Stochastic Portfolio Theory Optimization and the Origin of Rule-Based Investing*

Gianluca Oderda†

February 12, 2014

Abstract

Over the last few years, a number of rule-based portfolio allocation approaches deviating from market-cap weighting have been suggested to offer a superior return vs. risk tradeoff relative to a market cap weighted benchmark allocation. The list of proposed rule-based, non cap-weighted approaches to asset allocation includes risk-focused approaches, such as minimum variance, risk parity and maximum diversification; agnostic approaches, such as equal weighting, and fundamental-focused approaches. While many empirical results have been shown, supporting the argument that rule-based, non cap-weighted approaches should be more efficient, compared to a market cap weighted benchmark allocation, to our knowledge, no theoretical justification has been offered. We build on the stochastic portfolio theory framework of Fernholz, to study the evolution of portfolio wealth relative to a market-cap weighted index. We prove that, given a market capitalization weighted index, the determination of the optimal portfolio, maximizing relative logarithmic growth with respect to this index at fixed tracking error risk, generates two-fund separation. The optimal portfolio always corresponds to the linear combination of two risky portfolios: the market-capitalization-weighted index itself, and a second portfolio completely independent from market weights, composed by four sub-portfolios consisting of rule-based, non cap-weighted allocation schemes: the global minimum variance portfolio, the equally weighted portfolio, the risk parity portfolio and the high cash flow rate of return portfolio.

*We would like to thank two anonymous referees for valuable comments, which made it possible to improve the first version of the manuscript

†Head of Quantitative Investments, Ersel Asset Management, Piazza Solferino 11, 10121 Torino. Email:gianluca.oderda@ersel.it Phone: +39 (0)11 552 0622

Keywords: Stochastic Portfolio Theory, Market Diversity, Portfolio Construction, Risk-Based Investing, Minimum Variance, Risk Parity, Maximum Diversification, Equal-Weight, Low-Volatility Anomaly

JEL Classification. G11.

1 Introduction

Over the last few years, a number of rule-based approaches to passive investing have gained popularity by claiming to offer risk-adjusted performance superior to that of traditional market capitalization-weighted indices. Three main typologies of rule-based, non cap-weighted allocation approaches have been proposed in the literature.

In risk-focused approaches, portfolio weights are only function of specific risk properties of the constituents. A first example of risk-focused approach is given by the minimum variance portfolio, first suggested by Haugen and Baker [1991] (for a review of minimum variance portfolios, see Clarke, Silva and Thorley [2011]), which arises naturally as the left-most portfolio on Markowitz's efficient frontier, and in simple terms can be thought to be the fully-invested portfolio with minimum risk. A second example of risk-focused approach is represented by the risk parity portfolio, first introduced by Qian [2005] and Qian [2006], whose properties have been extensively studied by Maillard, Roncalli and Teiletche [2010]: it is defined as the portfolio in which the risk contribution from each asset is made equal on an ex-ante basis, maximizing risk diversification.

A third and more recent example of risk-focused approach is represented by the maximum diversification portfolio, introduced by Choueifaty and Coignard [2008]. This portfolio is designed to maximize the ratio between the portfolio-weighted sum of asset volatilities and portfolio volatility.

The second typology of rule-based, non cap-weighted allocation approaches is represented by the agnostic equally weighted portfolio, in which there is no apparent link between any market or risk-related information and portfolio weights. An equally weighted portfolio allocates a fraction $1/N$ of the portfolio to each of the N assets available for investment. DeMiguel, Garlappi and Uppal [2009] evaluate the out-of-sample performance of the equally weighted strategy relative to the portfolio policy based on the sample-based mean-variance portfolio model (and also that of some of its extensions, designed to reduce the impact of estimation error). They conclude that no mean-variance portfolio model is consistently better than the equally weighted strategy in terms of Sharpe ratio, certainty-equivalent return, or turnover. Out of sample, the gain from optimal diversification appears to be more than offset by estimation errors in expected returns and covariances.

The third and last typology of rule-based, non cap-weighted allocation approaches is represented by fundamental-focused approaches, in which portfolio weights are a function of some fundamentals of the constituents, for instance the expected rate of return due to cash flows, such as coupons or dividends. A popular example of fundamental-focused allocation approach is the high dividend yield portfolio. In the seminal paper on fundamental indexation Arnott, Hsu and Moore [2005] rank all companies by trailing five-year average gross dividends, select the top one thousand companies under this metric and include each of these companies in the high dividend fundamental index at its relative metric weight. They show that a fundamental index weighted by gross dividends substantially outperforms a cap-weighted index.

Different authors have come up with empirical studies analyzing whether any of the above rule-based, non cap-weighted allocation approaches can be considered superior from a return vs. risk perspective, in comparison with cap-weighted indices. Chow, Hsu, Kalesnik, and Little [2011] find that most rule-based, non cap-weighted allocation strategies outperform their cap-weighted counterparts because of exposure to value and size factors. Leote, Lu and Moulin [2012] compare different rule-based, non cap-weighted allocation strategies on an equity universe (equally weighted portfolio, two variants of risk parity portfolios, the minimum variance portfolio and the maximum diversification portfolio) and analyze the factors behind their risk and performance. They show that each of these strategies, irrespective of its underlying complexity, can be explained by few equity style factors: low beta, small cap, value, and low residual volatility. Recently, Gander, Leveau and Pfiffner [2013] show that investing in just one type of rule-based, non cap-weighted allocation methodology often leads to unwanted concentration and cluster risks. They claim that, in order to avoid this problem, it is crucial to diversify across the different rule-based, non cap-weighted allocation methods.

