\hat{\mathrm{b}}_{k+1} = \frac{\int_0^1 \mathrm{b} S_k(\mathrm{b}) d\mathrm{b}}{\int_0^1 S_k(\mathrm{b}) d\mathrm{b}} \quad \text{by} \quad \hat{\mathrm{b}}_{k+1} = \frac{\sum_{i=0}^{20} \frac{i}{20} S_k(\frac{i}{20})}{\sum_{i=0}^{20} S_k(\frac{i}{20})} \quad . \tag{107}$$

The resulting wealth factor

$$\hat{S}_n = \prod_{k=1}^{n} \hat{\mathbf{b}}_k^t \mathbf{x}_k \tag{108}$$

is calculated using

$$\hat{\mathrm{b}}_k = \sum_{i=0}^{20} \frac{i}{20} S_k(\frac{i}{20}) / \sum_{i=0}^{20} S_k(\frac{i}{20}) \quad . \tag{109}$$

Telescoping still takes place under this quantization and it can be verified that $\hat{S}_n$ in (108) can be expressed in the equivalent form

$$\hat{S}_n = \frac{1}{21} \sum_{i=0}^{20} S_n(\frac{i}{20}) \quad . \tag{110}$$

Thus $\hat{S}_n$ is the arithmetic average of the wealths associated with the constant rebalanced portfolios.

Finally, note the calculation of the portfolio $\hat{\mathbf{b}}_{n+1} = (\hat{\mathrm{b}}_{n+1}, 1 - \hat{\mathrm{b}}_{n+1})$ in this example. Merely compute the inner product of the b column and $S_n(\mathrm{b})$ column in Table 1 and divide by the sum of the $S_n(\mathrm{b})$ column to obtain $\hat{\mathrm{b}}_{n+1}$. Note in particular that the universal portfolio $\hat{\mathbf{b}}_{n+1}$ is not equal to the log optimal portfolio $\mathbf{b}^*(F_n) = (.55, .45)$ with respect to the empirical distribution of the past.

A similar anaylsis can be performed on Commercial Metals and Kin Ark over the same period. Here Commercial Metals increased by the factor 52.0203 and Kin Ark by the factor 4.1276 (Fig. 5) . It seems that an investor wouldn't want any part of Kin Ark with an alternative like Commercial Metals available. Not so. The optimal constant rebalanced portfolio is $\mathbf{b}^* = (.65, .35)$, and the universal portfolio achieves $\hat{S}_n = 78.4742$, outperforming each stock. See Table 2 .

Next we put Commercial Metals (52.0203) up against Mei Corp (22.9160) . Here $S_n^* = 102.95$ and $\hat{S}_n = 72.6289$ as shown in Figure 6 and Table 3 . In constrast to these examples, IBM and Coca Cola show a lockstep performance, and, indeed, $\hat{S}_n$ barely outperforms the constituent stocks, as shown in Figure 7.

A final example crudely models buying on 50 percent margin. Suppose we have 4 investment choices each day: Commercial Metals, Kin Ark, and these same two stocks on 50 percent margin. Margin loans are settled daily at a 6 percent annual interest rate. The stock vector on the ith day is

$$\begin{aligned} \mathbf{x}_i &= (x_i, \quad 2x_i - 1 - r, \;\; y_i, \;\; 2y_i - 1 - r) \;, \qquad (111) \\ r &= .000233 \;, \qquad (112) \end{aligned}$$

where x_i and y_i are the respective price relatives for Commercial Metals and Kin Ark on day i . Plunging on margin into Commercial Metals yields a factor 19.73, and plunging into Kin Ark a factor 0.0000 + . Good as these stocks are, they can't survive the down factors induced by the leverage. But the random sample of the simplex of portfolios listed in Table 4 reveals $\hat{S}_n = 98.4240$, while the optimal rebalanced portfolio $\mathbf{b}^* = (.2, .5, .1, .2)$ results in a factor $S_n^* = 262.4021$. Clearly 98.4 beats the factor of 78 achieved when margin is unavailable. Both factors exceed the performance 52.02 of the best stock.

We observe that $\hat{S}_n = 98.4$ exceeds the factor $\hat{S}_n = 78.47$ obtained for these stocks when margin is unavailable. This is borne out by the fact that $\mathbf{b}^*$ is positive in each component, calling for a small amount of leverage in the *a posteriori* optimal rebalanced portfolio.

