MEMORANDUM

No 21/2013

Ross-type Dynamic Portfolio Separation (almost) without Ito Stochastic Calculus

Nils Chr. Framstad

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/econ
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 20/13	John K. Dagsvik, Steinar Strøm and Marilena Locatelli *Compensated Discrete Choice with Particular Reference to Labor Supply*
No 19/13	Erik Biørn *Age-Cohort-Time Effects in Sickness Absence: Exploring a Large Data Set by Polynomial Regression*
No 18/13	Halvor Mehlum and Gry Østenstad *The Political Economy of Migration Politics in Oil-rich Gulf States*
No 17/13	Halvor Mehlum, Ragnar Torvik and Simone Valente *China's Savings Multiplier*
No 16/13	Finn R. Førsund *Measuring Efficiency and Effectiveness in the Public Sector*
No 15/13	Mads Greaker and Kristoffer Midttømme *Optimal Environmental Policy with Network Effects: Is Lock-in in Dirty Technologies Possible?*
No 14/13	Diderik Lund and Ragnar Nymoen *Comparative Statistics for Real Options on Oil: What Stylized Facts to Use?*
No 13/13	Nils-Henrik M. von der Fehr *Transparency in Electricity Markets*
No 12/13	Nils Chr. Framstad *When Can Environmental Profile and Emissions Reduction Be Optimized Independently of the Pollutant Level*
No 11/13	Nils Chr. Framstad and Jon Strand *Energy Intensive Infrastructure Investments with Retrofits in Continuous Time: Effects of Uncertainty on Energy Use and Carbon Emissions*

Previous issues of the memo-series are available in a PDF® format at:
http://www.sv.uio.no/econ/english/research/memorandum/

ROSS-TYPE DYNAMIC PORTFOLIO SEPARATION, (ALMOST) WITHOUT ITÔ STOCHASTIC CALCULUS*

Memo 21/2013–v1

(This version September 11, 2013).

Nils Chr. Framstad†‡

Abstract: While it is common knowledge that portfolio separation in a continuous-time lognormal market is due to the basic properties of the Gaussian distribution, the usual textbook exposition relies on dynamic programming and thus Itô stochastic calculus and the appropriate regularity conditions. This paper shows how Ross-type distributions-based separation properties in continuous-time and discrete-time models, are easily inherited from a single-period model, generalizing and simplifying an approach of Khanna and Kulldorff (Finance Stoch. 3 (1999), pp. 167–185) down to multivariate distributions theory, stochastic dominance and the definition of the Itô integral. In addition to (re-) covering the classical cases of elliptical distributions (with or without risk-free opportunity) and symmetric α-stables/substables, this paper also gives separation results for non-symmetric stable returns distributions under *no shorting*-conditions, this including new cases of *one fund* separation without risk-free opportunity. Applicability of the skewed cases to insurance and banking is discussed, as well as limitations.

Key words and phrases: Portfolio separation, mutual fund theorem, elliptical distributions, (Lévy-Pareto) α-stable distributions, Lévy processes, stochastic dominance, portfolio constraints, incomplete markets, risk management.

MSC (2010): 91G10, 91G80, 60E07, 60G5∗∗, 93E20, 49K45.

JEL classification: G11, C61, D81, D52.

*The author gratefully acknowledges: This manuscript includes results which appeared in the author's doctoral dissertation, concluding a project funded by the Research Council of Norway. The work was initiated at the Stockholm School of Economics, a stay supported by NorFA; thanks to Tomas Björk for introducing me to the idea and the Khanna/Kulldorff approach.

†Address: Department of Economics, University of Oslo, Box 1095 Blindern, NO-0317 Oslo, Norway. `mailto:n.c.framstad+research@econ.uio.no`

‡Also affiliated with The Financial Supervisory Authority of Norway. The usual disclaimer applies.

1 Introduction

The concept of portfolio separation can be briefly formulated as follows: under what conditions can a market of a large number of investment opportunities be replaced by a few market indices («funds») without the investors being worse off? The prototypical result is two-fund separation (Tobin, [54]) which is the property that two market indices suffice; it is frequently referred to as two-fund *monetary* separation if one can choose one of them as an risk-free investment opportunity (i.e. with certain return). Tobin's work builds on the mean–variance approach by Markowitz [24]; as a historical note, Markowitz was predated by more than a decade by de Finetti [7] (see Pressaco and Serafini [41] and Markowitz' account [25] where he also credits Roy [45]). Tobin obtained the result in a single-period model with either Gaussian returns or quadratic utility, and since then, the results on portfolio separation have mainly fallen into these categories:

(i) Characterization of the agents who – for «any» market (some modeling paradigma of course has to be chosen) – are satisfied with a fixed set of funds, irrespective of their wealth. Separation in this sense is a property of a given preference (typically, a utility function for expected utility). The funds will be common to those agents who share this utility function, so that changes in wealth due to profits and losses will not change the funds, only the allocation among them – this assuming that preferences and parameters are invariant over time.

A characterization of those utility functions which admit two-fund separation for a single-period model, was given in 1970 by Cass and Stiglitz [5]. Cass–Stiglitz type portfolio separation is still a topic of research, see e.g. Schachermayer et al. [50] in a continuous-time model using modern probabilistic methods.

(ii) Characterization of the returns distributions for which «any» agent under consideration will be satisfied with the same (few) funds – again, under some model paradigma, like expected utility maximization.

A characterization of those returns distributions which yield k-fund separation over all (greedy) expected utility maximizers in a single-period model, was given in 1978 by Ross [44] in terms of *first-order stochastic dominance.* As it is not trivial which distributions admit the construction, this Ross-type portfolio separation is still a topic for research; this author's recent contribution [13] does in essence boil down to the work of verifying the Ross criteria.

Portfolio separation properties may also be deduced from other types of models or assumptions, e.g. through the application of *risk measures*, as pointed out by this author [11] and recently independently by De Giorgi et al. [15]. Ross-type two-fund separation is used to establish CAPM under the appropriate distributional assumptions, but one can for that purpose take the portfolio separation property as an assumption by itself, without founding that one in precise conditions on distributions or choice.

1.1 Ross-type portfolio separation

This paper is about the distributions-based portfolio separation property, but in dynamic models. The setup is a consumption–portfolio optimization problem where there are either insurance profits/losses modelled by additive strong Markov processes – Lévy processes in continuous time – or there are traded opportunities, whose prices (assumed exogenously given, so the agents are (small) price takers) are geometric processes with such additive strong Markov / Lévy processes as driving noise. The market is free from transaction costs, although certain constraints on the portfolios will be a main focus of interest. The returns distributions will be assumed elliptical, α-stable or an independen scaling of the latter (generalizing so-called sub-stability). The portfolio separation results will be valid for any expected utility maximizer which can satisfy the portfolio constraints; some of the constraints may implicitely require e.g. nonnegative net wealth.

As it turns out, portfolio separation in continuous time may be derived by way of an analogous discrete-time model of the same distributions. The literature on the single-period model is therefore highly relevant to this paper, and we mention a few references, both before and after the aforementioned characterization by Ross [44], showing that it has certainly been a bumpy road. When Tobin [54] established the property for the Gaussian distribution, he conjectured that any two-parameter distribution would do – counterexamples were given by Samuelson [48], Borch [2] and Feldstein [10]. In the meantime, the α-stable distributions case was treated already by Fama's 1965 paper [9] in the case with mutually independent and symmetric noise sources, and the result can also be recovered from the Samuelson's treatment [47]. The defining properties of those distributions led Cass and Stiglitz to suggest that they were necessary for monetary two-fund separation (in the «Ross» sense), a conjecture rectified by Agnew [1]. In 1983, Owen and Rabinovitch [40] showed that any *elliptically distributed* vector would satisfy the Ross conditions, while Chamberlain [6] established that under square integrability, these would be precisely those distributions for which *every* risk-averse agent would be a mean–variance-optimizer – hence tying the knot back to the original mean–variance approach (indeed, the key properties were to be found as far back as Schoenberg [51], [52] in 1938, before modern portfolio theory).

For the continuous-time setup, two-fund separation in a continuous-time complete geometric Brownian (lognormal) market was obtained by Merton [26] (cf. [27]) by means of dynamic programming, which has prevailed as the standard tool. Fast-forwarding nearly two decades, Khanna and Kulldorff [20] employed a technique which turns out to be a clever way of fitting the Ross-type criteria more directly to continuous-time models. The Khanna and Kulldorff approach does, by remarkably simple methods, remove the completeness (which, as it turns out, should hardly be any issue for the matter) and risk aversion assumptions from the Merton approach, and also allows for «no short positions» constraints on the opportunities or on a subset of them. Khanna and Kulldorff do however assume geometric Brownian prices and the existence of a risk-free investment opportunity.

This paper removes a few of the assumptions of Khanna and Kulldorf [20], simplifying the approach, covering both discrete and continuous time, and admitting more general portfolio restrictions and probability distributions. This includes cases without risk-free investment opportunity, also adaptable to different lending and borrowing rates, as well as a wider class of distributions for the underlying stochastic process, most notable the α-stable ones as treated by Fama and Samuelson, although in more generality; we allow for stochastic dependence between the investment opportunities, which in the non-Gaussian case is a highly non-trivial matter. Due to some confusion in the literature, it should be noted that iid coordinates does not imply ellipticity (nor the converse), except the Gaussian case. Ellipticity means – modulo an affine transformation – rotational invariance (see e.g. [46, Proposition 2.5.2]), this fact answering an issue raised by Owen and Rabinovich [40, footnote 4] (see also e.g. also Ortobelli et al. [37] and [38], focusing on the elliptical stable distributions). We shall see that under no shorting restrictions, even the *skewed* α-stable distributions admit separation – however, a condition of «common skewness» must (usually) be satisfied, effectively constraining the portfolio except for the symmetric laws. It is common to disregard the skewed distributions, like in Bradley and Taqqu [3]; Fama makes a remark that the separation result still holds if the agents disregard skewness – however, that might be highly unreasonable, as in the most extreme cases it implies indifference between a positive and a negative excess return. In our case, the appropriate portfolio restrictions will lead to further separation results, including a *one fund separation* theorem which to the best of this author's knowledge is unpublished.

1.2 α-stable laws in empirical and practical applications

Although portfolio separation theorems are historically considered to be in the pure theory side of finance, the α-stable distributions have indeed been used as modeling alternative to the Gaussian – being precisely the distributions obtainable from the generalized central limit theorem, they are natural candidates when the Gaussian exhibits a too light tail. Historical references indicating tails heavier than the Gaussian in financial returns include works by Mandelbrot and Fama (see the collection [21, Chapters E1, E14 and E16]). Further contributors on α-stable distributions in finance include Rachev, Mittnik and coauthors ([32], [33], cf. also the collection [43]). It should be noted though, that while this paper, like Fama [9] and Ross [44], consider absolute returns, the empirical studies cited above concern log-returns. Unlike the Gaussian-noise case, where the exponential of a Brownian motion with drift and a geometric differential equation with Brownian noise, both lead to Gaussian log-returns, a geometric differential equation with stable driving does not yield stable log-returns. Works relating absolute returns to stable distributions include Mantegna [22], estimating α-values ranging from 1.00 ± 0.04 to 1.40 ± 0.04 for returns on the Milan stock exchange (assuming symmetry, hence zero expected return, which may be considered objectionable). Janicki et al. [19] treat the problem of pricing options on assets following geometric differential equations (as we do) with symmetric stable input noise, to explain the smile effect. As for skewness, empirical studies have indicated that skewnesses do vary considerably, even for returns distributed with close to similar α's, see e.g. [34] and [31].

The non-Gaussian α-stables do certainly have questionable properties; not only are they unbounded like the Gaussian (meaning e.g. no lower bound on absolute return), they also lack moments of order α and above. Furthermore, the associated geometric Lévy process changes sign, unlike the geometric Brownian, and could lead to losses exceeding the holdings in the opportunity (contrary to limited liability, but certainly desirable for other investments!). However, as long as the model is not taken to such an extent that objectionables are *applied,* they need not be any issue. For example, quantile measures are widespread in financial risk management (the so-called «value-at-risk»), despite theoretical shortcomings (e.g. [53] and the references therein). A quantile completely disregards the portfolio return distribution's far tail, and objections based on unboundedness (for the Gaussian) and lack of moments (for the α-stables) are less crucial. Indeed, there is no reason for a limited liability firm to worry over the distributional properties beyond bankruptcy per se, only to the extent that it affects e.g. their funding costs (as lenders may care for the losses given default).

1.3 Organization of the paper, and notation

The model for portfolio value will be a controlled Itô stochastic process of the form

$$\mathrm{d}Y(t) = \boldsymbol{v}(t)^\top \big[\boldsymbol{\mu}(t)\ \mathrm{d}t + \Sigma(t)\ \mathrm{d}\boldsymbol{Z}(t)\big] - \mathrm{d}C(t) \qquad \text{(for } t \geq t_0\text{, with } Y(t_0) \text{ given)} \qquad (\star)$$

and section 2 will justify this from the standard points of view (insurance profits and loss, or consumption-financing trading portfolio) – without any rigor, as it is textbook material, at least if $\boldsymbol{Z}$ were assumed to be a Brownian motion. Then it will review (subsection 2.1) the concept(s) of first-order stochastic dominance, then motivate which of them to use in the dynamic setup. Section 3 will then give a model with full assumptions, and the main proof technique to be applied in the subsequent two sections; section 4 and 5 will establish the separation results for the elliptical, resp. the α-stable laws. Section 6 will discuss the differences between these distribution classes' separation properties, discuss applicability and limitations of the new results obtained in the paper, and conclude.

To clarify the title, which promises an exposition with hardly any Itô calculus, we shall employ the Itô *integral* as the limit in probability of step integrals. We shall do without the calculus that involves the second-order term of the Itô change of variables rule, which in textbook expositions like e.g. [28] or [36] enters in the second-derivative term(s) in the Hamilton–Jacobi–Bellman equation – analogously, we will avoid the change of variable rule for α-stable Lévy processes (which in the integrable case requires a compensation term and cannot readily be treated using a piecewise ordinary chain rule). We shall from formula (1) alter the «consumption» choice process in order to avoid the change of variable, but that would be doable without the second-order term as long as we stick to the standard assumption of locally risk-free numéraire. In subsection 4.3 we will however need an auxilliary market state $M(t)$ to be well-defined in terms of a stochastic differential *equation* ($\star\star$), and thus, in the canonical existence and uniqueness proof by way of Picard iteration, a quadratic variation term from Itô calculus.

The paper will employ the following notation: Matrices will be denoted with capital Greek letters. Superscript $^\top$ denotes transpose. Vectors – by default columns – will be denoted in boldface ($\mathbf{1}$ will be the vector of ones), and only random vectors and vector-valued stochastic processes will be boldface capitals – however, the portfolio $\boldsymbol{v}$ which can be controlled at any time, in minuscle. The reader can note the agent can not influence anything in Greek letters unless specified as a function of choice variables. Stochastic processes will be denoted with or without the time argument. We shall use X or $\boldsymbol{X}$ for the generic random variable (not for processes!), and x or $\boldsymbol{x}$ for the generic free variable for e.g. static optimization. The $\sim$ symbol will denote equality in probability law (i.e. distribution, or for processes: finite-dimensional distributions). Sets will be denoted in blackboard bold, with $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$ denoting the closed positive, resp. closed negative, orthant. The differential « d» always denotes Itô type differential – there will be no other concept of stochastic integration.