In spite of the abundance of empirical works, to our knowledge, no authors have managed to show what is the theoretical reason why a particular rule-based, non cap-weighted allocation approach, or a combination thereof, should emerge as a superior portfolio construction methodology from a return vs. risk perspective, relative to a market-capitalization-weighted benchmark.

In this paper we attempt to do so, by building on the stochastic portfolio theory framework of Fernholz [2002], to study the evolution of portfolio wealth relative to a market index. The central assumption in stochastic portfolio theory, which is also crucial for the result shown in this paper, is diversity of the financial market, namely the fact that market capitalization can never be concentrated in a single asset. Under this reasonable assumption, when a financial market component (depending on the financial market we observe, this could be a single company, a country, a market sector, or an asset class) grows extremely large, its rate of growth must decline, to ensure that it will not end up dominating the market as a whole.

The derivation of our result starts from the decomposition of the logarithmic return of a generic portfolio relative to a market index. In the return decomposition formula we obtain, we identify two terms: a drift term, which we aim to maximize, and a noise term, which cannot be controlled, but can be shown to remain bounded, if the market remains diverse all the time. The solution to the drift maximization problem at fixed tracking risk budget generates a two-fund separation theorem [1]: the investor's optimal portfolio can be constructed by holding each of the market portfolio and of a risky portfolio fully independent from market weights, in a ratio depending on the tracking risk constraint only. More importantly, we show that the portfolio weight component independent from market weights, emerging from the maximization of the drift term at fixed tracking error risk level, is given by the linear combination of four sub-portfolios using rule-based, non cap-weighted allocation schemes: the global minimum variance portfolio, the equally weighted portfolio, the risk parity portfolio and the high cash flow rate of return portfolio. In particular, we find that the equally weighted and risk parity portfolio components emerge as solutions of the drift maximization problem for its variance-dependent component.

First, in Section 3, we derive our result by assuming myopic agents using logarithmic utility, aiming to maximize relative logarithmic wealth between a portfolio and a reference cap-weighted index. Then we derive the utility function of an investor aiming to maximize relative arithmetic return between portfolio and reference cap-weighted index, at fixed tracking risk budget. We show that also in this case the same two-fund separation theorem can be obtained: within the optimal solution, the risky portfolio component independent

[1] For a review of portfolio separation theorems, see Ingersoll [1987], Chapter 6.

from market weights conserves a structure identical to the one discussed for the relative logarithmic wealth maximization problem.

Jurczenko, Michel and Teiletche [2013] have recently shown that rule-based, non cap-weighted portfolio construction methodologies are special cases of a generic function defined by two parameters, one controlling for the sensitivity to covariance estimates, and the other setting the tolerance for individual total risk. Their result is elegant, because it allows to unify risk-based investment approaches under a single mathematical description. The difference with what we present in this paper, however, is quite significant: in their paper, risk-based investment approaches appear by minimizing portfolio variance with an ad hoc constraint on portfolio weights, depending on asset covariances and on the two parameters just mentioned. In our paper, risk-based investment approaches appear by maximizing a component of the drift term in the portfolio return decomposition, which can be derived from stochastic portfolio theory, following a descriptive approach, consistent with the observable characteristics of actual portfolios and markets.

We also emphasize that solving the maximization problem for the variance-dependent drift component is closely related to finding the maximum diversification portfolio of Choueifaty and Coignard [2008]. We add to the results of Choueifaty and Coignard [2008] by proving that their maximum diversification portfolio can be written as a linear combination of an equally weighted portfolio and of a risk parity portfolio.

The high cash flow rate of return portfolio is the solution of the drift maximization problem for its cash flow rate-dependent component, and allocates higher weights to assets whose expected dividend or coupon payments are higher relative to their prices, therefore exhibiting a "value" bias.

According to the results we provide, optimal portfolios always come from a combination of rule-based, non cap-weighted allocation approaches, never from a single methodology alone. In this respect, we provide a theoretical justification to the empirical findings in the work of Gander, Leveau and Pfiffner [2013]: a rule-based, non cap-weighted allocation strategy is optimal when it is well diversified not just across constituents, but also across different allocation methods.

The paper is organized as follows: in Section 2 we present the model we use for financial asset returns and introduce the value process for a portfolio, both in absolute terms and relative to a benchmark market-cap weighted index. Section 3 contains the main result of this paper: we derive the optimal portfolio maximizing relative drift with respect to a market-cap weighted index at fixed tracking error risk, proving the two-fund separation result. In Section 4 we draw our conclusions. Lengthy mathematical proofs are in Appendix A.

The content of this paper has been deliberately kept theoretical. In a forthcoming paper we will conduct an empirical analysis on the US financial market, identified by its three major asset classes (equities, government bonds and corporate bonds), as well as on other asset universes, to support the theoretical findings presented in this document.

