Stable Models for the Distribution of Equity Capital

Robert Fernholz

INTECH
One Palmer Square
Princeton, NJ 08542

February 21, 2001

Abstract

The distribution of capital, or equivalently, the distribution of firm size, is studied in the context of equity markets modeled in continuous time. The capital distribution is defined to be the set of market weights arranged in decreasing order, and for the market of exchange-traded U.S. stocks, this capital distribution appears to be reasonably stable over time. A model based on the continuous semimartingale representation for stock price processes is proposed, and it is shown that for a stable capital distribution in a market comprising n stocks, the model depends on $2n - 2$ constant parameters. Empirical tests and simulations show that the model provides an accurate representation of the short-term dynamics of the distribution of capital in the U.S. market.

Key words: Capital distribution, firm size, local time, ranked market weights.
JEL classification: G10, L11, C62.
MSC (1991) classification: 90A09, 60H30, 60G44.

1 Introduction

The distribution of equity capital, or equivalently, the distribution of the size of firms, has been studied extensively, using many methodologies (see, e.g., Gibrat (1931), Simon and Bonini (1958), Steindl (1965), Ijiri and Simon (1977)). Historically, the goal of the research in this area has been to find plausible mechanisms by which various models of corporate behavior would result in the observed distribution of firm size. In contrast to the historical literature, our goal in this paper is to construct an equity market model that is compatible with the observed capital distribution, without regard to the process which may have produced that distribution. In particular, we shall model the U.S. equity market over the period from 1990 to 1999.

In an equity market, the *market weight* of a stock is defined to be the capitalization of the underlying company divided by the total capitalization of the market, and the *capital distribution* of the market is defined as the family of market weights arranged in decreasing order. The *capital distribution curve* refers to the log-log plot of the market weights arranged in descending order, i.e., the logarithms of the market weights versus the logarithms of their respective ranks. Simon (1955) showed that for a market in which stocks are replaced at a constant rate, under certain simple conditions the capital distribution curve would be approximately a straight line. This weight distribution is called a *Pareto distribution*, and is discussed in detail in Steindl (1965). The Pareto distribution appears to apply to other size distributions (e.g., cities, ecological systems, etc.), and it has been suggested that its ubiquity is the result of a fundamental law of nature (see Bak (1996)).

How closely does the Pareto distribution conform to actual capital distributions? Ijiri and Simon (1974) observed that the log-log plots of actual capital distributions are usually concave, and hence are not accurately represented by a simple Pareto model. Since then, much of the literature has been devoted to explaining this observed discrepancy. Among the more current contributions, Jovanovic (1982) and Hopenhayn (1992) proposed dynamic models for corporate growth, Axtell (1999) used game-theoretic methods, and Hashemi (2000) applied the theory of evolutionary economics to explain the observed distribution of firm size.

Let us now consider some actual capital distribution curves, and see what they look like. In Figure 1 we see eight capital distribution curves for the U.S. stock market for the dates December 31, 1929, 1939, 1949, 1959, 1969, 1979, 1989, 1999. These curves are generated using capitalization data from the monthly stock database of the Center for Research in Securities Prices (CRSP) at the University of Chicago. The distribution curves include all NYSE, AMEX, and NASDAQ stocks after the removal of closed-end funds, REITs, and ADRs not included in the S&P 500 Index. Each successive curve contains a larger number of stocks than the previous one, and can be identified by this characteristic.

The curves in Figure 1 are evidently not straight lines, but instead are generally concave, as noted by Ijiri and Simon (1974). All the curves have a similar shape, so some type of stability appears to be present. We wish to develop a market model that is consistent with the observed stability of the capital distribution.

Market stability has been considered in various forms in the past. For example, Grandmont and Hildenbrand (1974) and Hellwig (1980) considered various forms of stationary economic equilibria, and Fernholz (1999) studied conditions under which market diversity, a measure of the spread of capital in the market, is maintained. Stability of the capital distribution is consistent with the conditions imposed in these references, however it is somewhat more restrictive over the long term. The model we present in this paper is a short-term model that captures the essential features of the structure of the capital distribution of the U.S. market, but cannot be expected to mimic the long-term behavior in which shifts in the distribution take place.

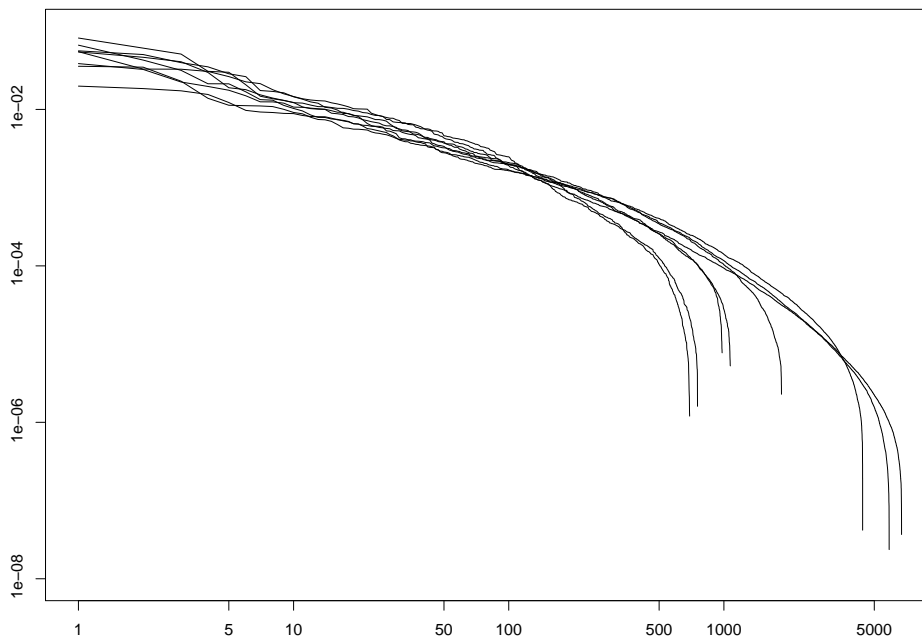We shall operate in the setting of continuous-time mathematical finance, with stock price pro-

Figure 1: Capital distribution curves: 1929–1999

cesses represented by continuous semimartingales (see, e.g., Duffie (1992) or Karatzas and Shreve (1998)). The representation of market weights in terms of continuous semimartingales is straight-forward, but in order to represent the ranked market weights, it is necessary to use semimartingale local times to capture the behavior when ranks change. The methodology for this analysis was developed in Fernholz (2001), and is outlined here in an Appendix. By using the representation of ranked market weights given in Fernholz (2001), we are able to determine the asymptotic behavior of the capital distribution. For a market with a stable capital distribution, this asymptotic behavior provides insight into the steady-state structure of the market.

We shall assume that we operate in a continuously-traded, frictionless market in which the stock prices vary continuously and the companies pay no dividends. We assume that companies neither enter nor leave the market, nor do they merge or break up, and that the total number of shares of a company remains constant. Shares of stock are assumed to be infinitely divisible, so we can assume without loss of generality that each company has a single share of stock outstanding.

Section 2 of the paper contains some basic definitions and results regarding the basic market model that we use. In Section 3 we present a model for a stable capital distribution, and we apply this model to the U.S. equity market in Section 4. Section 5 is a summary, and the Appendix contains some technical mathematical results that we need in the other sections.

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), but follows the logarithmic representation used in, e.g., Fernholz (1999).

Consider a family of n stocks represented by their price processes $X_1, \ldots, X_n$. 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 $i = 1, \ldots, n$. Here X_0^i, $i = 1, \ldots, n$, are positive constants and $W = \{W(t) = (W_1(t), \ldots, W_n(t)), \mathcal{F}_t, t \in [0, \infty)\}$ is a standard n-dimensional Brownian motion defined on a complete probability space $\{\Omega, \mathcal{F}, P\}$ where $\{\mathcal{F}_t\}$ is the P-augmentation of the natural filtration $\{\mathcal{F}_t^W = \sigma(W(s); 0 \le s \le t)\}$. The *growth rate processes* $\gamma_i = \{\gamma_i(t), \mathcal{F}_t, t \in [0, \infty)\}$, $i = 1, \ldots, n$, are measurable, adapted, and satisfy $\int_0^T |\gamma_i(t)| dt < \infty$, a.s., for all $T > 0$. For $i, \nu = 1, \ldots, n$, 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.

