# Normal Dividends*

ROBERT FERNHOLZ
INTECH
One Palmer Square
Princeton, NJ 08542
bob@enhanced.com

IOANNIS KARATZAS
Departments of Mathematics and Statistics
Columbia University
New York, NY 10027
ik@math.columbia.edu

December 8, 2003

**Abstract**

Dividend payments compatible with several market risk/return criteria are derived. The first criterion is that if large and small stocks have the same level of long-term risk, then the expected long-term return on these two classes of stock should be the same. The second criterion is that the market portfolio be optimal for the problem of maximizing logarithmic utility. The third criterion is that the equity market be mean-variance efficient, at least in an asymptotic sense. The dividend rates that satisfy each of these criteria in the context of a stable first-order model for the financial market are essentially the same, and suggest that normal dividend payments on a stock should exactly nullify the stock's rate of return on capital gains. For the larger stocks, these normal dividend rates appear to be much higher than actual dividend rates in the U.S. equity market.

*Key words:* Dividends, ranked market weights, small stocks, mean-variance efficiency.
*JEL classification:* G11, G35, L25, C62.
*MSC (2000) classification:* 91B28.

*This work was completed in the spring semester of 2002, while the second author was on sabbatical leave at the Cowles Foundation for Research in Economics, Yale University. He wishes to thank the Foundation for its hospitality.

# 1 Introduction

Miller and Modigliani (1961) showed that the value of a company as an investment should be independent of whether or not the company paid dividends. Hence, equity market models without dividends would be as representative of reality as those with dividends. Since the inclusion of dividends complicates the structure of a market model, it is not surprising that dividends are frequently omitted from the models used in mathematical finance.

Here we shall show that dividend payments are needed in order to satisfy certain risk/return criteria in the market. First, in a stable market without dividends small stocks will have higher long-term return than larger stocks, even though these two classes of stock have the same level of long-term risk. Second, dividends are necessary if the market portfolio is to be optimal for the problem of maximizing logarithmic utility over the long term. Third, for the market to be mean/variance efficient over the long term, dividend payments are needed.

What is perhaps surprising with the criteria we consider is that the dividend streams that satisfy them are essentially the same in each of these three cases: they cancel the rates of return from capital gains of the individual stocks. As a result, the total return of every stock in the market is the same. This radical result has the benefit of transforming the total return processes for the stocks into martingales.

Section 2 of the paper contains some basic definitions and notation regarding the market model that we use. In Section 3 we consider the problem of large stocks versus small stocks, and in Section 4 we estimate the dividend rates that nullify the return advantage of the smaller stocks. Section 5 discusses the dividend structure required for the market portfolio to be optimal for a logarithmic utility function. Section 6 is devoted to the mean-variance efficiency of an asymptotically stable market. In Section 7 we conclude with a proposal for a definition of the *normal* dividend rate for an equity market.

# 2 The market model

In this section we introduce the general market model that we shall use in the rest of the paper. This model is consistent with the usual market models of continuous-time mathematical finance found in, e.g., Duffie (1992) or Karatzas and Shreve (1998). The preliminary material of this section is presented in greater detail in Fernholz (2002).

Consider a *market* $\mathcal{M}$ consisting of $n$ stocks represented by their price processes $X_1, \dots, X_n$. We assume that there is a single share of each stock, so $X_i(t)$ represents the total capitalization of the $i$-th company at time $t$. The price processes evolve according to

$$X_i(t) = X_0^i \exp\left( \int_0^t \gamma_i(s)\, ds + \int_0^t \sum_{\nu=1}^n \xi_{i\nu}(s)\, dW_\nu(s) \right), \quad t \in [0, \infty), \tag{2.1}$$

for $i = 1, \dots, n$. Here $X_0^1, \dots, X_0^n$, are positive constants and $W(t) = (W_1(t), \dots, W_n(t))$, $t \in [0, \infty)$, is a standard $n$-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$ and adapted to a given filtration $\{\mathcal{F}_t\}$. The *growth rate processes* $\gamma_i = \{\gamma_i(t), \mathcal{F}_t, t \in [0, \infty)\}$, $i = 1, \dots, n$, are measurable, adapted, and satisfy $\int_0^T |\gamma_i(t)| dt < \infty$, a.s., for all $T > 0$. For $i, \nu = 1, \dots, n$, the *volatility processes* $\xi_{i\nu} = \{\xi_{i\nu}(t), \mathcal{F}_t, t \in [0, \infty)\}$ are measurable, adapted, and satisfy:

*i)* $\int_0^T \xi_{i\nu}^2(t) dt < \infty$, a.s., for all $T > 0$;

*ii)* $\lim_{t \to \infty} t^{-1} \xi_{i\nu}^2(t) \log\log t = 0$, a.s.;

*iii)* $\xi_{i1}^2(t) + \cdots + \xi_{in}^2(t) > 0$, $t \in [0, \infty)$, a.s.

We shall assume the stock price processes satisfy:

*iv*) for all $i \neq j$, the set $\{t : X_i(t) = X_j(t)\}$ has Lebesgue measure zero, a.s.;

*v*) for all $i < j < k$, the set $\{t : X_i(t) = X_j(t) = X_k(t)\} = \varnothing$, a.s.