2 Financial Markets and Portfolio Strategies in Stochastic Portfolio Theory

2.1 The Model for Financial Asset Returns

Our description of the financial market closely follows that in Fernholz [2002]. We consider a financial market where uncertainty is described by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, F_t)$ on which we define a $k-$dimensional Brownian motion W. We assume that F_t is the natural filtration of the Brownian motion. There are $2 \leq n \leq k$ risky securities with price process $X_i(t)$, $i \in \{1, ...n\}$, whose log risky prices evolve according to:

$$d \log X_i(t) = (\gamma_i(t) + \delta_i(t))\, dt + \xi_i^\top(t) dW(t), \tag{1}$$

where we define $\gamma_i(t)$ the logarithmic price growth rates, and $\delta_i(t)$ the cash flow rates of return, related to dividend or coupon payments. The reason why we keep cash flow rates of return separate from logarithmic price growth rates is that, in our context, dividend or coupon payments allow assets to have returns, without affecting their own market capitalizations (and therefore their market cap weights). In the above formula $\xi_i(t)$ is a $k-$dimensional vector, which measures the sensitivity of stock i to k Brownian diffusion processes. We call $X(t) := (X_1(t), \cdots, X_n(t))$ the column vector of risky prices, $\gamma(t) := (\gamma_1(t), \cdots, \gamma_n(t))$ the column vector of mean rates of returns, $\delta(t) := (\delta_1(t), \cdots, \delta_n(t))$ the column vector of cash flow rates of returns, $\xi(t) := (\xi_1^\top(t), \cdots, \xi_n^\top(t))$ the $n \times k-$dimensional diffusion matrix and $\sigma(t) := \xi(t)\xi^\top(t)$ the $n \times n-$dimensional covariance matrix. Applying straightforward Ito's calculus, risky prices evolve according to

$$dX_i(t) = X_i(t) \left[(\alpha_i(t) + \delta_i(t))\, dt + \xi_i^\top(t) dW(t) \right], \tag{2}$$

where we define $\alpha_i(t) := \gamma_i(t) + \frac{\sigma_i^2(t)}{2}$ the mean rates of returns. Notice that the growth rate $\gamma_i(t)$ is a better indicator of long-term price behavior than the rate of return $\alpha_i(t)$. In Fernholz [2002] it is proved [2] that, over a long time horizon, the logarithmic price of an asset is fully determined by the time integral of its growth rate, $\gamma_i(t)$, assuming that its volatility remains bounded.

[2] see Proposition 1.3.1 in Chapter 1

Assumption 1: *There exists a positive constant ϵ such that*

$$x^\top \hat{\sigma}(t) x \geq \epsilon ||x||^2 \tag{3}$$

This corresponds to a uniformly non-degenerate market (Fernholz [2002]), where the covariance matrix $\sigma(t)$ is nonsingular and positive definite at all times.

2.2 The Portfolio Value Process

We now define the portfolio process $\pi(t) := (\pi_1(t), \cdots, \pi_n(t))$, which corresponds to a trading strategy limited to investing only in risky securities. We further impose that it corresponds to a full investment, i.e. that $\pi^\top(t)e = 1$, where e is a $n-$dimensional vector of ones. The value process corresponding to a portfolio $\pi(t)$ is denoted $Z_\pi(t)$ and its logarithmic return is given by

$$d \log Z_\pi(t) = (\gamma_\pi(t) + \delta_\pi(t))\, dt + \xi_\pi^\top(t) dW(t), \tag{4}$$

where growth rate, cash flow rate and volatility of the value process are given by

$$\gamma_\pi(t) = \pi(t)^\top \gamma(t) + \gamma_\pi^\star(t), \tag{5}$$

$$\delta_\pi(t) = \pi(t)^\top \delta(t), \tag{6}$$

$$\xi_\pi(t) = \pi(t)^\top \xi(t). \tag{7}$$

The portfolio excess growth rate $\gamma_\pi^\star(t)$ is defined as follows:

$$\gamma_\pi^\star(t) = \frac{1}{2} \left[\pi(t)^\top \text{diag}(\sigma(t)) - \pi(t)^\top \sigma(t) \pi(t) \right], \tag{8}$$

where $\text{diag}(x)$ is the column vector whose components are the diagonal elements of the matrix x. The excess growth rate is the difference between the weighted sum of individual stock variances and the overall portfolio variance, and can therefore be interpreted as a diversification return.

The next result is central for our later derivation and is summarized in the following proposition.

Proposition 1: *The value process corresponding to a portfolio $\pi(t)$ can be written* [3] *as the sum of a martingale component, whose expected value is zero, and by a drift term, which depends on asset growth rates, on the portfolio excess growth rate and on the portfolio cash flow rate:*

$$\begin{aligned} d\log Z_\pi(t) = \pi(t)^\top \gamma(t)dt + \frac{1}{2}\left[\pi(t)^\top \mathrm{diag}(\sigma(t)) - \pi(t)^\top \sigma(t)\pi(t)\right]dt + \pi(t)^\top \delta(t)dt \\ + \xi_\pi^\top(t)dW(t) \end{aligned} \tag{9}$$

2.3 The Market Portfolio

A particular portfolio strategy is the market portfolio $\mu(t)$, where the weights invested in each stock correspond to their relative market capitalizations. Assuming a single infinitely divisible share per stock, the value process of the market portfolio is $Z_\mu(t) = X(t)^\top e$, and we define the market capitalization weight of a stock as follows

$$\mu_i(t) = \frac{X_i(t)}{Z_\mu(t)}. \tag{10}$$

We now introduce the notion of market diversity, which is central for the validity of our main results in the paper.