We shall use the following terminology: a random vector $\boldsymbol{X}$ is *symmetric* iff it is distributed like $-\boldsymbol{X}$, and *symmetric about* $\boldsymbol{\mu}$ iff $\boldsymbol{X}-\boldsymbol{\mu}$ is symmetric. As negative (semi-) definiteness will not be of any relevance to this paper, we shall use *semidefinite* for positive semidefinite matrices, and *merely semidefinite* for matrices that are positive semidefinite but not positive definite. We shall denote the positive semidefinite square root of a semidefinite Γ by $\Gamma^{1/2}$. A set $\mathbb{V}$ is *radial* if $\forall a>0,\ a\boldsymbol{x}\in\mathbb{V}$ iff $\boldsymbol{x}\in\mathbb{V}$.

2 The tradet market and insurance models

Central to our analysis will be the controlled stochastic process Y given by ($\star$). We shall outright *assume* this model from the next section on, but will give a non-rigorous discussion here. The reader familiar with the basics of mathematical finance – not restricted to geometric Brownian motions which remain positive! – can easily skip the rest of this section.

The arguably simplest model leading to ($\star$), is based on insurance policies/subportfolios with profits and loss flows $\mathrm{d}\breve{G}_i$, scaled up by the agent's exposure u_i; these are called *risky* opportunities. With $i=0$ denoting the opportunity of a *risk-free* investment with interest ρ, and extracting dividends or consumption, cumulatively denoted $\breve{C}$, the agent's wealth will develop according to

$$\mathrm{d}\breve{Y}(t)=\boldsymbol{v}(t)^\top\,\mathrm{d}\breve{\boldsymbol{G}}(t)+(\breve{Y}(t)-\boldsymbol{v}^\top\mathbf{1})\rho\,\mathrm{d}t-\,\mathrm{d}\breve{C}(t) \tag{1}$$

where we have stacked up the opportunities from $i=1$ in vector form, and put the remaining amount $\breve{Y}-\boldsymbol{v}^\top\mathbf{1}$ risk-free. Now if we represent $\breve{\boldsymbol{G}}$ as $\mathrm{d}\breve{\boldsymbol{G}}(t)=\breve{\boldsymbol{\mu}}\,\mathrm{d}t+\Sigma\,\mathrm{d}\boldsymbol{Z}$ (motivated by the usual approach of assuming $\boldsymbol{Z}$ to be a martingale – this assuming integrability conditions which may not hold in this paper!), we arrive at ($\star$) by putting

$$\mathrm{d}C=\,\mathrm{d}\breve{C}-\rho\breve{Y}\,\mathrm{d}t,\qquad\text{and}\qquad\mathrm{d}\boldsymbol{G}(t)=\boldsymbol{\mu}\,\mathrm{d}t+\Sigma\,\mathrm{d}\boldsymbol{Z}\qquad\text{with}\quad\boldsymbol{\mu}=\breve{\boldsymbol{\mu}}-\rho\mathbf{1} \tag{2}$$

and allowing C to be subject to the agent's choice. Thus we avoid the theory of stochastic differential *equations* – this is just a controlled Itô process. Furthermore, working with excess drifts $\boldsymbol{\mu}$ removes ρ from the model, and we shall treat cases where no risk-free opportunity exists – corresponding to imposing $\boldsymbol{v}^\top \mathbf{1} = Y$ for each agent. Also, discounting away the ρ accounts for the insignificance of the common additive risk component found elsewhere in the literature, e.g. the «y» in Ross [44, Theorems 1 to 3].

For applicability, the insurance model could reasonably have G_i evolving quite asymmetrically, for example with large (but rare) negative jumps – with amplitude possibly < -1, leaving the agent to cash out the loss. However, we shall assume that the portfolio lives on memoryless (strongly Markov) – the agent can only choose her exposure to it.

Now for a traded model, one assumes the investment opportunities having prices P_i. With N_i units held at time t, at market value $N_i P_i$, one assumes the self-financing property, *defined* by the ansatz that the value of the position evolves as $N_i(t)\,\mathrm{d}P_i$ (we shall assume N left-continuous, i.e. $N_i(t) = N_i(t^-)$). With the canonical form $\mathrm{d}P_i(t) = P_i(t^-)\,\mathrm{d}\breve{G}_i$ for price fluctations, market value of the position obeys $N_i(t)P_i(t^-)\,\mathrm{d}\breve{G}_i$, and putting $v_i = N_i P_i$ – the value exposed – we can aggregate up to (1), again assuming a risk-free opportunity at interest rate ρ. Formally, we are done. However, if G_i may have jumps of amplitude < -1, we are in a situation which violates limited liability. That is not a problem per se, as we have not made any assumption of limited liability, but one may for example ask: if G_i develops asymmetrically, yielding for P_i a high positive rate of return between jumps, but jumps by, say, -200 percent to $P_i(t) = -P_i(t^-)$ – is it reasonable that keeping holding the position will now incur high expenses but with possibly large positive jumps? The interpretation that resolves this, is to imagine it works like the insurance case: an event triggers a claim on you, and you unwind the position; however, the investment opportunity remains, memoryless, and you can immediately choose to invest in it again.

It should be noted, however, that the investment opportunity would have to live on after such jumps in order not to interfere with the separation properties this paper is all about. If it were not, then the opportunity set would change in a way that would be correlated with the returns, and the agent would have to optimize with regard to the joint distribution of return and change in optimization problem. That model feature is prone to destroy portfolio separation properties (cf. [12]).

2.1 Stochastic dominance and the Ross argument

This section will review first-order stochastic dominance (the key argument of Ross [44]) for random variables and point out what is the natural generalization to stochastic *processes*, for the purposes of this paper.

Recall that a real-valued random variable X^* first-order (weakly) stochastically dominates another, X, if any of the three following equivalent criteria hold:

(i) there exists some nonpositive random variable X_- such that $X \sim X^* + X_-$.

(ii) $\mathrm{CDF}_X \geq \mathrm{CDF}_{X^*}$

(iii) $\mathsf{E}[u(X^*)] \geq \mathsf{E}[u(X)]$ for every bounded nondecreasing function u.

A recent article by Østerdal [39] elaborates on the various concepts in more general frameworks. An ordering of probability distributions «preferring more to less» should rank X^*

at least as good as X; conversely, if neither of X and X^* first-order dominates the other, then there exist two expected utility maximizers which rank them oppsitely – indeed, by (ii) there are numbers $\bar{x}_1$ and $\bar{x}_2$ such that $1_{x \geq \bar{x}_1}$ and $1_{x \geq \bar{x}_2}$ disagree on the orderings. Note that if we were only to require (iii) to hold for *concave* functions, we would be led to what Ross [44] calls *weak* separation. However, concave functions are inconvenient for our purposes, as we consider possibly distributions with non-integrable negative part, for which expected concave utility must diverge.

2.2 Heuristics on dominance and separation in a single period

This subsection will review how stochastic dominance leads to separation in a single-period model, without specifying the model in full detail; the assumptions needed for proper theorems will be given in section 3, and for the purpose of the illustration, the reader can imagine returns being multinormal. For a single period, consider ($\star$) with constant coefficients on time $[0, 1]$, constrained to the portfolio being constant and chosen at time 0, and consumption only at time 0, consuming $C(0)$ and at time 1 where $Y(1)$ is consumed so that $C(1) = C(0) + Y(1)$. Suppose that for two portfolios $\boldsymbol{v}^*$ and $\boldsymbol{v}$ we have

$$\boldsymbol{v}^{*\top} \Sigma[\boldsymbol{Z}(1) - \boldsymbol{Z}(0)] \sim \boldsymbol{v}^\top \Sigma[\boldsymbol{Z}(1) - \boldsymbol{Z}(0)] \qquad \text{and} \qquad \boldsymbol{v}^{*\top}\boldsymbol{\mu} \geq \boldsymbol{v}^\top \boldsymbol{\mu}. \tag{3}$$

Now define C^* by $C^*(0) = C(0)$. Then any greedy expected utility maximizer will prefer to use $(C^*, \boldsymbol{v}^*)$ over $(C, \boldsymbol{v})$, as the return from using the portfolio $\boldsymbol{v}^*$ will dominate the return from using $\boldsymbol{v}$ (just choose $X_- = -(\boldsymbol{v}^* - \boldsymbol{v})^\top \boldsymbol{\mu}$); the agent can consume the same initially, and more – in distribution – at time 1. (Note the reason why we do not consume the excess up front; that would require borrowing before it accumulates up at the end of the period, and that will violate some of the portfolio restrictions we will treat later.)

Let us see how this works under the assumption of jointly Gaussian returns. Without loss of generality, assume $\boldsymbol{Z}$ to have zero mean and let $\boldsymbol{\mu}$ take care of the expected return – then the first part of (3) will hold if $\boldsymbol{v}^*$ and $\boldsymbol{v}$ have the same seminorm $(\boldsymbol{v}^{*\top} \Gamma \boldsymbol{v}^*)^{1/2} = (\boldsymbol{v}^\top \Gamma \boldsymbol{v})^{1/2}$, where Γ is the covariance matrix of $\Sigma[\boldsymbol{Z}(1) - \boldsymbol{Z}(0)]$. Assuming Γ to be positive definite, then monetary separation stems from the fact that if the risky fund $\boldsymbol{f}$ is chosen to maximize $\boldsymbol{f}^\top \boldsymbol{\mu}$ subject to $\boldsymbol{f}^\top \Gamma \boldsymbol{f} = 1$, then for every $\boldsymbol{v}$, the return using $(\boldsymbol{v}^\top \Gamma \boldsymbol{v})\ \boldsymbol{f}$ will dominate. That means that anyone who orders known probabilities consistent with first-order stochastic dominance – in particular, every greedy expected utility maximizer for which expectation converges – can do with investing only in $\boldsymbol{f}$ (common to all agents) and the risk-free opportunity, without any welfare loss. (As a note: it is common in the literature to minimize variance for given value of drift; that gives a nice interpretation of the Lagrange multipliers, and disposes of a one-line argument on the constraint qualification – but makes an assumption of risk aversion, which is unnecessary.)

Now the above construction yields (C^*, Y^*) equivalent in law to $(C, Y) + (K, 0)$ where $\mathrm{d}K \geq 0$. This exhibits the mass transfer definition (i) of stochastic dominance as the one to apply in the dynamic model. To that end, we shall first establish the dominance property in finitely many periods; those step strategies form a subset of (Itô) integrands, and we can pass to the limit for continuous time. The case of discrete time with infinitely many periods could be established by a similar procedure using somewhat simpler mathematical machinery, but we will not go that route – that case will automatically be covered as a special case of continuous time, by restricting to infinite-step strategies.

3 The dynamic model

The modelling framework will be a filtered probability space, with right-continuous filtration complete at initial time t_0. This formal framework will be notationally suppressed throughout the paper, using terms like «t_d-conditionally» to refer to the sigma-algebra of events known at time t_d – and also even «τ-conditionally» for conditioning on the events known at a stopping time τ. All differentials will be notation for Itô type stochastic integrals; whenever needed, we shall apply the full semimartingale-based definition as in Protter [42], in terms of the dense set of predictable simple (that is, with a nonstochastic bound on the number of steps!) integrands and semimartingale integrators. In order to apply Riemann sums, we shall for convenience make a standing assumption of piecewise continuity of the sample paths:

3.0.1 Assumption. All processes will be assumed adapted to the underlying filtration, all stochastic integrators – in particular the driving semimartingale noise $\boldsymbol{Z}$ – are right-continuous with left-hand limits, while all integrands are predictable (in particular, the control) and left-continuous with right-hand limits. □

On this space, we shall take as model the controlled (through *cumulative consumption* C and $\mathbb{R}^n$-valued *portfolio* $\boldsymbol{v}$) stochastic process

$$\mathrm{d}Y(t) = \boldsymbol{v}(t)^\top \left[\boldsymbol{\mu}(t)\ \mathrm{d}t + \Sigma(t)\ \mathrm{d}\boldsymbol{Z}(t)\right] - \mathrm{d}C(t) \qquad \text{(with } Y(t_0) \text{ given)} \qquad (\star)$$

where the driving noise $\boldsymbol{Z}$ takes values in $\mathbb{R}^m$, $\boldsymbol{\mu}$ in $\mathbb{R}^n$, while Σ takes matrix values in $\mathbb{R}^{n\times m}$.

As we shall utilize the limit transition from discrete to continuous time in the definition of the Itô integral, we shall consider both discrete and continuous time for the driving process $\boldsymbol{Z}$. Combinations of discrete and continuous dynamics will only introduce unnecessary notational unconvenience, and is left to the reader. Any (adapted) process living on discrete time, is a semimartingale when considered as objects on continuous time, and so the differential is indeed an Itô integral on continuous time as well – but compared to simply constraining the continuous-time case, the discrete-time assumptions do then do then allow for fixed jump times and for distributions which are not infinitely divisible:

3.0.2 Assumption. $\boldsymbol{Z}$, always common to all agents, is a strong Markov process with increments independent of the past and present; if the agent is permitted to trade on continuous time, we also assume it to be a Lévy process. We shall say that the process *lives on* a subset of $[t_0, \infty)$ if those are the only times under consideration and $(C, \boldsymbol{v})$ is required to be constant outside these times. Phrases like *all times*, *every time*, *each time*, etc. will only refer to the time set the process lives on. If the process does not, live beyond a (stopping) time τ, that is referred to as the *horizon.* □

The following assumption of constant coefficients will be applied in proofs, as the time-dependent version can be approximated by piecewise constants:

3.0.3 Assumption. The market coefficients $\boldsymbol{\mu}$ and Σ, and $\mathbb{V}$ defined in Defnition 3.0.4 below, will be nonrandom and constant in time, and common to all agents.
This assumption can however be relaxed to allow for deterministic time-dependency. □

The possibility that the market or opportunity set need not be the same for everyone, covers for example cases where there is no risk-free opportunity ($\boldsymbol{v}^\top \mathbf{1} = Y$, where Y is individual) or where interest rates are different, for example that agents who wish to be net borrowers could face a higher rate (cf. Corollary 4.2.5). It should be noted that while Definition 3.0.4 does not restrict the dynamics of the set-valued process $\mathbb{V}_t$, the next section will assume specific forms that enable the separation results.

Now we make assumptions on the opportunity set; recall that we only consider the times the process lives on.

