Stable Diffusions Interacting through Their Ranks, as Models of Large Equity Markets

IOANNIS KARATZAS
INTECH Investment Management, Princeton
Department of Mathematics, Columbia University

Joint work with Tomoyuki Ichiba, Adrian Banner, Vasileios Papathanakos and E. Robert Fernholz

Talk in Paris January 2011

Log-Log Capital Distribution Curves I

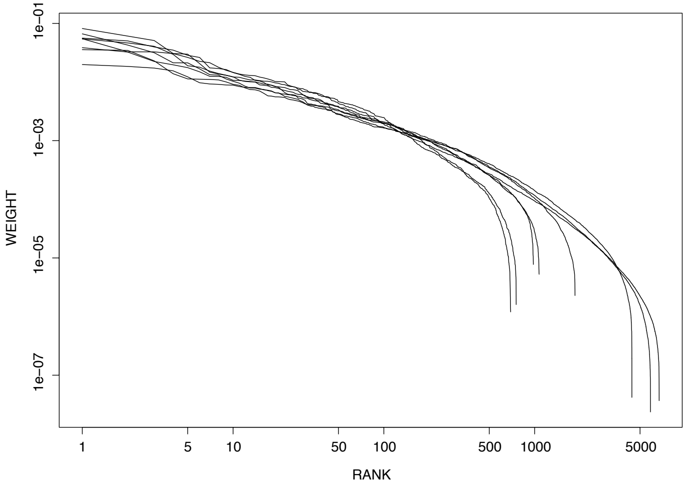


Figure: U.S. equity market, 1929-1999 (E.R. Fernholz (2002), p. 95)

Log-Log Capital Distribution Curves II

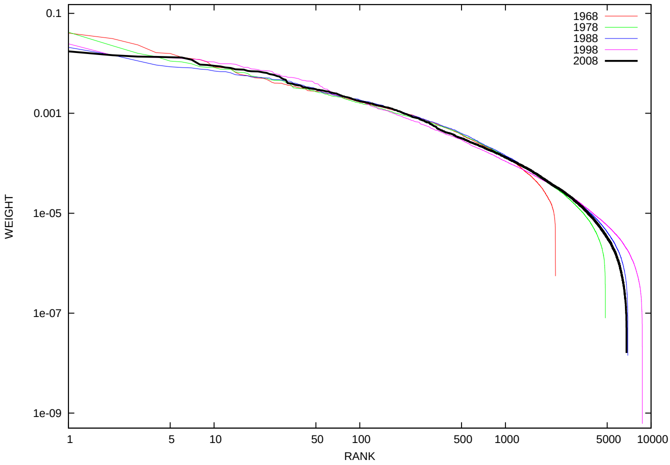


Figure: Capital distribution curves, U.S. equity market, 1968-2008

What kinds of models can describe this long-term stability?

Definition of Hybrid Atlas Model

- Capitalizations $\mathfrak{X} := \{(X_1(t), \ldots, X_n(t)), \;\; 0 \le t < \infty\}$.
- Descending Order Statistics (lexicographic tie-breaks):

$$\max_{1 \le i \le n} X_i(t) =: X_{(1)}(t) \ge X_{(2)}(t) \ge \cdots \ge X_{(n)}(t) := \min_{1 \le i \le n} X_i(t).$$

The curves of the previous slides are (smoothed) maps

$$\log k \longmapsto \frac{1}{T} \int_0^T \log \left(\frac{X_{(k)}(t)}{X_1(t) + \cdots + X_n(t)} \right) dt,$$

for $k = 1, 2, \cdots, n$ over different decades $[0, T]$ (for instance, Jan 1969 – Dec 1978; of course, each decade has its own, associated market "size" n).

Log-Capitalizations

Log-capitalizations $Y_i(t) := \log X_i(t)$.

Reverse-Order Statistics: $Y_{(1)}(t) \geq \cdots \geq Y_{(n)}(t)\,, \quad 0 \leq t < \infty\,.$

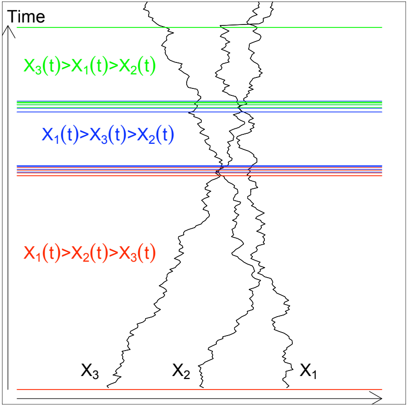
Dynamics of log-capitalizations.:

$$dY_i(t) = (\gamma + \gamma_i + g_k)\, dt + \sigma_k\, dW_i(t) \qquad \text{if} \quad Y_{(k)}(t) = Y_i(t)\,;$$

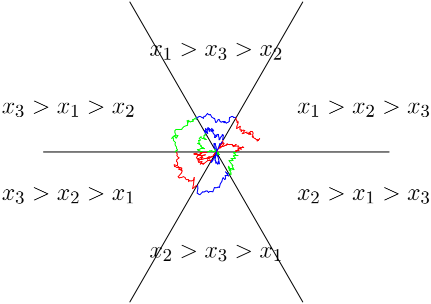
for $1 \leq i\,, k \leq n\,,\ 0 \leq t < \infty$, where $W(\cdot)$ is $n-$dim. B.M.
System of Brownian particles interacting through their ranks.
Unique weak solution (Bass & Pardoux, PTRF '87).

