The Annals of Applied Probability
2016, Vol. 26, No. 3, 1329–1361
DOI: 10.1214/15-AAP1118

DIVERSE MARKET MODELS OF COMPETING BROWNIAN PARTICLES WITH SPLITS AND MERGERS[1]

BY IOANNIS KARATZAS AND ANDREY SARANTSEV

Columbia University and University of Washington

We study models of regulatory breakup, in the spirit of Strong and Fouque [*Ann. Finance* **7** (2011) 349–374] but with a fluctuating number of companies. An important class of market models is based on systems of competing Brownian particles: each company has a capitalization whose logarithm behaves as a Brownian motion with drift and diffusion coefficients depending on its current rank. We study such models with a fluctuating number of companies: If at some moment the share of the total market capitalization of a company reaches a fixed level, then the company is split into two parts of random size. Companies are also allowed to merge, when an exponential clock rings. We find conditions under which this system is nonexplosive (i.e., the number of companies remains finite at all times) and diverse, yet does not admit arbitrage opportunities.

1. Introduction.

Stochastic Portfolio Theory (SPT) is a fairly recently developed area of mathematical finance. It tries to describe and understand characteristics of large, real-world equity markets using an appropriate stochastic framework, and to analyze this framework mathematically. It was introduced by Fernholz in the late 1990s, and was developed fully in his book [8]; a survey of somewhat more recent developments appeared in [12].

One feature of real-world markets that this theory tries to account for is diversity. A market is called *diverse*, if at no time is a single stock allowed to dominate almost the entire market in terms of capitalization. To be a bit more precise, let us define the *market weight* of a certain company as the ratio of its capitalization (stock price, times the number of shares outstanding) to the total capitalization of the entire market, across all companies. If no market weight ever exceeds a certain threshold, a fixed number between zero and one, then this market model is called *diverse*.

Such models have one very important feature: with a fixed number of companies and a strictly nondegenerate covariance structure, they allow arbitrage on certain fixed, finite time-horizons. The market portfolio can be outperformed in

Received April 2014; revised February 2015.

[1]Supported in part by NSF Grants DMS-09-05754, DMS-10-07563, DMS-13-08340 and DMS-14-05210.

MSC2010 subject classifications. Primary 60K35, 60J60; secondary 91B26.

Key words and phrases. Competing Brownian particles, splits, mergers, diverse markets, arbitrage opportunity, portfolio.

these models using fully invested, long-only portfolios. This was shown in [8], Chapter 3; further examples of portfolios outperforming the market are given in [11, 13], [12], Section 11. Some such models were constructed in [13, 24, 27, 28] and [12], Chapter 9; see also the related articles [1, 22].

Another feature of large equity markets that SPT tries to capture is that stocks with larger capitalizations tend to have smaller growth rates and smaller volatilities. In an attempt to model this phenomenon, the authors of [2] introduced a new model of *Competing Brownian Particles* (CBPs). Imagine a fixed, finite number of particles moving on the real line; at each time, they are ranked from top to bottom, and each of them undergoes Brownian motion with drift and diffusion coefficients depending on its current rank. From these random motions, one constructs a market model with a finite number of stocks: the logarithms of the companies' capitalizations evolve as a system of CBPs. Recently, these systems were studied extensively (see [6, 9, 14–18, 25, 26, 31]) and were generalized in several directions: [10, 18, 20, 29, 30, 32]. However, these market models are *not* diverse; see [2], Section 7 and Remark 8 below.

We would like to alter the CBP-type model a bit, in order to make it diverse. In real equity markets, diversity is in large part a consequence of anti-monopolistic legislation and regulation: when a company becomes dominant, a governmental agency (the "regulator") forcibly splits it into smaller companies. We implement this idea in our model.

In this paper, we construct a diverse model from the above CBP-based one. We fix a certain *threshold* between 0 and 1. When a company's market weight reaches this threshold, the regulatory agency enforces a breakup of the company into two (random) parts. We also allow for the opposite phenomenon: companies can merge at random times.

The mechanism for merging companies is as follows: immediately after a split or merger, we set an exponential clock whose rate depends on the number of extant companies. If the clock rings before any market weight has hit the threshold, the regulatory agency picks two companies at random as candidates for a possible merger, according to a certain rule described right below. If the planned action results in a company with market weight exceeding the threshold, then this putative merger is suppressed; otherwise, it is allowed to proceed.

We use the following rule for mergers: The company which currently occupies the highest capitalization rank is excluded from consideration, and two of the remaining $N - 1$ companies are chosen randomly, according to the uniform distribution over the $\binom{N-1}{2}$ possible choices. With this rule, and with a threshold sufficiently close to 1, the merger will always be allowed to proceed. In this manner, the process of capitalizations evolves as an exponentiated system of competing Brownian particles, until either (i) one of the market weights hits the threshold, or (ii) the exponential clock rings. In case (i), the number of companies will increase by one; in case (ii), it will decrease by one.

We refer the reader to the very interesting paper [34], which considers general (i.e., not just CBP-based) equity market models of regulatory breakup with a split when a market weight reaches the given threshold. Mergers in that paper obey a different rule than they do here: at the moment of any split, there is a simultaneous merger of the two smallest companies, so the total number of companies remains constant. We feel that this feature is a bit too restrictive, so in the model developed here mergers are allowed to happen independently of splits. This comes at a price, which is both "technical" and substantive: the number of companies in the model is now fluctuating randomly, in ways that need to be understood before any reasonable analysis can go through. The foundational theory for generic market models with a randomly varying number of stocks was developed by Strong in the important and very useful article [33].

It is also important to stress that the mechanisms enforcing diversity in the model studied here are quite different from those used in [13] or [12]. In those papers, the number of companies is fixed and strong repulsive drifts are imposed, in order to keep the configuration of market weights from reaching certain regions of the unit simplex; the resulting market weights, however, have continuous paths. Here, by contrast, the number of companies fluctuates due to breakups and mergers; and the resulting market weights exhibit discontinuities at such "event times." These differences have a rather drastic effect: relative arbitrage, which *does* exist with respect to the market portfolio in [13] and [12], is proscribed here.

1.1. *Preview.* The main results of this paper are as follows. First, we show that under certain conditions the process that counts the number of companies is *nonexplosive*: this number does not become infinite in finite time, so the model can be defined on infinite time horizons. Second, this model turns out to admit an equivalent martingale measure by means of a suitable Girsanov transformation: *although diverse, the model proscribes arbitrage.* This is in contrast with the models from [12], where splits/mergers are not allowed. Indeed, it was observed in [33] that in the presence of splits/mergers, diversity might not lead to arbitrage; in [34], Strong and Fouque established this for their models with a fixed number of companies. We establish the same result for our model, which allows the number of extant companies to fluctuate randomly.

The paper is organized as follows. Section 2 provides an informal yet somewhat detailed description of this model, and states the main results. Section 3 lays out the formal construction of the model. Section 4 is devoted to the proofs of our results. The Appendix develops a crucial technical result.

2. Informal construction and main results.

2.1. *Description of the model.* Consider a stock market with a variable number of companies

$$X(\cdot) = \{X(t), 0 \le t < \infty\}, \qquad X(t) = \big(X_1(t), \dots, X_{\mathcal{N}(t)}(t)\big)',$$

where $X_i(t) > 0$ is the capitalization of the company i at time $t \geq 0$, and $\mathcal{N}(t)$ is the number of companies in the market at that time. The integer-valued random function $t \mapsto \mathcal{N}(t)$ will be piecewise constant; we shall call it the *counting process* of our model, as it records the number of companies that are extant at any given time.

At each interval of constancy of this process, the logarithms $Y_i(\cdot) = \log X_i(\cdot)$, $i = 1, \dots, N$ behave like a system of Competing Brownian Particles (CBPs) with rank-dependent drifts and variances. More precisely, the kth largest among the N real-valued processes $Y_1(\cdot), \dots, Y_N(\cdot)$ behaves like a Brownian motion with local drift g_{Nk} and local variance σ_{Nk}^2, for $k = 1, \dots, N$. These g_{Nk} and $\sigma_{Nk} > 0$ with $N \geq 2$, $1 \leq k \leq N$, are given real constants. If two or more particles occupy the same position at the same time, then we break the tie and assign ranks according to the lexicographic order; more on this in Section 3. We call this model (with constant number of stocks) a *CBP-based model*.

When the *market weight*

$$\mu_i(t) = \frac{X_i(t)}{\mathcal{C}(t)}, \qquad \mathcal{C}(t) := X_1(t) + \cdots + X_{\mathcal{N}(t)}(t) \tag{1}$$

of some company $i = 1, \dots, \mathcal{N}(t)$ reaches a given, fixed threshold $1 - \delta$, a governmental regulatory agency splits this company into two new companies; one with capitalization $\xi X_i(t)$, and the other with capitalization $(1 - \xi)X_i(t)$. Here, the random variable ξ is independent of everything that has happened in the past, and has a given probability distribution F supported on $[1/2, 1)$; whereas $\delta \in (0, 1/2)$ is a given constant.

In addition, for every integer $N \geq 3$ there is an exponential clock with rate $\lambda_N \geq 0$ (a rate of zero means that the clock never rings); we take formally $\lambda_2 = 0$, cf. Remark 1 below. When this clock rings, two companies are chosen at random, as candidates to be merged and form one new company. The choice is made according to a certain probability distribution $\mathcal{P}_N(X(t))$ on the family of subsets of $\{1, \dots, N\}$ which contain exactly two elements, and this distribution depends on the current state $X(t)$ of the system. (One example of such dependence is given below, in Assumption 4; additional clarification is provided in Section 3.1.) If the so-amalgamated company has market weight larger than or equal to $1 - \delta$, the putative merger is suppressed; otherwise, the merger is allowed to proceed.

Within the framework of the model thus described in an informal way, and more formally in Section 3 below, we raise and answer the following questions:

(i) Are there explosions in this model (i.e., can the number of companies become infinite in finite time) with positive probability? Can this model be defined on an infinite time-horizon?

(ii) What is the concept of a portfolio in this model? Does the model admit (relative) arbitrage?

The answers are described in Theorems 2.1 and 2.2 below.

REMARK 1. We note that this model is free of *implosions*, by its construction: when there are only two companies, their putative merger would result in a company with market weight equal to 1 and would thus be suppressed. This is the reason we took at the outset $\lambda_2 = 0$, meaning that with only two companies present the merger clock never rings. As a result, at any given moment there are at least two companies in the equity market model under consideration; and we need not specify the rule for picking companies when there are only two of them, $N = 2$.

2.2. *Portfolios and wealth processes.* In the context of the above model, a *portfolio* is a process

$$\pi(\cdot) = \{\pi(t), 0 \le t < \infty\}, \qquad \pi(t) = \big(\pi_1(t), \dots, \pi_{\mathcal{N}(t)}(t)\big)'$$

for which there exists some real constant $K_\pi \ge 0$ such that $|\pi_i(t)| \le K_\pi$ holds for all $0 \le t < \infty$ and $i = 1, \dots, \mathcal{N}(t)$. The quantity $\pi_i(t)$ is called the *portfolio weight* assigned at time t by the portfolio $\pi(\cdot)$ to the company i; whereas

$$\pi_0(t) := 1 - \sum_{i=1}^{\mathcal{N}(t)} \pi_i(t), \qquad 0 \le t < \infty \tag{2}$$

represents the proportion of wealth invested at time t in a money market with zero interest rate. A portfolio is called *fully invested*, if it never touches the money market, that is, if $\pi_0(\cdot) \equiv 0$; it is called *long-only*, if $\pi_i(t) \ge 0$ holds for all $i = 0, 1, \dots, \mathcal{N}(t), 0 \le t < \infty$.

The prototypical fully invested, long-only portfolio is the *market portfolio* $\pi(\cdot) \equiv \mu(\cdot)$ of (1). At the other extreme stands the *cash portfolio* $\pi(\cdot) \equiv \kappa(\cdot)$ with $\kappa_i(t) = 0$ for all $i = 1, \dots, \mathcal{N}(t), 0 \le t < \infty$; this never touches the equity market, and keeps all wealth in cash at all times.

When the counting process $\mathcal{N}(\cdot)$ jumps up (after a split) or down (after a merger), the portfolio process behaves as follows:

(i) if two companies merge, the portfolio weight corresponding to the new company's stock is the sum of the portfolio weights corresponding to the two old stocks; whereas

(ii) if a company gets split into two new ones, its weight in the portfolio is partitioned in proportion to the weights of the newly minted companies.

The formal description of these actions is postponed to Section 3.

Suppose now that a small investor, whose actions cannot influence asset prices, starts with initial capital \$1 and invests in the stock market according to a portfolio rule $\pi(\cdot)$. The corresponding *wealth process* $V^\pi(\cdot) = \{V^\pi(t), 0 \le t < \infty\}$ takes then values in $(0, \infty)$, satisfies

$$\frac{\mathrm{d}V^\pi(t)}{V^\pi(t)} = \sum_{i=1}^{\mathcal{N}(t)} \pi_i(t) \frac{\mathrm{d}X_i(t)}{X_i(t)}, \tag{3}$$

and is not affected when the number of companies changes, that is, when the counting process $\mathcal{N}(\cdot)$ jumps. For a derivation of (3) with a fixed number of companies, see, for instance, [8], page 6.