From (2.1), we see that the stocks 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 $i = 1, \ldots, n$, and in this form it is evident that the stock price processes are continuous semi-martingales. We shall frequently refer to the stock price processes simply as *stocks*.

Consider the matrix-valued process ξ defined by $\xi(t) = (\xi_{i\nu}(t))_{1 \le i, \nu \le n}$ and define the *covariance process* σ where $\sigma(t) = \xi(t)\xi^T(t)$. Then $\sigma_{ij}(t)\, dt = d\langle \log X_i, \log X_j \rangle_t$, for all $t \in [0, \infty)$, a.s. The conditions on the volatility processes ensure that the $\sigma_{ij}(\cdot)$ are a.s. locally L^1 functions that satisfy

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

We shall need to impose a mild nondegeneracy condition on the families of stocks that we consider.

Definition 2.1. The processes $X_1, \ldots, X_n$ are *pathwise mutually nondegenerate* if:

i) for all $i \ne j$, $\quad \{t : X_i(t) = X_j(t)\}$ has Lebesgue measure zero, a.s.;

ii) for all $i < j < k$, $\quad \{t : X_i(t) = X_j(t) = X_k(t)\} = \varnothing$, a.s.

The components of multidimensional Brownian motion $(W_1, \ldots, W_n)$ are pathwise mutually nondegenerate, at least for $t > 0$. Condition (i) is well-known for Brownian motion, and condition (ii) follows from the fact that, with probability one, 2-dimensional Brownian motion never returns to the origin (see Karatzas and Shreve (1991)).

Definition 2.2. A *market* $\mathcal{M}$ is a family of stocks $X_1, \ldots, X_n$ that are pathwise mutually nondegenerate.

Definition 2.3. A *portfolio* of the stocks $X_1, \ldots, X_n$ in the market $\mathcal{M}$ is a measurable, adapted process π that is bounded on $[0, \infty) \times \Omega$ and satisfies $\pi_1(t) + \cdots + \pi_n(t) = 1$, for $t \in [0, \infty)$, a.s.

For each i, the process π_i represents the *proportion*, or *weight*, of X_i in the portfolio. A negative value for $\pi_i(t)$ indicates a short sale.

3

Suppose $Z_\pi(t)$ represents the value of an investment in the portfolio π at time t. Then $Z_\pi(t)$ satisfies

$$dZ_\pi(t) = Z_\pi(t) \sum_{i=1}^n \pi_i(t) \frac{dX_i(t)}{X_i(t)}, \quad t \in [0, \infty). \tag{2.3}$$

This equation and an initial value $Z_\pi(0) > 0$ determine the portfolio value through time (see Fernholz (1999)), so we shall call the process Z_π the *portfolio value process* for π. Two applications of Itô's lemma transform (2.3) into

$$d \log Z_\pi(t) = \sum_{i=1}^n \pi_i(t) \, d \log X_i(t) + \gamma_\pi^*(t) \, dt, \quad t \in [0, \infty), \quad \text{a.s.}, \tag{2.4}$$

where

$$\gamma_\pi^*(t) = \frac{1}{2} \left(\sum_{i=1}^n \pi_i(t) \sigma_{ii}(t) - \sum_{i,j=1}^n \pi_i(t) \pi_j(t) \sigma_{ij}(t) \right), \quad t \in [0, \infty), \tag{2.5}$$

is called the *excess growth rate* of π. Equation (2.4) is equivalent to

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

where the *portfolio growth rate process* γ_π 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 π is defined by

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

with

$$\langle \log Z_\pi \rangle_t = \int_0^t \sigma_{\pi\pi}(s) \, ds, \quad \text{a.s.} \tag{2.7}$$

The following proposition generalizes Proposition 2.2 of Fernholz (1999), and shows that the growth rate of a portfolio determines the portfolio's long-term behavior

Proposition 2.1. *For any portfolio π in $\mathcal{M}$,*

$$\lim_{T \to \infty} \frac{1}{T} \left(\log Z_\pi(T) - \int_0^T \gamma_\pi(t) \, dt \right) = 0, \quad \text{a.s.} \tag{2.8}$$

Proof. Equation (2.6) gives the canonical decomposition of the semimartingale $\log Z_\pi$. Since the proportions in π are bounded, condition (*ii*) for the stock volatility processes implies that

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

This equation, equation (2.7), and some elementary calculus, imply that Lemma A.3 can be applied to (2.8). The conclusion follows. $\qquad\square$

Since single stocks can be considered to be portfolios with a single non-zero proportion, Proposition 2.1 also applies to single stocks. There is one portfolio that will be of particular interest to us here.

Definition 2.4. The portfolio μ defined by

$$\mu_i(t) = \frac{X_i(t)}{X_1(t) + \cdots + X_n(t)}, \quad t \in [0, \infty), \tag{2.9}$$

for $i = 1, \dots, n$, is called the *market portfolio (process)*.

It can easily be verified that the weights μ_i defined by (2.9) satisfy the requirements of Definition 2.3, and that they are continuous semimartingales. If we let

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

then $Z(t)$ satisfies (2.3) with the weights μ_i, so with appropriate initial conditions, $Z_\mu = Z$, and the portfolio value process represents the combined capitalization of all the stocks in the market. Henceforth we shall let μ exclusively represent the market portfolio and Z represent its value process. The processes $\mu_1, \dots, \mu_n$ are called the *market weight processes*, or simply, *market weights*.

The market weight processes are quotient processes

$$\begin{aligned} d \log \mu_i(t) &= d \log\big(X_i(t)/Z(t)\big) \\ &= d \log X_i(t) - d \log Z(t), \end{aligned} \tag{2.11}$$

so the logarithmic change in a market weight represents the relative logarithmic return of the corresponding stock compared to the market. It follows from (2.11) that the variance rate of $\log \mu_i - \log \mu_j$ is the same as the variance rate of $\log X_i - \log X_j$, and hence can be represented by

$$\sigma_{i:j}(t) = \sigma_{ii}(t) - 2\sigma_{ij}(t) + \sigma_{jj}(t) \quad t \in [0, \infty). \tag{2.12}$$

With this notation,

$$\langle \log \mu_i - \log \mu_j \rangle_t = \int_0^t \sigma_{i:j}(s)\, ds, \quad t \in [0, \infty), \quad \text{a.s.}$$

Since the capital distribution of the market is composed of the weights μ_i, these results will be needed in the following sections.

3 The structure of stable capital distributions

In this section we construct a model for a market with a stable capital distribution. The capital distribution involves the ranked market weights, so we must start with a precise definition of rank processes. We then proceed with a representation for the ranked market weights that involves the use of semimartingale local times. Finally, we derive our market model, and we show how it relates to the classical models that result in the Pareto distribution.

Definition 3.1. Let $X_1, \dots, X_n$ be processes. For $1 \leq k \leq n$, the k-th *rank process* of $X_1, \dots, X_n$ is defined by

$$X_{(k)}(t) = \max_{i_1 < \cdots < i_k} \min\big(X_{i_1}(t), \dots, X_{i_k}(t)\big), \quad t \in [0, \infty),$$

where $1 \leq i_1$ and $i_k \leq n$. We shall adopt the convention that $X_{(0)}$ and $X_{(n+1)}$ are defined such that for $t \in [0, \infty)$, $X_{(0)}(t) > X_{(1)}(t)$ and $X_{(n+1)}(t) < X_{(n)}(t)$.

Note that, according to Definition 3.1, for $t \in [0, \infty)$,

$$\max_{1 \leq i \leq n} X_i(t) = X_{(1)}(t) \geq X_{(2)}(t) \geq \cdots \geq X_{(n)}(t) = \min_{1 \leq i \leq n} X_i(t),$$

so that at any given time, the values of the rank processes represent the values of the original processes arranged in descending order.