	company name i	k^{th} ranked company
Drift ("mean") Diffusion ("variance")	γ_i	g_k $\sigma_k > 0$

Illustration ($n = 3$) of Interactions through Rank: Linear and Caleidoscopic Views



Paths in $\mathbb{R}_+ \times$ time



A path in different wedges of $\mathbb{R}^n$

Permutations and Polyhedral Chambers

For $\mathbf{p} \in \Sigma_n$ (symmetric group on n elements), define wedge

$$\mathcal{R}_{\mathbf{p}} := \left\{ y \in \mathbb{R}^n : y_{\mathbf{p}(1)} \geq y_{\mathbf{p}(2)} \geq \cdots \geq y_{\mathbf{p}(n)} \right\},$$

a polyhedral chamber consisting of all points $y \in \mathbb{R}^n$ such that $y_{\mathbf{p}(k)}$ is ranked k^{th} among $y_1, \cdots, y_n$. We resolve ties "lexicographically", always in favor of the lowest index ("name") i. This results in a partition of $\mathbb{R}^n$ into pairwise-disjoint chambers. (To wit: $\mathbf{p}(k)$ is the "name" (index) of the particle that occupies the k^{th} rank in the permutation $\mathbf{p} \in \Sigma_n$.)

Define also the "coarser" chambers

$$Q_k^{(i)} := \left\{ y \in \mathbb{R}^n : y_i \text{ is ranked } k^{th} \text{ among } (y_1, \ldots, y_n) \right\}$$
$$= \bigcup_{\{\mathbf{p} \in \Sigma_n : \mathbf{p}(k) = i\}} \mathcal{R}_{\mathbf{p}} ; \quad 1 \leq i, k \leq n .$$

Vector Representation as a Diffusion

$$dY(t) = \mathbf{C}\big(Y(t)\big)\,dt + \mathbf{S}\big(Y(t)\big)\,dW(t)\,; \quad 0 \le t < \infty$$

$$\mathbf{C}(y) = \sum_{\mathbf{p}\in\Sigma_n} \big(g_{\mathbf{p}^{-1}(1)} + \gamma_1 + \gamma\,, \ldots, \, g_{\mathbf{p}^{-1}(n)} + \gamma_n + \gamma\big)' \cdot \mathbf{1}_{\mathcal{R}_\mathbf{p}}(y)$$

$$= \sum_{k=1}^{n} \Big(\big(g_k + \gamma_1 + \gamma\big) \cdot \mathbf{1}_{Q_k^{(1)}}(y)\,, \ldots\,, \, \big(g_k + \gamma_n + \gamma\big) \cdot \mathbf{1}_{Q_k^{(n)}}(y) \Big)',$$

$$\mathbf{S}(y) = \sum_{\mathbf{p}\in\Sigma_n} \underbrace{\operatorname{diag}\big(\sigma_{\mathbf{p}^{-1}(1)}\,, \ldots, \, \sigma_{\mathbf{p}^{-1}(n)}\big)}_{\mathfrak{s}_\mathbf{p}} \cdot \mathbf{1}_{\mathcal{R}_\mathbf{p}}(y)\,; \qquad y \in \mathbb{R}^n$$

$$= \operatorname{diag}\left(\sum_{k=1}^{n} \sigma_k \cdot \mathbf{1}_{Q_k^{(1)}}(y)\,, \ldots\,, \sum_{k=1}^{n} \sigma_k \cdot \mathbf{1}_{Q_k^{(n)}}(y) \right).$$

Semimartingale Representation of Ranked Processes

Recall $Y_{(1)}(t) \geq \cdots \geq Y_{(n)}(t)$, and denote by

$$\Lambda^{k,\ell}(t) := L^{Y_{(k)} - Y_{(\ell)}}(t)$$

the local time accumulated at the origin by the nonnegative semimartingale $Y_{(k)}(\cdot) - Y_{(\ell)}(\cdot)$ up to time t, for $1 \leq k < \ell \leq n$.

Lemma: For $k = 1, \ldots, n$, $0 \leq t \leq T$, we have

$$dY_{(k)}(t) = \Big(\gamma + g_k + \sum_{i=1}^{n} \gamma_i \, \mathbf{1}_{Q_k^{(i)}}(Y(t)) \Big) \, dt + \sigma_k \, dB_k(t) + \frac{1}{2} \Big[d\Lambda^{k,k+1}(t) - d\Lambda^{k-1,k}(t) \Big] ,$$

with the independent Brownian Motions (F.B. Knight)

$$B_k(\cdot) := \sum_{i=1}^{n} \int_0^{\cdot} \mathbf{1}_{Q_k^{(i)}}(Y(t)) \, dW_i(t) , \qquad k = 1, \cdots , n .$$

Reminder: The Local Time at the origin, accumulated up to time t by a continuous semimartingale $Y(\cdot) = Y(0) + M(\cdot) + V(\cdot)$, is

$$L^Y(t) \;:=\; Y^+(t) - Y^+(0) - \int_0^t \mathbf{1}_{\{Y(s)>0\}}\, dY(s)$$

$$= \lim_{\varepsilon \downarrow 0} \frac{1}{4\,\varepsilon} \int_0^t \mathbf{1}_{\{|Y(s)|<\varepsilon\}}\, d\langle M \rangle(s)\,.$$