Definition 1: *The financial market is diverse if there exist a finite constant $\kappa > 0$ such that*

$$\max_{1\leq i\leq n} \mu_i(t) \leq 1 - \kappa. \tag{11}$$

Assumption 2: *The market coefficients $\gamma(t)$ and $\xi(t)$ are such that diversity is maintained up to time T.*

While diversity seems like a reasonable assumption, as anti-trust law would prevent the existence of a single dominating corporation, it is at odds with common mathematical description of financial markets. For example, the Black-Scholes market with constant coefficient geometric Brownian motion is not diverse. Indeed, in such a setting, the stock price with largest constant growth rate γ_i ends up dominating the market with unit probability.

[3] see Proposition 1.1.5 in Chapter 1 of Fernholz [2002], for a formal proof

2.4 The Relative Value Process: Portfolio vs. Market

The logarithmic return of portfolio π relative to a reference portfolio ν is given by

$$d\log\left(\frac{Z_\pi(t)}{Z_\nu(t)}\right) = \sum_{i=1}^{n} \pi_i(t) d\log(X_i(t)/Z_\nu(t)) + \gamma_\pi^\star(t)dt + (\delta_\pi(t) - \delta_\nu(t))\,dt \tag{12}$$

In this formula $d\log(X_i(t)/Z_\nu(t))$ is the price log return of asset i relative to the reference portfolio ν. Remarkably [4] , the portfolio excess growth rate $\gamma_\pi^\star(t)$ can be written as in equation (8) using the covariance of stock return, $\sigma(t)$, or using the covariance of relative stock returns:

$$\begin{aligned} \gamma_\pi^\star(t) &= \frac{1}{2}\left[\pi(t)^\top \mathrm{diag}(\sigma(t)) - \pi(t)^\top \sigma(t)\pi(t)\right] \\ &= \frac{1}{2}\left[\pi(t)^\top \mathrm{diag}(\tau^\nu(t)) - \pi(t)^\top \tau^\nu(t)\pi(t)\right], \end{aligned} \tag{13}$$

where τ^ν denotes the relative covariance matrix with elements

$$\begin{aligned} \tau_{ij}^\nu(t) &= \frac{1}{dt} d\langle \log\frac{X_i(t)}{Z_\nu(t)}, \log\frac{X_j(t)}{Z_\nu(t)}\rangle \\ &= (\sigma_{ij}(t) - \sigma_{i\nu}(t) - \sigma_{j\nu}(t) + \sigma_{\nu\nu}(t)), \end{aligned} \tag{14}$$

and $\sigma_{j\nu}(t)$ denotes the instantaneous covariance between stock j and the reference portfolio ν.

A special reference portfolio is the market portfolio. In this case, the relative return equation simplifies significantly and can be expressed as a function of market capitalizations and relative covariances of stock prices with respect to the market portfolio. To see this, it is enough to replace the generic reference portfolio ν with the market portfolio μ in Eq.(12), and to remember the definition of market portfolio weight, Eq.(10). One finds:

$$d\log(X_i(t)/Z_\mu(t)) = d\log(\mu_i(t)) \tag{15}$$

The next Proposition describes the portfolio relative value process vs. the market.

Proposition 2: *The return of a portfolio π relative to the market portfolio μ can be*

[4] see Lemma 1.3.4 in Chapter 1 of Fernholz [2002], for a formal proof

written [5] *as the sum of two terms: the portfolio-weighted sum of logarithmic changes in holdings' market capitalization weights, plus a drift term, which depends on the portfolio excess growth rate and on the cash flow rate differential between the portfolio and the market. It is explicitly given by:*

$$d\log\left(\frac{Z_\pi(t)}{Z_\mu(t)}\right) = \frac{1}{2}\left[\pi(t)^\top \mathrm{diag}(\sigma(t)) - \pi(t)^\top\sigma(t)\pi(t)\right]dt + \left(\pi(t)^\top - \mu(t)^\top\right)\delta(t)dt + \sum_{i=1}^{n}\pi_i(t)d\log\mu_i(t). \tag{16}$$

Notice that asset logarithmic growth rates, $\gamma_i(t)$, present in the drift term of Eq.(9), which describes portfolio logarithmic wealth dynamics, do not appear in the drift term of Eq.(16), which describes relative logarithmic wealth dynamics of the same portfolio vs. a market cap weighted index. In reality, asset growth rates relative to the market are hidden in the asset market cap weight dynamics.

3 Optimal Relative Drift Portfolios in Stochastic Portfolio Theory: Two Fund Separation Theorem

In this Section we present the main result of our paper. We concentrate on the maximization of portfolio relative log wealth at a fixed tracking risk level. Our utility function, at each time t, can be written as:

$$U\left(\pi(t),\delta(t),\sigma(t),\mu(t)\right) = \left(\pi(t)^\top - \mu(t)^\top\right)\delta(t) + \frac{1}{2}\left[\pi(t)^\top \mathrm{diag}(\sigma(t)) - \pi(t)^\top\sigma(t)\pi(t)\right] + \\ -\lambda_1\left(\pi^T(t)e - 1\right) - \lambda_2\left[(\pi(t)^\top - \mu(t)^\top)\sigma(t)(\pi(t)-\mu(t)) - \chi^2\right], \tag{17}$$

We are trying to maximize the drift term in Eq.(16), while imposing the budget constraint, as well as the constraint that portfolio tracking error risk must be equal to a predetermined value χ. In the above equation λ_1 and λ_2 are Lagrange multipliers. The above problem is convex, since the matrix $\sigma(t)$ is positive definite, according to Assumption 1. By taking the gradient of the utility function with respect to the vector of weights $\pi(t)$, the solution