3.0.4 Definition. Consider for each t a given nonempty *control region* $\mathbb{V} = \mathbb{V}_t \subseteq \mathbb{R}^n$, henceforth also referred to as a *portfolio constraint*; the *unconstrained* case will refer to when $\mathbb{V} = \mathbb{R}^n$. A *strategy* is a predictable $\mathbb{R}^{n+1}$-valued stochastic process $(C, \boldsymbol{v})$ such that $(\star)$ is well-defined as an Itô integral. *No shorting* shall mean that $\mathbb{V} \subseteq \mathbb{R}^n_+$. Implicitely relaxing Assumption 3.0.3 by wealth being individual, *no borrowing* shall mean that $\boldsymbol{v}^\top \mathbf{1} \leq Y$ on $\mathbb{V}$ and *no risk-free opportunity* shall mean that $\boldsymbol{v}^\top \mathbf{1} = Y$ on $\mathbb{V}$. □

Now for the agents' choices, the following assumption adapts the stochastic dominance concept to our framework – although with a minor simplification which for our purposes turns out only to be an insignificant technical issue for the very particular case of Theorem 5.2.6 at the end of section 5. Again, recall that we only consider the times the process lives on:

3.0.5 Assumption. Each agent's set of preferences form a partial ordering for each stopping time $\tau \geq t_0$, predictable, such that whenever there is a predictable nondecreasing K for which two wealth–consumption pairs (Y^*, C^*) and (Y, C) satisfy

$$\{(Y(\,\cdot\,), C(\,\cdot\,) - C(\tau) + K(\,\cdot\,) - K(\tau))\}_{t \geq \tau} \sim \{(Y^*(\,\cdot\,), C^*(\,\cdot\,) - C^*(\tau))\}_{t \geq \tau} \tag{4}$$

for the τ-conditional probability laws – then the agent (weakly) prefers, at time τ, the wealth–consumption pair $\{(Y^*, C^*)\}_{t \geq \tau}$ to $\{(Y, C)\}_{t \geq \tau}$. □

In words, this means that an agent will accept an offer of additional future consumption (i.e. higher $\mathrm{d}C$, not merely higher C) for the same future wealth (i.e. Y) – this interpreted in probability law, and *viewed from any stopping time* τ. We shall largely suppress τ in the notation. Furthermore, we shall allow for the ambiguity of terminology in ordering strategies, inherited from the wealth–consumption process laws:

3.0.6 Definition. Consider two strategies $(C^*, \boldsymbol{v}^*)$ and $(C, \boldsymbol{v})$ with corresponding wealth processes Y^* and Y. We say that (Y^*, C^*) is *preferred to* (Y, C) (i.e. stated without regard to τ), if it is preferred at any stopping time $\tau \geq t_0$. In this case we also say that $(C^*, \boldsymbol{v}^*)$ is *preferred to* $(C, \boldsymbol{v})$. □

Free disposal can be interpreted as a special case, re-interpreting «consumption» accordingly. The term *optimal* we will used loosely whether as to mean a solution to an optimization problem at a given time, or for a given agent.

3.0.7 Remark. *Risk aversion* is not a main focus of this paper, but will be mentioned to relate some of the results to literature based on risk-aversion. However, as we consider possibly non-integrable distributions, we cannot readily employ the definition in terms of rejecting any conditionally zero-mean noise. The natural analogue for our purposes, will be rejecting any conditionally symmetric (about zero) noise, i.e. weakening $Y \sim Y^*$ into the property $\mathrm{d}Y \sim \mathrm{d}\bar{Y} + \mathrm{d}\tilde{Y}$ and $\mathrm{d}Y \sim \mathrm{d}\bar{Y} + r(t)\,\mathrm{d}\tilde{Y}$ where $\mathrm{d}\tilde{Y} \sim -\,\mathrm{d}\tilde{Y}$ is conditionally independent and $|r| \leq 1$ (everything viewed τ-conditionally, all stopping times τ). This will become more intuitive after Proposition 3.1.2, but we shall keep risk-aversion on an intuitive level and leave rigor to the interested reader. □

Note that in Definition 3.0.4, we can do without the common assumption that ($\star$) should possess a «solution» – as the strategies are merely stochastic processes, not feedback functions, there is no Y on the right-hand side. Yet, the term *strategy* as above, is not sufficiently restrictive to be applied to finance, if the full definition of the Itô integral is applied; then doubling strategies will lead to arbitrages (cf. Dudley's theorem [8] which shows that the Itô integral over Brownian motion could evaluate to any random variable measurable wrt. the sigma-algebra generated by the Brownian path – on a related note, see also subsection 3.2 concerning free lunches with vanishing risk). Thus one will in applications make some suitable kind of restriction admit only sufficiently «tame» portfolios. Portfolio separation is however not about solving for optimal strategy, but about restricting without welfare loss the opportunity set, and if an admissibility restriction rules out certain strategies in the first place, they simply need not be removed by the separation property. We shall therefore merely specify a property that an «admissibility» restriction should possess, in order to make sure that, cf. Definition 3.0.6, the strategy $(C^*, \boldsymbol{v}^*)$ is admissible *whenever* $(C, \boldsymbol{v})$ *is:*

3.0.8 Assumption. Given a restriction of strategies (in the sense of Definition 3.0.4) to a subset called *admissible* (possibly individual to each agent), the restriction is assumed to have the property that whenever $(C, \boldsymbol{v})$ is admissible, $C^* \sim C + K$ for some nondecreasing K and the processes $\int_{t_0}^t \boldsymbol{v}^\top \Sigma\,\mathrm{d}\boldsymbol{Z}$ and $\int_{t_0}^t \boldsymbol{v}^{*\top} \Sigma\,\mathrm{d}\boldsymbol{Z}$ have the same law then $(C^*, \boldsymbol{v}^*)$ is admissible as well. □

The reason for not specifying one single admissibility concept, is that we wish to establish a theorem valid for several of the definitions applied in the literature to ensure «tame» portfolios that do not lead to doubling strategy arbitrages. For example, the classical L^2 framework for the Itô integral is maybe the most technically convenient for Brownian motion, but not so when we shall consider even non-integrable driving noise – then we cannot even ad hoc assume the martingale property. And the assumption of an a.s. lower bound on wealth is natural with continuous processes, but requires more thought with unbounded jumps. Therefore, this paper is better off leaving admissibility flexible.

It will not be a problem for us to consider a strategy which in some application will be considered «not admissible»; the preference assumptions to be made next, will enable us to compare pairwise to falsify optimality of given strategies – which makes it uninteresting for *any* Assumption 3.0.8-compliant admissibility concept.

For this reason, we do not need to worry that the following assumption can not eliminate doubling strategy arbitrages, only immediate single-periode «static» arbitrages – a term that will be used subsequently – and also redundant opportunities. By the strong Markov property, we need not invoke the general approach of stochastic intervals where the agent waits for a stopping time when a free lunch occurs:

3.0.9 Assumption. The market parameters and the control region $\mathbb{V}$ are assumed to be such that, no non-null admissible strategy with $\mathrm{d}C = 0$ has the property that $\mathrm{d}Y$ (a.s., on a positive set in time) is

- ≥ 0 if there is risk-free opportunity
- ≥ 0 for any ρ if not.

Furthermore, we assume that if there is no risk-free opportunity and $\boldsymbol{\mu} = \mu_0\mathbf{1}$, then $\mu_0 = 0$. □

The last part shall turn out convenient for those degenerate but classical cases applicable, and the reader can verify that it represents no loss loss of generality for the purposes of this paper. Notice that the second bullet point ensures that no «free lunch» can be scaled up to a.s. match every fictitious interest rate. It does not make sense to call it an «arbitrage» just because the economy admits a «yardstick» positive minimum return (e.g. due to growth). This will only be relevant in Theorem 5.2.6, which is arguably not the most interesting parameter range. Notice furthermore the rationale behind forbidding even Y constant on an interval: if this is generated by a non-null portfolio, then this is a risk-free opportunity. If no such is assumed to exist, then that very assumption is violated; if there is on, then we have another and hence a redundant opportunity, and we should rebuild the model without it.

Thus far, the term «agent» has been used loosely. We now restrict the use of the phrase:

3.0.10 Definition. From now on, the term *agent* will be reserved to those for which an admissible strategy exists (cf. Assumption 3.0.8 and Definition 3.0.4), and whose preferences satisfy Assumption 3.0.5. □

This restriction may be substantial if for example if $\mathbb{V}$ is defined by no shorting and no borrowing; with positive prices, then $Y \geq 0$ is necessary in order to be an «agent». We simply assume everyone else to be removed from the market.

We are now ready to define separation. Let us note that we do not assume existence of an optimal strategy (only, tautologically, one admissible for each agent). In return, we have to stick to a slightly weaker concept of portfolio separation than usual, but in the presence of an optimal strategy, it coincides with the usual definition:

3.0.11 Definition (k-fund separation)**.** Suppose that there exist $k-1$ predictable $\mathbb{R}^n$-valued $\boldsymbol{f}_1(t), \ldots, \boldsymbol{f}_{k-1}(t)$ (common to all agents), such that for each strategy $(C, \boldsymbol{v})$ and each agent for whom this strategy is admissible, there exists an admissible $(C^*, \boldsymbol{v}^*)$ which every agent prefers to $(C, \boldsymbol{v})$, and with $\boldsymbol{v}^*(t)$ spanned by $\boldsymbol{f}_1(t), \ldots, \boldsymbol{f}_{k-1}(t)$ for every t. Then we say that the market admits

- $k-1$ *fund separation* if $\boldsymbol{v}^*(t)^\top \mathbf{1} = Y(t)$ identically for every agent, and
- k *fund separation* – in this case also k *fund monetary separation* – otherwise. □

This is Ross-type portfolio separation – a property not of the agents' specific utility functions, but one of the market and the probability law of the underlying process $\mathrm{d}\boldsymbol{G} = \boldsymbol{\mu}\ \mathrm{d}t + \Sigma\ \mathrm{d}\boldsymbol{Z}$.

3.0.12 Remark. A few notes:

(a) Definition 3.0.11 does not require the funds to be linearly independent – k-fund separation is a special case of $k+1$-fund separation (although it is vacuous to speak of n fund separation in a market with n opportunities including if applicable the risky).

(b) The way Definition 3.0.11 is formulated, it allows for the funds to vary, even stochastically as function of the state of the market – this is what Merton [28, section 15.7] refers to as *generalized separation,* to be covered in subsection 4.3.

(c) We shall also speak of «separation over risk-averse agents» in the same loose manner as of risk aversion. □

3.1 The main proof technique

We now give a key tool to the proofs for portfolio separation. We state it in terms of finite-step integrands on finite horizon – more specifically, *simple predictable* integrands – but we are going to *apply* it by starting with some given portfolio $\boldsymbol{v}$ for which the Itô integral ($\star$) is well-defined; convergence of the approximation will hold by the very definition of the Itô integral as the limit in probability of sums over the steps of a simple predictable; our integrators are assumed to be semimartingales, which form precisely the class of «good integrators», see Protter's book [42]. (As a note: This functional analysis view of semimartingales was the «new approach» subtitle of the first edition of the book. Arguably, the book has helped establish this approach so much that it was only apt to drop the subtitle. The machinery is somewhat hi-tech compared to the Wiener–Poisson integral, and the reader who wants merely a brief introduction to the idea, could read the AMS Bulletin review [55] of the first edition.)

The following might require $\mathbb{V}$ to be the entire space, but in subsequent applications we shall apply it merely for step strategies that *approximate* admissible portfolios. Note that Assumption 3.0.8 makes the strategies either both admissible or none of them.

3.1.1 Lemma. *Fix a $D \in \mathbb{N}$, a bounded horizon τ_D and a partition $\tau_D \geq \tau_{D-1} \geq \cdots \geq \tau_1 \geq \tau_0 = t_0$, relax Assumption 3.0.3 to coefficients constant on each (τ_{d-1}, τ_d), and impose* *the restriction that on these open intervals, $\boldsymbol{v}$ is constant. Fix a strategy $(C, \boldsymbol{v})$. Then for any portfolio process $\boldsymbol{v}^*$ with*

$$\Big\{\boldsymbol{v}^*(\tau_d)^\top \Sigma[\boldsymbol{Z}(\tau_{d+1}) - \boldsymbol{Z}(\tau_d)]\Big\}_{d=0,\ldots,D-1} \sim \Big\{\boldsymbol{v}(\tau_d)^\top \Sigma[\boldsymbol{Z}(\tau_{d+1}) - \boldsymbol{Z}(\tau_d)]\Big\}_{d=0,\ldots,D-1} \tag{5}$$

and there are τ_d-measurable $K_d \geq 0$, $d = 0, \ldots, D-1$ such that

$$\Big\{\boldsymbol{v}^*(\tau_d)^\top \boldsymbol{\mu}(\tau_d)\Big\}_{d=0,\ldots,D-1} \sim \Big\{\boldsymbol{v}(\tau_d)^\top \boldsymbol{\mu}(\tau_d) + K_d\Big\}_{d=0,\ldots,D-1} \tag{6}$$

then is a C^* with such that every agent prefers $(C^*, \boldsymbol{v}^*)$ to $(C, \boldsymbol{v})$; in continuous time, $\mathrm{d}C^* \sim \mathrm{d}C^* + K_d\,\mathrm{d}t$, while if the process only lives on the partition, the excess K_d is consumed at τ_{d+1}.

The proof uses the characteristic function $\mathsf{E}\exp\int_{t_0}^{\infty} \mathrm{i}\vartheta(t)Y(t)\,\mathrm{d}t$ of the process. This function takes as argument any sufficiently nice (smooth, bounded, compactly supported) *deterministic* function ϑ. Now for the vector-valued process $\Sigma\,\mathrm{d}\boldsymbol{Z}$ we would have used a vector-valued function, still deterministic; the key property in passing to a merely predictable $\vartheta\boldsymbol{v}$, is the strong Markov property. Note that in order to prove the equivalence of process pairs (C, Y) and (C^*, Y^*), we could use the characteristic function of that pair; however, having proved equivalence up to t_d, we will see that we can just proceed one step forward with Y only, as controls are predictable:

Proof. With horizon τ_D, and agents having preferences over terminal wealth, we can without loss of generality disregard terminal consumption. Suppose first $D = 1$. If (5) holds at initially, then put $K_0 = (\boldsymbol{v}^*(\tau_0) - \boldsymbol{v}(\tau_0))^\top \boldsymbol{\mu}(\tau_0)$, as the quantities are known (τ_{D-1}-measurability is the relevant property to generalize). Now use induction on D: Assuming true for a given D, then for $D+1$ consider the characteristic function of the process Y. Rewrite $\mathsf{E}\exp\int_{\tau_0}^{\tau_{D+1}} \mathrm{i}\vartheta(t)Y(t)\,\mathrm{d}t$ as

$$\mathsf{E}\exp\left(\mathrm{i}\int_{\tau_0}^{\tau_{D+1}} \vartheta(t)Y(t\wedge\tau_D)\,\mathrm{d}t + \mathrm{i}\int_{\tau_D}^{\tau_{D+1}} \vartheta(t)\Big[Y(t) - Y(\tau_D)\Big]\,\mathrm{d}t\right) \tag{7}$$

The first integral is τ_D-measurable. Condition and consider

$$\mathsf{E}_{\tau_D}\exp\left(\mathrm{i}\int_{\tau_D}^{\tau_{D+1}} \vartheta(t)\Big[Y(t) - Y(\tau_D)\Big]\,\mathrm{d}t\right) \tag{8}$$

$$\begin{aligned} = \ &\exp\left(\mathrm{i}\int_{\tau_D}^{\tau_{D+1}} \vartheta(t)\Big[C(\tau_D^-) - C(t) + (t-\tau_D)\boldsymbol{v}(\tau_D)^\top\boldsymbol{\mu}(\tau_D)\Big]\,\mathrm{d}t\right) \\ &\times \mathsf{E}_{\tau_D}\exp\left(\mathrm{i}\int_{\tau_D}^{\tau_{D+1}} \vartheta(t)\boldsymbol{v}(\tau_D)^\top\Sigma(\tau_D)\Big[\boldsymbol{Z}(t) - \boldsymbol{Z}(\tau_D)\Big]\right)\mathrm{d}t \end{aligned} \tag{9}$$

For the first line of (9), we note that the excess to consume in continuous time is the flow $\big(\boldsymbol{v}^*(\tau_D) - \boldsymbol{v}(\tau_D)\big)^\top\boldsymbol{\mu}(\tau_D)\,\mathrm{d}t$, which is positive by assumption; in discrete time, consume the excess at the end of the period except the last. By the strong Markov property of $\boldsymbol{Z}$, the latter line of (9) is the characteristic function of the increment, evaluated at the vector-valued function $\boldsymbol{\theta}(t) = \vartheta \cdot (t - \tau_D)\boldsymbol{v}(\tau_D)^\top \Sigma(\tau_D)$. We have thus written the characteristic function of Y in terms of τ_D-measurable quantities. Now use the induction hypothesis, and we are done. □

For most distributions, Lemma 3.1.1 will not be applicable. The question is, what distributions admit the construction to such an extent that it leads to the portfolio separation result? The canonical example is the Gaussian, which is the case covered by Khanna and Kulldorf [20]. If $\boldsymbol{Z}$ is a Brownian motion, with possibly correlated coordinates through a (constant) instantaneous covariance matrix $\Delta = (\mathrm{d}\boldsymbol{Z}\ \mathrm{d}\boldsymbol{Z}^\top)/\,\mathrm{d}t$, then for any constant vector $\boldsymbol{x}$, $t \mapsto \boldsymbol{x}^\top \boldsymbol{Z}(t)$ is a univariate Brownian motion with variance $\boldsymbol{x}^\top \Delta \boldsymbol{x}\,\mathrm{d}t$. This case, with geometric Brownian prices, is the canonical one, and let us therefore use that one to show how continuous trading reduces to Lemma 3.1.1. Now there is the Σ matrix which makes the Δ redundant, but let Σ take care of any time-variation and suppose Δ constant; for any vector $\boldsymbol{x} \in \mathbb{R}^n$, put

$$Q_{\boldsymbol{x}}(t) = \sqrt{\boldsymbol{x}^\top \Sigma(t) \Delta \Sigma^\top(t) \boldsymbol{x}} \tag{10}$$

The next lemma easily follows for Brownian motion by equivalence in law from textbooks (e.g. [36, Section 8.4]), but the key for the next sections is those properties that are shared by elliptical and stable distributions. We therefore include a proof, which invokes only those features.