The resulting process $L^Y(\cdot)$ is increasing, continuous, flat off the set $\{t \geq 0 : Y(t) = 0\}$. If $Y(\cdot) \geq 0$, this becomes

$$L^Y(t) \;=\; \int_0^t \mathbf{1}_{\{Y(s)=0\}}\, dY(s) \;=\; \int_0^t \mathbf{1}_{\{Y(s)=0\}}\, dV(s)\,.$$

- For continuous semimartingales $Y_1(\cdot), \cdots, Y_n(\cdot)$ we have for the local times at the origin (Yan, Ouknine; mid-80's):

$$L^{Y_1 \wedge Y_2}(t) + L^{Y_1 \vee Y_2}(t) \;=\; L^{Y_1}(t) + L^{Y_2}(t)\,, \qquad 0 \leq t < \infty$$

Banner & Ghomrasni (2008): More generally

$$\sum_{k=1}^{n} L^{Y_{(k)}}(t) = \sum_{i=1}^{n} L^{Y_i}(t), \qquad 0 \le t < \infty .$$

- Semimartingale representation for the ranked processes

$$dY_{(k)}(t) = \sum_{i=1}^{n} \mathbf{1}_{Q_k^{(i)}}(Y(t))\, dY_i(t)$$

$$+ \sum_{\ell=k+1}^{n} \frac{1}{\mathcal{N}_k(t)}\, d\Lambda^{k,\ell}(t) - \sum_{\ell=1}^{k-1} \frac{1}{\mathcal{N}_k(t)}\, d\Lambda^{k,\ell}(t) .$$

Please note the "upward pressure" coming from the lower ranks ($\ell = k+1, \cdots, n$), as well as the "downward pressure" from the upper ranks ($\ell = 1, \cdots, k-1$).

Here we keep track of the "size of the crowd" in rank k via

$$\mathcal{N}_k(t) := \#\{\, i : Y_i(t) = Y_{(k)}(t) \,\};$$

we also assume that all the semimatingales' bounded variation parts are absolutely continuous w.r.t. Lebesgue measure λ, and that for all (i, j) we have $\lambda(\{t \geq 0 : Y_i(t) = Y_j(t)\}) = 0$.

Idea of Proof of Lemma: For any three indices $1 \leq i, j, m \leq n$, the "rank-gap" process

$$\max_{\nu=i,j,m} Y_\nu(\cdot) - \min_{\nu=i,j,m} Y_\nu(\cdot)$$

dominates a Bessel process in dimension $\delta > 1$, and analysis of its local time shows that

$$L^{Y_{(k)} - Y_{(\ell)}}(\cdot) \equiv \Lambda^{k,\ell}(\cdot) \equiv 0\,, \qquad |\, k - \ell \,| \geq 2\,.$$

Serendipity: *even if* triple (or higher-order) collisions occur, they just *do not matter* for the respective local times.

These Local Times can be estimated...

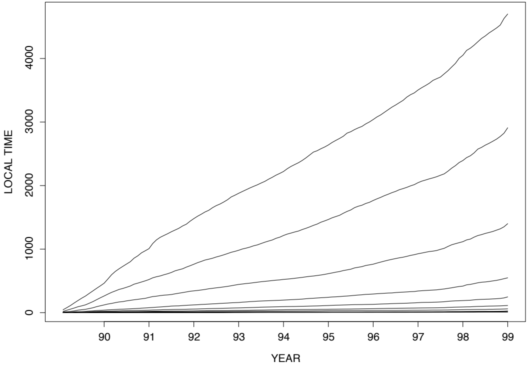


Figure: The estimated local time processes $\Lambda^{k,k+1}(t)$ for $k = 10$, 20, 40, ... , 5120; U.S. CRSP data, January 1990 to December 1999. (From E.R. Fernholz (2002) *Stochastic Portfolio Theory*, page 107.)

Discussion: Such estimation comes from the construction of rank-based portfolios that invest in an index-like fashion (according to relative capitalization) in, say, the top k stocks.

The performance of such a portfolio, relative to the entire market or to the submarket consisting of the bottom $n - k$ stocks, involves a “leakage” term proportional to the local time $\Lambda^{k,k+1}(\cdot)$ which measures the “turnover" between ranks k and $k + 1$; this can then be estimated based on observables.

- The linearity of the growth of local times is yet another indication of an underlying stability or ergodic behavior.

(Recall that for, say, Brownian motion, local time grows like $\sqrt{T}$; whereas for processes with an invariant distribution and stochastic stability, local time grows like T.)

What kinds of conditions can insure such stochastic stability?

Stability conditions

We shall assume, for every $k = 1, \dots, n-1,\ \mathbf{p} \in \Sigma_n$:

$$\sum_{k=1}^{n} g_k + \sum_{i=1}^{n} \gamma_i = 0\,, \qquad \sum_{\ell=1}^{k} \left(g_\ell + \gamma_{\mathbf{p}(\ell)}\right) < 0\,.$$

Very roughly speaking: *Assign big growth rates (and big variances) to the smallest stocks; then a stable capital distribution does indeed emerge.*

As Pal & Pitman (2008) remark, the stability conditions ensure that the “cloud of particles" will stick together: no sub-collection of particles can “form its own galaxy", as it were, and drift apart without ever again making contact with the rest.

Example 1 – Atlas model:

$$g_1 = \cdots = g_{n-1} = -\mathfrak{g} < 0;$$
$$g_n = (n-1)\mathfrak{g} > 0;$$
$$\gamma_1 = \cdots = \gamma_n = 0.$$

The company with the lowest capitalization provides all the growth (or support, as with the Titan of mythical lore) for the entire structure. (Here, companies are “anonymous" as far as their growth rates are concerned.)