From (2.1), we see that the stock price processes satisfy

$$d \log X_i(t) = \gamma_i(t)\, dt + \sum_{\nu=1}^{n} \xi_{i\nu}(t)\, dW_\nu(t), \quad t \in [0, \infty), \tag{2.2}$$

for $i = 1, \dots, n$. In this form it is evident that these processes are continuous semimartingales, and we shall frequently refer to them simply as *stocks*. The growth rate of a stock determines its long-term behavior, since for $i = 1, \dots, n$,

$$\lim_{T\to\infty} \frac{1}{T}\Big(\log X_i(T) - \int_0^T \gamma_i(t)\, dt\Big) = 0, \quad \text{a.s.} \tag{2.3}$$

A proof of this can be found in Fernholz (2002).

The market *covariance process* is the matrix-valued process $\sigma$ defined by

$$\sigma_{ij}(t) \triangleq \sum_{\nu=1}^{n} \xi_{i\nu}(t)\xi_{j\nu}(t) = \frac{d}{dt}\langle \log X_i, \log X_j\rangle_t, \quad t \in [0, \infty). \tag{2.4}$$

**Definition 2.1.** A *portfolio* of the stocks $X_1, \dots, X_n$ in the market $\mathcal{M}$ is a bounded, measurable, adapted process $\pi : [0, \infty) \times \Omega \to \mathbb{R}^n$ that satisfies $\pi_1(t) + \cdots + \pi_n(t) = 1$, for $t \in [0, \infty)$, a.s.

For each $i$, the process $\pi_i$ represents the *proportion*, or *weight*, of $X_i$ in the portfolio. A negative value for $\pi_i(t)$ indicates a short sale of the $i$-th stock. Suppose $Z_\pi(t)$ represents the value of an investment in the portfolio $\pi$ at time $t$. Then $Z_\pi(t)$ satisfies

$$\begin{aligned} \frac{dZ_\pi(t)}{Z_\pi(t)} &= \sum_{i=1}^{n} \pi_i(t) \frac{dX_i(t)}{X_i(t)} \\ &= \sum_{i=1}^{n} \pi_i(t)\Big(b_i(t)\, dt + \sum_{\nu=1}^{n} \xi_{i\nu}(t)\, dW_\nu(t)\Big), \quad t \in [0, \infty), \end{aligned} \tag{2.5}$$

where $b_i(t) \triangleq \gamma_i(t) + \frac{1}{2}\sigma_{ii}(t)$ is the *rate of return* of the $i$-th stock. This equation and an initial value $Z_\pi(0) > 0$ determine the portfolio value through time, so we shall call the process $Z_\pi$ the *portfolio value process* for $\pi$. Two applications of Itô's rule transform (2.5) into

$$d \log Z_\pi(t) = \sum_{i=1}^{n} \pi_i(t)\, d \log X_i(t) + \gamma_\pi^*(t)\, dt, \quad t \in [0, \infty), \quad \text{a.s.}, \tag{2.6}$$

where

$$\gamma_\pi^*(t) = \frac{1}{2}\Big(\sum_{i=1}^{n} \pi_i(t)\sigma_{ii}(t) - \sum_{i,j=1}^{n} \pi_i(t)\pi_j(t)\sigma_{ij}(t)\Big), \quad t \in [0, \infty), \tag{2.7}$$

is called the *excess growth rate process* of $\pi$. Equation (2.6) is equivalent to

$$d \log Z_\pi(t) = \gamma_\pi(t)\, dt + \sum_{i,\nu=1}^{n} \pi_i(t)\xi_{i\nu}(t)\, dW_\nu(t), \quad t \in [0, \infty), \quad \text{a.s.}, \tag{2.8}$$

where the *portfolio growth rate process* $\gamma_\pi$ is defined by

$$\gamma_\pi(t) = \sum_{i=1}^{n} \pi_i(t)\gamma_i(t) + \gamma^*_\pi(t), \quad t \in [0, \infty).$$

The *portfolio variance process* for $\pi$ is defined by

$$\sigma_{\pi\pi}(t) \triangleq \sum_{i,j=1}^{n} \pi_i(t)\pi_j(t)\sigma_{ij}(t), \quad t \in [0, \infty), \quad \text{a.s.,}$$

with

$$d\langle \log Z_\pi \rangle_t = \sigma_{\pi\pi}(t)\, dt, \quad \text{a.s.} \tag{2.9}$$

The portfolio $\mu$ defined by

$$\mu_i(t) = \frac{X_i(t)}{X_1(t) + \cdots + X_n(t)}, \quad t \in [0, \infty), \tag{2.10}$$

for $i = 1, \ldots, n$, is called the *market portfolio.* It can easily be verified that the weights $\mu_i$ of (2.10) satisfy the requirements of Definition 2.1, and that they are continuous semimartingales. With an appropriate initial value, the value $Z_\mu$ of the market portfolio satisfies

$$Z_\mu(t) = X_1(t) + \cdots + X_n(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{2.11}$$

The processes $\mu_1, \ldots, \mu_n$ are called the *market weight processes*, or simply, *market weights.* A market is called *coherent* if for $i = 1, \ldots, n$ we have

$$\lim_{t\to\infty} t^{-1} \log \mu_i(t) = 0, \quad \text{a.s.}$$

A necessary and sufficient condition for coherence is that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \big(\gamma_i(t) - \gamma_j(t)\big) dt = 0, \quad \text{a.s.,} \tag{2.12}$$

for all $1 \le i, j \le n$; see Fernholz (2002), Proposition 2.1.2. We shall assume henceforth that the market $\mathcal{M}$ is coherent.