For a market $\mathcal{M}$ with weight processes $\mu_1, \ldots, \mu_n$, the vector $(\mu_{(1)}(t), \ldots, \mu_{(n)}(t))$ is called the *capital distribution* of the market at time $t \in [0, \infty)$. For any $t \in [0, \infty)$, we can define p_t to be the random permutation of $\{1, \ldots, n\}$ such that for $k \in \{1, \ldots, n\}$,

$$\mu_{p_t(k)}(t) = \mu_{(k)}(t), \quad \text{and} \quad p_t(k) < p_t(k+1) \quad \text{if} \quad \mu_{(k)}(t) = \mu_{(k+1)}(t). \tag{3.1}$$

The permutation p_t is uniquely defined by (3.1), and associates each rank process with one of the original market weights that has the same value at time t.

Although a change in a market weight represents the relative return of the corresponding stock, changes in the ranked market weights do not have such a simple interpretation. In order to characterize the changes in the ranked market weights in terms of the relative returns of the stocks, we need to recall the concept of *semimartingale local time*, a measure of the amount of time a semimartingale spends near the origin.

Definition 3.2. Let X be a continuous semimartingale. Then the *local time* for X (at 0) is the process Λ_X defined for $t \in [0, \infty)$ by

$$\Lambda_X(t) = \frac{1}{2}\Big(|X(t)| - |X(0)| - \int_0^t \operatorname{sgn}(X(s))\,dX(s)\Big), \tag{3.2}$$

where $\operatorname{sgn}(x) = 2\,I_{(0,\infty)}(x) - 1$, with $I_{(0,\infty)}$ the indicator function of $(0, \infty)$.

For general background on local times, see Karatzas and Shreve (1991), and for background on their application in mathematical finance, see Fernholz (2001). Equation (3.2), which we use as a definition, is one of the *Tanaka-Meyer formulas* (see Karatzas and Shreve (1991), (3.7.9)). It can be shown that $\Lambda_X(t)$ is almost surely nondecreasing in t, and satisfies

$$I_{\{0\}}(X(t))\,d\Lambda_X(t) = d\Lambda_X(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{3.3}$$

(see Karatzas and Shreve (1991), 3.7.1). This implies, for example, that for one-dimensional Brownian motion B, Λ_B is a non-negative random measure on $[0, \infty)$ that almost surely has support contained in the set $\{t : B(t) = 0\}$, and hence is singular with respect to Lebesgue measure.

The ranked market weights can be represented using local times. Proposition A.1 implies that the ranked market weight processes $\mu_{(1)}, \ldots, \mu_{(n)}$ satisfy

$$d \log \mu_{(k)}(t) = \sum_{i=1}^n I_{\{0\}}\big(p_t(k) - i\big)d \log \mu_i(t) + \frac{1}{2}d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t)$$

$$-\frac{1}{2}d\Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{3.4}$$

(By convention, $\Lambda_{\log \mu_{(n)} - \log \mu_{(n+1)}}(t) = 0$ and $\Lambda_{\log \mu_{(0)} - \log \mu_{(1)}}(t) = 0$, for all $t \in [0, \infty)$.) If we define

$$g_k(t) = \gamma_{p_t(k)}(t) - \gamma_\mu(t), \quad t \in [0, \infty), \tag{3.5}$$

then,

$$d \log \mu_{(k)}(t) = g_k(t)\,dt + \frac{1}{2}d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) - \frac{1}{2}d\Lambda_{\log \mu_{(k-1)} - \log \mu_{(k)}}(t)$$

$$+ \sum_{\nu=1}^n \xi_{p_t(k)\nu}(t)\,dW_\nu(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{3.6}$$

6

We shall now use this representation to analyze the capital distribution. Consider the capital distribution curves in Figure 1. The evolution of these curves over the seven decades shown exhibits some form of stability, although the curves do not all have exactly the same shape. We need to capture these properties in a formal definition.

Definition 3.3. The capital distribution $(\mu_{(1)}, \ldots, \mu_{(n)})$ of the market $\mathcal{M}$ is *asymptotically stable* if:

i) $\lim_{t \to \infty} t^{-1} \log \mu_{(n)}(t) = 0$, a.s.,

ii) for $k = 1, \ldots, n-1$, $\quad \lim_{t \to \infty} t^{-1} \Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) = \boldsymbol{\lambda}_{k,k+1} > 0$, a.s.

iii) for $k = 1, \ldots, n-1$, $\quad \lim_{t \to \infty} t^{-1} \langle \log \mu_{(k)} - \log \mu_{(k+1)} \rangle_t = \boldsymbol{\sigma}_{k:k+1}^2 > 0$, a.s.

Condition (i) will be satisfied if, for example, the smallest weight $\mu_{(n)}$ does not tend to zero too quickly, a reasonable condition to impose on a stable capital distribution. Conditions (ii) and (iii) imply that asymptotic limits exist for the slopes of the processes $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}$ and $\langle \log \mu_{(k)} - \log \mu_{(k+1)} \rangle$, so both of these conditions are consistent with the heuristic idea of asymptotic stability. Condition (ii) appears to be reasonably consistent with Figure 2 below, a plot of the values of $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t)$ for the U.S. equity market. A plot of the estimated quadratic variation processes $\langle \log \mu_{(k)} - \log \mu_{(k+1)} \rangle_t$ is similar to Figure 2, so condition (iii) also appears to be consistent.

Proposition 3.1. *If* $(\mu_{(1)}, \ldots, \mu_{(n)})$ *is asymptotically stable, then for* $k = 1, \ldots, n$,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T g_k(t) \, dt = \frac{1}{2} \boldsymbol{\lambda}_{k-1,k} - \frac{1}{2} \boldsymbol{\lambda}_{k,k+1}, \quad \text{a.s.,} \tag{3.7}$$

Proof. Follows from Proposition A.2. $\qquad\square$

In practice, the values of the $\boldsymbol{\lambda}_{k,k+1}$ can be estimated from the asymptotic slopes of the local times $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}$. To be consistent with the notation (A.7) in the Appendix, we define

$$\boldsymbol{g}_k = \frac{1}{2} \boldsymbol{\lambda}_{k-1,k} - \frac{1}{2} \boldsymbol{\lambda}_{k,k+1}, \quad \text{a.s.,} \tag{3.8}$$

and this implies that for $k = 1, \ldots, n-1$,

$$\boldsymbol{\lambda}_{k,k+1} = -2(\boldsymbol{g}_1 + \cdots + \boldsymbol{g}_k), \tag{3.9}$$

as in (A.8). From (3.9) it follows that $\boldsymbol{g}_1 + \cdots + \boldsymbol{g}_k < 0$ for $k = 1, \ldots, n-1$, a.s., and that $\boldsymbol{g}_1 + \cdots + \boldsymbol{g}_n = 0$. Hence, an asymptotically stable capital distribution is characterized by the $2n-2$ parameters $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$ and $\boldsymbol{\sigma}_{1:2}^2, \ldots, \boldsymbol{\sigma}_{n-1:n}^2$, and these parameters will be called the *characteristic parameters* of the distribution.