Example 2 – Atlas model with stock-specific drifts:

$$g_1, \ldots, g_n \quad \text{as above;} \qquad \sum_{i=1}^{n} \gamma_i = 0\,, \quad \max_{1 \le i \le n} \gamma_i < \mathfrak{g}\,.$$

For instance:

$$\gamma_i = \mathfrak{g}\left(1 - \frac{2i}{n+1}\right), \;\; 1 \le i \le n\,.$$

Stochastic Stability

The average (center of gravity)

$$\overline{Y}(\cdot) := \frac{1}{n} \sum_{i=1}^{n} Y_i(\cdot)$$

of the log-capitalizations

$$d\,\overline{Y}(t) = \gamma\, dt + \frac{1}{n} \sum_{k=1}^{n} \sigma_k\, d\, B_k(t)$$

is Brownian motion with variance $\sum_{k=1}^{n} (\sigma_k / n)^2$, drift γ.

Recall here the independent Brownian Motions

$$B_k(\cdot) = \sum_{i=1}^{n} \int_0^{\cdot} \mathbf{1}_{Q_k^{(i)}}(Y(t))\, dW_i(t)\,, \qquad k = 1, \cdots, n\,.$$

The above stability conditions imply that the process of deviations from the center of gravity

$$\widetilde{Y}(\cdot) := \big(Y_1(\cdot) - \overline{Y}(\cdot), \ldots, Y_n(\cdot) - \overline{Y}(\cdot)\big)$$

is positive recurrent, uniformly over compact sets.

From the theory of R.Z. Khas'minskii (1960, 1980) we have then the following stochastic stability result:

Proposition: The process $\widetilde{Y}(\cdot)$ is stable in distribution; to wit, there is a unique invariant probability measure $\mu(\cdot)$ such that for every bounded, measurable $f : \Pi \to \mathbb{R}$ we have, with $\Pi := \{y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0\}$, the Strong Law of Large Numbers

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f\left(\widetilde{Y}(t)\right) dt = \int_\Pi f(y)\, \mu(dy)\,, \quad a.s.$$

Average Occupation Times

Setting $f(\cdot) = \mathbf{1}_{\mathcal{R}_{\mathbf{p}}}(\cdot)$ (respectively, $\mathbf{1}_{Q_k^{(i)}}(\cdot)$), we define the average occupation times of $X(\cdot)$ in the polyhedral chambers $\mathcal{R}_{\mathbf{p}}$ (respectively, $Q_k^{(i)}$):

$$\theta_{\mathbf{p}} := \mu(\mathcal{R}_{\mathbf{p}}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbf{1}_{\mathcal{R}_{\mathbf{p}}}(X(t))\, dt\,, \qquad \mathbf{p} \in \Sigma_n\,,$$

$$\vartheta_{k,i} := \mu(Q_k^{(i)}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbf{1}_{Q_k^{(i)}}(X(t))\, dt\,, \quad 1 \le k, i \le n\,.$$

Equilibrium Identity:

$$\gamma_i + \sum_{k=1}^{n} g_k\, \vartheta_{k,i} \;=\; 0\,; \quad i = 1, \ldots, n\,.$$

Example 2 – Atlas model with stock-specific drifts:

$$g_1 = \cdots = g_{n-1} = -\mathfrak{g} < 0; \quad g_n = (n-1)\mathfrak{g} > 0\,;$$

$$\sum_{i=1}^{n} \gamma_i = 0\,, \quad \max_{1 \le i \le n} \gamma_i < \mathfrak{g}\,.$$

• In this case, the proportions of time the various stocks occupy the lowest ("Atlas") rank are given by

$$\vartheta_{n,i} \;=\; \frac{1}{n}\left(1 - \frac{\gamma_i}{\mathfrak{g}}\right)\,, \qquad i = 1, \cdots, n\,.$$

We shall obtain more general formulas for these quantities in a short while ... .

Strong Laws of Large Numbers

Stability implies an *SLLN* for Local Times, $\forall\ k = 1, \cdots, n-1$:

$$\lim_{T \to \infty} \frac{1}{T} \Lambda^{k,k+1}(T) = -2 \sum_{\ell=1}^{k} \Big(g_\ell + \sum_{i=1}^{n} \vartheta_{\ell,i}\, \gamma_i \Big)$$

$$= -2 \sum_{\mathbf{p} \in \Sigma_n} \theta_{\mathbf{p}} \sum_{\ell=1}^{k} \Big(g_\ell + \gamma_{\mathbf{p}(\ell)} \Big) > 0\,, \qquad \text{a.s.}$$

Typically, this quantity increases with rank k, much like the picture we saw a moment ago: the higher the rank (to wit: the bigger the k, the smaller the stock in terms of capitalization), the bigger the intensity of “market turnover" around it.

- This will be the case, for instance, under the condition (satisfied in Examples 1, 2):

$$g_k + \gamma_i < 0 \,; \qquad \forall \;\; 1 \le k \le n-1 \,, \;\; 1 \le i \le n \,.$$

Together with

$$\sum_{k=1}^{n} g_k + \sum_{i=1}^{n} \gamma_i = 0 \,,$$

this condition implies stability.

What can be said about $\vartheta_{k,i}$ and μ ?