DEFINITION 1 (Relative arbitrage). We say that a given portfolio $\pi(\cdot)$ *represents an arbitrage opportunity relative to another portfolio* $\rho(\cdot)$ over the time horizon $[0, T]$, for some real number $T > 0$, if we have

$$\mathbf{P}\big(V^{\pi}(T) \geq V^{\rho}(T)\big) = 1, \qquad \mathbf{P}\big(V^{\pi}(T) > V^{\rho}(T)\big) > 0. \tag{4}$$

In words: over the time-horizon $[0, T]$, the portfolio $\pi(\cdot)$ performs at least as well as $\rho(\cdot)$ with probability one, and strictly better with positive probability. When $\rho(\cdot) \equiv \kappa(\cdot)$ is the cash portfolio, this reduces to the usual definition of arbitrage.

2.3. *Main results*. Let us impose some conditions on the parameters of this model. A salient feature of real-world markets is that stocks with smaller market weights tend to have larger drift coefficients (growth rates), so it is not unreasonable to impose the following condition.

ASSUMPTION 1.

$$g_{N1} \leq \min_{2 \leq k \leq N} g_{Nk} \qquad \text{for every } N \geq 2.$$

We shall also impose the following conditions: there exist constants $\overline{\sigma}$, $\underline{\sigma}$, $\overline{g}$ such that we have the following.

ASSUMPTION 2. We have $\delta \in (0, 1/6)$, as well as

$$0 < \underline{\sigma} \leq \sigma_{Nk} \leq \overline{\sigma} < \infty, \qquad |g_{Nk}| \leq \overline{g} \qquad \text{for all } N \geq 2 \text{ and } k = 1, \dots, N.$$

ASSUMPTION 3. The probability distribution F of the random variable ξ responsible for splitting companies is supported on the interval $[1/2, 1 - \varepsilon_0]$, where $\varepsilon_0 \in (0, 1/2)$. In other words,

$$\xi \backsim F \quad \Rightarrow \quad 1/2 \leq \operatorname{ess\,inf} \xi \leq \operatorname{ess\,sup} \xi \leq 1 - \varepsilon_0. \tag{5}$$

ASSUMPTION 4. The rule for picking companies to be merged is as follows: With $N \geq 3$, we exclude the company which occupies the highest rank in terms of capitalization and choose at random two of the remaining $N - 1$ companies according to the uniform distribution over the

$$m_N = \binom{N-1}{2} \tag{6}$$

possible such choices. If two or more companies are tied in terms of capitalization, we resolve the tie *lexicographically*, that is, always in favor of the lowest index i.

REMARK 2. Under Assumptions 2 and 4, mergers are never suppressed. Indeed, of the chosen companies, the one with the biggest market weight will occupy the second place at best, so its market weight will be no more than $1/2$; whereas the other will occupy the third place at best, so its market weight will not exceed $1/3$. Therefore, the market weight of the amalgamated company will not exceed $5/6$, a number smaller than $1 - \delta$ because we have $\delta < 1/6$ from Assumption 2, *and so the merger will not be suppressed.* Moreover, *all* of the new market weights will be bounded away from the threshold $1 - \delta$, so it will take some time for *any* company extant after the merger to hit this threshold.

ASSUMPTION 5. The rates of the exponential clocks satisfy, for some real constants c and $\alpha > 0$,

$$\lambda_N \asymp cN^\alpha, \qquad N \to \infty. \tag{7}$$

REMARK 3. This condition is perhaps the most significant one: it ensures that mergers happen with sufficient intensity, so that the number of companies in the model will not only not become infinite in a finite amount of time, but will also have a "tame" temporal growth (cf. Proposition 4.1 below).

As an illustration for Condition (7), suppose there are N companies; then, according to the rules of Assumption 4, there are m_N such possible mergers as in (6). If each pair of companies has its own merger exponential clock Ξ_i with the same parameter λ, and if $\Xi_1, \ldots, \Xi_{m_N}$ are independent, then the earliest merger will happen at the smallest of those exponential clocks; but

$$\min(\Xi_1, \ldots, \Xi_{m_N}) \backsim \mathcal{E}(m_N \lambda),$$

so $\lambda_N = m_N \lambda \asymp N^2$ as $N \to \infty$. That is, the requirement (7) holds in this case with $\alpha = 2$.

The following two theorems are our main results. They are proved in Section 4.

THEOREM 2.1. *Under Assumptions 1–5, the above market model is free of explosions and can thus be defined on an infinite time-horizon.*

THEOREM 2.2. *Under the Assumptions 1–5, no relative arbitrage is possible over any given time horizon* $[0, T]$ *of finite length.*

3. Formal construction.

3.1. *Notation.* We let $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$. For every integer $N \geq 2$, $\delta \in (0, 1)$, we let

$$\Delta_+^N := \{(z_1, \ldots, z_N) \in \mathbb{R}^N | z_1 > 0, \ldots, z_N > 0, z_1 + \cdots + z_N = 1\},$$

$$\Delta_+^{N,\delta} := \{(z_1, \ldots, z_N) \in \Delta_+^N | z_1 \leq 1 - \delta, \ldots, z_N \leq 1 - \delta\}.$$

We also denote

$$
\begin{aligned}
(8) \qquad & \mathcal{S} := \bigcup_{N=2}^{\infty} (0,\infty)^N, \qquad \widetilde{\mathcal{S}} := \bigcup_{N=3}^{\infty} (0,\infty)^N, \\
& \mathcal{M} := \bigcup_{N=2}^{\infty} \Delta_+^N, \qquad \mathcal{M}^{\delta} := \bigcup_{N=2}^{\infty} \Delta_+^{N,\delta}.
\end{aligned}
$$

For $x \in \mathcal{S}$, we denote by $\mathfrak{N}(x)$ the number of components of x, that is, the integer $N \geq 2$ for which $x \in (0,\infty)^N$; and for $N = \mathfrak{N}(x)$, we denote by $\mathfrak{z}(x) \in \Delta_+^N$ the vector with components

$$
\mathfrak{z}_i(x) := \frac{x_i}{x_1 + \cdots + x_N}, \qquad i = 1, \dots, N.
$$

The market-weight process $\mu(\cdot) = \mathfrak{z}(X(\cdot))$ of (1) is said to be *on the level N* at time t, if $\mu(t) \in \Delta_+^N$.

For any given vector $y \in \mathbb{R}^N$, we denote by $y_{(1)} \geq y_{(2)} \geq \cdots \geq y_{(N)}$ its *ranked components*; in this ranking, ties are resolved lexicographically, always in favor of the lowest index as in [2, 18] and in Assumption 4.

We denote by $C^r(A)$ the set of r times continuously differentiable functions $f : A \to \mathbb{R}$.

For every integer $N \geq 3$, we denote by $\mathcal{R}_N$ the family of subsets of $\{1, \dots, N\}$ which contain exactly two elements. We denote by $\mathcal{Q}_N$ the collection of all probability distributions on $\mathcal{R}_N$.

REMARK 4. Under the Assumption 4, the probability distributions $\{\mathcal{P}_{\mathfrak{N}(x)}(x)\}_{x \in \widetilde{\mathcal{S}}}$ in Section 2 are constructed thus: For any given $x \in \widetilde{\mathcal{S}}$, we let $N = \mathfrak{N}(x)$, rank lexicographically the components of the vector x, and consider the smallest index $j \in \{1, \dots, N\}$ such that $x_j \geq x_k$ holds for all $k = 1, \dots, N$. Then $\mathcal{P}_N(x) \in \mathcal{Q}_N$ is the uniform distribution on the family of subsets of $\{1, \dots, N\} \setminus \{j\}$ that contain exactly two elements [there are exactly m_N such subsets, as in (6)].

3.2. *Fixed number of companies*. First, let us formally introduce auxiliary CBP-based models with a fixed, finite number of particles. These will serve as building blocks for our ultimate model; in [34], similar preparatory models are referred to as "premodels."

Fix an integer $N \geq 2$, and consider a system of N particles moving on the real line, formally expressed as an $\mathbb{R}^N$-valued process

$$
Y(\cdot) = \{Y(t), 0 \leq t < \infty\}, \qquad Y(t) = (Y_1(t), \dots, Y_N(t))'.
$$

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$, $\mathbb{G} = \{\mathcal{G}(t)\}_{0 \leq t < \infty}$, where the filtration satisfies the *usual conditions* of right-continuity and augmentation by null sets, and let $W(\cdot) = \{W(t), 0 \leq t < \infty\}$ be a standard N-dimensional $(\mathbb{G}, \mathbf{P})$-Brownian motion.

DEFINITION 2. With $g_1, \dots, g_N$ given real numbers, and $\sigma_1, \dots, \sigma_N$ given positive real numbers, a *finite system of CBPs with symmetric collisions* is an $\mathbb{R}^N$-valued process governed by the system of stochastic differential equations

$$\mathrm{d}Y_i(t) = \sum_{k=1}^{N} \mathbf{1}_{\{Y_i(t)=Y_{(k)}(t)\}}\big(g_k\,\mathrm{d}t + \sigma_k\,\mathrm{d}W_i(t)\big), \qquad i = 1, \dots, N, 0 \le t < \infty.$$

Informally, such a model posits that the kth largest particle moves as a one-dimensional Brownian motion with local drift g_k and local variance σ_k^2. We denote the ranked (in decreasing order) statistics for the components of this system as

$$\max_{1\le i\le N} Y_i(\cdot) =: Y_{(1)}(\cdot) \ge Y_{(2)}(\cdot) \ge \cdots \ge Y_{(N)}(\cdot) := \min_{1\le i\le N} Y_i(\cdot). \tag{9}$$

Similarly, $\mu_{(k)}(t)$ is the kth ranked market weight at time t: namely, $\mu_{(1)}(t) \ge \cdots \ge \mu_{(N)}(t)$. We set $\Lambda_{(k,k+1)}(\cdot) = \{\Lambda_{(k,k+1)}(t), 0 \le t < \infty\}$ for the local time accumulated at the origin by the nonnegative semimartingales $Y_{(k)}(\cdot) - Y_{(k+1)}(\cdot) = \{Y_{(k)}(t) - Y_{(k+1)}(t), 0 \le t < \infty\}$ with $k = 1, \dots, N-1$ (for notational convenience, we set also $\Lambda_{(0,1)}(\cdot) \equiv \Lambda_{(N,N+1)}(\cdot) \equiv 0$ for all $t \in [0, \infty)$). Then the equation

$$\mathrm{d}Y_{(k)}(t) = g_k\,\mathrm{d}t + \sigma_k\,\mathrm{d}B_k(t) + \tfrac{1}{2}\,\mathrm{d}\Lambda_{(k,k+1)}(t) - \tfrac{1}{2}\,\mathrm{d}\Lambda_{(k-1,k)}(t), \qquad 0 \le t < \infty \tag{10}$$

describes the dynamics of the ranked semimartingales in (9), where the standard Brownian motions

$$B_k(\cdot) := \sum_{i=1}^{N} \int_0^{\cdot} \mathbf{1}_{\{Y_i(t)=Y_{(k)}(t)\}}\,\mathrm{d}W_i(t), \qquad k = 1, \dots, N \tag{11}$$

are independent by the P. Lévy theorem. We refer to [2], [18], Lemma 1 and [14], Section 3, for the derivation of (10), as well as to [4] for the existence and uniqueness in distribution of a weak solution to the CBP system of Definition 2. As shown in [16], pathwise uniqueness, and thus existence of a strong solution, also hold for this system up until the first time three particles collide—and the latter never happens if the mapping $k \mapsto \sigma_k^2$ is concave (cf. [15, 16, 30, 31]).

The *CBP-based market model* with a fixed number N of companies is defined as a collection of N real-valued, strictly positive stochastic processes $X_i(\cdot) = \{X_i(t), 0 \le t < \infty\}, i = 1, \dots, N$ with $X_i(t) := e^{Y_i(t)}$. The dynamics of these processes are given by

$$\mathrm{d}\log X_i(t) = \sum_{k=1}^{N} \mathbf{1}_{\{X_i(t)=X_{(k)}(t)\}}\big[g_k\,\mathrm{d}t + \sigma_k\,\mathrm{d}W_i(t)\big], \tag{12}$$

or equivalently

$$\frac{\mathrm{d}X_i(t)}{X_i(t)} = \sum_{k=1}^{N} \mathbf{1}_{\{X_i(t)=X_{(k)}(t)\}}\left[\left(g_k + \frac{\sigma_k^2}{2}\right)\mathrm{d}t + \sigma_k\,\mathrm{d}W_i(t)\right]. \tag{13}$$

In this model, the vector process $\mu(\cdot) = (\mu_1(\cdot), \dots, \mu_N(\cdot))' = \mathfrak{z}(X(\cdot))$ of the market weights $\mu_i(t) := X_i(t)/(X_1(t) + \cdots + X_N(t))$ with $i = 1, \dots, N, 0 \le t < \infty$ for its various companies, evolves as an N-dimensional diffusion governed by the system of SDEs

$$\begin{aligned}
\mathrm{d}\log\mu_i(t) = \Bigg[&\sum_{k=1}^{N} g_k \mathbf{1}_{\{\mu_i(t)=\mu_{(k)}(t)\}} - \sum_{k=1}^{N} g_k \mu_{(k)}(t) \\
&- \frac{1}{2}\sum_{k=1}^{N} \sigma_k^2\big(\mu_{(k)}(t) - \mu_{(k)}^2(t)\big)\Bigg]\mathrm{d}t \\
&+ \sum_{k=1}^{N} \sigma_k\big[\mathbf{1}_{\{\mu_i(t)=\mu_{(k)}(t)\}}\,\mathrm{d}W_i(t) \\
&- \mu_{(k)}(t)\sum_{\nu=1}^{N} \mathbf{1}_{\{\mu_\nu(t)=\mu_{(k)}(t)\}}\,\mathrm{d}W_\nu(t)\big], \\
&\qquad i = 1, \dots, N.
\end{aligned} \tag{14}$$

We derive this system from the general expression in equation (2.4) of [12], Section 2. Substituting in that expression the concrete values of drift and covariance coefficients for the CBP-based market model of (12) under consideration, we arrive at the dynamics of (14) for the $\log\mu_i(\cdot)$'s.