Coherence does not imply that $\lim_{T\to\infty} T^{-1} \int_0^T \gamma^*_\mu(t) dt = 0$, a.s. (i.e., that there is no long-term-average gain from diversification). This property holds under the condition $\gamma_i(t) = \gamma_j(t)$, a.s., for a $t \ge 0$, $i \ne j$, which is clearly considerably stronger than (2.12); see Proposition 2.2.3 in Fernholz (2002). The Atlas model of section 5.3 of that book provides an example of a coherent market for which $\lim_{T\to\infty} T^{-1} \int_0^T \gamma^*_\mu(t) dt$ exists a.s. and is positive.

The process $(\mu_{(1)}, \ldots, \mu_{(n)})$ is called the *capital distribution* process, and its component processes, the (reverse) order statistics

$$\mu_{(1)}(t) = \max_{1\le i\le n} \mu_i(t) \ge \mu_{(2)}(t) \ge \cdots \ge \mu_{(n)}(t) = \min_{1\le i\le n} \mu_i(t),$$

are called the *ranked market weights.* For $t \in [0, \infty)$, let $p_t$ be the random permutation of $\{1, \ldots, n\}$ such that for $k$ in $\{1, \ldots, n\}$,

$$\mu_{p_t(k)}(t) = \mu_{(k)}(t), \quad \text{and} \quad p_t(k) < p_t(k+1) \quad \text{if} \quad \mu_{(k)}(t) = \mu_{(k+1)}(t). \tag{2.13}$$

With this notation, it can be shown (see Fernholz (2002), p.81) that the ranked weight processes $\mu_{(k)}$ satisfy

$$d \log \mu_{(k)}(t) = \sum_{i=1}^{n} I_{\{i\}}(p_t(k))\, d \log \mu_i(t) + \frac{1}{2} d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) - \frac{1}{2} d\Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t), \tag{2.14}$$

for $t \in [0, \infty)$, a.s., where $\Lambda_X$ represents the local time at the origin of the continuous semimartingale $X$. By convention, $\Lambda_{\log \mu_{(0)} - \log \mu_{(1)}}(\cdot) \equiv 0 \equiv \Lambda_{\log \mu_{(n)} - \log \mu_{(n+1)}}(\cdot)$.

We now wish to introduce dividends into our model. Suppose we let $\Delta_i(t)$ represent the cumulative dividends paid by the $i$-th stock up to time $t$, and allow dividend payments to be either positive or negative. We shall call $\Delta_i$ the *dividend process* for $X_i$, and assume that all dividend processes are adapted, continuous, and of finite first variation, a.s. With this notation, the dividend process $\Delta_\pi$ for the portfolio $\pi$ satisfies

$$d\Delta_\pi(t) = \sum_{i=1}^{n} \pi_i(t)\, d\Delta_i(t), \quad t \in [0, \infty).$$

We define the *total return processes* $\widehat{X}_i$, $i = 1, \ldots, n$, by

$$d \log \widehat{X}_i(t) = d \log X_i(t) + d\Delta_i(t), \quad t \in [0, \infty), \tag{2.15}$$

and for a portfolio $\pi$, the corresponding total return process $\widehat{Z}_\pi$, where

$$d \log \widehat{Z}_\pi(t) = d \log Z_\pi(t) + d\Delta_\pi(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{2.16}$$

In some cases we shall have *dividend rate* processes $\delta_i$ such that

$$d\Delta_i(t) = \delta_i(t)\, dt, \quad t \in [0, \infty), \tag{2.17}$$

for $i = 1, \ldots, n$ In this case, for a stock $X_i$ with dividends, the *total rate of return* process is given by

$$\hat{b}_i(t) \triangleq \gamma_i(t) + \frac{1}{2}\sigma_{ii}(t) + \delta_i(t) = b_i(t) + \delta_i(t), \quad t \in [0, \infty). \tag{2.18}$$

An application of Itô's rule to the equation

$$d \log \widehat{X}_i(t) = \gamma_i(t)\, dt + \delta_i(t)\, dt + \sum_{\nu=1}^{n} \xi_{i\nu}(t)\, dW_\nu(t), \quad t \in [0, \infty),$$

that results from (2.2), (2.15), and (2.17) indicates that for $i = 1, \ldots, n$ we have

$$d\widehat{X}_i(t) = \widehat{X}_i(t) \Big( \hat{b}_i(t)\, dt + \sum_{\nu=1}^{n} \xi_{i\nu}(t)\, dW_\nu(t) \Big), \quad t \in [0, \infty), \quad \text{a.s.} \tag{2.19}$$

# 3 Small stocks versus large stocks

Over the long term, small stocks have a tendency to outperform large stocks (see Banz (1981) and Reinganum (1981)). Conventionally, this phenomenon has been explained by the putatively greater risk of smaller stocks, but an alternative explanation was proposed in Fernholz (1998, 2001). Here we follow this alternative explanation, and use it to determine dividend payment streams that will offset the small-stock advantage.