EXAMPLE: Equal Variances, $\gamma = \gamma_1 = \cdots = \gamma_n = 0$

Just a bunch of Brownian motions with drifts determined by their ranks. In this case the equations become

$$dY_i(t) = \Big(\sum_{k=1}^{n} g_k \mathbf{1}_{Q_k^{(i)}}(Y(t)) \Big) dt + dW_i(t) \;=\; D_i \Phi(Y(t)) dt + dW_i(t) \,.$$

A *conservative diffusion,* with drift given by a conservative vector field and continuous, piecewise smooth potential

$$\Phi(y) \;:=\; \sum_{k=1}^{n} g_k \, y_{(k)} \,, \qquad y \in \mathbb{R}^n \,.$$

The stability conditions imply $\Phi(0) = 0$, $\int_{\mathbb{R}^n} e^{2\Phi(y)} \, dy < \infty$ and

$$\Phi(y) \;=\; \sum_{k=1}^{n-1} \big(y_{(k)} - y_{(k+1)} \big) \Big(\sum_{\ell=1}^{k} g_\ell \Big) < 0 \,, \qquad y \in \mathbb{R}^n \setminus \{0\} \,.$$

Now standard theory shows the existence of an invariant measure for the process $Y(\cdot)$, with unnormalized probability density function in the form of a product-of-exponentials

$$e^{2\Phi(y)} = \exp\Big\{ -\sum_{k=1}^{n-1} \lambda_k \big(y_{(k)} - y_{(k+1)}\big)\Big\},$$

with (the stability conditions once again!)

$$\lambda_k := -2 \sum_{\ell=1}^{k} g_\ell > \mathbf{0}, \qquad k = 1, \cdots, n-1 .$$

(Independence of successive gaps. Reversibility.)

In reality: variances grow with rank (the smaller the stock, the more volatile it tends to be). And of course, growth rates should depend on name as well as rank... .

Linearly Growing Variances

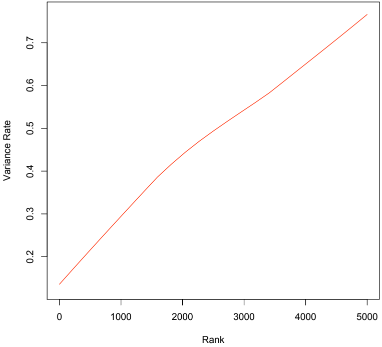


Figure: Smoothed variance by rank, U.S. Equity market, 1990-1999.

We shall assume that variances grow linearly by rank:

$$\sigma_2^2 - \sigma_1^2 = \sigma_3^2 - \sigma_2^2 = \cdots = \sigma_n^2 - \sigma_{n-1}^2 \geq 0\,.$$

Semimartingale Reflected Brownian Motions

Recall the ranked semimartingale decomposition

$$dY_{(k)}(t) = \sum_{i=1}^{n} \mathbf{1}_{Q_k^{(i)}}(Y(t))\, dY_i(t) + \frac{1}{2}\left[d\Lambda^{k,k+1}(t) - d\Lambda^{k-1,k}(t) \right].$$

The vector $\Xi(\cdot)$ of "**Gaps**" $\Xi_k(\cdot) := Y_{(k)}(\cdot) - Y_{(k+1)}(\cdot) \geq 0$ with

$$\Xi_k(\cdot) = \Xi_k(0) + \sum_{i=1}^{n} \int_0^{\cdot} \left(\mathbf{1}_{Q_k^{(i)}} - \mathbf{1}_{Q_{k+1}^{(i)}} \right)\big(Y(t)\big) \cdot dY_i(t)$$

$$-\frac{1}{2}\left[\Lambda^{k-1,k}(\cdot) + \Lambda^{k+1,k+2}(\cdot) \right] + \Lambda^{k,k+1}(\cdot)\,, \quad 1 \leq k \leq n-1$$

can be seen as a **semimartingale reflected Brownian motion** in the nonnegative orthant (Harrison, Reiman, Williams).

- Finally, we define the *indicator map* $\boxed{\mathbb{R}^n \ni y \mapsto \mathfrak{p}^y \in \Sigma_n}$

$$y_{\mathfrak{p}^y(1)} \geq y_{\mathfrak{p}^y(2)} \geq \cdots \geq y_{\mathfrak{p}^y(n)}\,, \quad \text{so that} \quad \mathfrak{p}^y = \mathbf{p} \iff y \in \mathcal{R}_{\mathbf{p}}\,,$$

where $\mathfrak{p}^y(k)$ is the name (index) of the coördinate that occupies the k^{th} rank among $y_1, \cdots, y_n$.

We introduce also the **Index process**

$$\boxed{\mathfrak{P}_t := \mathfrak{p}^{Y(t)}} \qquad 0 \leq t < \infty\,,$$

with values in the symmetric group Σ_n. The definition implies

$$Y_{\mathfrak{P}_t(1)} = Y_{(1)}(t) \geq \cdots \geq Y_{(n)}(t) = Y_{\mathfrak{P}_t(n)}\,, \qquad 0 \leq t < \infty\,.$$