REMARK 5. In the terminology of [34], Remark 2, the companies in models of this sort are "generic": The characteristics of their capitalizations' dynamics depend entirely on the ranks the companies occupy in the capitalization hierarchy; they are not idiosyncratic (i.e., name- or sector-dependent).

3.3. *Formal construction of the main model.* Let us begin the formal construction of our model. This will take the form of a process $X(\cdot) = \{X(t), 0 \le t < \infty\}$ on the state-space $\mathcal{S}$ of (8).

For every $N \ge 2$ and every $x = (x_1, \dots, x_N)' \in (0, \infty)^N$, we construct a probability space $(\Omega^{N,x}, \mathcal{F}^{N,x}, \mathbf{P}^{N,x})$ which contains countably many i.i.d. copies $Y^{N,x,n}(\cdot), n \in \mathbb{N}$ of the solution

$$Y^{N,x}(\cdot) = \{Y^{N,x}(t), 0 \le t < \infty\}, \qquad Y^{N,x}(t) = \big(Y_1^{N,x}(t), \dots, Y_N^{N,x}(t)\big)'$$

to the following system of stochastic differential equations:

$$
\begin{aligned}
\mathrm{d}Y_i^{N,x}(t) &= \sum_{k=1}^{N} \mathbf{1}_{\{Y_i^{N,x}(t)=Y_{(k)}^{N,x}(t)\}}\big(g_{Nk}\,\mathrm{d}t + \sigma_{Nk}\,\mathrm{d}W_i(t)\big), \\
&\qquad Y_i^{N,x}(0) = \log x_i
\end{aligned} \tag{15}
$$

for $i = 1, \ldots, N$. Here $W(\cdot) = \{W(t), t \geq 0\}$ is a standard N-dimensional Brownian motion, and the parameters g_{Nk}, σ_{Nk} satisfy the conditions of Assumptions 1 and 2.

For every $N \geq 3$, we fix a collection of probability distributions $\{\mathcal{P}_N(x)\}_{x \in \widetilde{\mathcal{S}}} \subseteq \mathscr{Q}_N$. This specification will provide the rule for choosing two out of the existing $N = \mathfrak{N}(x)$ companies to merge, when the system is in state $x \in \widetilde{\mathcal{S}}$ and an exponential clock rings.

- Consider another probability space $(\Omega', \mathcal{F}', \mathbf{P}')$ which contains:

  (a) countably many countably many i.i.d. copies $\xi_1, \xi_2, \xi_3, \ldots$ of a random variable ξ with given probability distribution F, which is supported on the interval $[1/2, 1 - \varepsilon_0]$;

  (b) for each $N \geq 2$, countably many copies $\eta_1^N, \eta_2^N, \ldots$ of an exponential clock η^N with rate λ_N, if this rate is positive (if the rate is zero, as we assume it is for $N = 2$, we let $\eta_1^N = \eta_2^N = \cdots = \infty$);

  (c) for each $N \geq 2$ and each probability distribution $\mathfrak{p} \in \mathscr{Q}_N$, countably many i.i.d. copies $\zeta_i(N, \mathfrak{p}), i = 1, 2, \ldots$ of a random element $\zeta(N, \mathfrak{p})$, which takes values in $\mathcal{R}_N$ and is distributed according to $\mathfrak{p}$. (Please recall here the notation of Section 3.1.)

- Let $(\Omega, \mathcal{F}, \mathbf{P})$ be the direct product of these probability spaces. Starting from a point $X(0) = x \in \mathcal{S}$, we let $N_0 = \mathfrak{N}(x)$ be the number of companies extant at $t = 0$, and construct a process $X(\cdot) = \{X(t), 0 \leq t < T\}$ and a random time T, the "lifetime" of $X(\cdot)$, as follows:

  Step (i): For $t \leq \tau_1 \wedge \eta_1^{N_0}$, we define the random vector $X(t) = (X_1(t), \ldots, X_{N_0}(t))'$ with values in $(0, \infty)^{N_0}$ as

$$
\begin{aligned}
&X_i(t) := \exp\big(Y_i^{N_0,x,1}(t)\big) \quad \text{and} \\
&\tau_1 := \inf\big\{t \geq 0 \,|\, \exists i = 1, \ldots, N_0 : \mu_i^{N_0,x}(t) = 1 - \delta\big\}
\end{aligned} \tag{16}
$$

  (we adopt here the usual convention $\inf \varnothing = \infty$), where the market weight of the company i is

$$
\begin{aligned}
&\mu_i(t) \equiv \mu_i^{N_0,x}(t) := \frac{X_i(t)}{X_1(t) + \cdots + X_{N_0}(t)}, \\
&\qquad 0 \leq t \leq \tau_1 \wedge \eta_1^{N_0}; i = 1, \ldots, N_0.
\end{aligned} \tag{17}
$$

Since $\delta \in (0, 1/2)$, there can be at most one index i with $\mu_i^{N_0,x}(\tau_1) = 1 - \delta$. Thus, the moment of the first jump, or "event," in this setup, is

$$T_1 := \tau_1 \wedge \eta_1^{N_0}.$$

Step (ii): If $\tau_1 \le \eta_1^{N_0}$, $\tau_1 < \infty$, we pick the unique $i \in \{1, \dots, N_0\}$ such that $\mu_i(\tau_1) = 1 - \delta$, and define the vector $X(\tau_1+) \in (0, \infty)^{N_0+1}$ as follows:

$$X_\nu(\tau_1+) = X_\nu(\tau_1), \qquad \nu = 1, \dots, i-1;$$
$$X_\nu(\tau_1+) = X_{\nu+1}(\tau_1), \qquad \nu = i, \dots, N_0 - 1;$$
$$X_{N_0}(\tau_1+) = \xi_1 X_i(\tau_1), \qquad X_{N_0+1}(\tau_1+) = (1 - \xi_1) X_i(\tau_1).$$

To wit: at the time τ_1 of (16), company i is split into two new companies, anointed with the names N_0 and $N_0 + 1$. These inherit the capitalization $X_i(\tau_1)$ of their progenitor in proportions ξ_1 and $1 - \xi_1$, respectively. Companies $1, \dots, i-1$ keep both their names and their capitalizations; whereas the companies formerly known as $i+1, \dots, N_0$ keep their capitalizations but change their names to $i, \dots, N_0 - 1$, respectively.

Step (iii): If $\tau_1 > \eta_1^{N_0}$, a subset with two elements $\{i, j\} \subseteq \{1, \dots, N_0\}$ is selected according to the random variable $\zeta_1(N_0, \mathcal{P}_{N_0}(X(\eta_1^{N_0})))$ whose distribution is $\mathcal{P}_{N_0}(X(\eta_1^{N_0})) \in \mathscr{Q}_{N_0}$.

On the event $\{\mu_i(\eta_1^{N_0}) + \mu_j(\eta_1^{N_0}) \ge 1 - \delta\}$, we proceed to step (iv), case B below. Otherwise, we define the vector $X(\eta_1^{N_0}+) \in (0, \infty)^{N_0-1}$ as follows, say with $i < j$:

$$X_\nu(\eta_1^{N_0}+) = X_\nu(\eta_1^{N_0}), \qquad \nu = 1, \dots, i-1;$$
$$X_\nu(\eta_1^{N_0}+) = X_{\nu+1}(\eta_1^{N_0}), \qquad \nu = i, \dots, j-2;$$
$$X_\nu(\eta_1^{N_0}+) = X_{\nu+2}(\eta_1^{N_0}), \qquad \nu = j-1, \dots, N_0 - 2;$$
$$X_{N_0-1}(\eta_1^{N_0}+) = X_i(\eta_1^{N_0}) + X_j(\eta_1^{N_0}).$$

Once again, companies $1, \dots, i-1$ keep both their names and their capitalizations. The erstwhile companies $i+1, \dots, j-1$ keep their capitalizations but change their names to $i, \dots, j-2$; whereas the erstwhile companies $j+1, \dots, N_0$ keep their capitalizations but change their names to $j-1, \dots, N_0-2$. The former companies i and j merge; they create a new company, anointed with the index (name) $N_0 - 1$, which inherits the sum total of their capitalizations.

Step (iv): We let $N_1 = \mathcal{N}(T_1+)$ be the new number of companies extant right after time $T_1 = \tau_1 \wedge \eta_1^{N_0}$, and note that there are three possibilities:

Case A: $N_1 = N_0 + 1$ on the event $\{\tau_1 \le \eta_1^{N_0}, \tau_1 < \infty\}$ of a split;

Case B: $N_1 = N_0$ on the event $\{\tau_1 > \eta_1^{N_0}, \mu_i(\eta_1^{N_0}) + \mu_j(\eta_1^{N_0}) \ge 1 - \delta\}$ of a "suppressed" merger;

Case C: $N_1 = N_0 - 1$ on the event $\{\tau_1 > \eta_1^{N_0}, \mu_i(\eta_1^{N_0}) + \mu_j(\eta_1^{N_0}) < 1 - \delta\}$ of a "successful" merger.

We define

$$X_i(t) := \exp\bigl(Y_i^{N_1,x_1,2}(t - T_1)\bigr) \qquad \text{for } T_1 < t \le T_1 + \bigl(\tau_2 \wedge \eta_2^{N_1}\bigr). \tag{18}$$

Here, $T_1 = \tau_1 \wedge \eta_1^{N_0}$, $x_1 = X(T_1)$, and

$$\tau_2 := \inf\bigl\{t > 0 | \exists i = 1, \dots, N_1 : \mu_i^{N_1,x_1}(T_1 + t) = 1 - \delta\bigr\},$$

where $\mu_i^{N_1,x_1}(t)$ is defined by analogy with (17), in terms of the capitalizations in (18), as

$$\mu_i(t) \equiv \mu_i^{N_1,x_1}(t) := \frac{X_i(t)}{X_1(t) + \cdots + X_{N_1}(t)},$$
$$T_1 < t \le T_1 + \bigl(\tau_2 \wedge \eta_2^{N_1}\bigr), i = 1, \dots, N_1.$$

Thus, the time of the second jump in the integer-valued process $\mathcal{N}(\cdot)$ is

$$T_2 := T_1 + \bigl(\tau_2 \wedge \eta_2^{N_1}\bigr).$$

Step (v): We similarly define the values of the capitalization processes after the second jump. On each next step, we use new independent copies of variables η_i^N and ξ_i. [If this jump corresponds to a merger, then we choose two companies to be merged according to the distribution $\zeta_2(N_1, \mathcal{P}_{N_1}(X(\eta_2^{N_1})))$.] Then we define their evolution until the moment T_3 of the third jump, etc.

REMARK 6. We already saw in Remark 2 that the specification of the probability distributions $\{\mathcal{P}_{\mathfrak{N}(x)}(x)\}_{x \in \widetilde{\mathcal{S}}}$ as in Assumption 4 and Remark 4 guarantees that, with three or more companies present, no merger is ever suppressed [i.e., that case B in (iv) above never occurs].

Step (vi): This construction leads to a strong Markov process $X(\cdot) = \{X(t), 0 \le t < T_\star\}$ with state space $\mathcal{S}$ and piecewise-continuous, LCRL (Left-Continuous with Right Limits) paths, defined on the time interval $[0, T_\star)$ with

$$T_\star := \lim_{m \to \infty} \uparrow T_m. \tag{19}$$

The resulting market-weight process $\mu(\cdot) = \mathfrak{z}(X(\cdot))$ of (1) has state space $\mathcal{M}^\delta$ as in (8). In particular, $\max_{1 \le i \le \mathcal{N}(t)} \mu_i(t) \le 1 - \delta$ holds for all $t \in [0, \infty)$, so the resulting market model is *diverse* in the terminology of [8], Chapter 2. We also note that, in all cases of the above construction, the total capitalization $\mathcal{C}(\cdot)$ in (1) is preserved at each "event-time" T_m, namely $\mathcal{C}(T_m+) = \mathcal{C}(T_m)$, $\forall m \in \mathbb{N}$.

DEFINITION 3. We say that the so-constructed model *admits explosions*, if $\mathbf{P}(T_\star = \infty) < 1$. Otherwise, the model is said to be *free of explosions*.