3.1.2 Lemma. *Suppose $\boldsymbol{Z}$ is a Brownian motion, with constant covariance matrix $\Delta = (\mathrm{d}\boldsymbol{Z}\ \mathrm{d}\boldsymbol{Z}^\top)/\,\mathrm{d}t$. Relax Assumption 3.0.3 to allow for deterministic time-dependent $\boldsymbol{\mu}$ and Σ and $\mathbb{V}_t$. Consider an arbitrary admissible strategy $(C, \boldsymbol{v})$, and let $\boldsymbol{v}^*$ be any strategy such that for each t, $\boldsymbol{v}^*(t)$ solves the problem*

$$\max_{\boldsymbol{x} \in \mathbb{V}_t} \boldsymbol{x}^\top \boldsymbol{\mu}(t) \qquad s.t. \qquad Q_{\boldsymbol{x}} = Q_{\boldsymbol{v}(t)} \tag{11}$$

Then there is a C^* such that $(C^*, \boldsymbol{v}^*)$ is preferred.

Proof. Suppose not; let $\boldsymbol{v}^*$ satisfy $\boldsymbol{v}^*(t) \in \mathbb{V}_t$, $\dot{K}(t) := (\boldsymbol{v}^* - \boldsymbol{v})^\top \boldsymbol{\mu}(t) \geq 0$, K not constant (put $K(0) = 0$), and $Q_{\boldsymbol{v}^*(t)}(t) = Q_{\boldsymbol{v}(t)}(t)$. Consider the characteristic function of the process Y: for any continuous ϑ with bounded support it is

$$\mathsf{E} \exp \int_{t_0}^{\infty} \mathrm{i}\vartheta(t) Y(t)\,\mathrm{d}t = \exp\left(\mathrm{i}Y_{t_0} \int_{t_0}^{\infty} \vartheta(t)\,\mathrm{d}t\right) \mathsf{E} \exp \int_{t_0}^{\infty} \mathrm{i}\vartheta(t) \int_{t_0}^{t} \mathrm{d}Y(t')\,\mathrm{d}t \tag{12}$$

The innermost integral is an Itô integral, and by definition the limit in probability of integrals over simple predictables. Consider such a partition ϖ; since $\vartheta(t)$ has bounded support by, say T, we can truncate the stopping times by T. For the step strategies $\boldsymbol{v}_\varpi$ and $\boldsymbol{v}^*_\varpi$, Lemma 3.1.1 applies, as the τ_d-conditional characteristic function of the increment $\boldsymbol{v}^\top(t)\Sigma(t)\,\mathrm{d}\boldsymbol{Z}(t)$ only depends on $Q(\tau_d)$. By assumption, the consumption from the $(C^*, \boldsymbol{v}^*)$

strategy dominates. Then up to the arbitrary T, $(C^* - K, Y^*) \sim (C, Y)$, because this equivalence holds on each partition, thus in the limit, and the limit transition also shows that the strategy is admissible by Assumption 3.0.8. □

3.1.3 Remark. Risk-aversion, though not rigorously defined herein, would require the agent to shun dispersion and in addition to restricting themselves to the maximizers in (11), also to the subset that minimizes $Q_{\boldsymbol{x}}$ given $\boldsymbol{x}^\top \boldsymbol{\mu}$.

3.2 A recipe for dominating strategies, Brownian Z

Consider now the following regularity condition of «definiteness on $\mathbb{V}$»:

$$Q_{\boldsymbol{x}}(t) > 0 \qquad \text{for all non-null } \boldsymbol{x} \in \mathbb{V}_t \tag{13}$$

If this fails on a positive set, then there will either be a redundant investment opportunity (if the null is attained only for portfolios that yield zero drift) or an arbitrage.

Now suppose that (13) holds for each t. Then Lemma 3.1.2 gives a recipe for eliminating strategies by constructing dominating improvements. For simplicity, assume constant coefficients.

- For each number $Q \geq 0$, solve the (static!) maximization problem

  $$\max_{\boldsymbol{x} \in \mathbb{V}} \boldsymbol{x}^\top \boldsymbol{\mu} \quad \text{subject to } \boldsymbol{x}^\top \varGamma \boldsymbol{x} = Q^2. \tag{14}$$

  (Under (13) and closed radial $\mathbb{V}$ such that the value Q^2 is attained for at least one $\boldsymbol{x} \in \mathbb{V}$, the extreme value theorem will apply – see however Remark 4.2.3.)
- For each strategy $(C, \boldsymbol{v})$, calculate $Q_{\boldsymbol{v}(t)}(t) = (\boldsymbol{v}(t)^\top \varGamma \boldsymbol{v}(t))^{1/2}$. Construct the strategy $(C^*, \boldsymbol{v}^*)$ by a $\boldsymbol{v}^*(t)$ which for each t solves (14) for $Q = Q_{\boldsymbol{v}(t)}(t)$ – and consume the excess.
- Then applying Lemma 3.1.2 cf. Lemma 3.1.1: every agent will then prefer $(C^*, \boldsymbol{v}^*)$ to $(C, \boldsymbol{v})$.

In other words, there is for Brownian $\boldsymbol{Z}$ no welfare loss to rule out strategies which do not *for every fixed time* solve problem (14). This mean–variance optimization problem deviates from the frequently occurring procedure of minimizing variance for given expected return, which would be the (possibly more restrictive) approach of risk-averse agents. Problem (14) leads to two-fund monetary separation in the unconstrained case, with the risky portfolio being a scaling of $\varGamma^{-1}\boldsymbol{\mu}$. We shall see in the next section that this is the case for all elliptically distributed Lévy processes $\boldsymbol{Z}$, and exploit a few forms of $\mathbb{V}$ generated by a radial constraint (recovering Khanna and Kulldorff [20]) and by linear constraints. Then section 5 will exploit the α-stable case, which does require some modifictions to fit into Lemma 3.1.1.

3.2.1 Remark. The above recipe assumes absence of «static» arbitrages, but disguises the issue that condition (13) is not sufficient to eliminate arbitrages by doubling strategies. And, even if those are ruled out by the admissibility definition in question (e.g. imposing some large upper bound on Q is mathematically convenient however artificial), condition (13) and Assumption 3.0.9 still do not rule out *free lunch with vanishing risk:* if $\boldsymbol{x}^\top \boldsymbol{\mu}(t)/Q_x(t)$ tends to infinity, say, as $t \nearrow \bar{t}$, then we can construct a sequence which over a small interval yields a drift of 1 for arbitrarily low variance – and by scaling up, an arbitrary high drift (to consume) for a given variance. This is an inconvenient situation, which could be remedied (if coefficients are non-random) by requiring that (13) holds for every t, and

$$\sup_{\boldsymbol{v}(t)\in\mathbb{V}_t, t\in(t_0,T)} \frac{\boldsymbol{v}(t)^\top \boldsymbol{\mu}(t)}{Q_{\boldsymbol{v}(t)}(t)} < \infty, \qquad \text{all } T > t_0 \tag{15}$$

On the surface, this looks like an admissibility restriction on $\mathbb{V}$, but as we have assumed that every simple predictable portfolio is admissible, this is effectively a restriction on the coefficients.

Note that both Lemma 3.1.2 and our definition of separation, are mathematically valid in the presence of free lunches with vanishing risk and even arbitrages; we can still try to look for basis vectors spanning set of dominating strategies. Again, it is not necessary for the purposes of this paper to do the job of restricting the construction to what is practically interesting. □

4 Separation with elliptical distributions

An elliptical distribution is characterized nicely by «mean» and «covariance» (suitably interpreted), and a univariate *radial* variable. The ellipticals constitute *the* class of distributions to do proper mean-variance trade-off on, and they are also the class of distributions for which regression is linear in dimension 3 and above (Hardin [18]). We will first review the essentials of the distribution class, before proceeding to the separation theorems.

4.1 Review of the basic properties

From e.g. Cambanis et al. [4], recall that an *elliptical*, also known as an *elliptically contoured* random vector $\boldsymbol{X}$ is one for which the characteristic function can be written as

$$\mathsf{E}[\exp(\mathrm{i}\boldsymbol{\vartheta}^\top \boldsymbol{X})] = \exp(\mathrm{i}\boldsymbol{\vartheta}^\top \boldsymbol{\mu})\ h(\boldsymbol{\vartheta}^\top \varDelta \boldsymbol{\vartheta}) \tag{16}$$

where $\varDelta$ is usually assumed to be a positive definite matrix; in Remark 4.2.3, we shall allow for the extension to semidefinite $\varDelta$. The notational similarity between the location $\boldsymbol{\mu}$ and the drift term in the wealth process, is deliberate and will be clear below. Such a distribution can be represented as

$$\boldsymbol{X} \sim \boldsymbol{\mu} + \varXi \boldsymbol{W} \sqrt{R} \tag{17}$$

where Ξ is a nonrandom matrix, not necessarily square, $\boldsymbol{W}$ is uniform on the unit sphere in the appropriate-dimensional Euclidean space, and the *radial* variable $\sqrt{R}$ is independent – then $\Delta = \Xi\Xi^\top$. For example, the multinormal distributions are generated by taking $\sqrt{R}$ to be the absolute value of a standard normal (equivalent to scaling with the standard normal itself; taking the radial to be nonnegative, is merely a convention as $\boldsymbol{W}$ is symmetric). Within the theory of elliptical distributions, it is common to refer to $\boldsymbol{\mu}$ and Δ as «mean» and «covariance» regardless of integrability; we have $\mathsf{E}[\boldsymbol{X}] = \boldsymbol{\mu}$ iff $\boldsymbol{X} \in \mathsf{L}^1$, and $\mathsf{E}[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^\top]$ is a scalar multiple of Δ iff $\boldsymbol{X} \in \mathsf{L}^2$.

The key to the separation property is the distribution of linear combinations. By rotational invariance, the distribution of $\boldsymbol{x}^\top \Xi \boldsymbol{W}$ depends only on the Euclidean norm of $\boldsymbol{x}^\top \Xi$. So is also the case after scaling by the independent R; the distribution is fully determined by $\boldsymbol{x}^\top \Delta \boldsymbol{x}$, if $\boldsymbol{\mu} = \boldsymbol{0}$; otherwise we get a translation by $\boldsymbol{x}^\top \boldsymbol{\mu}$.

4.2 Separation properties

In this section, we shall assume that the increments of $\boldsymbol{Z}$ are elliptical about the origin, i.e. with a mean of null. Then, speaking loosely in terms of infinitesimals (this is no problem under infinite divisibility!) like in the Brownian case, we will have the increments

$$\boldsymbol{\mu}\,\mathrm{d}t + \Sigma\ \mathrm{d}\boldsymbol{Z} \quad \text{elliptical, with mean } \boldsymbol{\mu}\,\mathrm{d}t \text{ and covariance matrix } \Gamma\ \mathrm{d}t = \Sigma\Delta\Sigma^\top\,\mathrm{d}t, \tag{18}$$

(again, Remark 4.2.3 will discuss the merely semidefinite case); and, the distribution of $\boldsymbol{v}(t)^\top \Sigma(t)\,\mathrm{d}\boldsymbol{Z}$ is completely characterized by instantaneous drift $\boldsymbol{v}(t)^\top \boldsymbol{\mu}(t)$ and instantaneous variance $\boldsymbol{v}(t)^\top \Gamma(t) \boldsymbol{v}(t)$.

4.2.1 Remark. We shall, without mentioning it explicitely, allow for the «relaxed» version of Assumption 3.0.3 where coefficients are time-dependent deterministic (requiring, of course, that the assumptions of the theorems hold for each t), but suppress this generality. The notational modifications required in the proofs, will be left to the reader. □

For the construction analogous to the Gaussian example given in subsection 3.2, it therefore suffices in the unconstrained case that the portfolio $\boldsymbol{v}^*(t)$ maximizes $\boldsymbol{x}^\top \boldsymbol{\mu}$ given that $\boldsymbol{x}^\top \Gamma \boldsymbol{x}$ equals $\boldsymbol{v}(t)^\top \Gamma \boldsymbol{v}(t)$ – the existence of maximizer is granted if there is no arbitrage (i.e. if no investment opportunity can be written in terms of the others plus excess drift). A precise condition for absence of arbitrage is not equally straightforward when e.g. shorting is forbidden on one or more of the opportunities, which is a special case of part (a) of the following theorem; we assume for sake of simplicity positive definite Γ and leave the more general case to Remark 4.2.3.

4.2.2 Theorem (Two-fund separation). *Consider the dynamic model* ($\star$) *with* $\boldsymbol{Z}$ *having elliptical increments as in* (18) *with* Γ *positive definite. Assume precisely one of the following three conditions:*

(a) $\mathbb{V}$ *is closed and radial and common to all agents, or*

(b) $\mathbb{V} = \{\boldsymbol{v};\ \boldsymbol{v}(t)^\top \boldsymbol{1} = Y(t)\}$ *(agent-dependent), i.e. no risk-free opportunity, or*

(c) *$\mathbb{V}$ is the intersection of a convex cone common to all agents (and with vertex at $\mathbf{0}$), and the agent-specific hyperplane in (b), and $\boldsymbol{\mu} = \mu_0 \mathbf{1} = \mathbf{0}$ cf. Assumption 3.0.9 last part.*

Then the market admits two-fund separation.

Proof. We shall apply the very same recipe of section 3.2 as in the Gaussian case. The transition to continuous time works by merely observing that the equivalence in law established in Lemma 3.1.2, cf. the approximation in Lemma 3.1.1, only uses the ellipticity property of the Gaussian distribution, i.e. that the distribution of the increments is determined by drift and the quadratic form: $Q_{\boldsymbol{v}^*(t)}(t) = Q_{\boldsymbol{v}(t)}(t)$ we have $\boldsymbol{v}^*\ \mathrm{d}\boldsymbol{Z} \sim \boldsymbol{v}^\top\ \mathrm{d}\boldsymbol{Z}$, and we have a dominating strategy as long as $(\boldsymbol{v}^*(t) - \boldsymbol{v}(t))^\top \boldsymbol{\mu}(t) \geq 0$, each t.