As in the Appendix, we can proceed from an asymptotically stable capital distribution to a stable capital distribution. Condition (iii) of Definition 3.3 is of the form of (A.10), so as in (A.12), we have a *stable model* for the differences between consecutive ranked weights,

$$d\big(\log \mu_{(k)}(t) - \log \mu_{(k+1)}(t)\big) = -\boldsymbol{\lambda}_{k,k+1} \, dt + d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) + dM_{k:k+1}(t), \tag{3.10}$$

for $k = 1, \ldots, n-1$. Here the $M_{k:k+1}$ are continuous martingales with $\langle M_{k:k+1} \rangle_t = \boldsymbol{\sigma}_{k:k+1}^2 t$, for $t \in [0, \infty)$, a.s. Equation (3.10) and (A.22) imply that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \big(\log \mu_{(k)}(t) - \log \mu_{(k+1)}(t)\big) dt = \frac{\boldsymbol{\sigma}_{k:k+1}^2}{2\boldsymbol{\lambda}_{k,k+1}}, \quad \text{a.s.,} \tag{3.11}$$

7

for $t \in [0, \infty)$ and $k = 1, \ldots, n-1$. To estimate the slope of the log-log plot of the capital distribution curve at the point above $\log k$, (3.11) can be used to obtain

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{\log \mu_{(k)}(t) - \log \mu_{(k+1)}(t)}{\log(k) - \log(k+1)} \, dt \approx -\frac{k\boldsymbol{\sigma}_{k:k+1}^2}{2\boldsymbol{\lambda}_{k,k+1}}, \quad \text{a.s.,} \tag{3.12}$$

for $k = 1, \ldots, n - 1$. From (3.9) and (3.12), we see that if the $\boldsymbol{\sigma}_{k:k+1}^2$ and $\boldsymbol{g}_k$ change by a common factor, then there is no change in the capital distribution. If the $\boldsymbol{\sigma}_{k:k+1}^2$ stay fixed, and the $\boldsymbol{g}_k$ increase by a common factor, then the slope of the capital distribution curve will decrease. This would seem to be reasonable, since this would lower the growth rates for the larger stocks and raise the growth rates for the smaller stocks.

In the next section, we shall test how closely the stable model (3.10) reproduces the capital distribution generated by actual data from the U.S. stock market, but let us first consider here an example of a market similar in structure to that considered by Simon (1955).

Example 3.1. Suppose that we have stock price processes $X_1, \ldots, X_n$ defined as in (2.2) that satisfy

$$d \log X_i(t) = \gamma_i(t) \, dt + \sigma \, dW_i(t), \quad t \in [0, \infty),$$

and assume that for a positive constant g, a.s., for all $t \in [0, \infty)$, $\gamma_{p_t(n)}(t) = ng$, and $\gamma_{p_t(i)}(t) = 0$ for $i = 1, \ldots, n - 1$. In this case, Proposition 2.1 implies that

$$\lim_{t \to \infty} t^{-1} \big(\log X_1(t) + \cdots + \log X_n(t) \big) = ng, \quad \text{a.s.,}$$

so, by symmetry it can be shown that

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

for $i = 1, \ldots, n$. Therefore,

$$\lim_{t \to \infty} t^{-1} \log X_{(k)}(t) = g, \quad \text{a.s.,} \tag{3.13}$$

for $i = k, \ldots, n$. From this it can be shown that

$$\lim_{t \to \infty} t^{-1} \log \big(X_1(t) + \cdots + X_n(t) \big) = g, \quad \text{a.s.,}$$

and therefore,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \gamma_\mu(t) \, dt = g, \quad \text{a.s.}$$

This is reasonable, since the smallest stock has growth rate ng and the other stocks have growth rate zero, and each stock is smallest about $1/n$ of the time, so the market should asymptotically grow at the rate g.

Proposition A.1 implies that

$$d \log X_{(n)}(t) = ng \, dt - \frac{1}{2} d\Lambda_{\log X_{(n-1)} - \log X_{(n)}}(t) + \sigma \, dB(t), \quad t \in [0, \infty), \quad \text{a.s.,}$$

where B is a Brownian motion. Hence, (3.13) and the strong law of large numbers for Brownian motion (Karatzas and Shreve (1991), 2.9.3) imply that

$$\boldsymbol{\lambda}_{n-1,n} = 2(n-1)g, \quad \text{a.s.,}$$

8

and inductively, that

$$\boldsymbol{\lambda}_{k,k+1} = 2kg, \quad \text{a.s.},$$

for $k = 1, \ldots, n-1$. Therefore, (3.8) implies that $\boldsymbol{g}_k = -g$, for $k = 1, \ldots, n-1$, and $\boldsymbol{g}_n = (n-1)g$. For $k = 1, \ldots, n-1$,

$$\boldsymbol{\sigma}_{k:k+1}^2 = 2\sigma^2, \quad \text{a.s.},$$

so the characteristic parameters for this example are $\boldsymbol{g}_1 = \cdots = \boldsymbol{g}_{n-1} = -g$ and $\boldsymbol{\sigma}_{1:2}^2 = \cdots = \boldsymbol{\sigma}_{n-1:n}^2 = 2\sigma^2$. It follows from (3.12) that the log-log slope between $\mu_{(k)}$ and $\mu_{(k+1)}$ is $-\sigma^2/2g$, for $k = 1, \ldots, n-1$, so the log-log plot of the asymptotic capital distribution is a straight line, just as in Simon (1955).

This example shows that a market in which all the stocks have the same growth rate, except for the smallest stock which has a higher growth rate, results in a Pareto distribution. The higher growth rate of the smallest stock simulates the replacement of stocks that occurs in the models of Simon (1955) and Steindl (1965). In Fernholz (1999) it was shown that for a market in which all the stocks share a common growth rate, the capital distribution asymptotically degenerates so that the capital becomes concentrated into the largest stock. Hence, some variation among the growth rates of the stocks is necessary in order for there to exist an asymptotically stable capital distribution.

4 Application to the U.S. equity market

We now have a model for a stable capital distribution, and we would like to see how closely this model conforms to reality. To determine this, in this section we apply the methodology we have developed to model the capital distribution of the U.S. equity market over the ten year period from January 1, 1990, to December 31, 1999. The data we use here are from the same CRSP monthly stock database used to for the capital distribution curves in Figure 1.

To carry out the calculations of the previous section, we must estimate the values of the parameters $\boldsymbol{\lambda}_{k,k+1}$ To do so, we need the values for the local time processes $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}$. Now, in Fernholz (2001), it was shown that for a portfolio ξ with weights

$$\xi_{(i)}(t) = \begin{cases} \dfrac{\mu_{(i)}(t)}{\mu_{(1)}(t) + \cdots + \mu_{(k)}(t)}, & i \leq k, \\ 0, & i > k, \end{cases}$$

for $t \in [0, \infty)$, the relative return of ξ versus the market satisfies

$$\begin{aligned} d\log\big(Z_\xi(t)/Z(t)\big) = \ & d\log\big(\mu_{(1)}(t) + \cdots + \mu_{(k)}(t)\big) \\ & - \frac{1}{2}\xi_{(k)}(t)\, d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t), \quad t \in [0, \infty), \quad \text{a.s.} \end{aligned} \tag{4.1}$$

Since the relative return on the left-hand side of (4.1) and the first term on the right-hand side can both be estimated easily, this equation can be used to determine the value of $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t)$. With these estimates, we can proceed to estimate the parameters $\boldsymbol{\lambda}_{k,k+1}$.

Figure 2 shows the local time processes $\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}$ for the values of $k = 10, 20, 40, \ldots, 5120$. Although the curves are not straight lines, nevertheless we shall use the slopes of the lines connecting the endpoints of the curves to estimate the values of the parameters $\boldsymbol{\lambda}_{k,k+1}$. Curves for the estimated quadratic variation processes $\langle \log \mu_{(k)} - \log \mu_{(k+1)} \rangle$ are similar to those of Figure 2, and the values
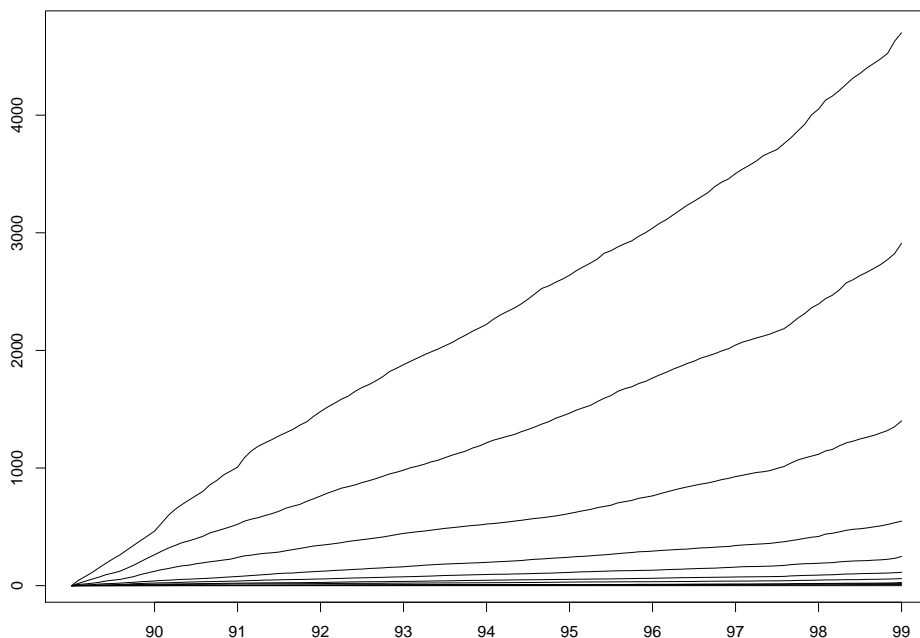
Figure 2: $\Lambda_{\log\mu_{(k)}-\log\mu_{(k+1)}}(t)$, $k = 10, 20, 40, \ldots, 5120$