Theorem 2.1 guarantees that $\mathbf{P}(T_\star = \infty) = 1$ holds under Assumptions 1–5. In the absence of explosions, the process $X(\cdot)$ is defined on all of $[0, \infty)$, and we denote by $\mathbb{F} = \{\mathcal{F}(t)\}_{0 \le t < \infty}$ the smallest filtration to which $X(\cdot)$ is adapted and which satisfies the usual conditions.

REMARK 7. In addition to being diverse, the model just constructed has intrinsic relative variance (equivalently, excess growth rate for market portfolio) which is bounded away from zero, namely

$$\gamma_*^\mu(t) := \frac{1}{2}\left(\sum_{k=1}^N \sigma_{Nk}^2 \mu_{(k)}(t)\big(1 - \mu_{(k)}(t)\big)\right)\bigg|_{N=\mathcal{N}(t)} \ge (\sigma_0^2 \delta)/2 > 0; \qquad 0 \le t < \infty.$$

We owe this observation to Dr. Robert Fernholz. See Proposition 3.1 in [11], or Example 11.1 in [12], for the significance of such a positive lower bound in the context of arbitrage relative to the market portfolio with a fixed number of companies.

3.4. *Portfolios and associated wealth processes*. Let us discuss portfolios and the wealth processes they generate. A *portfolio* $\pi(\cdot)$ is an $\mathbb{F}$-progressively measurable process $\pi(\cdot) = \{\pi(t), 0 \le t < \infty\}$, $\pi(t) = (\pi_1(t), \dots, \pi_{\mathcal{N}(t)}(t))'$ for which there exists some real constant $K_\pi \ge 0$ such that, almost surely: $|\pi_i(t)| \le K_\pi$ holds for all $0 \le t < \infty$ and $i = 1, \dots, \mathcal{N}(t)$.

When the integer-valued process $\mathcal{N}(\cdot)$ suffers a downward or upward jump, this portfolio must behave according to the rules described informally in Section 2. We formalize these rules presently.

(A) Assume that at time t, the ith and jth companies merge into one company; this is then anointed with index $N - 1$, where $N = \mathcal{N}(t)$ is the number of companies immediately before the merger. The new portfolio weights are

$$\pi_\nu(t+) = \pi_\nu(t), \qquad \nu = 1, \dots, i-1;$$
$$\pi_\nu(t+) = \pi_{\nu+1}(t), \qquad \nu = i, \dots, j-2;$$
$$\pi_\nu(t+) = \pi_{\nu+2}(t), \qquad \nu = j-1, \dots, N-2; \qquad \pi_{N-1}(t+) = \pi_i(t) + \pi_j(t).$$

In words: companies not involved in the merger are assigned the same portfolio weights, under their new appellations if necessary; whereas the newly minted company $N - 1$ inherits the sum of the portfolio weights formerly assigned to its two parent companies.

(B) Assume that at time t the ith company, with capitalization $X_i(t)$, is split into two companies [anointed with the indices N and $N + 1$, where $N = \mathcal{N}(t)$ is the number of companies immediately before the split]. The new portfolio weights are

$$\pi_\nu(t+) = \pi_\nu(t), \qquad \nu = 1, \dots, i-1;$$

$$\pi_\nu(t+) = \pi_{\nu+1}(t), \qquad \nu = i, \dots, N-1;$$

$$\pi_N(t+) = \pi_i(t)\frac{X_N(t+)}{X_i(t)}, \qquad \pi_{N+1}(t+) = \pi_i(t)\frac{X_{N+1}(t+)}{X_i(t)}.$$

Once again, companies not involved in the split keep their weights in the portfolio, under their new appellations if necessary; whereas each of the two newly created companies N and $N+1$ inherits the weight in the portfolio of the parent company, in proportion to its currently assigned capitalization.

The corresponding wealth process $V^\pi(\cdot) = \{V^\pi(t), 0 \le t < \infty\}$ is continuous: *it does not suffer a jump when a split or a merger happen.* It is $\mathbb{F}$-adapted, takes values in $(0, \infty)$, and is governed for each integer $m \in \mathbb{N}_0$ by the dynamics

$$\frac{\mathrm{d}V^\pi(t)}{V^\pi(t)} = \sum_{i=1}^{N_m} \pi_i(t)\frac{\mathrm{d}X_i(t)}{X_i(t)}, \qquad t \in (T_m, T_{m+1}), V^\pi(0) = 1 \tag{20}$$

and with $T_0 = 0$. Quite clearly, we have $V^\kappa(\cdot) \equiv 1$ for the cash portfolio; and $V^\mu(\cdot) \equiv \mathcal{C}(\cdot)/\mathcal{C}(0)$ for the market portfolio of (1). As mentioned before, the amount $\pi_0(t)V^\pi(t)$ in the notation of (2) is invested in the money market at time t.

4. Proofs.

4.1. *Subexponential tail.* We state and prove the following crucial proposition. This result postulates that the distribution of the maximum number of companies over any finite time-interval has a tail which is lighter than that of any exponential distribution.

PROPOSITION 4.1. *Under Assumptions 1–5, for any $T \in (0, \infty)$, we have*

$$\lim_{u\to\infty} \frac{1}{u}\Big[-\log \mathbf{P}_x\Big(\max_{0\le t\le T} \mathcal{N}(t) > u\Big)\Big] = \infty.$$

In particular, for all $c \in (0, \infty)$ we have

$$\mathbf{E}_x\Big[\exp\Big(c \max_{0\le t\le T} \mathcal{N}(t)\Big)\Big] < \infty. \tag{21}$$

Theorem 2.1 follows directly from this proposition: If the maximal number of companies over the time-horizon $[0, T]$ has this property, then it is a.s. finite, which is another way of saying that the counting process $\mathcal{N}(\cdot)$ does not explode. Theorem 2.2 also uses this fact, but in subtler ways; its proof is postponed until Section 4.5.

The rest of this section is organized as follows. In Section 4.2, we explain the main idea of the proof of Proposition 4.1. In Section 4.3, we derive some preliminary estimates and lay the groundwork for the rest of the proof. In Section 4.4, we carry out the proof of Proposition 4.1 in full detail. In Section 4.5, we use this Proposition to prove Theorem 2.2.

4.2. *Overview of the proof of Proposition* 4.1. We shall use the following notation: Consider the random sequence $\{N_m\}_{m\in\mathbb{N}_0}$ with $N_m = \mathcal{N}(T_m+)$, and the sequence of "event-times" $\{T_m\}_{m\in\mathbb{N}_0}$, as in (19) and (20). The quantity N_m is the level of the process $\mu(\cdot)$ of market weights; in other words, the number of companies extant during the time interval (T_m, T_{m+1}) between the mth and the $(m+1)$st jumps of the integer-valued process $\mathcal{N}(\cdot)$.

The idea of the proof of Proposition 4.1 is as follows. We say that a *double jump upward* happens at the step m, if $N_m = N_{m-1} + 1$ and $N_{m+1} = N_m + 1$. If $N_m = N$, then we say that this is a *double jump upward from level* $N - 1$ *to level* $N + 1$ at step m; we denote this event by

$$A(m, N) := \{N_{m+1} = N_m + 1, N_m = N_{m-1} + 1\}. \tag{22}$$

Suppose we start from the level N, and the maximal number of companies during the time interval $[0, T]$ is larger than or equal to $2L$, where $L > N$ is some large number. Then it takes time less than T to get from the level L to the level $2L$; this will require at least $L - 1$ double jumps upward, for instance, one from L to $L + 2$, another from $L + 1$ to $L + 3$, etc. Note that such double jumps upward may "overlap," when there are three or more consecutive jumps upward.

The crucial part in the proof of Proposition 4.1 is to show that the probability of a double jump upward from $N - 1$ to $N + 1$ is small for large N. This is done in Lemmas 4.3, 4.4 and 4.5.

Indeed, as we shall see in Lemma 4.2, immediately after the first upward jump, from $N - 1$ to N, the top market weight will be less than or equal to $1 - \delta_0$, where

$$\delta_0 := 1 - (1 - \delta)(1 - \varepsilon_0) > \delta. \tag{23}$$

In other words, the process $\mu(\cdot)$ of market weights will "stay away" from the threshold $1 - \delta$, which it must hit before the exponential clock $\mathcal{E}(\lambda_N)$ rings, for a double jump at level N to happen. But from Lemma A.1 in the Appendix, the probability of this event is at most

$$\begin{aligned} p_N &:= 2\left(\frac{(1-\delta_0)\vee(1/2)}{1-\delta}\right)^{\lambda_N^{1/2}/\overline{\sigma}} = 2\exp\bigl(-\alpha_0\lambda_N^{1/2}\bigr) \\ &\qquad\text{where } \alpha_0 := \frac{1}{\overline{\sigma}}\log\left(\frac{1-\delta}{(1-\delta_0)\vee(1/2)}\right). \end{aligned} \tag{24}$$

Then we fix the number of steps u and claim that it is unlikely for the process to perform $L - 1$ double jumps upward within u steps. But if the process gets to the level $2L$ in $M \geq u$ steps, then there will be a lot of jumps downward, at independent exponential random times. Since there will be a lot of these random times, we can apply the large deviation theory and argue that their sum is very likely to be greater than T.

4.3. *Preliminary remarks and estimates.*

4.3.1. *Leaving a given level.* The process $\mu(\cdot) = \mathfrak{z}(X(\cdot))$ of market weights evolves in the following manner: As long as it stays on the $(N-1)$-dimensional manifold $\Delta_+^{N,\delta}$ (i.e., "on the level N"), the process $\mu(\cdot)$ evolves as an N-dimensional diffusion governed by the system (14) of SDEs.

How does the process $\mu(\cdot)$ leave the level N? There are two possibilities:

(i) An exponential clock $\eta_j^N \backsim \mathcal{E}(\lambda_N)$ rings; by construction, the random variable η_j^N is independent of the diffusion given by the system of SDEs in (14). Then we choose randomly two companies to merge, *excluding the top one*. As mentioned in Remark 2, this requirement is essential for ensuring than the merger will not be suppressed [i.e., with this proviso we never find ourselves in case B(iv) of Section 3.3].

(ii) The market weight of one of these N companies, say of the ith one, hits at some time τ the level $1-\delta$; thus $\mu_i(\tau) = 1-\delta$ and $\sum_{j\neq i} \mu_j(\tau) = \delta$. Then we pick a random variable $\xi \backsim F$, independent of the past, and split the ith company into two new companies: these are assigned market weights $\xi\mu_i(\tau)$ and $(1-\xi)\mu_i(\tau)$.

Since $1/2 \leq \xi \leq 1-\varepsilon_0$ from Assumption 3, each of the resulting two new market weights is at most $(1-\delta)(1-\varepsilon_0) = 1-\delta_0$ as in (23); whereas all the other companies, those unaffected by the split, have market weights bounded from above by δ. Because $\varepsilon_0 \in (0, 1/2)$ and $\delta \in (0, 1/6)$, we have $\delta < (1-\delta)/2 < (1-\delta)(1-\varepsilon_0) = 1-\delta_0 < 1-\delta$, so again *all* of the new market weights are bounded away from the threshold $1-\delta$.

Let us state this observation in the form of a separate lemma.

LEMMA 4.2. *For the time $\boldsymbol{\tau}$ of any upward jump in the process $\mathcal{N}(\cdot)$ and with δ_0 as in* (23), *we have*

$$\mu(\boldsymbol{\tau}+) \in \Delta_+^{N^\star,\delta_0} \qquad \textit{for some integer } N^\star \geq 2.$$

In other words, immediately after any upward jump, the market weight process is in $\mathcal{M}^{\delta_0}$ and $\mu_{(1)}(\boldsymbol{\tau}+) \leq 1-\delta_0$ holds.

4.3.2. *Jumping upward, rather than downward.* Let us obtain an upper bound for the conditional probability $\mathbf{P}_x(\tau_m \leq \eta_m^{N_{m-1}} | N_{m-1} = N)$ that the mth jump will be upward rather than downward, given that during the time-interval (T_{m-1}, T_m) the market weight process $\mu(\cdot) = \mathfrak{z}(X(\cdot))$ is at a given level $N \in \mathbb{N}$.

On the event $\{N_{m-1} = N\}$ and during the time-interval (T_{m-1}, T_m) with the "event-time" $T_m = T_{m-1} + (\tau_m \wedge \eta_m^N)$, the process of log-capitalizations evolves as a system of competing Brownian particles with drifts $g_k = g_{Nk}$ and variances $\sigma_k^2 = \sigma_{Nk}^2$, for $k = 1, \dots, N$.