Suppose that we fix some integer $m$ in $\{2, \ldots, n-1\}$ and define a *large-stock* portfolio $\xi$ with

$$\xi_{(k)}(t) = \begin{cases} \dfrac{\mu_{(k)}(t)}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)} & \text{for } k = 1, \ldots, m, \\ 0 & \text{for } k = m+1, \ldots, n, \end{cases} \tag{3.1}$$

for $t \in [0, \infty)$. Similarly, we define a *small-stock* portfolio $\eta$ with

$$\eta_{(k)}(t) = \begin{cases} 0 & \text{for } k = 1, \dots, m, \\ \dfrac{\mu_{(k)}(t)}{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)} & \text{for } k = m+1, \dots, n. \end{cases} \tag{3.2}$$

With these two portfolios, it was shown in Fernholz (2002), p.87, that a.s., for $t \in [0, \infty)$,

$$\begin{aligned} d \log\big(Z_\eta(t)/Z_\xi(t)\big) &= d \log\bigg(\frac{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)}\bigg) \\ &\quad + \frac{1}{2}\big(\xi_{(m)}(t) + \eta_{(m+1)}(t)\big) d\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}(t) \\ &= d \log\bigg(\frac{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)}\bigg) + \frac{\mu_{(m)}(t) + \mu_{(m+1)}(t)}{4} \\ &\quad \times \bigg(\frac{1}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)} + \frac{1}{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)}\bigg) \\ &\quad \times d\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}(t). \end{aligned} \tag{3.3}$$

The problem that arises here is that if the ratio of the relative capitalizations of the large-stock and small-stock portfolios remains stable over time, as we might expect it would, then the logarithm on the right-hand side of (3.3) will remain bounded over time, but the local time term in (3.3) is increasing, and hence will eventually dominate. As a result, *over the long term the return on the small-stock portfolio will be greater than the return on the large-stock portfolio.*

An example of this phenomenon is presented in Fernholz (2002), pp. 133–136; it shows that over the period from 1939 to 1998, the stocks ranked 101 to 1000 in the U.S. market had average logarithmic return more that 1% a year greater than the stocks ranked 1 to 100. Moreover, (3.3) shows that this is a structural feature unrelated to the relative riskiness of the two portfolios, so it would be nice to have a structural solution to this "anomaly." Let us see if we can resolve the problem with dividends.

Suppose we have cumulative dividend processes that satisfy

$$\Delta_{(k)}(t) = \frac{1}{2}\big(\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) - \Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t)\big), \quad t \in [0, \infty), \tag{3.4}$$

for $k = 1, \dots, n$. With these dividend processes, the cumulative dividend process $\Delta_\xi$ for the large-stock portfolio $\xi$ of (3.1) will satisfy

$$\begin{aligned} d\Delta_\xi(t) &= \sum_{k=1}^{m} \xi_{(k)}(t)\, d\Delta_{(k)}(t) \\ &= \frac{1}{2}\sum_{k=1}^{m} \xi_{(k)}(t)\big(d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) - d\Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t)\big) \\ &= \frac{1}{2}\sum_{k=1}^{m} \big(\xi_{(k)}(t) d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) - \xi_{(k-1)}(t) d\Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t)\big) \\ &= \frac{1}{2}\xi_{(m)}(t)\, d\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}(t), \quad t \in [0, \infty), \end{aligned} \tag{3.5}$$

a.s. Note that (3.5) follows from the fact that the support of $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}$ lies within the set $\{t : \mu_{(k-1)}(t) = \mu_{(k)}(t)\}$, which is the same set as $\{t : \xi_{(k-1)}(t) = \xi_{(k)}(t)\}$. In a similar manner, we can show that

$$d\Delta_\eta(t) = -\frac{1}{2}\eta_{(m+1)}(t)\, d\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}(t), \quad t \in [0, \infty) \tag{3.6}$$

holds almost surely for the small-stock portfolio $\eta$ of (3.2). For the total return processes corresponding to those of (3.3) with cumulative dividends given by (3.4), it follows from (2.16), (3.3), (3.5), and (3.6) that

$$\begin{aligned} d\log\big(\widehat{Z}_\eta(t)/\widehat{Z}_\xi(t)\big) &= d\log\big(Z_\eta(t)/Z_\xi(t)\big) + d\Delta_\eta(t) - d\Delta_\xi(t) \\ &= d\log\bigg(\frac{\mu_{(m+1)}(t)+\cdots+\mu_{(n)}(t)}{\mu_{(1)}(t)+\cdots+\mu_{(m)}(t)}\bigg), \quad t\in[0,\infty), \end{aligned} \tag{3.7}$$

a.s., and we see that the small-stock advantage *has been exactly offset by the cumulative dividend processes of* (3.4). In particular, it is easily seen that coherence implies

$$\lim_{t\to\infty}\frac{1}{t}\log\big(\widehat{Z}_\eta(t)/\widehat{Z}_\xi(t)\big) = 0, \quad \text{a.s.} \tag{3.8}$$

Note also that, for $m=n$, (3.5) implies

$$\Delta_\mu(t) = 0, \quad t\in[0,\infty), \quad \text{a.s.,} \tag{3.9}$$

so $\widehat{Z}_\mu(t) = Z_\mu(t)$, $t\in[0,\infty)$, a.s. This means that in the market as a whole these dividend payments are merely a redistribution of capital with no net effect on the market return.

In the next section we shall estimate the values of the dividend payments corresponding to the processes in (3.4) in the context of a long-term-stable model for the U.S. equity market.