of $\boldsymbol{\sigma}^2_{k:k+1}$ were estimated similarly to $\boldsymbol{\lambda}_{k,k+1}$. From the $\boldsymbol{\lambda}_{k,k+1}$ and the $\boldsymbol{\sigma}^2_{k:k+1}$, the values of the $\boldsymbol{g}_k$ and the slopes were calculated using (3.8) and (3.12), respectively.

Figure 3 is a log-log plot of the average capital distribution for the subset of the U.S. market comprising the largest 5120 stocks on December 31 of each year from 1990 to 1999. The horizontal coordinate is $\log k$, and the vertical coordinate is $\log\mu_{(k)}(t)$. The broken lines show the tangents at the points corresponding to the values of $k = 10, 20, 40, \ldots, 5120$. The slopes of the tangent lines were calculated using (3.12). There appears to be some local convexity in the curve near $k = 10$, and this may be the cause of the tangent line being slightly below the curve in that neighborhood. In general, the tangent lines appear to be quite accurate, and hence to confirm the efficacy of the methodology developed in Section 3.

From the $2n - 2$ parameters of our model, it is possible to construct a simple "first-order" approximation to the stock price processes that generated the capital distribution curve in Figure 3.

Definition 4.1. Suppose that a market $\mathcal{M}$ has an asymptotically stable capital distribution, and that $\boldsymbol{g}_k$ and $\boldsymbol{\sigma}_{k:k+1}$, for $k = 1, \ldots, n-1$, are the characteristic parameters of the distribution. Let

$$\boldsymbol{\sigma}^2_k = \frac{1}{4}(\boldsymbol{\sigma}^2_{k-1:k} + \boldsymbol{\sigma}^2_{k:k+1}), \quad \text{for } k = 2, \ldots, n-1,$$
$$\boldsymbol{\sigma}^2_1 = \frac{1}{2}\boldsymbol{\sigma}^2_{1:2}, \quad \text{and} \quad \boldsymbol{\sigma}^2_n = \frac{1}{2}\boldsymbol{\sigma}^2_{n-1:n}, \tag{4.2}$$

and let q_t be the inverse of the permutation p_t in (3.1), for $t \in [0, \infty)$. Then the *first-order model* for the market is given by

$$d\log X_i(t) = \boldsymbol{g}_{q_t(i)}\,dt + \boldsymbol{\sigma}_{q_t(i)}\,dV_i(t), \quad t \in [0, \infty), \tag{4.3}$$

for $i = 1, \ldots, n$, where $(V_1, \ldots, V_n)$ is an n-dimensional Brownian motion.
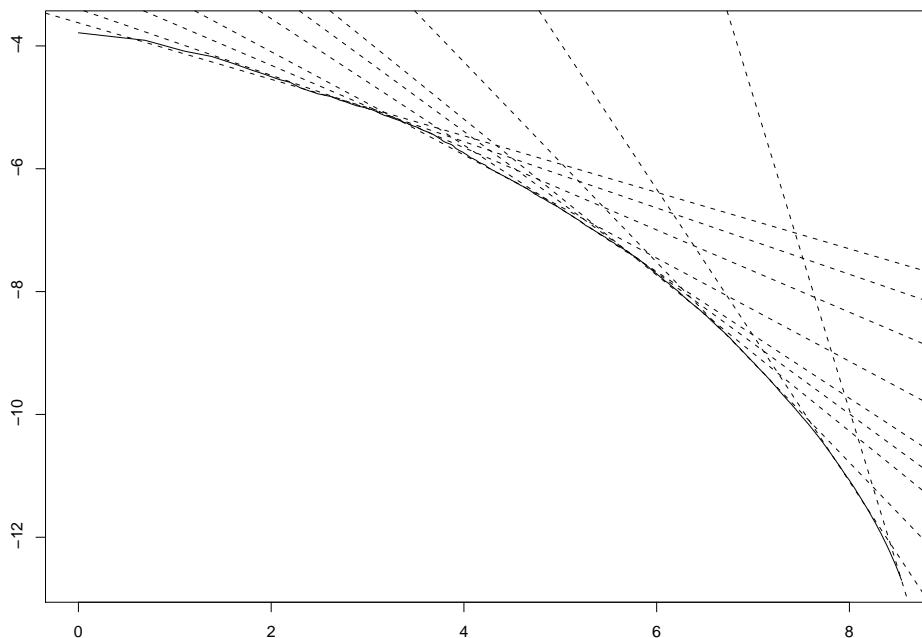
10

Figure 3: $\log \mu_{(k)}$ versus $\log k$, average value 1990 to 1999,
with tangent lines at $k = 10, 20, 40, \ldots, 5120$

Obviously, the first-order market model defined by (4.3) is much simpler that the general market model given by (2.2), for it depends on $2n - 2$ constant parameters rather than $n^2 + n$ stochastic processes. With this first-order model, (3.10) is replaced by

$$d\big(\log \mu_{(k)}(t) - \log \mu_{(k+1)}(t)\big) = -\boldsymbol{\lambda}_{k,k+1}\, dt + d\Lambda_{\log \mu_{(k)} - \log \mu_{(k+1)}}(t) + dN_{k:k+1}(t), \qquad (4.4)$$

for $t \in [0, \infty)$ and $k = 1, \ldots, n-1$, where the $N_{k:k+1}$ are continuous martingales with

$$\langle N_{k:k+1}\rangle_t = \frac{1}{4}\big(\boldsymbol{\sigma}_{k-1:k}^2 + 2\boldsymbol{\sigma}_{k:k+1}^2 + \boldsymbol{\sigma}_{k+1:k+2}^2\big)t, \quad t \in [0, \infty), \quad \text{a.s.,} \qquad (4.5)$$

for $k = 2, \ldots, n-2$. Hence the variance parameters for (4.4) will be a smoothed version of those for (3.10), so we cannot expect the simulated capital distribution to be exactly the same as the actual capital distribution. Moreover, this model approximates the martingales $M_{k:k+1}$ in (3.10) up to the total variation processes $\langle M_{k:k+1}\rangle$, but ignores the cross-variation processes $\langle M_{j:j+1}, M_{k:k+1}\rangle$ for $j \neq k$. The effect of these and other shortcomings of our first-order model will be better understood after we consider a test of the model on actual data.