First, we consider $m = 1$; by the comparison lemma from the Appendix, we get

$$\mathbf{P}_x\big(\tau_1 \le \eta_1^{N_0} | N_0 = N\big) = \mathbf{P}_x\big(\tau_1 \le \eta_2^{N_0}\big) \le \left(\frac{\mu_{(1)}(0) \vee (1/2)}{1-\delta}\right)^{\lambda_N^{1/2}/\overline{\sigma}},$$

because in the notation of Assumption 2, we have $\overline{\sigma} \ge \widetilde{\sigma} := \max_{1\le k\le N}(\sigma_{Nk})$. Now, back to the case of general m, the strong Markov property gives

$$\begin{aligned} \mathbf{P}_x\big(\tau_{m+1} \le \eta_{m+1}^{N_m} | N_m = N\big) &= \mathbf{E}_x\big(\mathbf{P}_{X(T_m)}\big(\tau_1 \le \eta_1^N\big)\big) \\ &\le 2\mathbf{E}_x\left(\frac{\mu_{(1)}(T_m+) \vee (1/2)}{1-\delta}\right)^{\lambda_N^{1/2}/\overline{\sigma}}. \end{aligned} \tag{25}$$

4.3.3. *A couple of auxiliary estimates.* In this subsection, we estimate the probability of a double jump from $N-1$ to $N+1$, and then show that this estimate is in some sense independent of the past: we can condition on an event which occurred before the first of these tandem jumps was completed (the jump from $N-1$ to N). Although this conditioning might influence the exact probability of the double jump, the estimate which is deduced in this subsection remains unchanged.

We recall the definition of $A(m, N)$ from (22).

LEMMA 4.3. *Fix $m \ge 1$, $N \ge 3$. Then, with p_N as in* (24), *we have the following estimate for the probability of a double upward jump*:

$$\begin{aligned} \mathbf{P}_x\big(A(m,N) | N_{m-1} = N-1\big) &\equiv \mathbf{P}_x(N_{m+1} = N+1 | N_m = N, N_{m-1} = N-1) \\ &\le p_N. \end{aligned}$$

PROOF. If $N_{m-1} = N-1$ and $N_m = N$, then by Lemma 4.2 we have $\mathfrak{z}(X(T_m+)) \in \Delta_+^{N,\delta_0}$, thus $\mu_{(1)}(T_m+) \le 1 - \delta_0$. It follows from (25) that $\mathbf{P}_x(N_{m+1} = N+1 | N_m = N, N_{m-1} = N-1)$ satisfies

$$\mathbf{P}_x\big(\tau_{m+1} \le \eta_{m+1}^{N_m} | N_m = N, N_{m-1} = N-1\big) \le 2\left(\frac{(1-\delta_0) \vee (1/2)}{1-\delta}\right)^{\lambda_N^{1/2}/\overline{\sigma}} = p_N$$

in the notation of (24). □

Let us show that the same estimate holds under conditioning on any event before the first jump. This will be necessary when we consider the probability of many double jumps. The events $A(m, N)$ and $A(m_1, N_1)$ are clearly *not independent*, but the estimate still holds.

LEMMA 4.4. *For every $m \geq 1$, $N \geq 3$ and $A \in \mathcal{F}(T_m)$, we have*

$$\begin{aligned}\mathbf{P}_x(N_{m+1} = N+1 | N_m = N, N_{m-1} = N-1, A) \\ = \mathbf{P}_x\big(A(m,N) | N_m = N, N_{m-1} = N-1, A\big) \leq p_N.\end{aligned}$$

PROOF. We claim that for $x_m \in (0,\infty)^N$ with $\mathfrak{z}(x_m) \in \Delta_+^{N,\delta_0}$, we have

$$\mathbf{P}_x\big(N_{m+1} = N+1 | X(T_m+) = x_m, N_m = N, N_{m-1} = N-1, A\big) \leq p_N.$$

This follows from the estimate of the previous lemma, and from the fact that T_{m+1} is a function of the initial condition $x_m = X(T_m+)$ and of the Brownian increments driving the system $Y^{N_m,x_m,m+1}(\cdot) = Y^{N,x_m,m+1}(\cdot)$; by construction, these increments are independent of $\mathcal{F}(T_m)$.

The strong Markov property now completes the proof. □

COROLLARY 4.5. *For fixed integers $m_1 < m_2 < \cdots < m_j < m$ and N_1, $N_2, \ldots, N_m, N$, we have*

$$\mathbf{P}_x\big(A(m,N) | A(m_1,N_1), A(m_2,N_2), \ldots, A(m_j,N_j)\big) \leq p_N.$$

PROOF. This is an immediate corollary of Lemma 4.4:

$$\begin{aligned}&\mathbf{P}_x\big(A(m,N) | A(m_1,N_1), A(m_2,N_2), \ldots, A(m_j,N_j)\big) \\ &\quad = \mathbf{P}_x\big(N_{m+1} = N+1, N_m = N, N_{m-1} = N-1 | \\ &\qquad A(m_1,N_1), A(m_2,N_2), \ldots, A(m_j,N_j)\big) \\ &\quad \leq \mathbf{P}_x\big(N_{m+1} = N+1 | N_m = N, N_{m-1} = N-1, \\ &\qquad A(m_1,N_1), A(m_2,N_2), \ldots, A(m_j,N_j)\big).\end{aligned}$$

But the event $A = A(m_1,N_1) \cap \cdots \cap A(m_j,N_j)$ belongs to $\mathcal{F}(T_m)$, because each event $A(m_i,N_i)$, $i = 1, \ldots, j$ depends on the state of the system at stopping times $T_{m_i-1}, T_{m_i}, T_{m_i+1}$, and all of these are dominated by T_m. An application of Lemma 4.4 completes the proof. □

We have the following consequence of Corollary 4.5.

COROLLARY 4.6. *For fixed $m_1 < \cdots < m_j$ and $N_1, \ldots, N_j$, we have the estimate*

$$\mathbf{P}_x\big(A(m_1,N_1) \cap \cdots \cap A(m_j,N_j)\big) \leq p_{N_1} \cdots p_{N_j}.$$

4.4. *The proof of Proposition* 4.1. With $N = \mathfrak{N}(x)$, let us estimate the probability that the market weight process $\mu(\cdot) = \mathfrak{z}(X(\cdot))$ rises during the time-interval $(0, T)$ from the level N to the level $2L$, where $L > N$. First, the process has to reach the level L; it will get there for the first time as the result of a split, and immediately after the jump it will be in Δ_+^{L,δ_0}. Then it will have time less than T to reach the level $2L$. For each $n = 2, 3, \dots$, the random variable

$$\Theta_n := \inf\{t \ge 0 : \mathcal{N}(t) = n\}$$

will denote the first time when the counting process $\mathcal{N}(\cdot)$ of our model hits the nth level (i.e., the first time n companies are extant).

Suppose we are able to establish the following estimate: *For every* $\beta > 0$*, there exist* $L_0 > N$ *and* $c_0 > 0$ *such that for every* $L > L_0$ *and* $y \in \Delta_+^{L,\delta_0}$ *we have*

$$\mathbf{P}_y(\Theta_{2L} \le T) \le c_0 e^{-\beta L}. \tag{26}$$

Then the rest of the proof will follow. Indeed, to get from x to the level $2L$ in time less than or equal to T, the process needs first to get to the level L by an upward jump. By Lemma 4.2, immediately after this jump, the process will be at some point $y \in \Delta_+^{L,\delta_0}$. Starting from this point, it has to reach the level $2L$ during the remaining time (which is of course smaller than T). Therefore, integrating over $y \in \Delta_+^{L,\delta_0}$ with respect to the distribution of $\mu(\Theta_L+)$ and using the strong Markov property, we will get then

$$\mathbf{P}_x\Bigl(\max_{0 \le t \le T} \mathcal{N}(t) \ge 2L\Bigr) = \mathbf{P}_x(\Theta_{2L} \le T) \le c_0 e^{-\beta L},$$

and the proof will be complete.

Thus, let us try to estimate the $\mathbf{P}_y$-probability of the event $\{\Theta_{2L} \le T\}$ in (26), for $y \in \Delta_+^{L,\delta_0}$. Suppose that it takes the process of market weights M jumps to reach the level $2L$; then for every real number $u > 0$ we have

$$\mathbf{P}_y(\Theta_{2L} \le T) = \mathbf{P}_y(\Theta_{2L} \le T, M > 3u) + \mathbf{P}_y(\Theta_{2L} \le T, M \le 3u). \tag{27}$$

We shall try to find a number $u > L$ such that the event $\{\Theta_{2L} \le T, M \le 3u\}$ is unlikely, and the event $\{\Theta_{2L} \le T, M > 3u\}$ is also unlikely. Let us introduce a couple of new pieces of notation:

$$\underline{\lambda}_L := \min(\lambda_{L+1}, \dots, \lambda_{2L-1}), \qquad \overline{\lambda}_L := \max(\lambda_3, \dots, \lambda_L).$$

First, we estimate the probability $\mathbf{P}_y(\Theta_{2L} \le T, M > 3u)$ on the right-hand side of (27). To get from the level L to the level $2L$ in M jumps, one needs to make L more upward than downward jumps. But the total number of these jumps is M, so the number of downward jumps is $(M - L)/2$. Therefore, on the event $\{\Theta_{2L} \le T, M > 3u\}$, there are

$$\frac{M - L}{2} > \frac{3u - u}{2} = u$$

downward jumps. If a jump proceeds from the level i to the level $i-1$, it takes time $\eta_i \backsim \mathcal{E}(\lambda_i)$ to make this jump happen (counting from the last one). All these exponential jump times are independent, so there exist i.i.d. $\mathcal{E}(1)$ random variables $\widetilde{\eta}_1, \dots, \widetilde{\eta}_u$ such that $\widetilde{\eta}_i = \lambda_i \eta_i$. We can write

$$\{\Theta_{2L} \le T, M > 3u\} \subseteq \{\eta_1 + \cdots + \eta_u \le T\} = \left\{\frac{\widetilde{\eta}_1}{\lambda_1} + \cdots + \frac{\widetilde{\eta}_u}{\lambda_u} \le T\right\}.$$

Since $\lambda_i \le \overline{\lambda}_{2L-1}, i = 3, \dots, 2L-1$, we have

$$\left\{\frac{\widetilde{\eta}_1}{\lambda_1} + \cdots + \frac{\widetilde{\eta}_u}{\lambda_u} \le T\right\} \subseteq \left\{\frac{\widetilde{\eta}_1 + \cdots + \widetilde{\eta}_u}{\overline{\lambda}_{2L-1}} \le T\right\} = \left\{\frac{\widetilde{\eta}_1 + \cdots + \widetilde{\eta}_u}{u} \le \frac{T\overline{\lambda}_{2L-1}}{u}\right\}.$$

We apply now techniques from the theory of large deviations, as in the book [7], Theorem 2.2.3 and Exercise 2.2.23(c), page 35. The rate function $\mathcal{H}$ for the exponential distribution $\mathcal{E}(1)$ is given by $\mathcal{H}(s) = s - 1 - \log s$, for $s > 0$ (this function is denoted by Λ^* in [7], Section 2.2). For $F = [0, s]$, according to the remark (c) on page 27 of the book [7] (immediately after the statement of Theorem 2.2.3), we have

$$\mathbf{P}_y\left\{\frac{\widetilde{\eta}_1 + \cdots + \widetilde{\eta}_u}{u} \le s\right\} \le 2\exp\Bigl(-u \inf_{v \in F} \mathcal{H}(v)\Bigr).$$

It is checked that the function $\mathcal{H}$ is decreasing on $(0, 1]$; therefore, for $s \in (0, 1)$, we have

$$\inf_{v \in F} \mathcal{H}(v) = \inf_{v \in [0,s]} \mathcal{H}(v) = \mathcal{H}(s).$$

Assuming that u is large enough, namely $u > L \vee (T\overline{\lambda}_{2L-1})$, we obtain

$$\mathbf{P}_y\left\{\frac{\widetilde{\eta}_1 + \cdots + \widetilde{\eta}_u}{u} \le \frac{T\overline{\lambda}_{2L-1}}{u}\right\} \le 2\exp\left(-u\mathcal{H}\left(\frac{T\overline{\lambda}_{2L-1}}{u}\right)\right),$$

therefore,

$$\begin{aligned} \mathbf{P}_y(\Theta_{2L} \le T, M > 3u) &\le \mathbf{P}_y\left\{\frac{\tilde{\eta}_1 + \cdots + \tilde{\eta}_u}{u} \le \frac{T\overline{\lambda}_{2L-1}}{u}\right\} \\ &\le 2\exp\left(-u\mathcal{H}\left(\frac{T\overline{\lambda}_{2L-1}}{u}\right)\right) =: \Sigma_1(u). \end{aligned} \tag{28}$$

Now, let us estimate the probability $\mathbf{P}_y(\Theta_{2L} \le T, M \le 3u)$ on the right-hand side of (27). In order to reach the level $2L$ starting from L in no more than $3u$ jumps, we need to have at least $L-1$ double jumps upward, as discussed in the preliminary remarks of Section 4.2.