Invariant Distribution for Adjacent Gaps and Indices

Proposition: Under the stability and linearly-growing-variance conditions, the invariant distribution $\nu(\cdot)$ of $(\Xi(\cdot), \mathfrak{P}.)$ is

$$\nu(A \times B) = \Big(\sum_{\mathbf{p} \in \Sigma_n} \prod_{k=1}^{n-1} \lambda_{\mathbf{p},k}^{-1} \Big)^{-1} \cdot \sum_{\mathbf{p} \in B} \int_A \exp\big(- \langle \lambda_{\mathbf{p}}, z \rangle \big)\, dz$$

for every measurable set $A \times B \in (\mathbb{R}_+)^{n-1} \times \Sigma_n$. Here we have set $\lambda_{\mathbf{p}} := (\lambda_{\mathbf{p},1}, \ldots, \lambda_{\mathbf{p},n-1})'$ to be the vector of components:

$$\lambda_{\mathbf{p},k} := \frac{-2 \sum_{\ell=1}^{k} \big(g_\ell + \gamma_{\mathbf{p}(\ell)} \big)}{(\sigma_k^2 + \sigma_{k+1}^2)/2} > 0\,; \quad \mathbf{p} \in \Sigma_n\,,\ 1 \le k \le n-1\,.$$

Please compare with expression on slide 24.

Discussion: The invariant measure $\nu(\cdot,\cdot)$ of $(\Xi(\cdot), \mathfrak{P}.)$ satisfies the **"Basic Adjoint Relationship"** (BAR) of Harrison & Williams (1987) (chamber-by chamber, then globally).

Its particular form leads to the density

$$\mathbb{P}\big(\Xi(t) \in A\big) = \Big(\sum_{\mathbf{p}\in\Sigma_n} \prod_{k=1}^{n-1} \lambda_{\mathbf{p},k}^{-1} \Big)^{-1} \cdot \sum_{\mathbf{p}\in\Sigma_n} \int_A \exp(-\langle \lambda_{\mathbf{p}}, z\rangle)\, dz$$

of sums-of-products-of-exponentials type, for the distribution of the semimartingale reflected Brownian motion process

$$\Xi(\cdot) := (\Xi_1(\cdot), \ldots, \Xi_{n-1}(\cdot))'$$

of adjacent gaps

$$\Xi_k(\cdot) := Y_{(k)}(\cdot) - Y_{(k+1)}(\cdot) \geq 0\,, \qquad k = 1, \ldots, n-1$$

under the invariant measure $\nu(\cdot\,,\cdot)$.

Discussion (cont'd): The assumption of *linearly growing variances* is crucial in the Proposition.

It guarantees that the structural **"Skew-Symmetry Condition"** (SSC) is satisfied, and that the process of adjacent gaps

$$\Xi(\cdot) = (\Xi_1(\cdot), \ldots, \Xi_{n-1}(\cdot))'$$

actually never visits the nonsmooth part of the boundary of the positive orthant (R. Williams (1987)).

Special case of a theory developed by T. Ichiba (2009) in his dissertation, concerning the absence of triple collisions.
Also: earlier work by Cépa & Lépingle (2007) with unbounded (electrostatic repulsion) drifts; here drifts are bounded.

This implies the absence of triple collisions for the components of the original process $Y(\cdot)$.

Comment: With $\mathfrak{D}$ the diagonal matrix of the covariance matrix $\mathfrak{A} = \{\mathrm{a}_{k\ell}\}_{1\le k,\ell\le n-1}$ with

$$\mathrm{a}_{k\ell} := \left(\sigma_k^2 + \sigma_{k+1}^2\right) \mathbf{1}_{\{\ell=k\}} - \sigma_k^2 \, \mathbf{1}_{\{\ell=k-1\}} - \sigma_{k+1}^2 \, \mathbf{1}_{\{\ell=k+1\}} ,$$

and with the $(n-1)\times(n-1)$ "reflection matrix" (slide 25)

$$\mathfrak{R} := \begin{pmatrix} 1 & -1/2 & & & \\ -1/2 & 1 & -1/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -1/2 & 1 & -1/2 \\ & & & -1/2 & 1 \end{pmatrix},$$

the **Skew-Symmetry Condition** (SSC) mandates

$$\boxed{2\left(\mathfrak{D} - \mathfrak{A}\right) = \left(\mathbf{I} - \mathfrak{R}\right)\mathfrak{D} + \mathfrak{D}\left(\mathbf{I} - \mathfrak{R}\right).}$$

It is satisfied in the case of linearly growing variances.

The components of the vector $\mathfrak{r}_k$ of the matrix $\mathfrak{R}$ provide the directions of reflection, when the face of the boundary

$$\mathfrak{F}_k := \left\{ (z_1, \cdots, z_{n-1})' \,|\, z_k = 0 \right\}, \qquad k = 1, \ldots, n-1$$

of the state-space $\mathfrak{S} = (\mathbb{R}_+)^{n-1}$ is hit – non-tagentially! – and the k^{th} component of $\Lambda(\cdot)$ increases. The BAR is

$$\boxed{\int_{\mathfrak{S} \times \Sigma_n} \big[\mathcal{A}(\mathbf{p}) \, f \big](z) \, d\nu(z, \mathbf{p}) + \frac{1}{2} \sum_{k=1}^{n-1} \int_{\mathfrak{F}_k} \langle \mathfrak{r}_k \,, \nabla f(z) \rangle (z) \, d\nu_{0k}(z) = 0}$$

for $f \in \mathcal{C}^2(\mathfrak{S})$, where

$$\big[\mathcal{A}(\mathbf{p}) \, f \big](z) := \frac{1}{2} \sum_{k=1}^{n-1} \sum_{\ell=1}^{n-1} \mathrm{a}_{k,\ell} \, \frac{\partial^2 f(z)}{\partial z_k \partial z_\ell} + \sum_{k=1}^{n-1} b_k(\mathbf{p}) \, \frac{\partial f(z)}{\partial z_k} \,,$$

$$b_k(\mathbf{p}) := \big(g_k + \gamma_{\mathbf{p}^{-1}(k)} \big) - \big(g_{k+1} + \gamma_{\mathbf{p}^{-1}(k+1)} \big) \,.$$