To test this first-order model, we considered the same 10 years of U.S. equity market data that we used for Figures 2 and 3. We first estimated the $\boldsymbol{\lambda}_{k:k+1}$ in the same manner as we did earlier, and then used (3.8) to calculate the values of the parameters $\boldsymbol{g}_k$. Then we used the market weights and (3.11) to estimate the $\boldsymbol{\sigma}_{k:k+1}$. The smoothed values of these parameters are shown in Figure 4. (The smoothing in Figure 4 was 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 values of the $\boldsymbol{g}_k$, for $k = 1, \ldots, 5119$, are shown by the solid line with the scale on the left. The values of the $\boldsymbol{g}_k$ go from negative for the larger stocks to positive for the smallest stocks. This curve shows considerable

11

structure, which might not be stable over time, so this is an area for possible future investigation. The values of the $\boldsymbol{\sigma}^2_{k:k+1}$ are shown in Figure 4 by the broken line using the scale on the right. This curve is approximately a straight line, so the smoothing effect for the first-order model noted in (4.5) is probably inconsequential in this case.
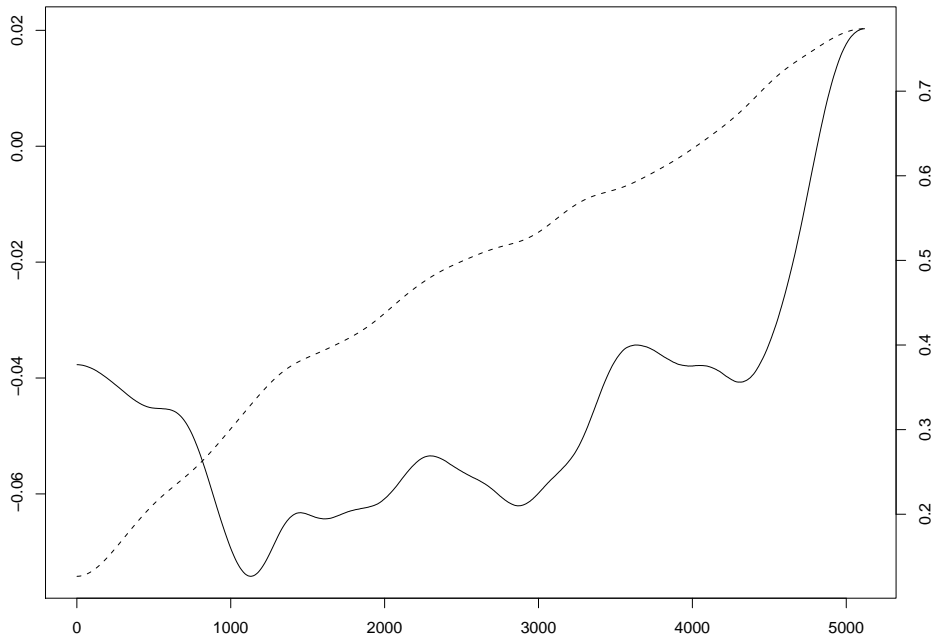

Figure 4: Smoothed values of $\boldsymbol{g}_k$ (solid line, left scale) and $\boldsymbol{\sigma}^2_{k:k+1}$ (broken line, right scale) for $k = 1, \dots, 5119$.

With the first-order model (4.3) based on the estimated characteristic parameters, $\boldsymbol{g}_k$ and $\boldsymbol{\sigma}^2_{k:k+1}$, we simulated 10 years of monthly data, and compared the results of this simulation to the original 10 years of data from which the parameters were estimated. In Figure 5, the upper and lower envelopes of the actual 10 year-end capital distribution curves are shown as solid lines. These envelopes are not themselves capital distribution curves, but rather show at each point on the horizontal axis the maximum and minimum values of the 10 curves at that point. The upper and lower envelopes of the 10 year-end simulated capital distribution curves are shown in Figure 5 as broken lines. (The last few points on each of the simulated curves were deleted because of distortion caused in the simulation by the large value of $\boldsymbol{g}_n$.) As can be seen, the actual curves showed a larger range of values than the simulated curves, particularly for the smaller weights. The average values of the actual curves and the simulated curves are not shown in Figure 5, but were quite close together (mean absolute log-difference $= .017$, as measured on the vertical axis of Figure 3).

Figure 5 suggests that the characteristic parameters $\boldsymbol{g}_k$ and $\boldsymbol{\sigma}^2_{k:k+1}$ vary somewhat over time, and that this variation causes the spread in the actual curves that is not present in the simulated curves. In fact, the actual upper and lower envelopes in Figure 5 conform quite closely to the capital distribution curves generated by raising the average capital distribution over the period to the 1.1 and .9 powers. These two curves cross at the narrow point between the upper and lower envelopes. Equation (3.11) implies that raising the weights to these powers has approximately the same effect as multiplying the parameters $\boldsymbol{g}_k$ by .9 and 1.1, respectively. This level of variation appears to be consistent with the observed deviation from straight lines of the local time curves in Figure 2. Hence,
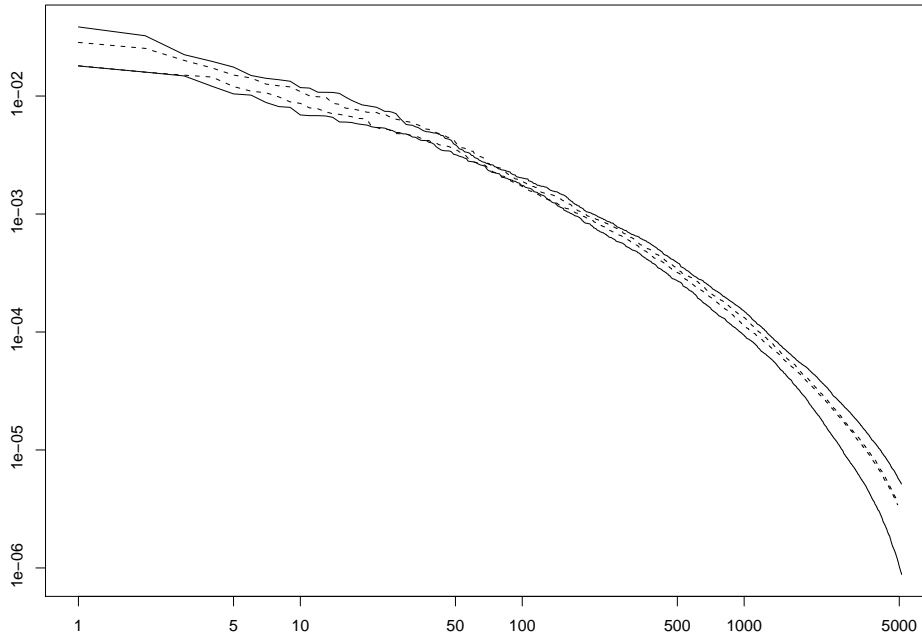
Figure 5: Upper and lower envelopes of the capital distribution curves,
actual data (solid lines), simulated data (broken lines)

while the stable model defined by (4.3) may capture the shape of the capital distribution curve at a given moment, it apparently cannot replicate the observed level of variation of the curve over time.

5   Conclusions

We have developed a model for stable capital distributions in equity markets. For a market of n stocks with a stable capital distribution, the stock price processes can be simulated by processes that depend on $2n - 2$ constant parameters. This model is significantly simpler that the general market model, which depends on $n^2 + n$ random processes. While the stable model and the general model may have the same capital distribution, it is not known to what extent the stable model replicates other behavioral characteristics of the general model.

The stable model was tested on the U.S. equity market over the period from January 1, 1990, to December 31, 1999, and it accurately reflected the short-term structure of the capital distribution. However, the stable model failed to capture the long-term variation that was present in capital distribution of the actual U.S.market.

The results presented here suggest several possibilities for future research: It would be interesting to improve our first-order model by, e.g., taking the cross-variation processes into account. The rank growth rate curve in Figure 4 seems to have considerable structure; is this structure stable, and how does it affect market behavior? It would be interesting to gain more understanding of the long-term variability of the capital distribution of the U.S. market.

A   Appendix: The dynamics of stable families of ranked continuous semimartingales

In this appendix we collect some technical results that we need on stable families of ranked continuous semimartingales. The general model that we derive here is used for the results in Section 3. As we saw in Section 3, we need local times to represent the ranked market weights, and to use local times effectively, we shall need the following regularity condition.

Definition A.1. Let X be a continuous semimartingale with canonical decomposition

$$X(t) = X(0) + M_X(t) + V_X(t), \quad t \in [0, \infty), \quad \text{a.s.,} \tag{A.1}$$

where M_X a continuous local martingale and V_X is a continuous process of bounded variation. We say that X is *absolutely continuous*, if the random signed measures dV_X and $d\langle X \rangle = d\langle M_X \rangle$ are both almost surely absolutely continuous with respect to Lebesgue measure on $[0, \infty)$.