One of these double jumps is from level L to level $L+2$, occurring at step m_1. Another is from level $L+1$ to level $L+3$, occurring at step m_2, etc., up to a double jump upward from level $2L-2$ to level $2L$, occurring at step m_{L-1}. The subset

$\{m_1, \dots, m_{L-1}\} \subseteq \{1, \dots, 3u-1\}$ with $1 \le m_1 < m_2 < \dots < m_{L-1} \le 3u-1$ can be chosen in

$$\binom{3u-1}{L-1} \le \binom{3u}{L} \le \frac{(3u)^L}{L!}$$

different ways. For a given subset $\{m_1, \dots, m_{L-1}\} \subseteq \{1, \dots, 3u-1\}$, Corollary 4.6 states that the probability $\mathbf{P}_y(\Theta_{2L} \le T, M \le 3u)$ is no more than

$$\begin{aligned} &\mathbf{P}_y\big(A(m_1, L+1), A(m_2, L+2), \dots, A(m_{L-1}, 2L-1)\big) \\ &\quad \le p_{L+1} \cdots p_{2L-1} \\ &\quad \le 2^{L-1} \exp\big(-\alpha_0(\lambda_{L+1}^{1/2} + \dots + \lambda_{2L-1}^{1/2})\big) \le 2^{L-1} \exp\big(-\alpha_0(L-1)\underline{\lambda}_L^{1/2}\big), \end{aligned}$$

and thus

$$\mathbf{P}_y(M \le 3u, \Theta_{2L} \le T) \le \frac{(3u)^L}{L!} 2^{L-1} \exp\big(-\alpha_0(L-1)\underline{\lambda}_L^{1/2}\big) =: \Sigma_2(u). \tag{29}$$

It follows from the estimates in (28) and (29) that the probability of the event $\{\Theta_{2L} \le T\}$ which we would like to estimate, as in (26) and (27), is

$$\begin{aligned} \mathbf{P}_y(\Theta_{2L} \le T) &= \mathbf{P}_y(\Theta_{2L} \le T, M \le 3u) + \mathbf{P}_y(\Theta_{2L} \le T, M > 3u) \\ &\le \Sigma_1(u) + \Sigma_2(u). \end{aligned}$$

Here, $u > L$. Note that $\lambda_L \asymp cL^\alpha$ as $L \to \infty$, so $\overline{\lambda}_{2L-1} \asymp 2^\alpha c L^\alpha$ and $\underline{\lambda}_L \asymp c 2^\alpha L^\alpha$ as $L \to \infty$.

We need now to fix the undetermined parameter u: we shall let $u = \lceil kL^{\alpha \vee 1} \rceil$ for large enough $k > 0$. The function $\mathcal{H}(\cdot)$ satisfies $\mathcal{H}(s) \ge -(1/2)\log s$ for $s \in (0, s_0)$ for some constant $s_0 \in (0, 1)$. Therefore, for large enough L and k we have the following estimates for these two summands. First, let us estimate $\Sigma_1(u)$ of (28): we have

$$\mathcal{H}\left(\frac{T\overline{\lambda}_{2L-1}}{u}\right) \ge -\frac{1}{2}\log\left(\frac{T\overline{\lambda}_{2L-1}}{u}\right) \ge \frac{1}{2}\log\frac{k}{T2^\alpha c} =: k_0,$$

therefore,

$$\Sigma_1(u) = 2\exp\left(-u\mathcal{H}\left(\frac{T\overline{\lambda}_{2L-1}}{u}\right)\right) \le 2\exp\big(-kk_0 L^{\alpha\vee 1}\big).$$

By taking k large enough, we can make kk_0 as large as we want. Since $\alpha \vee 1 \ge 1$, this proves that $\Sigma_1(u)$ decreases faster than any exponential function as $u \to \infty$. This completes the proof for this summand $\Sigma_1(u)$. The other summand

$$\Sigma_2(u) = \frac{(3u)^L}{L!} 2^{L-1} \exp\big(-\alpha_0(L-1)\underline{\lambda}_L^{1/2}\big)$$

from (29) decreases faster than any $e^{-\beta L}$ as $L \to \infty$ for any fixed $\beta > 0$, because $\underline{\lambda}_L \asymp cL^\alpha$ and

$$\log \Sigma_2(u) = L \log(3u) - \log(L!) + (L-1)\log 2 - \alpha_0 (L-1)\underline{\lambda}_L^{1/2}$$

is asymptotically smaller as $L \to \infty$ than

$$(\alpha \vee 1) L \log L - L \log L - \alpha_0 c (L-1) L^{(1/2)\vee(\alpha/2)} \asymp -\alpha_0 c L^{1+(\alpha\vee 1)/2}.$$

This establishes the bound of (26), so the proof of the proposition is complete.

4.5. *The proof of Theorem* 2.2. The general philosophy of the proof is as follows: For any given real number $T > 0$ we shall try to find an *equivalent martingale measure*, that is, a probability measure $\mathbf{Q}_T$ on $\mathcal{F}(T)$ with the following properties:

(i) $\mathbf{Q}_T \sim \mathbf{P}$ on $\mathcal{F}(T)$;
(ii) for every portfolio $\pi(\cdot)$, the wealth process $V^\pi(t), 0 \le t \le T$ is a $\mathbf{Q}_T$-martingale.

Suppose this is done; take two portfolios $\pi(\cdot)$ and $\rho(\cdot)$, and assume for a moment $\pi(\cdot)$ allows an arbitrage opportunity relative to $\rho(\cdot)$ on a given time horizon $[0, T]$ with $T \in (0, \infty)$. Then the conditions of (4) hold with respect to the measure $\mathbf{P}$ and, therefore, with respect to the measure $\mathbf{Q}_T$ as well. But

$$V^\pi(t) - V^\rho(t), \qquad 0 \le t \le T$$

is a $\mathbf{Q}_T$-martingale with initial value zero, therefore, $\mathbf{E}^{\mathbf{Q}_T}(V^\pi(T) - V^\rho(T)) = 0$ holds in contradiction to (4). This contradiction completes the proof, that arbitrage is not possible.

For a model with a given, fixed number of companies, an equivalent martingale measure is constructed thus: a Girsanov change of measure ensures that each capitalization process is a martingale with respect to the new measure, and the wealth process is a stochastic integral with these processes as integrators. As a result, the wealth process is also a martingale with respect to the new measure. But here the number of extant companies fluctuates, so we shall carry out a Girsanov construction up to the first jump, then carry out the same construction with the new number of stocks up to the second jump, and so on. We do this in a number of steps, as follows.

Step 1: First, as a preliminary step, let us consider the CBP-based market model from Section 3.2 with the dynamics of (15) and under an appropriate filtration $\mathbb{G} = \{\mathcal{G}(t)\}_{0 \le t < \infty}$ that satisfies the usual conditions.

Consider the processes $\Upsilon_i(\cdot) = \{\Upsilon_i(t), 0 \le t < \infty\}, i = 1, \dots, N$, given by

$$\Upsilon_i(t) := \int_0^t \frac{\mathrm{d}X_i(s)}{X_i(s)} = \sum_{k=1}^N \int_0^t \mathbf{1}_{\{X_i(t) = X_{(k)}(t)\}} \big(g_k \,\mathrm{d}t + \sigma_k \,\mathrm{d}W_i(t)\big), \tag{30}$$

$$0 \le t < \infty.$$

Over each interval $[0, T]$ with $T > 0$ a given real number, each $\Upsilon_i(\cdot \wedge T)$ can be turned into a martingale by the change of probability measure

$$\mathbf{Q}_T(A) = \mathbf{P}\big(Z(T)\mathbf{1}_A\big), \qquad A \in \mathcal{G}(T).$$

Here,

$$Z(t) = \exp\big(-M(t) - \tfrac{1}{2}\langle M\rangle(t)\big), \qquad 0 \le t < \infty,$$

and the $\mathbf{P}$-martingale $M(\cdot) = \{M(t), 0 \le t < \infty\}$ is given by

$$M(t) = \sum_{k=1}^{N}\left(\frac{g_k}{\sigma_k}\right)\sum_{i=1}^{N}\int_0^t \mathbf{1}_{\{X_i(u)=X_{(k)}(u)\}}\,\mathrm{d}W_i(u), \qquad 0 \le t < \infty.$$

The quadratic variation of this martingale $M(\cdot)$ is

$$\langle M\rangle(t) = \sum_{k=1}^{N}\left(\frac{g_k}{\sigma_k}\right)^2\sum_{i=1}^{N}\int_0^t \mathbf{1}_{\{X_i(u)=X_{(k)}(u)\}}\,\mathrm{d}u \le tN \max_{1\le k\le N}\left(\frac{g_k}{\sigma_k}\right)^2. \tag{31}$$

Using the Novikov condition ([21], Proposition 3.5.12), we see that $Z(\cdot)$ is a $\mathbf{P}$-martingale, and, therefore, $\mathbf{Q}_T$ a probability measure on $\mathcal{F}(T)$. Also, the quadratic variations of the processes $\Upsilon_i(\cdot)$ from (30) are given by

$$\langle \Upsilon_i\rangle(t) = \sum_{k=1}^{N}\sigma_k^2\int_0^t \mathbf{1}_{\{X_i(u)=X_{(k)}(u)\}}\,\mathrm{d}u \le t \max_{1\le k\le N}\sigma_k^2, \qquad i = 1, \dots, N, \tag{32}$$

whereas the independence of the Brownian motions $W_i(\cdot)$ and $W_j(\cdot)$ gives that

$$\langle \Upsilon_i, \Upsilon_j\rangle(\cdot) \equiv 0 \qquad \text{holds for } 1 \le i \ne j \le N. \tag{33}$$

It is clear from this discussion that $X_i(\cdot \wedge T) = \exp(Y_i(\cdot \wedge T) - (1/2)\langle Y_i\rangle(\cdot \wedge T))$, $i = 1, \dots, N$ are martingales (with zero cross-variations) under the probability measure $\mathbf{Q}_T$, which thus earns the appellation of Equivalent Martingale Measure (EMM) for the model of Section 3.2.

REMARK 8. It follows now easily, that the CBP-based market model of Section 3.2 is not diverse. For if this model were diverse on some time-horizon $[0, T]$ of finite length, Proposition 6.2 of [12] would proscribe for it EMMs, such as the probability measure $\mathbf{Q}_T$ just constructed. See also an illuminating discussion in [2], Section 7.

Step 2: Now, let $M^{N,x,n}(\cdot) = \{M^{N,x,n}(t), 0 \le t < \infty\}$ be the same martingale $M(\cdot)$ for the copy of a CBP-based market model

$$\big(\exp\big(Y_1^{N,x,n}(\cdot)\big), \dots, \exp\big(Y_N^{N,x,n}(\cdot)\big)\big)'$$

from Section 3.3. This model has parameters $g_k = g_{Nk}, \sigma_k = \sigma_{Nk}, k = 1, \dots, N$, the initial condition is x, and all the processes $\{M^{N,x,n}(\cdot)\}_{n\in\mathbb{N}}$ are independent. Also, denote by

$$\Upsilon_i^{N,x,n}(\cdot), \qquad i = 1, \dots, N,$$

the processes $\Upsilon_i(\cdot)$ of (30) for this copy of the market model. Slightly abusing notation, we define

$$M(t) = \sum_{m\in\mathbb{N}_0} M^{N_m,x_m,m+1}(t \wedge T_{m+1} - t \wedge T_m), \qquad 0 \le t < \infty \tag{34}$$

in the notation of Section 3.3. This is an $\mathbb{F} = \{\mathcal{F}(t)\}_{t\ge 0}$-local martingale, with localizing sequence $\{T_m\}_{m\in\mathbb{N}_0}$ and quadratic variation

$$\langle M\rangle(t) = \sum_{m\in\mathbb{N}_0} \langle M^{N_m,x_m,m+1}\rangle(t \wedge T_{m+1} - t \wedge T_m).$$

Step 3: Let us verify the Novikov condition

$$\mathbf{E}_x\bigl[\exp\bigl(\tfrac{1}{2}\langle M\rangle(T)\bigr)\bigr] < \infty, \qquad 0 \le T < \infty \tag{35}$$

of [21], Proposition 3.5.12. The expression (31) leads to the estimate

$$\langle M\rangle(t) \le \sum_{m\in\mathbb{N}_0} (t \wedge T_{m+1} - t \wedge T_m) N_m \max_{1\le k\le N_m} \left(\frac{g_{N_m k}}{\sigma_{N_m k}}\right)^2.$$

Since Assumption 2 implies that

$$\frac{|g_{N_m k}|}{\sigma_{N_m k}} \le \frac{\overline{g}}{\underline{\sigma}} =: C < \infty \qquad \text{holds for all } m \ge 0, k \ge 1,$$

we get

$$\langle M\rangle(T) \le C^2 T \max_{0\le t\le T} \mathcal{N}(t), \tag{36}$$

and so the left-hand side in (35) can be estimated as

$$\mathbf{E}_x\left(\exp\left[\frac{C^2 T}{2} \cdot \max_{0\le t\le T} \mathcal{N}(t)\right]\right).$$

But this quantity is finite for any given real number $T \in (0, \infty)$ because of (21) from Proposition 4.1, establishing the Novikov condition (35). Thus, the stochastic exponential

$$Z(\cdot) = \bigl\{Z(t), 0 \le t < \infty\bigr\}, \qquad Z(t) = \exp\bigl[-M(t) - \tfrac{1}{2}\langle M\rangle(t)\bigr] \tag{37}$$

is a $\mathbf{P}$-martingale, and we can define a new probability measure $\mathbf{Q}_T$ that satisfies

$$\mathrm{d}\mathbf{Q}_T = Z(T)\,\mathrm{d}\mathbf{P} \qquad \text{on } \mathcal{F}(T), \text{ for each given } T \in [0, \infty). \tag{38}$$

Step 4: We shall show now that, for any given real number $T > 0$, the wealth process $V^\pi(\cdot \wedge T)$ is a martingale for any given portfolio $\pi(\cdot)$, under the new measure $\mathbf{Q}_T$ just constructed. [This is very clearly the case for the cash portfolio, as $V^\kappa(\cdot) \equiv 1$.]