# 4 Dividend rates that offset the small-stock advantage

We would like to estimate the value of dividend *rate* processes that correspond closely to the cumulative dividend processes defined by (3.4). Let us assume that the asymptotic values

$$\boldsymbol{\lambda}_{k,k+1} = \lim_{t\to\infty} t^{-1}\Lambda_{\log\mu_{(k)}-\log\mu_{(k+1)}}(t) \tag{4.1}$$

exist a.s. for $k=1,\ldots,n$. (It was shown in Fernholz (2002), Section 5.4, that this assumption appears to be valid for the U.S. equity market.) It is tempting to consider dividend rates of the form

$$\delta_{(k)}(t) = \frac{1}{2}\big(\boldsymbol{\lambda}_{k,k+1} - \boldsymbol{\lambda}_{k-1,k}\big),$$

for $k=1,\ldots,n$, however this does not capture the dynamics of the dividends $\Delta_{(k)}$ defined in (3.4). The cumulative dividend processes in (3.4) are increasing when $\mu_{(k)}$ is near $\mu_{(k+1)}$ and decreasing when $\mu_{(k)}$ is near $\mu_{(k-1)}$, and unless the dividend rates are positive when $\mu_{(k)}$ is near $\mu_{(k+1)}$ and negative when $\mu_{(k)}$ is near $\mu_{(k-1)}$, equations analogous to (3.5) and (3.6) will not be valid.

Now, the cumulative dividend process is increasing when $\mu_{(k)}(t)$ is near $\mu_{(k+1)}(t)$, so if we consider the average of these two weights, this should approximate their common value when positive dividend payments take place. Hence, these positive payments correspond to positive dividend rates

$$\bigg(\frac{\mu_{(k)}(t)+\mu_{(k+1)}(t)}{2\mu_{(k)}(t)}\bigg)\boldsymbol{\lambda}_{k,k+1}, \quad t\in[0,\infty), \tag{4.2}$$

and similar reasoning leads to negative dividend rates of

$$-\bigg(\frac{\mu_{(k)}(t)+\mu_{(k-1)}(t)}{2\mu_{(k)}(t)}\bigg)\boldsymbol{\lambda}_{k-1,k}, \quad t\in[0,\infty), \tag{4.3}$$

which correspond to negative dividend payments. Accordingly, the dividend rates will be chosen as

$$\delta_{(k)}(t) = \frac{1}{2}\left(\frac{\mu_{(k)}(t) + \mu_{(k+1)}(t)}{2\mu_{(k)}(t)}\right)\boldsymbol{\lambda}_{k,k+1} - \frac{1}{2}\left(\frac{\mu_{(k-1)}(t) + \mu_{(k)}(t)}{2\mu_{(k)}(t)}\right)\boldsymbol{\lambda}_{k-1,k}, \tag{4.4}$$

for all $t \in [0, \infty)$, for $k = 1, \ldots, n$. With these dividend rates, it is not difficult to see that an equation similar to (3.5) will hold, so that

$$d\Delta_\xi(t) = \frac{1}{2}\left(\frac{\xi_{(m)}(t) + \xi_{(m+1)}(t)}{2}\right)\boldsymbol{\lambda}_{m,m+1}\, dt, \quad t \in [0, \infty), \quad \text{a.s.}, \tag{4.5}$$

and similarly,

$$d\Delta_\eta(t) = -\frac{1}{2}\left(\frac{\eta_{(m)}(t) + \eta_{(m+1)}(t)}{2}\right)\boldsymbol{\lambda}_{m,m+1}\, dt, \quad t \in [0, \infty), \quad \text{a.s.} \tag{4.6}$$

Since $\mu_{(m)}$ and $\mu_{(m+1)}$ are equal on the support of $\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}$, it follows that (3.5) and (4.5) are likely to be close in value, at least over the long term, and that the same is true for (3.6) and (4.6). Hence, the dividend rates given by (4.4) should approximately offset the small-stock advantage, as in (3.7). In particular, the augmented portfolio-value processes $\widehat{Z}_\eta$ and $\widehat{Z}_\xi$ of the small- and large-stock portfolios that correspond to the dividend rate processes of (4.4) satisfy

$$\begin{aligned} d\log\big(\widehat{Z}_\eta(t)/\widehat{Z}_\xi(t)\big) = d\log\left(\frac{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)}\right) + \frac{\mu_{(m)}(t) + \mu_{(m+1)}(t)}{4} \\ \times \left(\frac{1}{\mu_{(1)}(t) + \cdots + \mu_{(m)}(t)} + \frac{1}{\mu_{(m+1)}(t) + \cdots + \mu_{(n)}(t)}\right) \\ \times \big(d\Lambda_{\log \mu_{(m)} - \log \mu_{(m+1)}}(t) - \boldsymbol{\lambda}_{m,m+1}\, dt\big). \end{aligned}$$

In conjunction with coherence and (4.1), this implies

$$\lim_{t\to\infty} \frac{1}{t} \log\big(\widehat{Z}_\eta(t)/\widehat{Z}_\xi(t)\big) = 0, \quad \text{a.s.},$$

as in (3.8).

# 5 The stable first-order model

A *first-order model* for the market was proposed in Fernholz (2002), Section 5.5, and this model can be used to provide an estimate of the dividend rates given by (4.4) in the context of the U.S. equity market. The first-order model is a model of the asymptotic structure of the market, and hence gives a representation of long-term behavior. Let us recall the structure of this model.