Stock price processes and market weight processes are absolutely continuous semimartingales. Note that the sample paths of absolutely continuous semimartingales are usually not absolutely continuous functions of t: consider, e.g., Brownian motion. The proof of the next lemma is straightforward.

Lemma A.1. *Let X be an absolutely continuous semimartingale such that the set $\{t : X(t) = 0\}$ has Lebesgue measure 0, almost surely. Then*

$$\Lambda_{|X|}(t) = 2\Lambda_X(t), \quad t \in [0, \infty), \quad \text{a.s.}$$

Our development of ranked market weights depends on the following representation of the rank processes $X_{(1)}, \dots, X_{(n)}$, presented in Fernholz (2001).

Proposition A.1. *Let $X_1, \dots, X_n$ be pathwise mutually nondegenerate absolutely continuous semimartingales, and for $t \in [0, \infty)$, let p_t be the random permutation of $\{1, \dots, n\}$ such that for $k = 1, \dots, n$,*

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

Then the rank processes $X_{(k)}$, $k = 1, \dots, n$, are continuous semimartingales such that a.s., for all $t \in [0, \infty)$,

$$dX_{(k)}(t) = \sum_{i=1}^{n} I_{\{0\}}\big(p_t(k) - i\big) dX_i(t) + \frac{1}{2} d\Lambda_{X_{(k)} - X_{(k+1)}}(t) - \frac{1}{2} d\Lambda_{X_{(k-1)} - X_{(k)}}(t). \tag{A.2}$$

Under the hypotheses of Proposition A.1 each X_i has canonical decomposition

$$X_i(t) = V_i(t) + M_i(t), \quad t \in [0, \infty), \quad \text{a.s.,} \tag{A.3}$$

with

$$dV_i(t) = \gamma_i(t) \, dt, \quad t \in [0, \infty), \quad \text{a.s.,}$$

where γ_i is a measurable, adapted process. Let us introduce the notation

$$g_k(t) = \gamma_{p_t(k)}(t), \quad t \in [0, \infty), \tag{A.4}$$

for $k = 1, \ldots, n$. In this case, (A.2) becomes a.s., for all $t \in [0, \infty)$,

$$dX_{(k)}(t) = g_k(t)\, dt + \frac{1}{2}d\Lambda_{X_{(k)}-X_{(k+1)}}(t) - \frac{1}{2}d\Lambda_{X_{(k-1)}-X_{(k)}}(t) + \sum_{i=1}^{n} I_{\{0\}}\big(p_t(k) - i\big)dM_i(t), \quad \text{(A.5)}$$

where the processes M_i are the local martingales in the decomposition (A.3).

We now introduce some asymptotic stability conditions, and develop a model for rank processes that is consistent with these conditions.

Proposition A.2. *Let $X_1, \ldots, X_n$ be pathwise mutually nondegenerate absolutely continuous semimartingales, and suppose that there exist numbers $g, \lambda_{1,2}, \ldots, \lambda_{n-1,n}$ such that:*

i) for $k = 1, \ldots, n$, $\lim_{t\to\infty} t^{-2}\langle X_{(k)}\rangle_t \log\log t = 0$,   a.s.,

ii) for $k = 1, \ldots, n$, $\lim_{t\to\infty} t^{-1}X_{(k)}(t) = g$,   a.s.,

iii) for $k = 1, \ldots, n-1$, $\lim_{t\to\infty} t^{-1}\Lambda_{X_{(k)}-X_{(k+1)}}(t) = \lambda_{k,k+1} > 0$,   a.s.

Then for $k = 1, \ldots, n$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T g_k(t)\, dt = g + \frac{1}{2}\lambda_{k-1,k} - \frac{1}{2}\lambda_{k,k+1} \quad \text{a.s.} \tag{A.6}$$

Proof. Condition (i) and Lemma A.3 imply that (A.6) follows from (A.5) and conditions (ii) and (iii). $\qquad\square$

It is convenient to use the notation

$$g_k = g + \frac{1}{2}\lambda_{k-1,k} - \frac{1}{2}\lambda_{k,k+1}. \tag{A.7}$$

From this it follows that for $k = 1, \ldots, n-1$,

$$\lambda_{k,k+1} = 2kg - 2(g_1 + \cdots + g_k), \tag{A.8}$$

and that

$$g = \frac{1}{n}\sum_{k=1}^{n} g_k.$$

We are interested in the behavior of the differences $X_{(k)}(t) - X_{(k+1)}(t)$, and from (A.2) it follows that a.s., for all $t \in [0, \infty)$,

$$d\big(X_{(k)}(t) - X_{(k+1)}(t)\big) = \big(g_k(t) - g_{k+1}(t)\big)dt - \frac{1}{2}d\Lambda_{X_{(k-1)}-X_{(k)}}(t) - \frac{1}{2}d\Lambda_{X_{(k+1)}-X_{(k+2)}}(t)$$
$$+ d\Lambda_{X_{(k)}-X_{(k+1)}}(t) + \sum_{i=1}^{n}\big(I_{\{0\}}\big(p_t(k) - i\big) - I_{\{0\}}\big(p_t(k+1) - i\big)\big)dM_i(t). \tag{A.9}$$

Suppose that for $k = 1, \ldots, n-1$,

$$\lim_{t\to\infty} t^{-1}\langle X_{(k)} - X_{(k+1)}\rangle_t = \sigma_{k:k+1}^2 > 0, \quad \text{a.s.} \tag{A.10}$$

We can use (A.6) and (A.7) to construct a *stable* version of (A.9) such that a.s., for all $t \in [0, \infty)$,

$$d\big(X_{(k)}(t) - X_{(k+1)}(t)\big) = \big(g_k - g_{k+1} - \frac{1}{2}\lambda_{k-1,k} - \frac{1}{2}\lambda_{k+1,k+2}\big)dt$$
$$+ d\Lambda_{X_{(k)}-X_{(k+1)}}(t) + dM_{k:k+1}(t), \tag{A.11}$$

15

where $M_{k:k+1}$ is a continuous local martingale with $\langle M_{k:k+1} \rangle_t = \sigma_{k:k+1}^2 t$, for $t \in [0, \infty)$, a.s. By (A.7), (A.11) is equivalent to

$$d\big(X_{(k)}(t) - X_{(k+1)}(t)\big) = -\lambda_{k,k+1}\, dt + d\Lambda_{X_{(k)} - X_{(k+1)}}(t) + dM_{k:k+1}(t), \quad t \in [0, \infty), \tag{A.12}$$

for $k = 1, \ldots, n-1$, and this is the model we use to represent a stable capital distribution in Section 3.

To understand the behavior of (A.12), let us first consider the stochastic differential equation

$$dX(t) = -\alpha \operatorname{sgn}(X(t))\, dt + \sigma\, dW(t), \quad t \in [0, \infty), \tag{A.13}$$

where α and σ are positive constants. A solution X to this equation will be an absolutely continuous semimartingale, and (A.13) can be transformed into the simplified version

$$dZ(t) = -\operatorname{sgn}(Z(t))\, dt + dB(t), \quad t \in [0, \infty), \tag{A.14}$$

by the change of variables

$$Z(t) = \frac{\alpha}{\sigma^2} X\Big(\frac{\sigma^2 t}{\alpha^2}\Big),$$

with B the Brownian motion process defined by

$$B(t) = \frac{\alpha}{\sigma} W\Big(\frac{\sigma^2 t}{\alpha^2}\Big).$$

A weak solution for (A.14) is given in Karatzas and Shreve (1991), 5.3.12, along with the probability density of the solution (Karatzas and Shreve (1991), 6.3.5), from which it follows that, asymptotically, the solution X of (A.13) has the exponential density

$$\frac{\alpha}{\sigma^2} \exp\Big(-\frac{2\alpha|u|}{\sigma^2}\Big).$$

With this asymptotic density, it can be shown that for any constant $p > 0$,

$$\lim_{T \to \infty} \frac{X^p(T)}{T} = 0, \quad \text{a.s.} \tag{A.15}$$

The next lemma relates these results to (A.12).