We write the equation for $V^\pi(\cdot)$ in the form

$$\int_0^t \frac{\mathrm{d}V^\pi(u)}{V^\pi(u)} = \sum_{m \in \mathbb{N}_0} \sum_{i=1}^{N_m} \int_0^{t \wedge T_{m+1} - t \wedge T_m} \pi_i(T_m + u)\, \mathrm{d}\Upsilon_i^{N_m, x_m, m+1}(u), \qquad 0 \le t < \infty.$$

Here, the processes $\Upsilon_i^{N_m, x_m, m+1}(\cdot \wedge T)$ are defined via

$$\Upsilon_i^{N_m, x_m, m+1}(t) := \int_0^{t \wedge T_{m+1} - t \wedge T_m} \frac{\mathrm{d}X_i^{N_m, x_m, m+1}(s)}{X_i^{N_m, x_m, m+1}(s)}, \qquad i = 1, \dots, N_m,$$

and are local martingales under the probability measure $\mathbf{Q}_T$ in (38). Therefore, the process $L^\pi(\cdot \wedge T)$, defined via

$$L^\pi(t) := \int_0^t \frac{\mathrm{d}V^\pi(u)}{V^\pi(u)} = \sum_{m \in \mathbb{N}_0} \sum_{i=1}^{N_m} \int_0^{t \wedge T_{m+1} - t \wedge T_m} \pi_i(T_m + u)\, \mathrm{d}\Upsilon_i^{N_m, x_m, m+1}(u),$$

is an $(\mathbb{F}, \mathbf{Q}_T)$-local martingale. The value process $V^\pi(\cdot)$ is the stochastic exponent of $L^\pi(\cdot)$, namely

$$V^\pi(\cdot) = \exp\big(L^\pi(\cdot) - \tfrac{1}{2}\langle L^\pi \rangle(\cdot)\big).$$

If we can establish the Novikov condition

$$\mathbf{E}^{\mathbf{Q}_T}\big[\exp\big(\tfrac{1}{2}\langle L^\pi \rangle(T)\big)\big] < \infty, \tag{39}$$

then it will turn out that $V^\pi(\cdot \wedge T)$ is an $(\mathbb{F}, \mathbf{Q}_T)$-martingale, as indeed we set out to show at the start of this proof.

Indeed, from (32) and (33) we see that the processes $\Upsilon_i^{N_m, x_m, m+1}(\cdot)$, $i = 1, \dots, N_m$ have zero cross-variations, and quadratic variations

$$\begin{aligned}\langle \Upsilon_i^{N_m, x_m, m+1} \rangle(t) &= \sum_{k=1}^{N_m} \sigma_k^2 \int_0^{t \wedge T_{m+1} - t \wedge T_m} \mathbf{1}_{\{X_i^{N_m, x_m, m+1}(u) = X_{(k)}^{N_m, x_m, m+1}(u)\}}\, \mathrm{d}u \\ &\le t \max_{1 \le k \le N_m} \sigma_{N_m k}^2.\end{aligned}$$

Therefore, the process $\Upsilon_\pi^{N_m, x_m, m+1}(\cdot) = \{\Upsilon_\pi^{N_m, x_m, m+1}(t), 0 \le t < \infty\}$ given by

$$\Upsilon_\pi^{N_m, x_m, m+1}(t) := \sum_{i=1}^{N_m} \int_0^t \pi_i(T_m + u)\, \mathrm{d}\Upsilon_i^{N_m, x_m, m+1}(u)$$

has quadratic variation $\langle \Upsilon_\pi^{N_m, x_m, m+1}\rangle(T) \le T K_\pi^2 \cdot \max_{1\le k\le N_m} \sigma_{N_m k}^2$, because $|\pi_i(t)| \le K_\pi < \infty$ holds for all $0 \le t < \infty$, $i = 1, \dots, N_m$. This gives

$$\langle L^\pi\rangle(T) = \sum_{k\in\mathbb{N}_0} \langle \Upsilon_\pi^{N_k, x_k, k+1}\rangle(T \wedge T_{k+1} - T \wedge T_k) \le T K_\pi^2 \cdot \max_{\substack{N\ge 2\\ 1\le k\le N}} \sigma_{Nk}^2 \le T K_\pi^2 \overline{\sigma}^2,$$

and property (39) is proved.

4.6. *Some open questions.* (I) The above proof used the boundedness of the portfolio $\pi(\cdot)$ in a crucial way. It would be very interesting to see whether arbitrage in this (or in a related) model with splits and mergers might exist with more general, unbounded portfolios.

(II) The estimate (36) also gives the bound

$$\langle M\rangle(T) - \langle M\rangle(t) \le C^2 T \cdot \max_{t\le\theta\le T} \mathcal{N}(\theta)$$

for every $t \in [0, T]$. From the theory of Bounded Mean Oscillation (BMO) Martingales as developed, for instance, in the book [23], in order to show that the stochastic exponential $Z(\cdot)$ of (37), (34) is a martingale, it suffices to show that the process $\mathbf{E}(\langle M\rangle(T) - \langle M\rangle(t)|\mathcal{F}(t))$, $0 \le t \le T$ is uniformly bounded. Thus, on the strength of the last display, it is enough to show that the process

$$\mathbf{E}\Big(\max_{t\le\theta\le T} \mathcal{N}(\theta)\Big|\mathcal{F}(t)\Big), \qquad 0 \le t \le T$$

is uniformly bounded. If this can be done, it might obviate the need to establish the sub-exponential bound of Proposition 4.1.

(III) It would be very interesting to decide whether absence of arbitrage, say with respect to the market portfolio, survives when one begins to constrain the splits and/or mergers that can happen over a given period of time, or along genealogies of companies produced by any given split (one could mandate, e.g., that the resulting companies cannot be touched for a certain amount of time). We believe not, but this issue remains to be settled.

(IV) It would be interesting to extend the above generic analysis, by allowing for some "idiosyncratic" features in the model. These can take the form of allowing the growth rates and the local co-variation rates for the different assets to depend, not only on the rank, but also on the name of the particular company (e.g., as in Ichiba et al. [18] or the so-called *second-order model* from [10]). They could also take the form of giving strategic control to the various companies, on decisions such as whether to engage in a merger or not. Our present model does not allow such features.

APPENDIX: A COMPARISON LEMMA

LEMMA A.1. *Consider a CBP-based market model as described in Definition 2 of Section 3.2. Assume in the manner of* (1) *that* $g_1 \le \min_{2\le k\le N} g_k$, *and let*

$$\tau := \inf\{t \ge 0 : \exists i = 1, \dots, N, \text{s.t. } \mu_i(t) = 1 - \delta\}, \qquad \widetilde{\sigma} := \max_{1\le k\le N} \sigma_k.$$

Then for an independent random variable η, *exponentially distributed with parameter* $\lambda > 0$, *we have*

$$\mathbf{P}(\tau \le \eta) \le 2\left(\frac{\mu_{(1)}(0) \vee 1/2}{1-\delta}\right)^{\widetilde{\sigma}^{-1}\lambda^{1/2}}. \tag{40}$$

The idea behind the argument of the proof is as follows. We can rewrite the stopping time τ as $\tau = \inf\{t \ge 0 : \mu_{(1)}(t) = 1 - \delta\}$; indeed, whenever a market weight reaches $1-\delta$, then it also gets to be the largest market weight, because $1 - \delta > 1/2$. The process $\log \mu_{(1)}(\cdot)$ "reflects off" $\log \mu_{(2)}(\cdot)$, as made precise in (10); it behaves like an Itô process, but when it collides with $\log \mu_{(2)}(\cdot)$ a positive local time term emerges as a result of the collision.

Now, we would like to replace $\log \mu_{(1)}(\cdot)$ by something larger. Consider a similar process $U(\cdot)$, now reflected at $\log(1/2)$; the second top market weight never gets above $1/2$, and for this new reflection pattern the logarithm of the top market weight will reflect earlier, so the resulting reflected process $U(\cdot)$ will be larger. In particular, we shall have $\tau \ge \widetilde{\tau} := \inf\{t \ge 0 : U(t) = 1 - \delta\}$; and the probability that the exponential clock (which is responsible for mergers) rings later than τ (the time of a split), will be smaller than the probability that this exponential clock rings later than $\widetilde{\tau}$. But $U(\cdot)$ is reflected Brownian motion, so we can calculate this latter probability explicitly.

PROOF OF LEMMA A.1. Let us derive an equation for the dynamics of $\log \mu_{(1)}(\cdot)$. We recall the expression for the dynamics of $\log \mu_i(\cdot)$ from (14), and denote by $\Lambda_{(k,\ell)}(\cdot) = \{\Lambda_{(k,\ell)}(t), t \ge 0\}$ the local time at the origin of the continuous semimartingale

$$\log \mu_{(k)}(\cdot) - \log \mu_{(\ell)}(\cdot) = Y_{(k)}(\cdot) - Y_{(\ell)}(\cdot) \qquad \text{for } 1 \le k < \ell \le N.$$

We have $\Lambda_{(k,\ell)}(\cdot) \equiv 0$ if $\ell - k \ge 2$, see [18], Lemma 1, as well as

$$\mathrm{d}\log \mu_{(1)}(t) = \sum_{i=1}^{N} \mathbf{1}_{\{\mu_i(t)=\mu_{(1)}(t)\}}\, \mathrm{d}\log \mu_i(t) + \frac{1}{2}\,\mathrm{d}\Lambda_{(1,2)}(t) \tag{41}$$

from [3]. We also note from [18] that the set $\{t \ge 0 | \mu_{(k)}(t) = \mu_{(1)}(t)\} = \{t \ge 0 | Y_{(k)}(t) = Y_{(1)}(t)\}$ has zero Lebesgue measure, for $k = 2, \dots, N$. Introduce the following notation:

$$\beta(t) := g_1 - \sum_{k=1}^{N} g_k \mu_{(k)}(t) - \frac{1}{2}\sum_{k=1}^{N} \sigma_k^2 \big(\mu_{(k)}(t) - \mu_{(k)}^2(t)\big),$$

$$a(t) := \sigma_1^2 \big(1 - \mu_{(1)}(t)\big)^2 + \sum_{k=2}^{N} \sigma_k^2 \mu_{(k)}^2(t) > 0.$$

Using (14), we can rewrite (41) in the notation of (11) [denoting by $V(\cdot) = \{V(t), t \geq 0\}$ yet another one-dimensional standard $\{\mathcal{F}(t)\}_{t\geq 0}$-Brownian motion] as

$$\begin{aligned}\mathrm{d}\log \mu_{(1)}(t) &= \beta(t)\,\mathrm{d}t + \sigma_1\,\mathrm{d}B_1(t) - \sum_{k=1}^{N} \sigma_k \mu_{(k)}(t)\,\mathrm{d}B_k(t) + \frac{1}{2}\,\mathrm{d}\Lambda_{(1,2)}(t) \\ &= \beta(t)\,\mathrm{d}t + \sqrt{a(t)}\,\mathrm{d}V(t) + \frac{1}{2}\,\mathrm{d}\Lambda_{(1,2)}(t).\end{aligned}$$

For the coefficient $a(t)$, we get the following estimate: $a(t) \leq \sigma_1^2 + \max_{2\leq k\leq N} \sigma_k^2 \leq 2\tilde{\sigma}^2$. Also, the coefficient $a(t)$ is bounded away from zero, at least until the moment τ, because for $t \leq \tau$, $\mu_{(1)}(t) \leq 1 - \delta$, and $a(t) \geq \sigma_1^2 \delta^2$. Since $\mu_{(k)}(t) \in [0, 1]$, we get: $\mu_{(k)}(t) - \mu_{(k)}^2(t) \geq 0$. It is easy to get the following estimate for $\beta(t)$:

$$\begin{aligned}\beta(t) &\leq g_1\big(1 - \mu_{(1)}(t)\big) - \sum_{k=2}^{N} g_k \mu_{(k)}(t) \leq g_1\big(1 - \mu_{(1)}(t)\big) - \min_{2\leq k\leq N} g_k \cdot \sum_{k=2}^{N} \mu_{(k)}(t) \\ &= \Big(g_1 - \min_{2\leq k\leq N} g_k\Big)\big(1 - \mu_{(1)}(t)\big) \leq 0.\end{aligned}$$

Let us make a random time change, using Lemma 2 from [28] (for $\sigma = 1$ in the notation of this lemma). The time change is as follows:

$$t = T(s) = \inf\{t \geq 0 \,|\, \Delta(t) \geq s\}, \qquad s = \Delta(t) := \int_0^t a(v)\,\mathrm{d}v, t \in [0, \tau].$$

Denoting by $\overline{V}(\cdot) = \{\overline{V}(s), s \geq 0\}$ yet another standard Brownian motion, and

$$\begin{aligned} &Z(\cdot) = \{Z(s), s \geq 0\}, && Z(s) = \log \mu_{(1)}\big(T(s)\big), \\ &\overline{\Lambda}(\cdot) = \{\overline{\Lambda}(s), s \geq 0\}, && \overline{\Lambda}(s) = \tfrac{1}{2}\Lambda_{(1,2)}\big(T(s)\big), \\ &\gamma(s) = \frac{\beta(T(s))}{a(T(s))} \end{aligned}$$

we get the following equation:

$$\mathrm{d}Z(s) = \gamma(s)\,\mathrm{d}s + \mathrm{d}\overline{V}(s) + \mathrm{d}\overline{\Lambda}(s) \qquad \text{for } s \leq \Delta(\tau).$$

One important remark: $\gamma(s) \le 0$ for all $s \le \Delta(\tau)$. Let us establish one useful property of $\overline{\Lambda}$:

$$\text{if } Z(s) = \log \mu_{(1)}\big(T(s)\big) \ge \log(1/2) \qquad \text{then } \mathrm{d}\overline{\Lambda}(s) = 0. \tag{42}$$

Indeed, if $\mu_{(1)}(t) > \mu_{(2)}(t)$, then $\mathrm{d}\Lambda(t) = 0$; in words, Λ is constant in a neighborhood of t. Therefore, if $\mu_{(1)}(T(s)) > \mu_{(2)}(T(s))$, then $\mathrm{d}\overline{\Lambda}(s) \equiv \mathrm{d}\Lambda(T(s)) = 0$. In particular, if $\mu_{(1)}(T(s)) > 1/2$, then $\mu_{(1)}(T(s)) > 1/2 \ge \mu_{(2)}(T(s))$, and $\mathrm{d}\overline{\Lambda}(s) = 0$.