Suppose that the limits in (4.1) exist, as well as the limits

$$\boldsymbol{\sigma}^2_{k:k+1} = \lim_{t\to\infty} t^{-1}\langle \log \mu_{(k)} - \log \mu_{(k+1)}\rangle_t,$$

for $k = 1, \ldots, n-1$. In this case we say that the market is *asymptotically stable*, and it was shown in Fernholz (2002), p.102, that in an asymptotically stable market,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \big(\log \mu_{(k)}(t) - \log \mu_{(k+1)}(t)\big)dt = \frac{\boldsymbol{\sigma}^2_{k:k+1}}{2\boldsymbol{\lambda}_{k,k+1}} \quad \text{a.s.}, \tag{5.1}$$

for $k = 1, \ldots, n-1$. We shall henceforth assume that the market is asymptotically stable.

For $k = 1, \ldots, n$, let us define the parameters

$$\boldsymbol{g}_k = \frac{1}{2}\boldsymbol{\lambda}_{k-1,k} - \frac{1}{2}\boldsymbol{\lambda}_{k,k+1}. \tag{5.2}$$

The $2n-2$ parameters $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}, \boldsymbol{\sigma}^2_{1:2}, \ldots, \boldsymbol{\sigma}^2_{n-1:n}$ are called the *characteristic parameters* of the market. Consider the quantities $\boldsymbol{\sigma}^2_1, \ldots, \boldsymbol{\sigma}^2_n$ defined by

$$\begin{aligned}\boldsymbol{\sigma}^2_k &= \frac{1}{4}(\boldsymbol{\sigma}^2_{k-1:k} + \boldsymbol{\sigma}^2_{k:k+1}), \quad k = 2, \ldots, n-1,\\ \boldsymbol{\sigma}^2_1 &= \frac{1}{2}\boldsymbol{\sigma}^2_{1:2}, \quad \text{and} \quad \boldsymbol{\sigma}^2_n = \frac{1}{2}\boldsymbol{\sigma}^2_{n-1:n}.\end{aligned} \tag{5.3}$$

Then the model given by

$$d\log X_i(t) = \boldsymbol{g}_{q_t(i)}\,dt + \boldsymbol{\sigma}_{q_t(i)}\,dV_i(t), \quad t \in [0, \infty), \tag{5.4}$$

where $q_t$ is the inverse of $p_t$ and $V_1, \ldots, V_n$ are independent Brownian motions, is called the *first-order model* for the market. It was shown in Fernholz (2002) that the first-order model defines an asymptotically stable market with characteristic parameters $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$, $\big(3\boldsymbol{\sigma}^2_{1:2}+\boldsymbol{\sigma}^2_{2:3}\big)/4$, $\big(\boldsymbol{\sigma}^2_{k-1:k} + 2\boldsymbol{\sigma}^2_{k:k+1} + \boldsymbol{\sigma}^2_{k+1:k+2}\big)/4$ for $k = 2, \ldots, n-2$, and $\big(\boldsymbol{\sigma}^2_{n-2:n-1} + 3\boldsymbol{\sigma}^2_{n-1:n}\big)/4$. Since this is a smoothed version of the original variance parameters, if the original parameters are sufficiently smooth, the first-order parameters will match them closely. A more complete discussion of this relationship can be found in Fernholz (2002), Section 5.5.

With the first-order model and (5.2), the dividend rate in (4.4) becomes

$$\begin{aligned}\delta_{(k)}(t) &= -\boldsymbol{g}_k + \frac{1}{2}\bigg(\frac{\mu_{(k)}(t) + \mu_{(k+1)}(t)}{2\mu_{(k)}(t)} - 1\bigg)\boldsymbol{\lambda}_{k,k+1} - \frac{1}{2}\bigg(\frac{\mu_{(k-1)}(t) + \mu_{(k)}(t)}{2\mu_{(k)}(t)} - 1\bigg)\boldsymbol{\lambda}_{k-1,k}\\ &= -\boldsymbol{g}_k + \frac{1}{4}\bigg(\frac{\mu_{(k+1)}(t)}{\mu_{(k)}(t)} - 1\bigg)\boldsymbol{\lambda}_{k,k+1} + \frac{1}{4}\bigg(\frac{\mu_{(k-1)}(t)}{\mu_{(k)}(t)} - 1\bigg)\boldsymbol{\lambda}_{k-1,k}.\end{aligned}$$

for $t \in [0, \infty)$, a.s. Since $\log x \cong x - 1$ for $x$ sufficiently close to 1, this becomes

$$\delta_{(k)}(t) \cong -\boldsymbol{g}_k + \frac{1}{4}\log\bigg(\frac{\mu_{(k+1)}(t)}{\mu_{(k)}(t)}\bigg)\boldsymbol{\lambda}_{k,k+1} - \frac{1}{4}\log\bigg(\frac{\mu_{(k-1)}(t)}{\mu_{(k)}(t)}\bigg)\boldsymbol{\lambda}_{k-1,k}, \quad t \in [0, \infty), \quad \text{a.s.}$$

From (5.1) and (5.3), over the long term we have the approximation

$$\begin{aligned}\delta_{(k)}(t) &\cong -\boldsymbol{g}_k - \frac{1}{8}\boldsymbol{\sigma}^2_{k:k+1} - \frac{1}{8}\boldsymbol{\sigma}^2_{k-1:k}\\ &= -\boldsymbol{g}_k - \frac{1}{2}\boldsymbol{\sigma}^2_k,\end{aligned} \tag{5.5}$$

for $t \in [0, \infty)$, a.s. This will hold for $k = 2, \ldots, n-1$, and we can define it to hold for $k = 1$ and $n$. In terms of rates of return in (2.18) for the first-order model, this means that

$$\hat{b}_{(k)}(t) \cong 0, \quad t \in [0, \infty), \quad \text{a.s.}, \tag{5.6}$$

for $k = 1, \ldots, n$. But $\hat{b}_i(\cdot) \equiv 0$ makes the process in (2.19) a martingale, and means that under the first-order model with these dividend rates, all portfolios will have the same null rate of return.