9 The General Universal Portfolio.

If the best rebalanced portfolio $\mathbf{b}_n^*$ lies in the interior of a boundary k-face then only k stocks are active in the best rebalanced portfolio. Thus we expect to obtain the previous bounds on $\frac{\hat{S}_n}{S_n^*}$ with m replaced by k . This is achieved if we start with some mass on each face. To accomplish this, we let μ_S be the measure corresponding to the uniform distribution on $B(S) = \{\mathbf{b} \epsilon R^m : \sum \mathrm{b}_i = 1 \;, \mathrm{b}_i = 0 \;, i \epsilon S^c\}$, where $S \subseteq \{1, 2, \ldots, m\}$. Thus μ_S puts unit mass on the $|S|$-dimensional face of the portfolio simplex.

Let μ be the mixture of these measures given by

$$\mu = \frac{1}{2^m - 1} \sum \mu_S \tag{113}$$

where the sum is over all $S \neq \phi$, $S \subseteq \{1, 2, \ldots, m\}$.

The generalized universal portfolio now becomes

$$\hat{\mathbf{b}}_{n+1} = \frac{\int \mathbf{b} S_n(\mathbf{b}) \mu(d\mathbf{b})}{\int S_n(\mathbf{b}) \mu(d\mathbf{b})} \tag{114}$$

with

$$S_n(\mathbf{b}) = \prod_{i=1}^{n} \mathbf{b}^t \mathbf{x}_i \ , \ S_0(\mathbf{b}) = 1 \ . \tag{115}$$

To state the results we define $J_n^{(k)}(F_n)$ to be the $k \times k$ sensitivity matrix with respect to the active stocks S , $\mid S \mid = k$, where S is the smallest set of stocks such that all optimal rebalanced portfolios $\mathbf{b}^*(F)$ are in the interior of $B(S)$. Then

$$\frac{\hat{S}_n}{S_n^*} \sim \frac{(k-1)!}{2^m - 1} \left(\frac{2\pi}{n} \right)^{\frac{k-1}{2}} / \mid J_n^{(k)}(F_n) \mid^{1/2} \tag{116}$$

will be the asymptotic behavior of $\hat{S}_n / S_n^*$.

10 Concluding Remarks.

We now try to be sensible and ask how the universal portfolio works in practice. Of course the examples are encouraging, as the universal portfolio outperforms the constituent stocks. However, we have ignored trading costs. In practice we would not trade daily, but only when the current empirical holdings were far enough from the recommended $\hat{\mathbf{b}}_k$. (A rule of thumb might be to trade only if the increase in W is greater than the logarithm of the normalized transaction costs.)

We are really interested in whether $\hat{S}_n$ will "take off", leaving the stocks behind. We first discuss the target wealth S_n^* . The best rebalanced portfolio $\mathbf{b}^*(F_n)$ based on prior knowledge of the stock sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ yields wealth $S_n^* = \mathrm{e}^{nW_n^*}$. Now S_n^*

grows exponentially fast to infinity under mild conditions. For example, if one of the constituent stocks is a risk free asset with interest rate $r > 0$, then $W_n^* \geq \ln(1+r) > 0$, for all n , and $S_n^* \geq (1+r)^n \to \infty$. Since the universal portfolio yields

$$\hat{S}_n = \mathrm{e}^{n(W_n^* - 0(\frac{\ln n}{n}))} \quad , \tag{117}$$

it follows that $\hat{S}_n$ will tend to infinity, and $\hat{S}_n$ will have the same exponent as S_n^* , differing only in terms of order $(\ln n)/n$.

What state of affairs do we expect in the real world? Certainly we expect the stock sequence to be of full dimension m for n slightly greater than m . However, we don't expect all stocks to be active. But we do expect that two or more stocks will be active. This is important because it guarantees that the target growth rate W_n^* will be strictly greater than the growth rate of the constituent stocks. Consequently we believe that the universal portfolio will achieve

$$\frac{\hat{S}_n}{S_n(\mathbf{e}_i)} \to \infty \ , i = 1, 2, \ldots \ m \ ,$$

exponentially fast, where $S_n(\mathbf{e}_i)$ is the wealth relative of the ith stock at time n . However, n may need to be quite large before this exponential dominance manifests itself. In particular, we need n large enough that the difference in exponents between S_n^* and the stocks overcomes the $O((\ln n)/n)$ penalties incurred by universality. We conclude that $\hat{S}_n$ will leave the constituent stocks exponentially behind if there are at least 2 strictly active stocks in the best rebalanced portfolio.