So we are left with solving the static problem (14). In the radial case (a), the extreme value theorem grants a solution $\boldsymbol{f}$ for the case $Q = 1$, and $Q\boldsymbol{f}$ will solve for general $Q \geq 0$; construct then the portfolio $\boldsymbol{v}^*$ by $\boldsymbol{v}^*(t) = [\boldsymbol{v}(t)^\top \varGamma \boldsymbol{v}(t)]\, \boldsymbol{f}$. In the no risk-free case (b), (14) is the Lagrange problem

$$\max\ \boldsymbol{x}^\top \boldsymbol{\mu} \quad \text{subject to} \qquad \boldsymbol{x}^\top \varGamma \boldsymbol{x} = Q^2, \qquad \boldsymbol{x}^\top \mathbf{1} = Y, \tag{19}$$

where we only have to consider those Q for which there is a feasible point. When $Q^2 = Y^2$ there is only one feasible point, namely the «minimum variance portfolio»

$$Y\boldsymbol{f}_1 = Y \operatorname*{argmin}_{\boldsymbol{x}^\top \mathbf{1} = 1}\ \boldsymbol{x}^\top \varGamma \boldsymbol{x} \tag{20}$$

and this is the only point where the constraint qualification for the Lagrange problem may possibly fail. The Lagrange first-order condition reads

$$\boldsymbol{\mu} = 2\ell_0 \varGamma \boldsymbol{x} + \ell \boldsymbol{\pi} \tag{21}$$

$\varGamma$ is assumed invertible, so as long as $\ell_0 \neq 0$, we have $\boldsymbol{x}$ spanned by $\varGamma^{-1}\boldsymbol{\mu}$ and $\boldsymbol{f}_1 = \varGamma^{-1}\mathbf{1}/(\mathbf{1}^\top \varGamma^{-1}\mathbf{1})$. Should ℓ_0 vanish, we cannot use $\varGamma^{-1}\boldsymbol{\mu}$ for $\boldsymbol{f}_2$, as $\boldsymbol{\mu}$ must be $= \mu_0 \mathbf{1}$; then the maximization degenerates, and one can choose the second fund to be any vector $\boldsymbol{f}_2$ such that $\boldsymbol{f}_2^\top \varGamma \boldsymbol{f}_1 = 0$ and $\boldsymbol{f}_2^\top \varGamma \boldsymbol{f}_2 = 1$. So the funds can be chosen invariant over Y and Q. Case (c) works like the latter except that we cannot merely choose any $\boldsymbol{f}_2$. If $\boldsymbol{f}_2^\top \varGamma \boldsymbol{f}_2$ is bounded, choose the the maximizer (by convexity, $\boldsymbol{f}_1 + \boldsymbol{f}_2 \cdot q/(\boldsymbol{f}_2^\top \varGamma \boldsymbol{f}_2)$ will be admissible for all $\in [0, 1]$); if unbounded, choose one in a linear direction from $\boldsymbol{f}_1$ where arbitrary high volatility is attained. □

4.2.3 Remark. We make a couple of remarks to variants of the problem:

(a) For the two latter cases, the ones without risk-free opportunity, the $\boldsymbol{f}_2$ fund is not often mentioned in the literature, as it is «pure volatility» not contributing to expected return. Indeed, the classical case referred to as «one fund separation» (e.g. Ross [44, section 1]) – i.e. the special case of (c) where the «cone» is the entire space – is valid for risk-averse agents, who are satisfied with the minimum-variance portfolio as there is no additional return to higher volatility. The «pure volatility fund» will be needed to achieve equivalence in law in our construction, and satisfy risk takers.

(b) If the market is arbitrage-free but Γ is merely semidefinite, we can reduce the problem by passing to the eigendecomposition $\Gamma = \Pi\Lambda\Pi^\top$, and for convenience we order Λ with the nonzero eigenvalues for the n' first main diagonal entries, and zeroes on the bottom-right $n - n'$ main diagonal entries. If we rephrase problem (14) in terms of $\tilde{\boldsymbol{x}} = \Pi^\top \boldsymbol{x}$, $\tilde{\boldsymbol{\mu}} = \Pi^\top \boldsymbol{\mu}$, we notice that the set $\Pi^\top \mathbb{V}$ is radial if $\mathbb{V}$ is; in case (b), the absence of risk-free opportunity translates to the hyperplane constraint $\tilde{\boldsymbol{x}}^\top \tilde{\boldsymbol{\pi}} = Y$, where $\tilde{\boldsymbol{\pi}} = \Pi^\top \mathbf{1}$. The constraint set will still keep this form after transforming again by by scaling the first n coordinates by the square root of the respective eigenvalues, and then split by coordinates, letting the « ˆ » accent signify vectors with zeroes on the last $n - n'$ coordinates and « ˇ » denote vectors with zeroes on the first n', and normalize the n first by scaling with the square roots of the eigenvalues; problem (14) can then be written on the form

$$\max_{\hat{\boldsymbol{x}}+\check{\boldsymbol{x}}\in U'} (\hat{\boldsymbol{x}}^\top \hat{\boldsymbol{\mu}} + \check{\boldsymbol{x}}^\top \check{\boldsymbol{\mu}}) \quad \text{subject to } \hat{\boldsymbol{x}}^\top \hat{\boldsymbol{x}} = Q^2.$$

for a suitable set $\mathbb{V}'$ which is radial in case (a) and a hyperplane in case (b). For the radial case (a), the appropriate arbitrage-forbidding constraint is then

$$\boldsymbol{x}^\top \boldsymbol{\mu} \le 0 \qquad \text{for all } \boldsymbol{x} \in \mathbb{V} \text{ for which } \boldsymbol{x}^\top \Gamma \boldsymbol{x} = 0. \tag{22}$$

Applying this condition to choose $\check{\boldsymbol{x}} = \mathbf{0}$, the problem in $\hat{\boldsymbol{x}}$ is positive definite.

We merely sketch how the case (b) can be treated with this transformation: the static problem becomes

$$\max\ (\hat{\boldsymbol{x}}^\top \hat{\boldsymbol{\mu}} + \check{\boldsymbol{x}}^\top \check{\boldsymbol{\mu}}) \quad \text{subject to } \hat{\boldsymbol{x}}^\top \hat{\boldsymbol{x}} = Q^2, \qquad \hat{\boldsymbol{x}}^\top \hat{\boldsymbol{\pi}} + \check{\boldsymbol{x}}^\top \check{\boldsymbol{\pi}} = Y. \tag{23}$$

The conditions for finite drift for each Q now follow by routine calculations, which will lead to portfolio separation. It is necessary that $\check{\boldsymbol{\mu}}$ is a scaling $a\check{\boldsymbol{\pi}}$, otherwise there is unbounded drift for the minimal variance. We skip the details.

(c) The following is a curious (arti-)fact stemming from the generality of preferences compatible with first-order stochastic dominance. A constant will not first-order dominate a full-support variable, so consider a market with a «static» arbitrage opportunity, in the sense of a zero-variance portfolio with positive drift violating Assumption 3.0.9. Then, a transaction of borrowing to buy this, will be a safe positive amount, but not first-order stochastially dominating one with positive variance. To satisfy agents who *reject* the opportunity to trade away a fixed volatility for any finite size of a risk-free lunch, the construction of Lemma 3.1.1 will actually require a third fund – namely, one of positive variance. □

Having once and for all established the connection between the single period and the dynamic model, we shall from now on proceed by treating only the static optimization problems. We can cover no-borrowing constraints, or even different lending and borrowing interest rates, at the cost of additional funds. Let the «$\lesseqgtr$» symbol denote either $\leq$ or $=$, and assume that the portfolio constraint $\mathbb{V}$ is given in terms of linear constraints

$$\boldsymbol{v}(t)^\top \boldsymbol{\zeta}_j \lessdot z_j, \qquad j = 1, \dots, \bar{k} \tag{24}$$

where the $\boldsymbol{\zeta}_j$ are common to all agents, but the z_j may be individual (and time-dependent). The Lagrangian associated to the static optimization problem now becomes

$$\boldsymbol{x}^\top (\boldsymbol{\mu} - \sum_{j\geq 1} \ell_j \boldsymbol{\zeta}_j) - \ell_0 \boldsymbol{x}^\top \Gamma \boldsymbol{x}. \tag{25}$$

If the constraint qualification fails, then analogously to Theorem 4.2.2 we must be at the singleton where the ellipsoid is tangent to the convex polyhedron defined by the linear constraints. Again, this will be a limiting case in Q for the formula for the stationary point of the Lagrangian. If the Lagrange conditions hold, and Γ is definite, we have

$$\ell_0 \boldsymbol{x} = \Gamma^{-1}(\boldsymbol{\mu} - \sum_{j\geq 1} \ell_j \boldsymbol{\zeta}_j), \tag{26}$$

which leads to:

4.2.4 Theorem ($k+2$-fund separation)**.** *Assume the dynamics to follow* ($\star$)*, with* $\boldsymbol{Z}$ *being elliptical about the origin, and the portfolio constrained to satisfy* (24)*, with the* z_j *allowed to be individual. Suppose furthermore that* Γ *is positive definite, and let* $k+1$ *be the number of linearly independent vectors in* $\{\boldsymbol{\mu}, \boldsymbol{\zeta}_1, \cdots, \boldsymbol{\zeta}_{\bar{k}}\}$ *linearly independent (otherwise, remove constraints). Then we have* $k+2$*-fund separation (*$k+1$*-fund separation if* (24) *forbids risk-free opportunity).*

Proof. Consider the $k+1$ vectors $\Gamma^{-1}\boldsymbol{\mu}$ and $\Gamma^{-1}\boldsymbol{\zeta}_j$ on the right-hand side of (26); these are common to all agents. Unless $\ell_0 = 0$, then any portfolio not in the span of these vectors, can, using the construction of Lemma 3.1.1, be improved upon by one which is. This gives $k+2$-fund separation, with the last fund being the risk-free. In the case where $\ell_0 = 0$, the Lagrangian reduces to 0 and there is no optimization, merely a requirement to fulfill $\boldsymbol{v}(t) \in \mathbb{V}$ and the volatility level Q^2; the fund $\Gamma^{-1}\boldsymbol{\mu}$ will be redundant, as it must be in the span of the $\{\Gamma^{-1}\boldsymbol{\zeta}_j\}_j$, and we replace it by a vector orthogonal to the span in order to achieve the full range of possible volatility levels. □

In order to accommodate different rates for lending and borrowing, corresponding to $\boldsymbol{v}(t)^\top \mathbf{1}$ being $\leq$ resp. $\geq$ than Y, we simply assume an excess interest premium $\hat{\rho}(b)$ on $b = \boldsymbol{x}^\top \mathbf{1}$; this is of course just an extra constraint. Leverage-dependent interest rates are obtained by putting $\boldsymbol{\zeta}_0 = \mathbf{1}$ in the following, which is phrased slightly informally, as it is an adaptation of Theorem 4.2.4. As it admits $\hat{\rho}(b)$ to vary over agents, it also covers the cases where different agents see different interest rates in the market:

4.2.5 Corollary (Different interest rates require one more fund)**.** *Consider the setting of Theorem 4.2.4, but with a cost rate* $\hat{\rho}(b(t))\,\mathrm{d}t$ *(possibly individual function* $\hat{\rho}$ *to each agent) for the agent's choice* $b(t) = \boldsymbol{v}(t)^\top \boldsymbol{\zeta}_0$*. We then have* $k+3$*-fund separation (in particular,* 3*-fund separation if there are no constraints in* (24)*).*

Proof. Again, a static argument suffices. Suppose the agent chooses b, thereby obtaining an instantaneous drift after cost of $\boldsymbol{x}^\top \boldsymbol{\mu} - \hat\rho(b)\boldsymbol{x}^\top \boldsymbol{\zeta}_0 = -b\hat\rho(b) + \boldsymbol{x}^\top \boldsymbol{\mu}$ to maximize, subject to (24) augmented with $\boldsymbol{x}^\top \boldsymbol{\zeta}_0 = b$, which gives just another fund with b-dependent Lagrange multiplier. □

This result is of course also applicable for the case with leverage constraints applying to some of the agents – say, if only some of the agents are allowed to borrow, and some of these only up to $\bar b$ times their net wealth.

4.3 Generalized separation; market parameters driven by Z

Thus far we have relaxed Assumption 3.0.3 no longer than to deterministic time-dependence. Khanna and Kulldorff [20] point out that coefficients may be stochastic, as long as stochastically independent of $\boldsymbol{Z}$. The reason why dependence is prone to destroy separation properties, is that the agent will have to consider the *joint* distribution of the (tomorrow's wealth, tomorrow's maket) pair. Indeed, the dynamics of the underlying investment opportunities assumed so far – geometric for the traded market justifiction – ensures that neither wealth nor state of the market, depends on any of the prices levels P_i. If one parameter were to depend on the cardinal level of prices, then we would have to enter this as a state variable.

Merton [28, section 15.7] allows for a single state variable and deduces the need for another fund. His choice of the interest rate as this state variable is done for interpretability, and as we shall see, we can pick an arbitrary state variable – but not an arbitrary driving noise. To see how the setup can be adapted, one could assume it to be one of the investment opportunities opportunities – however, if we like, with $\mathbb{V}$ imposing null exposure to it. For notational convenience, we shall give it a letter, $M(t)$, and denote its parameters by overbars; it is not necessary to introduce any new noise sources, as we have not made any assumptions on the dimensionality of $\boldsymbol{Z}$. Thus, we assume (wealth, state) to obey the dynamics

$$\mathrm{d}\begin{pmatrix} Y(t) \\ M(t) \end{pmatrix} = (\boldsymbol{v}^\top,\ 1)\left[\begin{pmatrix} \boldsymbol{\mu}(t, M(t)) \\ \bar\mu(t, M(t)) \end{pmatrix}\mathrm{d}t + \begin{pmatrix} \Sigma(t, M(t)) \\ \bar\sigma^\top(t, M(t)) \end{pmatrix}\mathrm{d}\boldsymbol{Z}\right] - \begin{pmatrix} \mathrm{d}C(t) \\ 0 \end{pmatrix} \qquad (\star\star)$$

For given time and state the conditional increment using a step portfolio $\boldsymbol{x}$ is given by the drift and the instantaneous covariance, the latter taking the form

$$\begin{pmatrix} \boldsymbol{x}^\top \Gamma \boldsymbol{x} & \boldsymbol{x}^\top \boldsymbol{\gamma} \\ \boldsymbol{x}^\top \boldsymbol{\gamma} & \bar\sigma^\top \bar\sigma \end{pmatrix}\mathrm{d}t \qquad (27)$$

where all entities are allowed to depend on time and M. The analogue to Lemma 3.1.2 will now have to ensure that the agent is indifferent over identical joint distributions, but prefers more consumption; the maximization problem therefore gets another constraint $\boldsymbol{x}^\top \boldsymbol{\gamma} = \bar z$ where $\bar z$ is the value of $\boldsymbol{v}^\top \boldsymbol{\gamma}$ for the strategy to be dominated. Linear constraints were treated in Theorem 4.2.4, which we adapt to the following, skipping the proof and noting that the generalization to higher-dimensional M is straightforward:

4.3.1 Theorem (Generalized separation)**.** *Modify the setup of Theorem 4.2.4 by assuming* ($\star\star$) *in place of* ($\star$)*. Then Theorem 4.2.4 applies as if* (24) *is augmented with the additional constraint* $\boldsymbol{x}^\top \boldsymbol{\gamma} = \bar{z}$ *(where* $\bar{z}$ *is subject to the agent's choice), and with funds depending on* $(t, M(t))$ *as well.* □

Notice that even if M is uncorrelated with the return (not necessarily independent!), the additional fund vanishes – this in line with the properties we have used of elliptical distributions.

