Beta-Arbitrage strategies: when do they work, and why?

BERRADA, Tony Nicolas, et al.
Beta-Arbitrage strategies: when do they work, and why?∗

Tony Berrada† Reda Jurg Messikh‡ Gianluca Oderda§ Olivier Pictet¶

September 6, 2012

Abstract

Contrary to what traditional asset pricing would imply, a strategy that bets against beta, i.e. long in low beta stocks and short in high beta stocks, tends to out-perform the market. This puzzling empirical fact can be explained through the concept of relative arbitrage. Considering a market in which diversity is maintained, i.e. no single stock can dominate the entire market, we show that beta-arbitrage strategies out-perform the market portfolio with unit probability in finite time. We use the theoretical decomposition of beta-arbitrage excess return to provide empirical support to our explanation on equity country indices, equity sectors and individual stocks. Finally we show how to construct optimal beta-arbitrage strategies that maximize the expected return relative to a given benchmark.

Keywords: Relative arbitrage, Market diversity, Beta

JEL Classification. G11.

∗We thank Olivier Scaillet and conference participants at the Mathematical Finance Days 2011 in Montreal for comments. Financial support through NCCR Finrisk project A4 is greatfully acknowledged.
†University of Geneva and Swiss Finance Institute, GFRI, Bd Pont d’Arve 42, CH-1211 Geneva 4, Switzerland. Email:tony.berrada@unige.ch Phone: +41 (0)22 379 81 26
‡Pictet Asset Management, Rte des Acacias 60 CH-1211 Geneva 73, Switzerland. Email:rmessikh@pictet.com Phone: +41 (0)58 323 2323
§Ersel Asset Management SGR S.p.A. Piazza Solferino, 11 - 10121 Torino, Italia. Email:gianluca.oderda@ersel.it
¶Pictet Asset Management, Rte des Acacias 60 CH-1211 Geneva 73, Switzerland. Email:opictet@pictet.com Phone: +41 (0)58 323 2323
1 Introduction

The main conclusion of the Capital Asset Pricing Model developed by Sharpe [1964], Lintner [1965] and Mossin [1966] is that firms with higher systematic risk should on average earn higher returns than firms with low systematic risk. Despite its fundamental theoretical foundation, this result lacks empirical support. In an attempt to reconcile empirical evidence with theory, Black [1972] and Black [1993] introduce the so called beta factor, by creating a market neutral long-short portfolio that is long in low-beta stocks and short in high-beta stocks. He shows that the beta factor generates significant positive excess return, thereby contradicting the CAPM theory.

In this paper we develop a theoretical framework based on stochastic portfolio theory (Fernholz [2002]) to study the behavior of beta weighted portfolios. We contribute to the existing literature on three levels. First, we derive a return decomposition formula for beta weighted portfolios. We show that portfolios that overweight high beta stocks are dominated by the market portfolio and provide an explanation for the positive excess return of the beta factor. Second, we derive an optimal low beta portfolio strategy that maximizes the instantaneous excess return relative to the market portfolio. This strategy is dynamic and reacts to changes in the stocks covariance structure. Finally we conduct an empirical analysis of the return decomposition that supports our theoretical findings.

The central assumption in stochastic portfolio theory is diversity of the financial market, namely the fact that market capitalization can never be concentrated in a single security. Under this reasonable assumption, when a company grows extremely large, its rate of growth must decline sharply to ensure that it will not end up dominating the market. Therefore, holding the stock with the largest market capitalization is detrimental to any portfolio. In particular, the market portfolio must hold that stock, and is strictly dominated by an adequately chosen dynamic strategy. This is the intuition behind the notion of relative arbitrage first proposed in Fernholz [1999]. We show that this mechanism also applies to portfolios constructed on the basis of a stock’s beta.

We derive a decomposition of the return of a beta-weighted portfolio relative to the market portfolio, in which we identify three main components. Assuming that the market is diverse, the first two components can be shown to be strictly bounded. In other words,
the long term evolution of the relative return of the beta-weighted portfolio is entirely defined by the last term in the return decomposition. This term has no diffusion component and its sign depends exclusively on the choice of beta under/overweight at the portfolio level. A portfolio strategy that underweights low beta stocks and overweights high beta stocks generates a negative drift term. Conversely, over weighting low beta stocks generates a positive drift term.

The explicit structure of the return decomposition allows us to derive an optimal dynamic beta-weighted portfolio, which maximizes the drift term. We show that the strategy is market neutral when the beta dispersion is low, while it is short in the high beta stocks and long in the low beta stocks when beta dispersion is high.

Equipped with the return decomposition, we test our model using three distinct datasets covering different sample periods and asset categories. We consider US stocks, country stock indices and sector indices. In all three cases we find strong support for the model. As predicted by the theory, the portfolios that underweight low beta stocks and overweight high beta stocks under-perform relative to the market portfolio. In addition, the drift term of their log-return relative to the market is strongly negative when the dispersion in beta is largest. We study the empirical performance of the optimal dynamic beta-weighted portfolio on the US stocks dataset and find that it significantly out-performs the market portfolio.

Our result supports a long series of empirical studies which, since the early 70’s, have been questioning the positive linear relationship between stock beta’s and their subsequent returns. Black, Jensen and Scholes [1972] documents the out performance of low-beta stocks relative to high-beta stocks. Miller and Scholes [1972] shows that the security market line over a sample of NYSE stocks is much flatter than theory would predict. Malkiel [1995] studies the relationship between beta and return for mutual funds and finds the two variables to be almost unrelated.