The values in (5.5) can be estimated for actual equity markets. In Fernholz (2002), estimates were given for each of the terms in (5.5) for the U.S equity market over the period from 1990 to 1999. From those estimates, we have calculated the dividend rates $\delta_{(k)}(t)$, and the results are shown in Figure 1. The values are smoothed by convolution with a Gaussian kernel with $\pm 3.16\sigma$ spanning 1000 units on the horizontal axis, with reflection at the ends of the data. The rates are normalized so that none of them are negative. The actual dividend rates over the period are also presented (broken line), with the same smoothing performed.

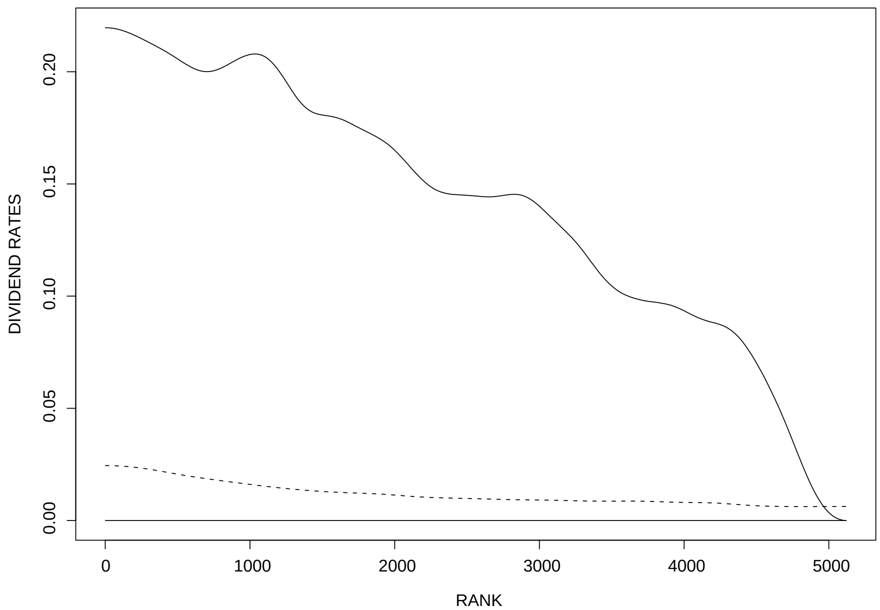


Figure 1: Dividend rates for U.S. equity market, 1990–1999.
Calculated rates (solid line), actual (broken line).

As is obvious from Figure 1, for the larger stocks, the calculated dividend rates are much higher than the corresponding actual rates. Liang and Sharpe (1999) suggest that over the period we consider, the rate of share repurchase by large companies, corrected for the issuance of stock options, might have almost doubled the effective dividend rate for these large companies. However, from Figure 1 it is clear that even with this adjustment, for the period under consideration the dividend rates of the large companies would not have been nearly high enough to offset the small-stock advantage.

# 6 Log-utility optimality of the market portfolio

We shall see in this section that the dividend rate structure of (5.5) is suggested by a rather different consideration: the optimality of the market portfolio (2.10) when maximizing the expected logarithmic utility in the context of the first-order model (5.4).

Suppose that the $i$-th stock pays dividends at the rate $\delta_i$ as in (2.17). Then for the total return process $\widehat{Z}_\pi$ of a portfolio $\pi$, equation (2.5) becomes

$$\frac{d\widehat{Z}_\pi(t)}{\widehat{Z}_\pi(t)} = \sum_{i=1}^{n} \pi_i(t)\Big(\big(\gamma_i(t) + \sigma_{ii}(t) + \delta_i(t)\big)\,dt + \sum_{\nu=1}^{n} \xi_{i\nu}(t)\,dW_\nu(t)\Big), \quad t \in [0,\infty).$$

It is well known (see, e.g., Karatzas and Shreve (1998)) that a portfolio $\pi$ is optimal for the problem of maximizing expected logarithmic utility $E\big(\log \widehat{Z}_\pi(T)\big)$ if

$$\sum_{j=1}^{n} \sigma_{ij}(t)\pi_j(t) = \gamma_i(t) + \frac{1}{2}\sigma_{ii}(t) + \delta_i(t), \quad t \in [0,\infty), \tag{6.1}$$

for $i = 1,\ldots,n$. *What dividend rates are required in order to have $\pi(\cdot) \equiv \mu(\cdot)$?* In the context of the first-order model (5.4), equation (6.1) reads $\boldsymbol{\sigma}_k^2\mu_{(k)}(t) = \boldsymbol{g}_k + \frac{1}{2}\boldsymbol{\sigma}_k^2 + \delta_{(k)}(t)$, for $k = 1,\ldots,n$, and leads to

$$\delta_{(k)}(t) \triangleq \delta_{p_t(k)}(t) = \boldsymbol{\sigma}_k^2\Big(\mu_{(k)}(t) - \frac{1}{2}\Big) - \boldsymbol{g}_k \cong -\boldsymbol{g}_k - \frac{1}{2}\boldsymbol{\sigma}_k^2, \tag{6.2}$$

for $k = 1,\ldots,n$. We are using here the approximation $\mu_{(k)}(t) \ll \frac{1}{2}$, which is sensible in a large, well-regulated market as in the U.S. Clearly, the same dividend structure emerges as in (5.5).