[5] see Proposition 1.2.5 in Chapter 1 of Fernholz [2002], for a formal proof

to the utility maximization problem can be written as:

$$\pi(t) = (1 - A)\mu(t) + A\sigma^{-1}(t)\delta(t) + \frac{A}{2}\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) + B\sigma^{-1}(t)e, \tag{18}$$

with the coefficients A and B given by:

$$A = \frac{1}{1 + 2\lambda_2}$$
$$B = \frac{\lambda_1}{1 + 2\lambda_2} = \lambda_1 A.$$

The coefficients A and B can be derived by imposing the budget constraint and the tracking error risk constraint. Eq. (18) can be rewritten as:

$$\pi(t) = (1 - A)\mu(t) + A\left[\sigma^{-1}(t)\delta(t) + \frac{1}{2}\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) + \lambda_1\sigma^{-1}(t)e\right], \tag{19}$$

We define $\eta(t)$:

$$\eta(t) = \sigma^{-1}(t)\delta(t) + \frac{1}{2}\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) + \lambda_1\sigma^{-1}(t)e. \tag{20}$$

By imposing the budget constraint, it is easily shown that $\eta(t)^\top e = 1$, which allows to obtain λ_1. The determination of optimal portfolio weights generates a two-fund separation theorem: the optimal solution is given by the linear combination of only two distinct risky portfolio, the market portfolio, and a portfolio completely independent from market portfolio weights:

$$\pi(t) = (1 - A)\mu(t) + A\eta(t)\,, \tag{21}$$

with $\eta(t)$ given by:

$$\begin{aligned}\eta(t) = {} & \sigma^{-1}(t)\delta(t) + \frac{1}{2}\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) - \\ & - \frac{e^T\sigma^{-1}(t)\delta(t) + \frac{1}{2}e^T\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) - 1}{e^T\sigma^{-1}(t)e}\sigma^{-1}(t)e\,.\end{aligned} \tag{22}$$

The parameter A is determined by the tracking risk constraint: if such constraint becomes very stringent, $\lambda_2 \to \infty$, and the optimal portfolio reduces to the market-cap weighted portfolio. It remains to be proved that Eq. (16), under the optimal $\pi(t)$ choice, consists of a drift term and of a bounded noise term. We start by noticing that:

$$d\log\mu_i(t) = \frac{d\mu_i(t)}{\mu_i(t)} - \frac{1}{2}\tau_{ii}(t)dt\,. \tag{23}$$

As a consequence,the $\mu(t)$-dependent part of the optimal weight solution in Eq. (21) gives rise to a pure drift term:

$$\begin{aligned}(1-A)\sum_{i=1}^{n}\mu_i(t)d\log\mu_i(t) &= (1-A)\left[\sum_{i=1}^{n}d\mu_i(t) - \frac{1}{2}\sum_{i=1}^{n}\mu_i(t)\tau_{ii}(t)dt\right] = \\ = -\frac{(1-A)}{2}\sum_{i=1}^{n}\mu_i(t)\tau_{ii}(t)dt &= -\frac{(1-A)}{2}\left[\sum_{i=1}^{n}\mu_i(t)\sigma_{ii}(t) - \sum_{i,j=1}^{n}\mu_i(t)\mu_j(t)\sigma_{ij}(t)\right]dt.\end{aligned} \tag{24}$$

On the other hand, we notice that, since $\eta(t)$ is fully independent from market weights, we can certainly write:

$$A\sum_{i=1}^{n}\eta_i(t)d\log\mu_i(t) = d\left[A\sum_{i=1}^{n}\eta_i(t)\log\mu_i(t)\right]\,. \tag{25}$$

The function $A\sum_{i=1}^{n}\eta_i(t)\log\mu_i(t)$ is bounded, if market diversity is assumed to hold at each time t. By replacing the optimal weight solution, Eq. (21) into Eq. (16), after some algebra, we find:

$$\begin{aligned}d\log\left(\frac{Z_\pi(t)}{Z_\mu(t)}\right) &= \frac{1}{2}\left[A\eta(t)^\top\mathrm{diag}(\sigma(t)) - A^2\eta(t)^\top\sigma(t)\eta(t)\right]dt - \\ &- A(1-A)\left[\mu(t)^\top\sigma(t)\eta(t) - \frac{1}{2}\mu(t)^\top\sigma(t)\mu(t)\right]dt + \\ &+ A\left(\eta(t)^\top - \mu(t)^\top\right)\delta(t)dt + Ad\left[\sum_{i=1}^{n}\eta_i(t)\log\mu_i(t)\right]\,.\end{aligned} \tag{26}$$

The above expression proves that the relative logarithmic return between the optimal portfolio we have found and the market index consists indeed of a (maximized) drift term and of a bounded noise term. In conclusion, we have seen from Eqs. (21) and (22) that the determination of the optimal portfolio maximizing relative drift vs. a market cap weighted index at fixed tracking risk budget originates a two-fund separation theorem: the optimal solution is the linear combination of the market portfolio and of a risky portfolio independent from market weights and consisting itself of three sub-components: a high cash flow-rate of return portfolio, an asset variance-dependent portfolio, and a global minimum variance portfolio. The optimal portfolio solution deviates from the market cap weighted benchmark depending on the tracking risk constraint, but always builds on these rule-based, non cap-weighted allocation building blocks.