Average Occupation Times

Corollary: The long-term-average occupation times are

$$\boxed{\theta_{\mathbf{p}} \;=\; \mu(\mathcal{R}_{\mathbf{p}}) \;=\; \nu(\mathfrak{S}, \{\mathbf{p}\}) \;=\; \left(\sum_{\mathbf{q}\in\Sigma_n} \prod_{k=1}^{n-1} \lambda_{\mathbf{q},k}^{-1} \right)^{-1} \cdot \prod_{k=1}^{n-1} \lambda_{\mathbf{p},k}^{-1}}$$

for each chamber $\mathcal{R}_{\mathbf{p}}$ $(\mathbf{p} \in \Sigma_n)$, and

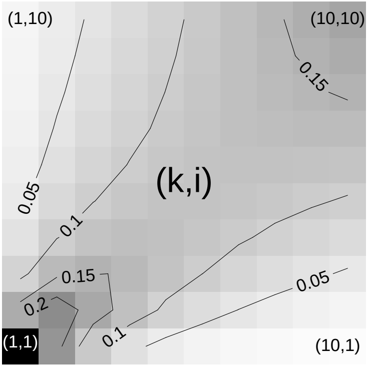
$$\underbrace{\vartheta_{k,i} \;=\; \sum_{\{\mathbf{p}\in\Sigma_n \,:\, \mathbf{p}(k)=i\}} \theta_{\mathbf{p}}}\;, \qquad i = 1, \ldots, n\,.$$

These DO satisfy (sanity check) the equilibrium identities

$$\gamma_i + \sum_{k=1}^{n} g_k\, \vartheta_{k,i} \;=\; 0\,; \quad i = 1, \ldots, n\,.$$

- If all $\gamma_i = 0$, then $\vartheta_{k,i} = \frac{1}{n}$ for $1 \le k, i \le n$ (first-order model of BFK (2005), includes the simple Atlas model as a special case).

- Heat map of $\vartheta_{k,i}$ when $n = 10$, $\sigma_k^2 = 1 + k$, $g_k = -1$ for $1 \le k \le 9$, $g_{10} = 9$, and $\gamma_i = 1 - (2i)/(n+1)$ for $i = 1, \ldots, 10$.



Distribution of Ranked Market Weights

Corollary: The invariant distribution of ranked market weights

$$M_{(k)}(\cdot) := \frac{X_{(k)}(\cdot)}{X_1(\cdot) + \cdots + X_n(\cdot)} \;; \quad k = 1, \ldots, n$$

has probability density function $\wp(m_1, \ldots, m_{n-1})$ given by

$$\boxed{\sum_{\mathbf{p} \in \Sigma_n} \theta_{\mathbf{p}} \frac{\lambda_{\mathbf{p},1} \cdots \lambda_{\mathbf{p},n-1}}{m_1^{\lambda_{\mathbf{p},1}+1} \cdot m_2^{\lambda_{\mathbf{p},2}-\lambda_{\mathbf{p},1}+1} \cdots m_{n-1}^{\lambda_{\mathbf{p},n-1}-\lambda_{\mathbf{p},n-2}+1} m_n^{-\lambda_{\mathbf{p},n-1}+1}}} \,,$$

$$0 < m_n \leq m_{n-1} \leq \ldots \leq m_1 < 1 \,, \;\; m_n := 1 - (m_1 + \cdots + m_{n-1}) \,.$$

- This is a distribution of ratios of Pareto type.

In the special case

$$\gamma_1 = \cdots = \gamma_n = 0$$

it takes the far simpler form

$$\wp(m_1, \ldots, m_{n-1}) = \prod_{k=1}^{n-1} \lambda_k \, m_k^{1+\lambda_k-\lambda_{k-1}} \cdot \big(1 - m_1 - \cdots - m_{n-1}\big)^{1-\lambda_{n-1}}$$

with

$$\lambda_k := \frac{(-4)\sum_{\ell=1}^{k} g_\ell}{\sigma_k^2 + \sigma_{k+1}^2}, \qquad \lambda_0 = \lambda_n = 0\,.$$

Capital Distribution Curves

$$\wp(m_1, \ldots, m_{n-1}) =$$

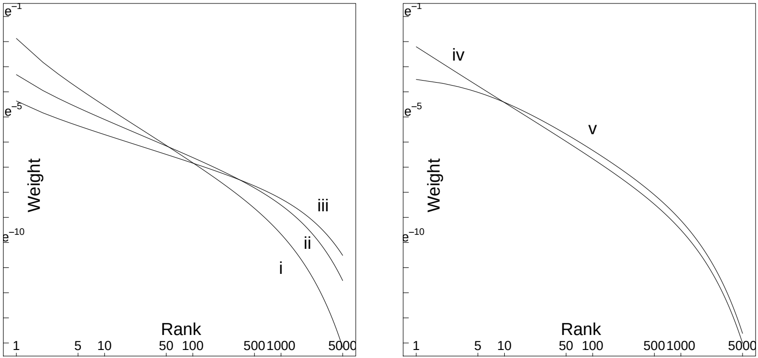
$$= \sum_{\mathbf{p} \in \Sigma_n} \theta_{\mathbf{p}} \frac{\lambda_{\mathbf{p},1} \cdots \lambda_{\mathbf{p},n-1}}{m_1^{\lambda_{\mathbf{p},1}+1} \cdot m_2^{\lambda_{\mathbf{p},2}-\lambda_{\mathbf{p},1}+1} \cdots m_{n-1}^{\lambda_{\mathbf{p},n-1}-\lambda_{\mathbf{p},n-2}+1} m_n^{-\lambda_{\mathbf{p},n-1}+1}}$$