To sum up, the following generalizations behave quite uncomplicated under ellipticity: Introducing a(nother) linear constraint $\rightsquigarrow$ another fund. Introducing a common (correlated) risk factor $\rightsquigarrow$ another fund. It is also shown in [13] – in a single period model – that partial information on one of the univariate projections (i.e. linear combinations of coordinates) of $\boldsymbol{Z}$, also leads to one more fund being required.

5 Separation with α-stable distributions

As in the previous section, we shall first review the basics of the distribution class, before proving separation. We largely follow Samorodnitsky and Taqqu [46, Chapter 1], but as of writing, the reader can also consult Nolan's online chapter [35] for a brief introduction to the univariate case; for an exposition within the theory of Lévy processes and infinite divisibility, see Sato's book [49, chapter 3]. Throughout this section, disregarded distributions with a marginal being a Dirac point mass.

5.1 Review of the basic properties

Recall that a *stable* $\mathbb{R}^n$-valued random variable $\boldsymbol{X}$ is one for which for two independent copies $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ of $\boldsymbol{X}$, and any two positive numbers a_1 and a_2, there exist non-random $a \geq 0$ and $\boldsymbol{d} \in \mathbb{R}^n$ such that

$$a_1\boldsymbol{X}_1 + a_2\boldsymbol{X}_2 \sim a\boldsymbol{X} + \boldsymbol{d}. \tag{28}$$

The variable is *strictly stable* if one can take $\boldsymbol{d} = \boldsymbol{0}$. There exists a unique *index of stability* $\alpha \in (0, 2]$ such that $a = (a_1^\alpha + a_2^\alpha)^{1/\alpha}$. The Gaussian is of course when $\alpha = 2$; otherwise, $\mathsf{E}[\sum |X_i|^p] < \infty$ for $p < \alpha$ but not for $p = \alpha$. The stable laws are precisely the ones attainable through the general(ized) CLT. They are unimodal, absolutely continuous with infinitely differentiable densities – and infinitely divisible, hence can be the distribution of Lévy processes. Simple closed-form formulae for densities are known only in a few cases: the Gaussian, the Cauchy ($\alpha = 1$, the distribution of a (Gaussian)$^{-1}$) and the Lévy distribution (the distribution of a (Gaussian)$^{-2}$ and for the hitting time of Brownian motion – this has $\alpha = 1/2$; notice that the phrase «Lévy distributions» is sometimes in the literature used to denote the full class of stable distributions.)

Let us first give some properties for the univariate case, where the stable distributions form a 4-parameter family $S_\alpha(\sigma, \beta, \mu)$, parametrizing the characteristic function by

$$\ln \mathsf{E} \exp(\mathrm{i}\vartheta X) = \mathrm{i}\mu\vartheta - |\sigma\vartheta|^\alpha \big[1 + \mathrm{i}\beta w(|\vartheta|, \alpha) \operatorname{sign} \vartheta\big], \qquad \text{where} \tag{29a}$$

$$w(|\vartheta|, \alpha) = \frac{2}{\pi} \ln|\vartheta| \quad \text{if } \alpha = 1, \text{ and} \quad = \tan(-\alpha\pi/2) \quad \text{otherwise.} \tag{29b}$$

(with $0 \ln 0$ interpreted as 0). $\sigma > 0$ is called the *scale parameter;* X/σ has scale parameter 1. $\mu \in \mathbb{R}$ is called the *location parameter*; $X - \mu$ is located at zero, and the integrables have mean equal to μ. We have a *skewness parameter* $\beta \in [-1, 1]$; $a_0 + a_1 X$ has skewness $\beta \operatorname{sign} a_1$. A non-Gaussian stable is then symmetric about μ iff $\beta = 0$. This parametrization has a fairly intuitive interpretation for $\alpha > 1$ (beware though that the standard Gaussian has $\sigma = 1/2$, not 1!) and the symmetric cases. It has a somewhat different interpretation when $\alpha < 1$, when a variable with $\beta = 1$ is supported by $[\mu, \infty)$ and the general variable is a linear combination of one such and one symmetric about μ. (All other cases than $\alpha < 1 = |\beta|$, have support $= \mathbb{R}$.) The skewed 1-stable case is an oddball for which «scale» and «location» might be somewhat confusing terminology; it is the only which can *not* be made strictly stable by translating. Also, the above parametrization has a discontinuity in distribution there, while it is continuous in distribution when the skewed 1-stables are removed. The reader should be warned that there are numerous parametrizations, leading to what Hall [17] called a «comedy of errors»; see also the discussion by Nolan [35, Section 1.3] who chooses to enumerate the parametrizations and index by four parameters plus the parametrization number (the above is his number 1).

For the *multivariate* α-stables, we have the representation analogue to (29a), cf. (29b):

$$\ln \mathsf{E} \exp(\mathrm{i}\boldsymbol{\vartheta}^\top \boldsymbol{X}) = \mathrm{i}\boldsymbol{\vartheta}^\top \boldsymbol{\mu} - \int_{\mathbb{S}} |\boldsymbol{\vartheta}^\top \boldsymbol{s}|^\alpha \big[1 + \mathrm{i}w(|\boldsymbol{\vartheta}^\top \boldsymbol{s}|, \alpha) \operatorname{sign}(\boldsymbol{\vartheta}^\top \boldsymbol{s})\big] \varkappa(\,\mathrm{d}\boldsymbol{s}) \tag{30}$$

in terms of a finite *spectral measure* $\varkappa$, living on the unit sphere $\mathbb{S}$ in $\mathbb{R}^n$, and which is unique for $\alpha < 2$, and so is the vector $\boldsymbol{\mu}$ – the notational similarity to the drift vector is intentional. $\boldsymbol{X}$ is strictly stable iff every component is, and $\boldsymbol{X}$ is symmetric iff both $\varkappa$ is symmetric and $\boldsymbol{\mu} = \boldsymbol{0}$. $\boldsymbol{\mu}$ equals the vector of the marginals' location parameters except for the non-symmetric 1-stables. See [46, Sections 2.3–2.4]. Any non-degenerate linear combination $\boldsymbol{x}^\top \boldsymbol{X}$ is then univariate α-stable:

$$\boldsymbol{x}^\top \boldsymbol{X} \sim S_\alpha(\sigma_{\boldsymbol{x}}, \beta_{\boldsymbol{x}}, \mu_{\boldsymbol{x}}) \qquad \text{with} \tag{31a}$$

$$\sigma_{\boldsymbol{x}} = \Big(\int \big|\boldsymbol{x}^\top \boldsymbol{s}\big|^\alpha \; \varkappa(\,\mathrm{d}\boldsymbol{s})\Big)^{1/\alpha}, \tag{31b}$$

$$\beta_{\boldsymbol{x}} = \sigma_{\boldsymbol{x}}^{-\alpha} \int (\boldsymbol{x}^\top \boldsymbol{s})^{<\alpha>} \; \varkappa(\,\mathrm{d}\boldsymbol{s}) \tag{31c}$$

$$\mu_{\boldsymbol{x}} = \boldsymbol{x}^\top \boldsymbol{\mu} - \frac{2}{\pi} \int \boldsymbol{x}^\top \boldsymbol{s} \cdot \ln\big|\boldsymbol{x}^\top \boldsymbol{s}\big| \; \varkappa(\,\mathrm{d}\boldsymbol{s}) \; \chi_{\{\alpha=1\}}, \tag{31d}$$

where $x^{<\alpha>}$ denotes the *signed power* $|x|^\alpha \operatorname{sign} x$.

In the Gaussian case $\alpha = 2$, where $\sigma_{\boldsymbol{x}}^2 = \int |\boldsymbol{x}^\top \boldsymbol{s}|^2 \; \varkappa(\,\mathrm{d}\boldsymbol{s}) = \boldsymbol{x}^\top\big(\int \boldsymbol{s}\boldsymbol{s}^\top \varkappa(\,\mathrm{d}\boldsymbol{s})\big)\boldsymbol{x}$, the dependence structure reduces nicely to the covariance matrix, and the ellipticals behave the same although a non-Gaussian vector spherical about the origin is not a vector of iid's. Also, covered by the previous section, there are symmetric α-stable elliptical vectors (just take the radial $\sqrt{R}$ of (17) to be the absolute value of a symmetric α-stable). With zero location, this has characteristic function of the form $\exp(-\mathrm{i}\sigma^\alpha ||\boldsymbol{\vartheta}||_2^\alpha)$, where the subscript «2» denotes the Euclidean norm. Contrast this with a vector of iid's, with characteristic function of the form $\exp(-\mathrm{i}\sigma^\alpha ||\boldsymbol{\vartheta}||_\alpha^\alpha)$, involving the L^α (quasi-) norm – indeed, the literature does not agree on which one is the natural multivariate extension, leaving terms like «multivariate Cauchy» ambiguous. The dependence structure of stable vectors is in the general case a highly non-trivial matter. For this reason, there is hardly any use for the Σ matrix used in the model, except to introduce time-inhomogeneity.

5.1.1 Remark (Sub-stable distributions)**.** The radial scaling in the elliptical distribution has an analogue here as well: we can make a radial scaling $R^{1/\alpha}$ of an α-stable and still have the projections characterized by the same parameters:

$$\mathsf{E}\exp(\mathrm{i}\vartheta\boldsymbol{x}^\top\boldsymbol{X}R^{1/\alpha}) = \mathsf{E}\exp\Big(\mathrm{i}\vartheta\mu_{\boldsymbol{x}}R^{1/\alpha} - |\vartheta\sigma_{\boldsymbol{x}}|^\alpha\cdot\Big[1+\mathrm{i}\beta_{\boldsymbol{x}}w(\alpha,\vartheta)\,\mathrm{sign}\,\vartheta\Big]R\Big) \tag{32}$$

These distributions are more general than what is usually called «sub-stability» in the literature, where typically one assumes symmetry, i.e. $\mu_{\boldsymbol{x}} = \beta_{\boldsymbol{x}} = 0\ \forall\boldsymbol{x}\in\mathbb{R}^n$ ([46, Section 3.8] also assumes the radial variable to be positive stable (i.e. with index $< 1 =$ skewness) and located at zero). For the purposes of separation, this will lead to easy ramifications provided one can fit those distributions to our framework. With $\mu_{\boldsymbol{x}} = 0$, (32) yields infinite divisibility if R is infinitely divisible. Additional results concerning infinite divisibility is given by for Misiewicz [29] and [30] (assuming symmetry). The α-stables themselves are infinitely divisible, and it easily follows by adding up increments that the scale parameter to the power α aggregates like $\mathrm{d}t$ (the so-called self-similarity property), directly generalizing the « $\sqrt{\mathrm{d}t}$ » behaviour of Brownian motion. □

5.2 Separation properties

For convenience we shall – analogous to the elliptical case – only carry out the analysis for the case where the driving noise $\boldsymbol{Z}$ has α-stable increments (disregarding the ellipticals, which are already covered):

5.2.1 Remark. Although theorems will be stated proven assuming that

$$\boldsymbol{x}^\top\big(\boldsymbol{\mu}+\Sigma[\boldsymbol{Z}(t+1)-\boldsymbol{Z}(t)]\big) \sim S_\alpha\big(\sigma_{\boldsymbol{x}},\beta_{\boldsymbol{x}},\mu_{\boldsymbol{x}}\big), \quad \text{as in equations (31),} \tag{33}$$

the results will admit the independent radial scaling $R^{1/\alpha}$ under the following conditions:

$$\boldsymbol{x}^\top\big(\boldsymbol{\mu}+\Sigma[\boldsymbol{Z}(t+1)-\boldsymbol{Z}(t)]\big) \sim \boldsymbol{x}^\top\boldsymbol{X}R^{1/\alpha} \qquad \text{where } \boldsymbol{x}^\top\boldsymbol{X}\sim S_\alpha\big(\sigma_{\boldsymbol{x}},\beta_{\boldsymbol{x}},\mu_{\boldsymbol{x}}\big) \tag{34a}$$

or alternatively: $(\alpha-1)\beta_{\boldsymbol{x}} = 0 \;\forall\; \boldsymbol{x}$, and

$$\boldsymbol{x}^\top \Sigma[\boldsymbol{Z}(t+1) - \boldsymbol{Z}(t)] \sim \boldsymbol{x}^\top \boldsymbol{X} R^{1/\alpha} \qquad \text{where } \boldsymbol{x}^\top \boldsymbol{X} \sim S_\alpha\big(\sigma_{\boldsymbol{x}}, \beta_{\boldsymbol{x}}, 0\big) \tag{34b}$$

provided, in continuous time, infinite divisibility. The results will also allow for the «relaxed» version of Assumption 3.0.3 where coefficients are time-dependent deterministic (requiring, though, that the assumptions of the theorems hold for each t); notice though that the Lévy process assumption requires α constant (although relaxing this assumption too – to so-called *stable-like* additive processes – would be straightforward). The notational modifications required in the proofs, will be left to the reader. □

Under the notational assumptions just made, we can and will without loss of generality assume $\boldsymbol{Z}$ to be n-variate and

$$\Sigma = \text{ the identity.} \tag{35}$$

In order to approach like in subsection 3.2, we give an analogue of Lemma 3.1.2. It should be noted though, that the free lunch issue addressed in Remark 3.2.1 calls for a condition like (15) – indeed, the analogue follows by using $\sigma_{\boldsymbol{x}}(t)$ in place of $Q_{\boldsymbol{x}}(t)$. However, in the case $\alpha < 1$, we must also assume location of the opposite sign of $\beta_{\boldsymbol{x}}$ whenever $|\beta_{\boldsymbol{x}}| = 1$; should then location hit zero, then we are facing a not only an arbitrage (a «static» one), but, even, one with infinite mean. The following result is nevertheless true even when a free lunch makes it less interesting:

5.2.2 Lemma. *Consider the dynamic model* $(\star)$ *with* $\boldsymbol{Z}$ *having stable increments satisfying* (33)*. Consider an arbitrary admissible strategy* $(C, \boldsymbol{v})$*, and let* $\boldsymbol{v}^*$ *be any strategy such that for each* t*,* $\boldsymbol{v}^*(t)$ *solves the problem*

$$\max_{\boldsymbol{x}\in\mathbb{V}} \;\Big\{\boldsymbol{x}^\top\boldsymbol{\mu} - \frac{2}{\pi}\int \boldsymbol{x}^\top\boldsymbol{s}\cdot\ln\big|\boldsymbol{x}^\top\boldsymbol{s}\big|\;\varkappa(\,\mathrm{d}\boldsymbol{s})\;\chi_{\{\alpha=1\}}\Big\} \qquad s.t. \qquad \sigma_{\boldsymbol{x}} = \sigma_{\boldsymbol{v}(t)} \quad and \quad \beta_{\boldsymbol{x}} = \beta_{\boldsymbol{v}(t)} \tag{36}$$

Then there is a C^* *such that* $(C^*, \boldsymbol{v}^*)$ *is preferred.*

Proof. The proof goes like Lemma 3.1.2 except that the increment – now $\boldsymbol{v}^\top\,\mathrm{d}\boldsymbol{Z}$ – depends on both $\sigma_{\boldsymbol{v}}$ and $\beta_{\boldsymbol{v}}$, both of which are now given. □