Our derivation so far is based on the maximization of relative logarithmic wealth and would seem to be valid for myopic agents only. We believe that our results could be more general than that. Let us take the case of an investor seeking to maximize relative arithmetic return between portfolio and cap-weighted benchmark at fixed tracking error risk. Such an investor would look at the expected relative return, $E\left[\frac{dZ_\pi(t)}{Z_\pi(t)} - \frac{dZ_\mu(t)}{Z_\mu(t)}\right]$. By means of standard Ito calculus, we can write relative arithmetic return as:

$$\frac{dZ_\pi(t)}{Z_\pi(t)} - \frac{dZ_\mu(t)}{Z_\mu(t)} = d\log\left(\frac{Z_\pi(t)}{Z_\mu(t)}\right) + \frac{1}{2}\left[\pi(t)^\top\sigma(t)\pi(t)dt - \frac{1}{2}\mu(t)^\top\sigma(t)\mu(t)\right]dt\,. \qquad (27)$$

The utility function which the investor would maximize would be given by a modification of Eq. (17), based on Eq. (27):

$$\tilde{U}\left(\pi(t),\delta(t),\sigma(t),\mu(t)\right) = \left(\pi(t)^\top - \mu(t)^\top\right)\delta(t) + \frac{1}{2}\left[\pi(t)^\top\mathrm{diag}(\sigma(t)) - \mu(t)^\top\sigma(t)\mu(t)\right] + \\ - \lambda_1\left(\pi^T(t)e - 1\right) - \lambda_2\left[(\pi(t)^\top - \mu(t)^\top)\sigma(t)(\pi(t) - \mu(t)) - \chi^2\right]. \qquad (28)$$

By repeating the same steps shown above, the determination of the optimal portfolio generates again a two-fund separation theorem:

$$\pi(t) = \mu(t) + \tilde{A}\tilde{\eta}(t)\,, \qquad (29)$$

where, by the budget constraint, the portfolio $\tilde{\eta}(t)$ is the zero weight, long/short portfolio variant of $\eta(t)$, satisfying $e^\top \tilde{\eta}(t) = 0$:

$$\begin{aligned}\tilde{\eta}(t) = \sigma^{-1}(t)\delta(t) + \frac{1}{2}\sigma^{-1}(t)\mathrm{diag}(\sigma(t)) - \\ - \frac{e^T\sigma^{-1}(t)\delta(t) + \frac{1}{2}e^T\sigma^{-1}(t)\mathrm{diag}(\sigma(t))}{e^T\sigma^{-1}(t)e}\sigma^{-1}(t)e\,.\end{aligned} \tag{30}$$

In conclusion, our result is not limited to myopic agents with logarithmic utility, but can be further generalized.

In the next subsection we turn our attention to the asset variance dependent term in the optimal weight solution, proving that it can be written as the linear combination of an equally weighted portfolio and a risk parity portfolio. This the central contribution of our paper.

3.1 The Asset Variance Dependent Term in the Optimal Portfolio Solution

In Eq.(22) we focus on the term $\sigma^{-1}(t)\mathrm{diag}(\sigma(t))$. We start by observing that the covariance matrix $\sigma(t)$ can be written as:

$$\sigma(t) = \Gamma(t) \circ R(t) = \Gamma(t) \circ (R_0(t) + \Delta R(t))\,, \tag{31}$$

where the symbol $\circ$ denotes the Hadamard product between matrices. The element i, j of the matrix $\Gamma(t)$ is given by the product of volatilities for assets i and j:

$$\Gamma_{ij}(t) = \sigma_i(t)\sigma_j(t). \tag{32}$$

The matrix $R(t)$ is the true correlation matrix, whereas the matrix $R_0(t)$ is its equal correlation approximation, given by:

$$R_0(t) = \begin{pmatrix} 1 & \bar{\rho}(t) & \bar{\rho}(t) & \dots & \bar{\rho}(t) \\ \bar{\rho}(t) & 1 & \bar{\rho}(t) & \dots & \bar{\rho}(t) \\ \bar{\rho}(t) & \bar{\rho}(t) & 1 & \dots & \bar{\rho}(t) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\rho}(t) & \bar{\rho}(t) & \bar{\rho}(t) & \dots & 1 \end{pmatrix}, \tag{33}$$

where $\bar{\rho}(t)$ is the average correlation between any pair of assets:

$$\bar{\rho}(t) = \frac{1}{N(N-1)} \sum_{i \neq j}^{N} \rho_{ij}(t). \tag{34}$$

The matrix $\Delta R(t)$ is given by the difference of the true correlation matrix and its equal correlation approximation:

$$\Delta R(t) = \begin{pmatrix} 0 & \rho_{12}(t) - \bar{\rho}(t) & \rho_{13}(t) - \bar{\rho}(t) & \dots & \rho_{1n}(t) - \bar{\rho}(t) \\ \rho_{21}(t) - \bar{\rho}(t) & 0 & \rho_{23}(t) - \bar{\rho}(t) & \dots & \rho_{2n}(t) - \bar{\rho}(t) \\ \rho_{31}(t) - \bar{\rho}(t) & \rho_{32}(t) - \bar{\rho}(t) & 0 & \dots & \rho_{3n}(t) - \bar{\rho}(t) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{n1}(t) - \bar{\rho}(t) & \rho_{n2}(t) - \bar{\rho}(t) & \rho_{n3}(t) - \bar{\rho}(t) & \dots & 0 \end{pmatrix}. \tag{35}$$