The invariant probability density for the ranked market weights from the from the previous slide, allows us to describe the long term average (and "expected") **slope** of the capital distribution curve at the various ranks k, thus also its shape:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{\log M_{(k+1)}(t) - \log M_{(k)}(t)}{\log(k+1) - \log k} \, dt =$$

$$\mathbb{E}^\nu \left(\frac{\log M_{(k+1)} - \log M_{(k)}}{\log(k+1) - \log k} \right) = \frac{-\mathbb{E}^\nu(\Xi_k)}{\log(1+k^{-1})} = -\frac{\sum_{\mathbf{p} \in \Sigma_n} \theta_{\mathbf{p}} \, \lambda_{\mathbf{p},k}^{-1}}{\log(1+k^{-1})} \, .$$

Illustrations



- $n = 5000$, $g_n = c_*(2n-1)$, $g_k = 0$, $1 \le k \le n-1$, $\gamma_1 = -c_*$, $\gamma_i = -2c_*$, $2 \le i \le n$, $\sigma_k^2 = 0.075 + 6k \times 10^{-5}$, $1 \le k \le n$. (i) $c_* = 0.02$, (ii) $c_* = 0.03$, (iii) $c_* = 0.04$.
- (iv) $c_* = 0.02$, $g_1 = -0.016$, $g_k = 0$, $2 \le k \le n-1$, $g_n = (0.02)(2n-1) + 0.016$,
- (v) $g_1 = \cdots = g_{50} = -0.016$, $g_k = 0$, $51 \le k \le n-1$, $g_n = (0.02)(2n-1) + 0.8$.

Connections

This theory allows us to compute growth-optimal and universal portfolios, for long-term money management (almost tailor-made for the "Empirical Bayes" Cover-Jamshidian theory of universal portfolios; this is another story, and talk...).

The theory we presented relies heavily on

- the "semimartingale reflecting Brownian motion" analysis of Queueing Networks in their heavy traffic limit approximation (J.M. Harrison, M. Reiman, R. Williams),

and has strong connections with

• the combinatorial analysis of interacting diffusions based on Coxeter groups (Chatterjee and Pal);

• discrete-time models of competing particle systems in Statistical Mechanics, such as Sherrington-Kirkpatrick models of spin glasses, with similar invariant distributions (M. Aizenman, A. Ruzmaikina, P.L. Arguin).

As the number $n \to \infty$ of particles increases to infinity, the empirical measure of the "configuration of particles" is characterized by evolution equations of the McKean-Vlasov type, and by partial differential equations of the porous medium form (very recent preprint by Mykhaylo Shkolnikov at Stanford).

References

BANNER, A., FERNHOLZ, E.R., & KARATZAS, I. (2005) Atlas models of equity markets. *Annals of Applied Probability* **15**, 2296-2330.

ICHIBA, T. & KARATZAS, I. (2010) On collisions of Brownian particles. *Annals of Applied Probability* **20**, 951-977.

ICHIBA, T., PAPATHANAKOS, V., BANNER, A., KARATZAS, I. & FERNHOLZ, E.R. (2011) Hybrid Atlas Models. *Annals of Applied Probability,* to appear.

HARRISON, J.M. & WILLIAMS, R. (1987) Brownian models of open queueing networks with homogeneous customer populations. *Stochastics* **22**, 77-115.

HARRISON, J.M. & WILLIAMS, R. (1987) Multidimensional reflected Brownian motions having exponential stationary distributions. *Annals of Probability* **15**, 115-137.

WILLIAMS, R. (1987) Reflected Brownian motion with skew symmetric data in a polyhedral domain. *Probability Theory and Related Fields* **75**, 459-485.

KHAS'MINSKII, R. (1980) *Stochastic Stability of Differential Equations.* Sijthoff and Noordhoff, The Netherlands.

YAN, J.A. (1985) A formula for local times of semimartingales. *Northeast Mathematical Journal* **1**, 138-110.

OUKNINE, Y. (1990) Local times of the product and the maximum of two semimartingales. *Lecture Notes in Mathematics* **1426**, 477-479.

BANNER, A. & GHOMRASNI, R. (2008) Local times of ranked continuous semimartingales. *Stochastic Processes and Their Applications* **118**, 1244-1253.

BASS, R. & PARDOUX, E. (1987) Uniqueness for diffusions with piecewise constant coëfficients. *Probability Theory and Related Fields* **76**, 557-572.

PAL, S. & PITMAN, J. (2008) One-dimensional Brownian particle systems with rank-dependent drifts. *Annals of Applied Probability* **18**, 2179-2207.

CHATTERJEE, S. and PAL, S. (2010) A phase transition behavior for Brownian motions interacting through their ranks. *Probability Theory and Related Fields* **147**, 123-159.

CHATTERJEE, S. and PAL, S. (2011) A combinatorial analysis of interacting diffusions. *Journal of Theoretical Probability,* to appear.