So, compared to the elliptical case, there is one more parameter to fix, namely the skewness of the portfolio return. The usual treatment in the literature assumes symmetry, and that is the only case in which all *unconstrained* portfolios yield the same skewness, as $\beta_{-\boldsymbol{x}} = -\beta_{\boldsymbol{x}}$. However, a construction like Theorem 4.2.2 will work if $\mathbb{V}$ restricts to portfolios which lead to skewness being one common number. This might actually be meaningful; the simplest non-symmetric example might be iid coordinates with common skewness and $\mathbb{V} \subseteq \mathbb{R}^n_+$, forbidding shorting. More generally, observing that $\{\boldsymbol{x};\ \beta_{\boldsymbol{x}} = \beta\}$ is – disregarding the origin – closed and radial (possibly empty), a one-line calculation yields the following:

5.2.3 Lemma. *Suppose that $\mathbb{H}$ is an orthant (possibly rotated about the origin) that $\hat{\varkappa}$ is a finite measure supported by the unit sphere intersected with $\mathbb{H}$, and that the spectral measure $\varkappa$ of $\boldsymbol{X}$ takes the form*

$$\varkappa(\mathbb{G}) = \begin{cases} \gamma_+ \hat{\varkappa}(\mathbb{G}) & \text{for } \mathbb{G} \subset \mathbb{H} \\ \gamma_- \hat{\varkappa}(-\mathbb{G}) & \text{for } \mathbb{G} \subset -\mathbb{H} \\ 0 & \text{if } \mathbb{G} \cap (-\mathbb{H} \cup \mathbb{H}) = \varnothing \end{cases} \tag{37}$$

where $-\mathbb{H}$ denotes $\{\boldsymbol{x};\ -\boldsymbol{x} \in \mathbb{H}\}$ and likewise for $-\mathbb{G}$. Then

$$\beta_{\boldsymbol{x}} = \frac{\gamma_+ - \gamma_-}{\gamma_+ + \gamma_-} \qquad \forall\ \boldsymbol{x} \in \mathbb{H} \tag{38}$$

so that $\beta_{\boldsymbol{x}}$ is constant on $\mathbb{H}$, and constant on $-\mathbb{H}$. □

Restricting portfolios then facilitates separation for nonsymmetric distributions. The following includes those results applicable for integrable returns:

5.2.4 Theorem (Two-fund separation)**.** *Consider the dynamic model ($\star$) with $\boldsymbol{Z}$ having stable increments satisfying* (33)*, let $\beta_{\boldsymbol{x}} = \beta$ (constant) on some closed radial $\mathbb{V}^{(\beta)}$ common to all agents, and assume that one of the following conditions applies:*

(a) $\mathbb{V} = \mathbb{V}^{(\beta)}$

(b) $\mathbb{V}^{(\beta)}$ *is a convex cone,* $\mathbb{V} = \mathbb{V}^{(\beta)} \cap \{\boldsymbol{x}^\top \mathbf{1} = Y\}$ *and* $\boldsymbol{\mu} = \mu_0 \mathbf{1} = \mathbf{0}$ *cf. Assumption 3.0.9 last part, and the distribution is symmetric if* $\alpha = 1$.

(c) $\alpha = 1$ *and* $\mathbb{V} = \mathbb{V}^{(\beta)} \cap \{\boldsymbol{x}^\top \boldsymbol{\lambda} \in \mathbb{A}\}$, *where $\mathbb{A}$ is a possibly agent-dependent set and*

$$\boldsymbol{\lambda} = \int \boldsymbol{s}\ \mathrm{d}\varkappa. \tag{39}$$

Then the market admits two-fund (monetary) separation.

Proof. Lemma 5.2.2 takes care of the transition from the static to the dynamic model, just as in the elliptical case. We can therefore consider merely the static location–scale optimization problem. Assume the distribution is not nonsymmetric 1-stable. For part (a) consider the corresponding static problem

$$\max_{\boldsymbol{x} \in \mathbb{V}} \boldsymbol{x}^\top \boldsymbol{\mu} \quad \text{subject to} \quad \sigma_{\boldsymbol{x}} = 1 \tag{40}$$

– involving no skewness parameter, as it is by assumption constant on $\mathbb{V}$. Like for the elliptical distributions, this problem has a solution $\boldsymbol{f}$. For an arbitrary strategy $(C, \boldsymbol{v})$ define $(C^*, \boldsymbol{v}^*)$ by $\boldsymbol{v}^*(t) = \sigma_{\boldsymbol{v}(t)}\ \boldsymbol{f}$ (having the same skewness), and Lemma 5.2.2 applies. For part (b), choose $\boldsymbol{f}_1$ to minimize $\sigma_{\boldsymbol{x}}$. If $\mathbb{V}$ is bounded, choose $\boldsymbol{f}_2 - \boldsymbol{f}_1$ to maximize. As $\boldsymbol{x} \mapsto \sigma_{\boldsymbol{x}}$ is continuous (by dominated convergence), $\{\boldsymbol{f}_1 + r\boldsymbol{f}_2\}_{r \in [0,1]}$ will attain all possible returns distributions. If $\mathbb{V}$ is unbounded, then just as for the elliptical distributions, we choose $\boldsymbol{f}_2$ such that $\{\boldsymbol{f}_1 + r\boldsymbol{f}_2\}_{r\geq 0}$ attains every possible value of scale in $[\sigma_{\boldsymbol{f}_1}, \infty)$.

Assume now $\alpha = 1$ and non-symmetry. For part (a), the instantaneous optimization problem – for arbitrary scale $q > 0$ – now becomes

$$\max_{\boldsymbol{x}\in\mathbb{V}} \{\boldsymbol{x}^\top \boldsymbol{\mu} - \frac{2}{\pi}\int \boldsymbol{x}^\top \boldsymbol{s} \cdot \ln|\boldsymbol{x}^\top \boldsymbol{s}| \;\mathrm{d}\varkappa\} \qquad \text{s. t.} \quad \int |\boldsymbol{x}^\top \boldsymbol{s}| \;\mathrm{d}\varkappa = q \tag{41}$$

Replace $\boldsymbol{x}$ by $q\boldsymbol{x}$, and note that $q\boldsymbol{x} \in \mathbb{V} \Leftrightarrow \boldsymbol{x} \in \mathbb{V}$ since $\mathbb{V}$ radial and $q > 0$:

$$q\max_{\boldsymbol{x}\in\mathbb{V}} \{\boldsymbol{x}^\top \boldsymbol{\mu} - \frac{2}{\pi}\int \boldsymbol{x}^\top \boldsymbol{s} \cdot (\ln q + \ln|\boldsymbol{x}^\top \boldsymbol{s}|) \;\mathrm{d}\varkappa\} \quad \text{s. t.} \quad \int |\boldsymbol{x}^\top \boldsymbol{s}| \;\mathrm{d}\varkappa = 1 \tag{42}$$

Now $q\int \boldsymbol{x}^\top \boldsymbol{s}\;\mathrm{d}\varkappa$ is in fact the skewness, constant by assumption, so $q\boldsymbol{f}$ solves (41) iff $\boldsymbol{f}$ solves it for the particular value $q = 1$. Finally, for part (c), then by the assumption of fixed level of skewness, the additional constraint only restricts the possible choices of scale level q. □

5.2.5 Remark.

(a) In part (b), then risk-averse agents will only want to use $\boldsymbol{f}_1$ if $\alpha > 1$, recovering the classical degeneracy of one-fund separation for those agents. The same is true if $\alpha = 1 = 1 - \beta_{\boldsymbol{v}}$, and if $\alpha < 1$ and the skewness imposed for $\beta_{\boldsymbol{v}}$ is ≤ 0. However if $\alpha < 1$ and $\beta_{\boldsymbol{v}} > 0$, then increased scale is equivalent in law to adding an independent symmetric noise *and* and independent positive infinite-mean variable. See Theorem 5.2.6 below.

(b) The two-fund separation result prevails if we restrict a radial $\mathbb{V}$ by simply disallowing certain scales, and the proof of Theorem 5.2.4 shows that such a condition is not necessarily too farfetched for $\alpha = 1$ – see the one-fund separation result in Theorem 5.2.7 below.

(c) The cases $\alpha \leq 1$ distinguish themselves from the integrable ones also for the symmetric case. The α-unit ball is not convex for $\alpha < 1$, and parallel-shifting out a hyperplane $\boldsymbol{x}^\top \boldsymbol{\mu}$ to the maximum extent possible, would lead to corner solution. $\alpha = 1$ will have corner solution as well. If we assume symmetry, $\mathbb{V} = \mathbb{R}^n$ and independent marginals of $\mathrm{d}\boldsymbol{Z}$; then the agents will not diversify. They can do with one coordinate only, namely the one with highest absolute value of the drift, shorting it if negative. If shorting is forbidden and we have symmetry or, more generally, Lemma 5.2.3 applying (pinning down skewness), the highest drift will be chosen. Notice however that this does depend on the (in-) dependence structure; an elliptical vector yields diversification even when the radial variable is stable with $\alpha < 1$. □

As noted in Remark 5.2.5, the totally skewed distributions with index $\alpha < 1$ are somewhat special in many respects. On the surface our preference assumptions are unreasonably general and fail to order variables that first-order dominance does order: if drift is fixed in part (b), and skewness is +1 on $\mathbb{V}$, the prospect of increasing scale leaves the agent with the choice of accepting / rejecting an independent wealth increase – positive, but not instantaneously, so our strategy of consuming excess is strictly speaking not applicable. We could modify the preference assumption 3.0.5 to cover this particular case, but the following is likely even simpler: for $\alpha < 1 = \beta_{\boldsymbol{x}}$, the process $\boldsymbol{x}^\top \boldsymbol{Z}$ is an increasing pure jump process, for which we are controlling the jump intensity (the Lévy measure). Therefore, in the choice of intensity ν and intensity $\nu^* \geq \nu$, the agent can draw an independent lottery at jump time as to whether consume it immediately (with probability $(\nu^* - \nu)/\nu$) or accumulate it to wealth. Then the wealth increments corresponding to ν and to ν^* follow the same law. Technically, as we require adapted strategies, we have to assume that these draws are measurable. We can then follow the lines of Lemma 5.2.2 to eliminate suboptimal strategies which do not maximize skewness for given *symmetric contribution to* scale (this contribution should be positive in the absence of arbitrage). A variable with skewness β and scale σ, can be decomposed into a symmetric of scale $(1 - |\beta|)^{1/\alpha}\sigma$ and one of skewness sign β and scale $|\beta|^{1/\alpha}\sigma$. Thus for fixed value of symmetric contribution $(1 - |\beta_{\boldsymbol{x}}|)\sigma_{\boldsymbol{x}}^\alpha$ to scale, the agent will then have to trade-off skew scale contribution $\beta_{\boldsymbol{x}}\sigma_{\boldsymbol{x}}^\alpha$ vs. drift – for both of which more is better. The trade-off may of course degenerate, and the following result contains some of those cases, arguably less interesting than Theorem 5.2.4:

5.2.6 Theorem (Two-fund separation, $\alpha < 1$)**.** *Assume $\alpha < 1$ and that the agent can at each jump time draw a uniform $[0, 1]$ variable independent of everything else. Let $\mathbb{V}$ be a closed radial set (common to all agents), such that $\mu_{\boldsymbol{x}} = \mathbf{0}$ on $\mathbb{V}$ or $|\beta_{\boldsymbol{x}}| = 1$ on $\mathbb{V}$. Then the market admits two-fund separation.*

Proof. The agent will maximize $\beta_{\boldsymbol{x}}^{<1/\alpha>}\sigma_{\boldsymbol{x}}$ subject to precisely one of the respective constraints $(1 - |\beta_{\boldsymbol{x}}|)\sigma_{\boldsymbol{x}}^\alpha = q^\alpha$ or $\boldsymbol{x}^\top \boldsymbol{\mu} = d$. In any case the problem is homogeneous. □

For $\alpha \leq 1$ we have cases of *one-fund separation* which in contrast to the classical results mentioned in Remark 5.2.5 part (a) do not assume risk aversion. To the best of this author's knowledge, Theorem 5.2.7 is new to the literature:

5.2.7 Theorem (Some nonintegrable cases of one-fund separation and two-fund separation without risk-free opportunity)**.**

(a) Assume $\alpha = 1$, and that the conditions of Theorem 5.2.4 part (c) to hold, that the skewness imposed is nonzero, that $\boldsymbol{\lambda} = \lambda_0 \mathbf{1} \neq \mathbf{0}$ (cf. (39)) and that $\mathbb{A}$ is the singleton $\{Y/\lambda_0\}$.

(b) Assume $\alpha < 1$ and that the agent can at each jump time draw a uniform $[0, 1]$ variable independent of everything else. Let $\beta_{\boldsymbol{x}} = 1$ on some closed radial $\mathbb{V}^{(1)}$ common to all agents, and suppose that $\mu_{\boldsymbol{x}}$ constant on $\mathbb{V} = \mathbb{V}^{(1)} \cap \{\boldsymbol{x}^\top \mathbf{1} = Y\}$.

Then we have one-fund separation: for some $\boldsymbol{f}$, each agent holds $Y\boldsymbol{f}/\boldsymbol{f}^\top\mathbf{1}$. If the conditions on $\beta_{\boldsymbol{x}}$ are relaxed by replacing $\beta_{\boldsymbol{x}}$ by $|\beta_{\boldsymbol{x}}|$, we have two-fund separation where all positive-wealth agents hold $Y\boldsymbol{f}_1/\boldsymbol{f}_1^\top\mathbf{1}$ and all negative-wealth agents hold $Y\boldsymbol{f}_2/\boldsymbol{f}_2^\top\mathbf{1}$.

Proof. For part (b) – note we can assume $\boldsymbol{\mu} = \mathbf{0}$, cf. Assumption 3.0.9 last part – the agent simply wants to maximize $\sigma_{\boldsymbol{x}}$ subject to budget. The maximand is is proportional to Y. Consider part (a). By assumption, we have only one possible level β of skewness in the first place, and as $\boldsymbol{\lambda}$ is a non-null scaling of $\mathbf{1}$, the absence of risk-free opportunity is a restriction of the form $\boldsymbol{v}^\top\boldsymbol{\lambda} \in \mathbb{A}$ with $\mathbb{A}$ being a singleton. (As $\mathbb{V}$ imposes non-zero skewness, zero-wealth agents must choose the null portfolio.) So Theorem 5.2.4 part (c) applies.
For the last assertion consider β and $-\beta$ separately. E.g. for part (b) the agent who faces $\beta_{\boldsymbol{x}} = -1$ wants to *minimize* $\sigma_{\boldsymbol{x}}$ subject to budget. □

Notice that in particular, the skewness assumption of part (a) holds if $\mathbb{V}^{(\beta)} = \mathbb{R}^n_+$ and Lemma 5.2.3 applies with $\gamma_+ \neq \gamma_-$ – and for part (b) if furthermore $\gamma_+\gamma_- = 0$. These constraints require nonnegative wealth if dictated by the imposed value of $\beta_{\boldsymbol{x}}$, which is the rationale behind stating the two-fund version of the theorem.