11 Acknowledgement.

I wish to thank Hal Stern for his invaluable contributions in obtaining and analyzing the stock market data. I also wish to thank the referee for helpful comments.

References

[1] Algoet, P. and T.M. Cover (1988): "Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment," *Annals of Probability,* 16, 2, 876-898.

[2] Arrow, K. (1974): *Essays in the Theory of Risk Bearing,* North Holland/American Elsevier.

[3] Barron, A. and T.M. Cover (1988): "A Bound on the Financial Value of Information," *IEEE Trans. on Information Theory* 34, 5, 1097-1100.

[4] Bell, R. and T.M. Cover (1980): "Competitive Optimality of Logarithmic Investment," *Math. of Operations Research,* 5, 2, 161-166.

[5] Bell, R. and T.M. Cover (1988): "Game-Theoretic Optimal Portfolios," *Management Science*, 34, 6, 724-733.

[6] Blackwell, D. (1956a): "Controlled Random Walks," *J. Proceedings of International Congress of Mathematics* III, Amsterdam, North Holland, 336-338.

[7] Blackwell, D. (1956b): "An Analog of the Minimax Theorem for Vector Payoffs," *Pacific Journal of Mathematics,* VI, 1-8.

[8] Breiman, L.(1961): "Optimal Gambling Systems for Favorable Games," *Fourth Berkeley Symposium on Mathematical Statistics and Probability,* 1, 65-78.

[9] Cover, T.M. (1984): "An Algorithm for Maximizing Expected Log Investment Return," *IEEE Trans. on Information Theory,* IT-30, 2, 369-373.

[10] Cover, T.M. and D. Gluss (1986): "Empirical Bayes Stock Market Portfolios," *Advances in Applied Mathematics,* 7, 170-181, (A summary also appears in *Proceedings of Conference Honoring Herbert Robbins,* Springer-Verlag, 1986.)

[11] Cover, T.M. and R. King (1978): "A Convergent Gambling Estimate of the Entropy of English," *IEEE Transactions on Information Theory*, 24, 4, 413-421.

[12] Finkelstein, M. and R. Whitley (1981): "Optimal Strategies for Repeated Games," *Adv. Appl. Prob.,* 13, 415-428.

[13] Hakansson, N. (1979): "A Characterization of Optimal Multiperiod Portfolio Policies," in *Portfolio Theory, 25 Years After: Essays in Honor of Harry Markowitz,* ed. Elton and Gruber, TIMS Studies in the Management Sciences, 11, North-Holland, Amsterdam; New York and Oxford, 169-177.

[14] Hannan, J.F. and H. Robbins (1955): "Asymptotic Solutions of the Compound Decision Problem," *Ann. Math. Stat.,* 16, 37-51.

[15] Kelly, J.L. (1956): "A New Interpretation of Information Rate," *Bell Systems Technical Journal*, 917-926.

[16] Markowitz, H. (1952): "Portfolio Selection," *The Journal of Finance* VIII, 1, 77-91.

[17] Markowitz, H. (1976): "Investment for the Long Run: New Evidence for an Old Rule," *The Journal of Finance*, XXXI, 5, 1273-1286.

[18] Merton, R.C. and P.A. Samuelson (1974): "Fallacy of the Log-Normal Approximation to Optimal Portfolio Decision-Making Over Many Periods," *Journal of Financial Economics,* 1, 67-94.

[19] Mossin, J. (1968): "Optimal Multiperiod Portfolio Policies," *The J. of Business,* 41, 215-229.

[20] Robbins, H. (1951): "Asymptotically Subminimax Solutions of Compound Statistical Decision Problems," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,* Berkeley: University of California Press.

[21] Samuelson, P.A. (1967): "General Proof that Diversification Pays," *J. of Financial Quantitative Analysis II,* 1-13.

[22] Samuelson, P.A. (1969): "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics,* 239-246.

[23] Samuelson, P.A. (1979): "Why We Should not Make Mean Log of Wealth Big Though Years to Act Are Long," *Journal of Banking and Finance,* 3, 305-307.

[24] Sharpe, W.F. (1963): "A Simplified Model for Portfolio Analysis," *Management Science,* 9, 277-293.

[25] Thorp, E. (1971): "Portfolio Choice and the Kelly Criterion," *Business Econom. Statist. Proc. Amer. Stat. Assoc.* 215-224.