Lemma A.2. *Let X be a solution to*

$$d|X(t)| = d\Lambda_{|X|}(t) - \alpha\, dt + \sigma\, dB(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{A.16}$$

Then

$$\lim_{t \to \infty} t^{-1} \Lambda_{|X|}(t) = \alpha, \quad \text{a.s.}, \tag{A.17}$$

and

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T |X(t)|\, dt = \frac{\sigma^2}{2\alpha}, \quad \text{a.s.} \tag{A.18}$$

Proof. Suppose that X is a (weak) solution to (A.13). Then, by the definition of local time, we have

$$d|X(t)| = 2\, d\Lambda_X(t) - \alpha\, dt + \sigma \operatorname{sgn}(X(t))\, dW(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{A.19}$$

16

The process B defined by

$$B(t) = \int_0^t \text{sgn}(X(s)) \, dW(s), \quad t \in [0, \infty),$$

is a Brownian motion, and (A.19) is equivalent to (A.16) by Lemma A.1. Therefore, a (weak) solution to (A.16) will also be a (weak) solution to (A.13), so (A.15) holds.

From (A.16) we have

$$\lim_{t \to \infty} t^{-1} |X(t)| = \lim_{t \to \infty} t^{-1} \Lambda_{|X|}(t) - \alpha + \sigma \lim_{t \to \infty} t^{-1} B(t), \quad \text{a.s.} \tag{A.20}$$

The first limit in (A.20) is zero by (A.15), and the last limit in (A.20) is zero since we can apply the strong law of large numbers for Brownian motion. This proves (A.17).

Itô's rule applied to (A.13) implies that a.s., for all $t \in [0, \infty)$,

$$dX^2(t) = 2X(t) \, dX(t) + \sigma^2 dt$$
$$= -2\alpha |X(t)| \, dt + \sigma^2 dt + 2\sigma X(t) \, dW(t).$$

Hence, a.s.,

$$\lim_{T \to \infty} \frac{X^2(T)}{T} = -2\alpha \lim_{T \to \infty} \frac{1}{T} \int_0^T |X(t)| \, dt + \sigma^2 + 2\sigma \lim_{T \to \infty} \frac{1}{T} \int_0^T X(t) \, dW(t). \tag{A.21}$$

The first limit in (A.21) is zero by (A.15), and the last limit in (A.21) is zero by (A.15) and Lemma A.3, below. This proves (A.18). □

Equation (A.16) shows that the movement of $|X(t)|$ in (A.16) determined by three components: a reflective impulse $d\Lambda_{|X|}(t)$ at the origin; a compressive drift $-\alpha \, dt$; and a random vibration $\sigma \, dB(t)$. Equation (A.17) shows that equilibrium exists between the reflective impulse and the compressive drift. Equation (A.18) shows that the time average of $|X(t)|$ is proportional to the variance of the random vibration, and inversely proportional to the compressive drift.

Equation (A.12) is of the form of (A.16) with $\alpha = \boldsymbol{\lambda}_{k,k+1}$ and $\sigma^2 = \boldsymbol{\sigma}^2_{k:k+1}$, so Lemma A.2 implies that (A.12) is consistent with condition (*iii*) of Proposition A.2. Lemma A.2 also implies that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \left(X_{(k)}(t) - X_{(k+1)}(t) \right) dt = \frac{\boldsymbol{\sigma}^2_{k:k+1}}{2\boldsymbol{\lambda}_{k,k+1}}, \quad \text{a.s.}, \tag{A.22}$$

and this is used in (3.11) to determine the shape of a stable capital distribution curve.

The final lemma we prove is of the form of a strong law of large numbers, and is used in the proof of Lemma A.2 as well as Proposition 2.1.

Lemma A.3. *Let M be a continuous local martingale such that*

$$\lim_{t \to \infty} t^{-2} \langle M \rangle_t \log \log t = 0, \quad \text{a.s.} \tag{A.23}$$

Then

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

Proof. By modifying the measure space Ω if necessary, we can construct a one-dimensional Brownian motion W_0 independent of M, and then we can define

$$M_0(t) = M(t) + W_0(t), \quad t \in [0, \infty).$$

17

Then M_0 is a continuous local martingale with

$$\langle M_0 \rangle_t = \langle M \rangle_t + t, \quad t \in [0, \infty), \quad \text{a.s.,} \tag{A.24}$$

and (A.23) and (A.24) imply that

$$\lim_{t \to \infty} t^{-2} \langle M_0 \rangle_t \log \log t = 0, \quad \text{a.s.} \tag{A.25}$$

From (A.24) we see that

$$\lim_{t \to \infty} \langle M_0 \rangle_t = \infty, \quad \text{a.s.,} \tag{A.26}$$

so the time change theorem for local martingales (Karatzas and Shreve (1991), 3.4.6) can be applied to show that there exists a Brownian motion B such that

$$B(\langle M_0 \rangle_t) = M_0(t), \quad t \in [0, \infty), \quad \text{a.s.} \tag{A.27}$$

Due to (A.26), we can apply the law of the iterated logarithm for Brownian motion (Karatzas and Shreve (1991), 2.9.23), which, along with (A.27), implies that

$$\limsup_{t \to \infty} \frac{|M_0(t)|}{\sqrt{2 \langle M_0 \rangle_t \log \log \langle M_0 \rangle_t}} = 1, \quad \text{a.s.} \tag{A.28}$$

From (A.25) it follows that

$$\lim_{t \to \infty} t^{-2} \langle M_0 \rangle_t \log \log \langle M_0 \rangle_t = 0, \quad \text{a.s.,}$$

and this and (A.28) imply that $\lim_{t \to \infty} t^{-1} M_0(t) = 0$, a.s. Since the strong law of large numbers for Brownian motion (Karatzas and Shreve (1991), 2.9.3) implies that $\lim_{t \to \infty} t^{-1} W_0(t) = 0$, a.s., the proposition follows. $\qquad \square$

References

Axtell, R. (1999, June). The emergence of firms in a population of agents: local increasing returns, unstable Nash equilibria, and power law size distributions. Technical report, Center on Social and Economic Dynamics, Brookings Institution.

Bak, P. (1996). *How Nature Works*. New York: Springer-Verlag.

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

Fernholz, R. (1999). On the diversity of equity markets. *Journal of Mathematical Economics 31* (3), 393–417.

Fernholz, R. (2001). Equity portfolios generated by functions of ranked market weights. *Finance and Stochastics*, to appear.

Gibrat, R. (1931). *Les Inégalités Économiques*. Paris: Sirey.

Grandmont, J. and W. Hildenbrand (1974). Stochastic processes of temporary equilibria. *Journal of Mathematical Economics 1*, 247–278.

Hashemi, F. (2000). An evolutionary model of the size distribution of firms. *Journal of Evolutionary Economics 10*, 507–521.

Hellwig, M. (1980). Stochastic processes of temporary equilibria, a note. *Journal of Mathematical Economics 7*, 287–299.

Hopenhayn, H. (1992). Entry, exit, and firm dynamics in long-run equilibrium. *Econometrica 60*, 1127–1150.

Ijiri, Y. and H. Simon (1974). Interpretations of departures from the Pareto curve firm-size distributions. *Journal of Political Economy 82*, 315–331.

Ijiri, Y. and H. Simon (1977). *Skew Distributions and the Sizes of Business Firms*. Amsterdam: North-Holland.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica 50*, 649–670.

Karatzas, I. and S. Shreve (1991). *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

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

Simon, H. (1955). On a class of skew distribution functions. *Biometrica 52*, 425–440.

Simon, H. and C. Bonini (1958). The size distribution of business firms. *American Economic Review 48*, 607–617.

Steindl, J. (1965). *Random Processes and the Growth of Firms: A Study of the Pareto Law*. New York: Hafner.