We shall show the comparison $Z(\cdot) \le Z_0(\cdot)$, where $Z_0(\cdot) = \{Z_0(s), s \ge 0\}$ is a one-dimensional Brownian motion with zero drift and unit dispersion, starting from $\log \mu_{(1)}(0)$ and reflected at $\log(1/2)$, namely

$$\mathrm{d}Z_0(s) = \mathrm{d}\overline{V}(s) + \mathrm{d}\Lambda^0(s).$$

Here, $\Lambda^0(\cdot) = \{\Lambda^0(s), s \ge 0\}$ is the local time of this reflecting Brownian motion at the site $\log(1/2)$. More precisely, let us show that

$$Z(s) \le Z_0(s), \qquad s \le \Delta(\tau). \tag{43}$$

The proof of (43) proceeds along the same lines as in [19], Chapter 6, Theorem 1.1, but with some adjustments which are necessary because of the local time terms. We define $\psi(x) := x_+^3$ for $x \in \mathbb{R}$; then $\psi \in C^2(\mathbb{R})$, and for $s \le \Delta(\tau)$

$$\begin{aligned} \psi\big(Z(s) - Z_0(s)\big) = {}& \int_0^s \psi'\big(Z(u) - Z_0(u)\big)\gamma(u)\mathrm{d}u \\ & + \int_0^s \psi'\big(Z(u) - Z_0(u)\big)\big(\mathrm{d}\overline{\Lambda}(u) - \mathrm{d}\Lambda^0(u)\big). \end{aligned} \tag{44}$$

This does not contain stochastic integrals, because in the expression for $Z(s) - Z_0(s)$ they cancel out. Let us show that the right-hand side of (44) is nonpositive. Indeed, $\gamma(s) \le 0$, and $\psi'(x) = 3x_+^2 \ge 0$, so the first integral in the right-hand side of (44) is nonpositive. Also, when $Z(s) > Z_0(s)$, we have: $Z(s) > Z_0(s) \ge \log(1/2)$. Therefore, for these $s \le \Delta(\tau)$, from (42) we get: $\mathrm{d}\overline{\Lambda}(s) = 0$, and so $(\mathrm{d}\overline{\Lambda}(u) - \mathrm{d}\Lambda^0(u)) \le 0$. Once again using the fact that $\psi'(x) = 3x_+^2 \ge 0$, we get that the second integral in (44) is also nonpositive. So the whole expression $\psi(Z(s) - Z_0(s)) \le 0$, and this implies $Z(s) \le Z_0(s)$, $s \le \Delta(\tau)$.

Let $\tau_0 := \inf\{t \ge 0 | Z_0(t) = \log(1-\delta)\}$. Since τ is the hitting time by $\mu_{(1)}$ of $1 - \delta$, or, in other words, by $\log \mu_{(1)}$ of $\log(1-\delta)$, we get that $\Delta(\tau)$ is the hitting time by $Z(\cdot) = \log \mu_{(1)}(T(\cdot))$ of the level $\log(1-\delta)$. Therefore, by comparison (43), we have $\tau_0 \le \Delta(\tau)$. But $\Delta'(t) = a(t) \le 2\widetilde{\sigma}^2$, so

$$\Delta(t) = \int_0^t a(s)\,\mathrm{d}s \le 2\widetilde{\sigma}^2 t, \qquad t \le \tau.$$

In particular, $\Delta(\tau) \le 2\widetilde{\sigma}^2\tau$, and $\tau \ge \tau_0/(2\widetilde{\sigma}^2)$. Therefore,

$$\mathbf{P}(\tau \le \eta) \le \mathbf{P}\big(\tau_0 \le 2\widetilde{\sigma}^2\eta\big).$$

Using [5], Part II, Section 3, formula 1.1.2, and the fact that $2\widetilde{\sigma}^2\eta$ is exponentially distributed with parameter $\lambda/(2\widetilde{\sigma}^2)$, we get

$$\mathbf{P}(\tau_0 \le 2\widetilde{\sigma}^2\eta) = \mathbf{P}_{\widetilde{x}}\Bigl(\sup_{0\le s\le 2\widetilde{\sigma}^2\eta} |\overline{B}(s)| \ge y\Bigr) = \frac{\operatorname{ch}(\widetilde{x}\sqrt{\lambda}\widetilde{\sigma}^{-1})}{\operatorname{ch}(\widetilde{y}\sqrt{\lambda}\widetilde{\sigma}^{-1})}$$

with $\widetilde{y} = \log(1-\delta) - \log(1/2)$, $\widetilde{x} = (\log(\mu_{(1)}(0)) - \log(1/2))^+$. From the elementary inequality $(e^z/2) \le \operatorname{ch} z \equiv (e^z + e^{-z})/2 \le e^z$, $z \ge 0$ we conclude

$$\frac{\operatorname{ch}(\widetilde{x}\sqrt{\lambda}\widetilde{\sigma}^{-1})}{\operatorname{ch}(\widetilde{y}\sqrt{\lambda}\widetilde{\sigma}^{-1})} \le 2\exp\bigl(-(\widetilde{y}-\widetilde{x})\sqrt{\lambda}\widetilde{\sigma}^{-1}\bigr),$$

and it is then straightforward to rewrite the right-hand side as (40). This completes the proof. □

Acknowledgements. The authors would like to thank E. Robert Fernholz, Jean-Pierre Fouque, Soumik Pal, Vasileios Papathanakos, Walter Schachermayer, Johannes Ruf and Philip Whitman for very helpful discussions, and Johannes Ruf for his detailed comments. They are obliged to Winslow Strong for an updated version of the manuscript [33]. They are deeply indebted to the referee for reading of the manuscript very carefully, and for making many incisive comments and suggestions.

REFERENCES

[1] AUDRINO, F., FERNHOLZ, E. R. and FERRETTI, R. G. (2007). A forecasting model for stock market diversity. *Ann. Finance* **3** 213–240.

[2] BANNER, A. D., FERNHOLZ, E. R. and KARATZAS, I. (2005). Atlas models of equity markets. *Ann. Appl. Probab.* **15** 2296–2330. MR2187296

[3] BANNER, A. D. and GHOMRASNI, R. (2008). Local times of ranked continuous semimartingales. *Stochastic Process. Appl.* **118** 1244–1253. MR2428716

[4] BASS, R. F. and PARDOUX, É. (1987). Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields* **76** 557–572. MR0917679

[5] BORODIN, A. N. and SALMINEN, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd ed. Birkhäuser, Basel. MR1912205

[6] CHATTERJEE, S. and PAL, S. (2010). A phase transition behavior for Brownian motions interacting through their ranks. *Probab. Theory Related Fields* **147** 123–159. MR2594349

[7] DEMBO, A. and ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability* **38**. Springer, Berlin. MR2571413

[8] FERNHOLZ, E. R. (2002). *Stochastic Portfolio Theory. Applications of Mathematics (New York)* **48**. Springer, New York. MR1894767

[9] FERNHOLZ, E. R., ICHIBA, T. and KARATZAS, I. (2013). Two Brownian particles with rank-based characteristics and skew-elastic collisions. *Stochastic Process. Appl.* **123** 2999–3026. MR3062434

[10] FERNHOLZ, E. R., ICHIBA, T. and KARATZAS, I. (2013). A second-order stock market model. *Ann. Finance* **9** 439–454. MR3082660

[11] FERNHOLZ, E. R. and KARATZAS, I. (2005). Relative arbitrage in volatility-stabilized markets. *Ann. Finance* **1** 149–177.

[12] FERNHOLZ, E. R. and KARATZAS, I. (2009). Stochastic portfolio Theory: An Overview. *Handb. Numer. Anal.* **15** 89–167.

[13] FERNHOLZ, E. R., KARATZAS, I. and KARDARAS, C. (2005). Diversity and relative arbitrage in equity markets. *Finance Stoch.* **9** 1–27. MR2210925

[14] ICHIBA, T. (2009). Topics in multi-dimensional diffusion theory: Attainability, reflection, ergodicity and rankings. Ph.D. thesis, Columbia University, ProQuest LLC, Ann Arbor, MI. MR2717746

[15] ICHIBA, T. and KARATZAS, I. (2010). On collisions of Brownian particles. *Ann. Appl. Probab.* **20** 951–977. MR2680554

[16] ICHIBA, T., KARATZAS, I. and SHKOLNIKOV, M. (2013). Strong solutions of stochastic equations with rank-based coefficients. *Probab. Theory Related Fields* **156** 229–248. MR3055258

[17] ICHIBA, T., PAL, S. and SHKOLNIKOV, M. (2013). Convergence rates for rank-based models with applications to portfolio theory. *Probab. Theory Related Fields* **156** 415–448. MR3055264

[18] ICHIBA, T., PAPATHANAKOS, V., BANNER, A., KARATZAS, I. and FERNHOLZ, E. R. (2011). Hybrid atlas models. *Ann. Appl. Probab.* **21** 609–644. MR2807968

[19] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[20] KARATZAS, I., PAL, S. and SHKOLNIKOV, M. (2016). Systems of Brownian particles with asymmetric collisions. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 323–354.

[21] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[22] KARDARAS, C. (2008). Balance, growth and diversity of financial markets. *Ann. Finance* **4** 369–397.

[23] KAZAMAKI, N. (1994). *Continuous Exponential Martingales and BMO*. *Lecture Notes in Math.* **1579**. Springer, Berlin. MR1299529

[24] OSTERRIEDER, J. R. and RHEINLANDER, T. (2006). Arbitrage opportunities in diverse markets via a non-equivalent measure change. *Ann. Finance* **2** 287–301.

[25] PAL, S. and PITMAN, J. (2008). One-dimensional Brownian particle systems with rank-dependent drifts. *Ann. Appl. Probab.* **18** 2179–2207. MR2473654

[26] PAL, S. and SHKOLNIKOV, M. (2014). Concentration of measure for Brownian particle systems interacting through their ranks. *Ann. Appl. Probab.* **24** 1482–1508. MR3211002

[27] RUF, J. and RUNGGALDIER, W. (2013). A systematic approach to constructing market models with arbitrage. Preprint. Available at arXiv:1309.1988.

[28] SARANTSEV, A. (2014). On a class of diverse market models. *Ann. Finance* **10** 291–314. MR3193785

[29] SARANTSEV, A. (2015). Infinite systems of competing Brownian particles. Preprint. Available at arXiv:1403.4229.

[30] SARANTSEV, A. (2015). Multiple collisions in systems of competing Brownian particles. Preprint. Available at arXiv:1309.2621.

[31] SARANTSEV, A. (2015). Triple and simultaneous collisions of competing Brownian particles. *Electron. J. Probab.* **20** 1–28. MR3325099

[32] SHKOLNIKOV, M. (2011). Competing particle systems evolving by interacting Lévy processes. *Ann. Appl. Probab.* **21** 1911–1932. MR2884054

[33] STRONG, W. (2014). Fundamental theorems of asset pricing for piecewise semimartingales of stochastic dimension. *Finance Stoch.* **18** 487–514. MR3232014

[34] STRONG, W. and FOUQUE, J.-P. (2011). Diversity and arbitrage in a regulatory breakup model. *Ann. Finance* **7** 349–374. MR2879921

DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027
USA
AND
INTECH INVESTMENT MANAGEMENT LLC
ONE PALMER SQUARE, SUITE 441
PRINCETON, NEW JERSEY 08542
USA
E-MAIL: ik@math.columbia.edu
ik@enhanced.com

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
BOX 354350
SEATTLE, WASHINGTON 98195
USA
E-MAIL: ansa1989@math.washington.edu