Notice that in the case of two assets, there is no need of the matrix $\Delta R(t)$. The inverse of the Hadamard product of matrices can be written as:

$$\sigma^{-1}(t) = \Gamma^{-1}(t) \circ \left(R_0(t) + \Delta R(t)\right)^{-1}, \tag{36}$$

where the element i, j of the matrix $\Gamma^{-1}(t)$ is given by:

$$\Gamma_{ij}^{-1}(t) = \frac{1}{\sigma_i(t)\sigma_j(t)}. \tag{37}$$

The sum of matrices $(R_0(t) + \Delta R(t))$ can be inverted by using the fact that $R_0(t)$ is invertible, and by using the formula:

$$[R_0(t) + \Delta R(t)]^{-1} = R_0^{-1}(t) - \left[\mathbf{1} + R_0^{-1}(t)\Delta R(t)\right]^{-1} R_0^{-1}(t)\Delta R(t) R_0^{-1}(t). \tag{38}$$

The inverse of the correlation matrix is therefore equal to the inverse of the equal correlation matrix plus a perturbative expansion in the matrix $\Delta R(t)$. We prove in Appendix A that:

$$\left[\Gamma^{-1}(t) \circ R_0^{-1}(t)\right] \mathrm{diag}(\sigma(t)) = \frac{1}{\bar{\rho}(t) - 1} e - \frac{\bar{\rho}(t) \sum_{i=1}^{N} \sigma_i(t)}{(\bar{\rho}(t) - 1)\left[N\bar{\rho}(t) + (\bar{\rho}(t) - 1)\right]} \begin{pmatrix} \frac{1}{\sigma_1(t)} \\ \frac{1}{\sigma_2(t)} \\ \vdots \\ \frac{1}{\sigma_N(t)} \end{pmatrix} \tag{39}$$

Looking at Eq.(38), we conclude that the expression

$$\Gamma^{-1}(t) \circ [R_0(t) + \Delta R(t)]^{-1} \, \mathrm{diag}(\sigma(t)) \tag{40}$$

yields the linear combination of an equally weighted portfolio (the weight vector e) and of a risk parity portfolio built by assuming a constant correlation matrix, $R_0(t)$. It is well known that the risk parity solution equilibrating a covariance matrix with an underlying constant correlation matrix is a vector of weights inversely proportional to volatilities. From Eqs.(38) and (39), we see that there will also be a correction term, depending on the matrix $\Delta R(t)$, which is the result of the action of the operator

$$\left[\mathbf{1} + R_0^{-1}(t)\Delta R(t)\right]^{-1} R_0^{-1}(t)\Delta R(t)$$

on the linear combination of equally weighted and risk parity portfolio. For the case $N = 2$ assets, there is no correction term.

4 Conclusions

In this paper we have shown that rule-based, non cap-weighted allocation strategies arise naturally from the optimization of the drift term in a stochastic portfolio theory description of portfolio return relative to a market index. Our results allow us to obtain, for each level of tracking error risk, the portfolio with the maximum tradeoff between expected drift, relative to the market index, and tracking error risk.

The portfolio with the maximum expected drift relative to a market index at fixed tracking error risk is the linear combination of only two distinct risky portfolios ("two-fund separation theorem"): the market portfolio itself and a risky portfolio, completely independent of market weights, which consists of a set of rule-based, non cap-weighted allocation approaches. The determination of optimal portfolio weights does not need any predictions on expected asset logarithmic growth rates, which are buried in the unpredictable asset market weight dynamics. We show that, if the market remains diverse, the asset market weight dynamics is stochastic but bounded.

The central finding of our paper is that the maximization of the asset variance-dependent component of relative drift leads to the linear combination of two well known rule-based, non cap-weighted allocation strategies, the equally weighted portfolio and the risk parity portfolio. Whereas in the literature these strategies have always been considered heuristic, our approach allows us to derive them from first principles.

Our result is consistent with the recent empirical research of Gander, Leveau and Pfiffner [2013], who have shown that, investing in just one type of rule-based, non cap-weighted allocation methodology can be risky. In order to avoid unwanted concentration and cluster risks, it is therefore crucial to diversify across different rule-based, non cap-weighted allocation methods, in line with our theoretical optimal result.

In a forthcoming paper we will conduct an empirical analysis on the US financial market, identified by its three major asset classes (equities, government bonds and corporate bonds), as well as on other asset universes, to support the theoretical findings presented in this document.

References

Clarke, R. , de Silva H. , Thorley S. "Minimum-Variance Portfolio Composition." Journal of Portfolio Management, 37,2 (2011), pp. 31-45

Maillard, S. , Roncalli, T. , and Teiletche, J. "The Properties of Equally Weighted Risk Contribution Portfolios." *The Journal of Portfolio Management* , 36,4 (2010), pp. 60-70.

DeMiguel, V. , Garlappi, L. , and Uppal, R. "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *Review of Financial Studies*, 22(5) (2009), pp. 1915-1953.

Arnott, R. ,Hsu J. , and Moore, P. "Fundamental Indexation." *Financial Analysts Journal* 61, 2 (2005), pp. 83.99.

Ingersoll, J. "Theory of Financial Decision Making." Roman & Littlefield Publishers, Inc. 1987.

Chow, T. , Hsu, J. , Kalesnik, V. , and Little, B. Morgan, "A Survey of Alternative Equity Index Strategies." *Financial Analyst Journal*, 67, 5 (2011) pp. 37-57.