# 7 Mean-variance efficiency of the first-order model

Markowitz (1952) defined a portfolio to be *mean-variance efficient* if, among all portfolios with the same rate of return, there is no portfolio with lower variance. Sharpe (1964) showed that under certain ideal circumstances the market portfolio will be mean-variance efficient. With the understanding that the first-order model represents merely a stable version of the market, and not necessarily the market itself, let us nevertheless determine the dividend rates compatible with mean-variance efficiency for the first-order model.

Suppose we start with the dividend rates (5.5), and perturb these rates rates by $\delta'_{(k)}(t)$, for $k = 1,\ldots,n$. In this case,

$$\sigma_\mu^2(t) = \sum_{k=1}^{n} \mu_{(k)}^2(t)\boldsymbol{\sigma}_k^2 \tag{7.1}$$

will be minimum under the constraints

$$\sum_{k=1}^{n} \mu_{(k)}(t)\big(b_{(k)}(t) + \delta'_{(k)}(t)\big) = \text{constant} \tag{7.2}$$

and

$$\sum_{k=1}^{n} \mu_{(k)}(t) = 1. \tag{7.3}$$

By (5.3), (7.2) reduces to

$$\sum_{i=1}^{n} \mu_{(k)}(t)\delta'_{(k)}(t) = \text{constant}.$$

Since (7.1) represents a minimum under the constraints (7.2) and (7.3), we have

$$\mu_{(k)}(t)\boldsymbol{\sigma}_k^2 = \lambda_1\delta'_{(k)}(t) + \lambda_2,$$

where $\lambda_1$ and $\lambda_2$ are constants. If we multiply this by $\mu_{(k)}(t)$ and sum over $k$, then we obtain

$$\sigma_\mu^2(t) = \lambda_1\delta'_\mu(t) + \lambda_2, \tag{7.4}$$

where

$$\delta'_\mu(t) = \sum_{k=1}^{n} \mu_{(k)}(t)\delta'_{(k)}(t).$$

Equation (7.4) will be satisfied if $\lambda_1 = 1$, $\lambda_2 = 0$, and $\delta'_\mu(t) = \sigma^2_\mu(t)$, from which we can conclude that

$$\delta'_{(k)}(t) = \mu_{(k)}(t)\boldsymbol{\sigma}^2_k.$$

How large are the perturbations $\delta'_{(k)}(t)$? We have $\delta'_{(1)}(t) \cong .0014$, and the rest of the $\delta'_{(k)}(t)$ decline rapidly with increasing $k$. Hence, the Figure 1 remains essentially unchanged with the perturbed dividend rates.

# 8 Conclusion: normal dividend rates

As we have seen in Sections 4, 5, and 6, the dividend rates that eliminate the small-stock advantage, those that maximize the log-utility of the first-order model of the market, and those that make the first-order model mean-variance efficient, are all essentially the same. These dividend rates cause all the stocks in the market to have the same rate of return, which normalizes to zero. We shall call these dividend rates the *normal* dividend rates for the market. With normal dividend rates, the stock price processes for the first-order market model become martingales.

From empirical data, we have seen that for the larger stocks in the market, the normal dividend rates are much higher than actual historical rates. Since the first-order model is an asymptotic model, it provides insight into long-term market behavior. The insight provided here would appear to be that unless the larger companies start to pay significantly higher dividends, small stocks are likely to be a better long-term investment than large stocks.

# References

Banz, R. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9*, 3–18.

Duffie, D. (1992). *Dynamic Asset Pricing Theory*. Princeton, NJ: Princeton University Press.

Fernholz, R. (1998, May/June). Crossovers, dividends, and the size effect. *Financial Analysts Journal 54*(3), 73–78.

Fernholz, R. (2001). Equity portfolios generated by functions of ranked market weights. *Finance and Stochastics 5*, 469–486.

Fernholz, R. (2002). *Stochastic Portfolio Theory*. New York: Springer-Verlag.

Karatzas, I. and S. E. Shreve (1998). *Methods of Mathematical Finance*. New York: Springer-Verlag.

Liang, J. N. and S. A. Sharpe (1999). Share repurchases and employee stock options and their implications for S&P 500 share retirements and expected returns. Technical report, Division of Research and Statistics, Federal Reserve Board, `http://www.federalreserve.gov/pubs/feds/1999/`.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance 7*, 77–91.

Miller, M. H. and F. Modigliani (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business 34*, 411–433.

Reinganum, M. (1981). Misspecification of capital asset pricing: empirical anomalies based on earnings yields and market values. *Journal of Financial Economics 9*, 19–46.

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance 19*, 425–442.