We conclude this section by pointing out, for completeness, that we cannot hope to allow skewness as a free choice at the cost of merely one additional fund, except in special cases (e.g. blockwise independence, to be discussed in the next section):

5.2.8 Counterexample (No three-fund monetary separation with skewness free of choice)**.** Assume we have risk-free opportunity and three independent risky α-stable with common $\alpha \in (1, 2)$, all with positive drift, all normalized to unit scale, and their respective skewnesses being 1, 1/2 and 0. Obviously, agents who want skewness ± 1, must invest solely in #1. It is easy to verify that agents who want skewness = 1/2 will choose a scaling of $(1, 1, 1)^\top$ if $\mu_1 + \mu_3 = 2^{(2-\alpha)/\alpha}\mu_2$ holds. For three-fund separation to apply, all agents must therefore choose $v_2 = v_3$. Now assume $\mu_2 > 0$, normalize it to 1 and write $\mu_1 = \eta \cdot 2^{(2-\alpha)/\alpha}$ and $\mu_3 = (1 - \eta) \cdot 2^{(2-\alpha)/\alpha}$, where $\eta \leq 1$ (possibly negative). The agent who wants scale = 1 and symmetry, must choose $v_1 = -2^{-1/\alpha}v_2$, and maximize $(1 - \eta)x_3 + (1 - 2^{-1/\alpha}\eta)x_2$ subject to $\frac{3}{2}|x_2|^\alpha + |x_3|^\alpha = 1$. The solution has $x_2 = x_3$ only iff $\eta = (3 \cdot 2^{1/\alpha} - 2)^{-1}$. □

Notice though that if we introduce yet another independent symmetric opportunity, with the same α, the agent can replace the two symmetrics by one «fund». This example also shows that, while it may be tempting in order to stick to «one skewness», we have no hope that rewriting all opportunities as differences between totally skewed ($\beta = 1$) variables could resolve the situation even for those vectors which admit decomposition into a matrix times a vector of iid's.

6 Discussion

This section will first discuss cases that do not, at least not easily, generalize from ellipticals to stable distributions, and then possible applications of generalizations that do carry over.

6.1 Cases not covered: why not?

For each separation theorem valid for the ellipticals, it is natural to ask whether it generalizes to the stables, symmetric or not. However, som have no straigthforward counterparts.

First, the reader might have noticed that there was no general two-fund separation theorem stated for the stable returns without risk-free opportunity, nor with general linear constraints. Not only does the matrix algebra-based proof fail, but the result also fails to carry over to the stable distributions, except special cases. A result of [14] implies the counterexample that with iid coordinates, then three funds do not suffice when $\alpha = 3/2$ (while e.g. four funds do when $\alpha = 4/3$).

For the analogous reason, there is no straightforward generally valid *generalized separation* theorem in the sense of subsection 4.3, as that was also based on the dependence structure being conveniently representable by matrix algebra and a linear constraint.

A different point to make, is that we have have not explicitely used α-stability of the increments of $\boldsymbol{Z}$; we have only used the characterizations of their projections' univariate distributions in terms of three parameters. Disregarding the radial scaling, what about merely assuming univariate α-stability of the portfolio return? As established by Marcus [23], there are non-stable distributions with all projections stable; however, none of these examples can be integrable, none infinitely divisible (hence fit for Lévy processes), none symmetric stable and none can have the property that the projections $\boldsymbol{x}^\top \boldsymbol{X}$ possess location on the linear $\mu_{\boldsymbol{x}} = \boldsymbol{x}^\top \boldsymbol{\mu}$ form. See [16, Theorem 2.3] and [46, Section 2.2, Theorem 2.1.5 and Theorem 3.1.2].

6.2 Possible consequences for applications

The extension of two-fund monetary separation to α-stability, can be regarded as a robustness result; if one is reasonably close to applying the central limit theorem, and the distributions are symmetric, we could expect the mutual fund property to be a reasonable model; as shown herein, the property is robust to skewness as long as that is one and the same number (implying, under asymmetry, the appropriate portfolio constraint). This makes the theorem possibly applicable to insurers who do not see symmetry; for example assume that profits and loss follow totally skewed distributions with a light upper tail and a heavy downside tail, and the assumption that insurance companies do not reinsure themselves short might be reasonable. Likewise, for a loan portfolio might we might get a better fit by a light upper and a heavy lower tail, than by a Gaussian. Loan portfolios deviate from the usual interpretation of the mutual fund theorem, in that the agents do not hold fractions of each separate opportunity – while large corporates do issue bonds to be traded, a home owner will keep mortgage and home insurance concentrated on a few firms. Nevertheless, a separation theorem is consistent with agents choosing portfolios that – modulo scale – have the same distribution; say, that financial conglomerates consisting of an insurance company and a bank (in the traditional savings/lending business) would choose more or less the same risk profile when it comes to exposure to home insurance, car insurance, corporate lending, retail lending etc. – modulo scale and capital base, as they may have different risk appetites.

Now what about an insurance company's asset/liability management? What about a financial conglomerate exposed also to market risk, trading in assets which have plausibly symmetric returns, and plausibly with different tail heaviness / index of stability? Say, there is an insurance portfolio with asymmetric profits/loss as above, and the firm invests its assets in part in a stock market with symmetric returns, and holding the rest risk-free? From Counterexample 5.2.8 we have no hope for a two-fund theorem even if these markets have the same α, but we have «two-fund monetary separation in each marke» if they are stochastically independent and the theorems given apply on each block. With stochastic dependencies, on the other hand, we are in difficulties already for non-elliptical α-stable vectors (i.e. one single α), and it is certainly no simpler if e.g. the stock market is a geometric Brownian Black–Scholes market while the insurance portfolios is non-elliptical.

6.3 Concluding remarks

This paper has shown how Ross-type portfolio separation in dynamic models is easily inherited from static models without dynamic programming. For Lévy processes we have to make the additional assumption of infinite divisibility, making it in some sense less general than the static case. This model bases preferences on first-order stochastic dominance, thus avoiding issues of convergence of expected utility. The transition from static to dynamic model to dynamic model can be done with little if any stochastic calculus – not even the second-order term from the Itô formula needs to be introduced. Rather, one utilizes basic tools of probability theory: the theory of the distributions in question, characteristic functions for stochastic processes, simple manipulation of double expectations, and, from analysis, the standard way of defining an integral by simple functions and extension by analytic continuation. If one for the purposes of an exposition is willing to make a shortcut like, e.g. *«* $\boldsymbol{v}^\top \mathrm{d}\boldsymbol{Z} \sim \sqrt{\boldsymbol{v}^\top \varGamma \boldsymbol{v}}\ \mathrm{d}Z_0$ *for a standard Brownian* Z_0 *»,* then one can do without subsection 3.1 (which proves this equivalence!) and then the classical two-fund separation theorem for geometric Brownian Black–Scholes markets is shown with quite a minimum of machinery. The modest requirements are in part due to the modest level of ambition, of merely reducing the dimensionality of the portfolio optimization problem rather than solving it completely.

In this paper, we have used the method to easily recover the separation property for classical cases, and extend these under no-shorting portfolio restrictions, making it possible (and easy!) to allow for certain non-symmetric returns distributions. While this may be of interest for modelling other financial activities than the trading portfolio perspective (like, lending and insurance), there are also challenges due to the rich dependence structures possible, and possibilities for future research on generalizations of this celebrated theorem.

References

[1] R. Agnew, *Counter-examples to an assertion concerning the normal distribution and a new stochastic price fluctuation model.*, Rev. Econ. Stud, 38 (1971), pp. 381–383.

[2] K. Borch, *A note on uncertainty and indifference curves*, Rev. Econ. Stud, 36 (1969), pp. 1–4.

[3] B. O. Bradley and M. S. Taqqu, *Financial risk and heavy tails*, in Rachev [43], pp. 35–103.

[4] S. Cambanis, S. Huang, and G. Simons, *On the theory of elliptically contoured distributions*, J. Multivariate Anal., 11 (1981), pp. 368–385.

[5] D. Cass and J. E. Stiglitz, *The structure of investor preferences and asset returns, and separability in portfolio allocation: a contribution to the pure theory of mutual funds*, J. Econom. Theory, 2 (1970), pp. 122–160.

[6] G. Chamberlain, *A characterization of the distributions that imply mean–variance utility functions*, J. Econom. Theory, 29 (1983), pp. 185–201.

[7] B. de Finetti, *Il problema dei "pieni".*, Giorn. Ist. Ital. Attuari, 11 (1940), pp. 1–88.

[8] R. M. Dudley, *Wiener functionals as Itô integrals*, Ann. Probability, 5 (1977), pp. 140–141.

[9] E. F. Fama, *Portfolio analysis in a stable Paretian market*, Management Sci., 11 (1965), pp. 404–419.

[10] M. S. Feldstein, *Mean-variance analysis in the theory of liquidity preference and portfolio selection*, Rev. Econ. Stud, 36 (1969), pp. 5–12.

[11] N. C. Framstad, *Coherent portfolio separation—inherent systemic risk?*, Int. J. Theor. Appl. Finance, 7 (2004), pp. 909–917.

[12] ——, *On portfolio separation in the Merton problem with bankruptcy or default*, in Proceedings of the International Conference on Stochastic Analysis and Applications, Dordrecht, 2004, Kluwer Acad. Publ., pp. 249–265.

[13] ——, *Portfolio separation properties of the skew-elliptical distributions, with generalizations*, Statist. Probab. Lett., 81 (2011), pp. 1862–1866.

[14] ——, *Portfolio separation with α-symmetric and pseudo-isotropic distributions.* Memorandum no. 12/2011, Department of Economics, University of Oslo, 2011.

[15] E. D. Giorgi, T. Hens, and J. Mayer, *A note on reward-risk portfolio selection and two-fund separation*, Finance Research Letters, 8 (2011), pp. 52–58.

[16] A. K. Gupta, T. T. Nguyen, and W.-B. Zeng, *Conditions for stability of laws with all projections stable*, Sankhyā: The Indian Journal of Statistics, Series A (1961-2002), 56 (1994), pp. pp. 438–443.

[17] P. Hall, *A Comedy of Errors: The Canonical Form for a Stable Characteristic Function*, Bull. Lond. Math. Soc., 13 (1981), pp. 23–27.

[18] C. D. Hardin, Jr., *On the linearity of regression*, Z. Wahrsch. Verw. Gebiete, 61 (1982), pp. 293–302.

[19] A. W. Janicki, I. Popova, P. H. Ritchken, and W. A. Woyczynski, *Option pricing bounds in an α-stable security market*, Comm. Statist. Stochastic Models, 13 (1997), pp. 817–839. Heavy tails and highly volatile phenomena.

[20] A. Khanna and M. Kulldorff, *A generalization of the mutual fund theorem*, Finance Stoch., 3 (1999), pp. 167–185.

[21] B. B. Mandelbrot, *Fractals and scaling in finance*, Springer-Verlag, New York, 1997. Discontinuity, concentration, risk, Selecta Volume E, With a foreword by R. E. Gomory.

[22] R. N. Mantegna, *Lévy walks and enhanced diffusion in the Milan stock exchange*, Phys. A, 179 (1991), pp. 232–242.

[23] D. J. Marcus, *Nonstable laws with all projections stable*, Z. Wahrsch. Verw. Gebiete, 64 (1983), pp. 139–156.

[24] H. Markowitz, *Portfolio selection*, J. Finance, 7 (1952), pp. 77–91.

[25] ——, *de Finetti scoops Markowitz*, Journal of Investment Management, 4 (2006), pp. 3–18.

[26] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J. Econom. Theory, 3 (1971), pp. 373–413.

[27] ——, *Erratum: "Optimum consumption and portfolio rules in a continuous-time model" (J. Econom. Theory 3 (1971), no. 4, 373–413)*, J. Econom. Theory, 6 (1973), pp. 213–214.

[28] ——, *Continuous-Time Finance*, Blackwell, 1992. Revised edition.

[29] J. K. Misiewicz, *Infinite divisibility of sub-stable processes. I. Geometry of sub-spaces of L_α-space*, Stochastic Process. Appl., 56 (1995), pp. 101–116.

[30] ——, *Infinite divisibility of sub-stable processes. II. Logarithm of probability measure*, J. Math. Sci., 81 (1996), pp. 2970–2979. Proceedings of the XVII Seminar on Stability Problems for Stochastic Models, Part I (Kazan′, 1995).

[31] S. Mittnik and M. S. Paolella, *Prediction of financial downside-risk with heavy-tailed conditional distributions*, in Rachev [43], pp. 385–404.

[32] S. Mittnik, S. Rachev, and E. Schwartz, *Value-at-risk and asset allocation with stable return distributions*, Allg. Stat. Arch., 86 (2002), pp. 53–67.

[33] S. Mittnik and S. T. Rachev, *Stable Paretian models in finance*, Wiley, New York, 2000.

[34] S. Mittnik, S. T. Rachev, and M. S. Paolella, *Stable Paretian modeling in finance: some empirical and theoretical aspects*, in A practical guide to heavy tails (Santa Barbara, CA, 1995), R. J. Adler, R. E. Feldman, and M. S. Taqqu, eds., Birkhäuser Boston, Boston, MA, 1998, pp. 79–110. Statistical techniques and applications, Papers from the workshop held in Santa Barbara, CA, December 1995.

[35] J. P. Nolan, *Stable Distributions – Models for Heavy Tailed Data*, Book in progress, 2011. Retrieved from `http://academic2.american.edu/~jpnolan`.

[36] B. Øksendal, *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, fifth ed., 1998. An introduction with applications.

[37] S. Ortobelli, I. Huber, S. T. Rachev, and E. S. Schwartz, *Portfolio choice theory with non-Gaussian distributed returns*, in Rachev [43], pp. 547–594.

[38] S. Ortobelli, I. Huber, and E. Schwartz, *Portfolio selection with stable distributed returns*, Math. Methods Oper. Res., 55 (2002), pp. 265–300. Special issue on mathematical models in market and credit risk.

[39] L. P. Østerdal, *The mass transfer approach to multivariate discrete first order stochastic dominance: Direct proof and implications*, J. Math. Econom., 46 (2010), pp. 1222–1228.

[40] J. Owen and R. Rabinovitch, *On the class of elliptical distributions and their applications to the theory of portfolio choice*, J. Finance, 38 (1983), pp. 745–52.

[41] F. Pressacco and P. Serafini, *The origins of the mean-variance approach in finance: revisiting de Finetti 65 years later*, Decis. Econ. Finance, 30 (2007), pp. 19–49.

[42] P. E. Protter, *Stochastic integration and differential equations*, vol. 21 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[43] S. T. Rachev, ed., *Handbook of Heavy Tailed Distributions in Finance*, Elsevier, February 2003.

[44] S. A. Ross, *Mutual fund separation in financial theory—the separating distributions*, J. Econom. Theory, 17 (1978), pp. 254–286.

[45] A. D. Roy, *Safety first and the holding of assets*, Econometrica, 20 (1952), pp. 431–449.

[46] G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian random processes*, Chapman & Hall, New York, 1994. Stochastic models with infinite variance.

[47] P. A. Samuelson, *Efficient portfolio selection for Pareto-Lévy investments*, Journal of Financial and Quantitative Analysis, 2 (1967), pp. 107–122.

[48] ——, *General proof that diversification pays*, Journal of Financial and Quantitative Analysis, 2 (1967), pp. 1–13.

[49] K.-i. Sato, *Lévy processes and infinitely divisible distributions*, vol. 68 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[50] W. Schachermayer, M. Sîrbu, and E. Taflin, *In which financial markets do mutual fund theorems hold true?*, Finance Stoch., 13 (2009), pp. 49–77.

[51] I. J. Schoenberg, *Metric spaces and completely monotone functions*, Ann. of Math. (2), 39 (1938), pp. 811–841.

[52] ——, *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc., 44 (1938), pp. 522–536.

[53] G. P. Szegő, *No more VaR (this is not a typo)*, Journal of Banking & Finance, 26 (2002), pp. 1247–1251.

[54] J. Tobin, *Liquidity preference as behavior toward risk*, Rev. Econ. Stud., 25 (1958), pp. 65–86.

[55] R. J. Williams, *Book Review: Stochastic integration and differential equations—a new approach*, Bull. Amer. Math. Soc. (N.S.), 25 (1991), pp. 170–180.