Gander, P. , Leveau, D. , and Pfiffner, T. , "Diversification - A Multi-Facetted Concept." *1741 Asset Management Research Note Series 1/2013* Available at SSRN: http://ssrn.com/abstract=2204317 or http://dx.doi.org/10.2139/ssrn.2204317.

Choueifaty, Y. , and Coignard, Y. , "Towards Maximum Diversification." *The Journal of Portfolio Management*, Vol. 35 (2008), No. 1, pp. 40-51.

Fernholz, R. E. *Stochastic Portfolio Theory.* Springer Science + Business Media, 2002, and references therein.

Haugen R. , and Baker, N. ,"The Efficient Market Inefficiency of Capitalization-Weighted Stock portfolios.", *Journal of Portfolio Management* 17 (1991), 35-40.

Leote R. , Lu X. , and Moulin, P. , "Demystifying Equity Risk-Based Strategies: A Simple Alpha plus Beta Description.", *Journal of Portfolio Management* 38 (2012), 56-70.

Jurczenko, E. , Michel, T. , and Teiletche, J. , "Generalized Risk-Based Investing." (March 2013). Available at SSRN: http://ssrn.com/abstract=2205979 or http://dx.doi.org/10.2139/ssrn.2205979

Qian, E. , "Risk Parity Portfolios: Efficient Portfolios through True Diversification", Working paper, Panagora Asset Management (2005).

Qian, E. , "On the Financial Interpretation of Risk Contributions: Risk Budgets Do Add Up.", *Journal of Investment Management* 4 (2006), 1-11.

A Proof of Equation 39

We would like to analytically evaluate:

$$\left[\Gamma^{-1}(t) \circ R_0^{-1}(t)\right] \mathrm{diag}(\sigma(t)), \tag{41}$$

where the matrix $\Gamma(t)$ is defined in Eq.(32), its inverse in Eq.(37) and the constant correlation matrix $R_0(t)$ in Eq.(33) We start by observing that the matrix $R_0(t)$ can be written as:

$$R_0(t) = \bar{\rho}(t) e e^T + (\bar{\rho}(t) - 1) \mathbb{1}, \tag{42}$$

where e is a $n-$dimensional vector of ones. Then the inverse of the matrix $R_0(t)$ can be written as:

$$R_0^{-1}(t) = \phi e e^T + \psi \mathbb{1}. \tag{43}$$

Considering that:

$$\begin{aligned} e e^T e e^T &= N e e^T \\ e e^T \mathbb{1} &= e e^T, \end{aligned} \tag{44}$$

we have:

$$\mathbb{1} = R_0^{-1}(t) R_0(t) = \left[N \phi \bar{\rho}(t) + \psi \bar{\rho}(t) + \phi \left(\bar{\rho}(t) - 1\right)\right] e e^T + \psi \left(\bar{\rho}(t) - 1\right) \mathbb{1}. \tag{45}$$

In order to satisfy the previous equation, the two coefficients ϕ and ψ must be equal to:

$$\begin{aligned} \psi &= \frac{1}{\bar{\rho}(t) - 1} \\ \phi &= -\frac{\bar{\rho}(t)}{\left(\bar{\rho}(t) - 1\right)\left[N \bar{\rho}(t) + \left(\bar{\rho}(t) - 1\right)\right]}. \end{aligned} \tag{46}$$

We therefore have:

$$R_0^{-1}(t) = -\frac{\bar{\rho}(t)}{\left(\bar{\rho}(t) - 1\right)\left[N \bar{\rho}(t) + \left(\bar{\rho}(t) - 1\right)\right]} e e^T + \frac{1}{\bar{\rho}(t) - 1} \mathbb{1}. \tag{47}$$

By looking at the previous equation, at Eq.(41), and at the definition of the matrix $\Gamma^{-1}(t)$, Eq.(37), we observe that:

$$\left[\Gamma^{-1}(t) \circ \mathbb{1}\right] \ \mathrm{diag}(\sigma(t)) = e, \tag{48}$$

and that

$$\left[\Gamma^{-1}(t) \circ ee^T\right] \ \mathrm{diag}(\sigma(t)) = \sum_{i=1}^{N} \sigma_i(t) \begin{pmatrix} \frac{1}{\sigma_1(t)} \\ \frac{1}{\sigma_2(t)} \\ \vdots \\ \frac{1}{\sigma_N(t)} \end{pmatrix}. \tag{49}$$

Therefore, using Eq.(47), we can write:

$$\left[\Gamma^{-1}(t) \circ R_0^{-1}(t)\right] \mathrm{diag}(\sigma(t)) = \frac{1}{\bar{\rho}(t)-1} e - \frac{\bar{\rho}(t) \sum_{i=1}^{N} \sigma_i(t)}{(\bar{\rho}(t)-1)\left[N\bar{\rho}(t) + (\bar{\rho}(t)-1)\right]} \begin{pmatrix} \frac{1}{\sigma_1(t)} \\ \frac{1}{\sigma_2(t)} \\ \vdots \\ \frac{1}{\sigma_N(t)} \end{pmatrix}. \tag{50}$$

We have therefore proved that the result of the action of the operator $\left[\Gamma^{-1}(t) \circ R_0^{-1}(t)\right]$ on the vector of asset variances is the linear combination of an equally weighted portfolio and of a risk parity portfolio.