As already mentioned, Black [1972] and Black [1993] introduce the so called beta factor along with a theoretical explanation for its statistical significance. Borrowing restrictions, as well as borrowing reluctance, force investors wanting a high degree of market risk to bid up the prices for high-beta stocks, making low-beta stocks more attractive investment opportunities. More recently, Frazzini and Pedersen [2010] present a model in which some
investors are prohibited from using leverage, and other investors’ leverage is limited by margin requirements. They find that, at equilibrium, tighter portfolio constraints flatten the security market line, by increasing the importance of the intercept relative to the slope. They also show that a market-neutral portfolio, which is long leveraged low-beta securities and short higher-beta securities, should earn a positive expected return on average.

This article contributes to the literature by showing how beta-arbitrage opportunities naturally emerge from the structure of a diverse financial market, where the total capitalization is never concentrated in one single stock. The seemingly harmless requirement of diversity provides a set of tools that allow us to clearly identify the source of beta-arbitrage in the context of our model. In section 2 we present the model of the financial market and derive our central theoretical result. In section 3 we derive the optimal dynamic beta-weighted portfolio and in section 4 we use the relative return decomposition formula to test empirically the validity of our explanation. All proofs are in the appendix.

2 The Model

Our description of the financial market in section 2.1, 2.2 and 2.3 closely follows that in Fernholz and Karatzas [2005]. Our main result on beta-weighted portfolios appears in section 2.4. All equalities and inequalities in the following discussion are assumed to hold almost surely and/or almost everywhere depending on the context. All proofs are collected in the Appendix.

2.1 Financial market and portfolio strategies

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 $X_i(t), i \in \{1, \ldots, n\}$ and we let $X_0(t)$ denote the price process of the riskfree asset. Risky prices evolve according to

$$dX_i(t) = X_i(t) \left( a_i(t)dt + \xi_i^T(t)dW(t) \right),$$
and the locally riskfree asset evolves according to

\[ dX_0(t) = X_0(t)r_t dt. \]

In these equations we define \( a_i(t) \) the mean rates of returns, \( r(t) \) the riskfree rate and \( \xi(t) \) 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, \( a(t) := (a_1(t), \cdots, a_n(t)) \) the column vector of mean rates of returns, \( \xi(t) := (\xi_1^T(t), \cdots, \xi_n^T(t)) \) the \( n \times k \)--dimensional diffusion matrix and \( \sigma(t) := \xi(t)\xi^T(t) \) the \( n \times n \)--dimensional covariance matrix.

**Assumption 1:** There exists a positive constant \( \epsilon \) such that

\[ x^\top \hat{\sigma}(t)x \geq \epsilon ||x||^2 \]

This corresponds to a uniformly non-degenerate market (Fernholz [2002]).

The logarithmic returns of risky stock prices are described by the following equation

\[ d\log X_i(t) = \gamma_i(t) dt + \xi_i^T(t)dW(t) \]

where \( \gamma_i(t) := a_i(t) - \frac{\xi_i^T(t)\xi_i(t)}{2} \) is the growth rate process.

**Assumption 2:** There exists a \( k \)--dimensional market price of risk process \( \theta(t) \) such that

\[ \xi(t)\theta(t) = a(t) \quad \text{and} \quad \int_0^T \theta^\top(t)\theta(t)dt < \infty. \]

Let us define the stochastic exponential \( H(T) \)

\[ H(T) := \exp \left( -\int_0^T \theta^\top(t)dW(t) - \frac{1}{2} \int_0^T \theta^\top(t)\theta(t)dt \right). \]

From assumption 2 we know that \( H(t) := \exp \left( -\int_0^t \theta^\top(s)dW(s) - \frac{1}{2} \int_0^t \theta^\top(s)\theta(s)ds \right) \) is a local martingale and a true martingale if \( E[H(T)] = 1 \). Similarly to other relative arbitrage strategies identified in Fernholz [2002] and Fernholz and Karatzas [2005], the existence
of beta-arbitrage strategies necessarily implies that $H(t)$ is a strict local martingale and that therefore there does not exist an equivalent martingale measure in the context of our financial market.

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 $\pi^\top(t)\mathbf{1}_n = 1$ where $\mathbf{1}_n$ 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

$$d \log Z_\pi(t) = \gamma_\pi(t)dt + \xi_\pi^\top(t)dW(t)$$

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

$$\gamma_\pi(t) = \pi(t)^\top \gamma(t) + \gamma^*_\pi(t)$$
$$\xi_\pi(t) = \pi(t)^\top \xi(t)$$

The portfolio excess growth rate $\gamma^*_\pi(t)$ plays an important role in the derivation of the beta-arbitrage and is defined as follows

$$\gamma^*_\pi(t) = \frac{1}{2} \left[ \pi(t)^\top \text{diag}(\sigma(t)) - \pi(t)^\top \sigma(t) \pi(t) \right]$$

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.\(^1\)

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

\(^1\)Erb and Harvey [2006] show the potential of portfolio excess growth rate in the domain of commodity indices. Regularly rebalancing a portfolio in which asset variances are high and correlations are low, as Eq.(1) clearly shows, is one of the few high-confidence ways investors can boost portfolio return. More recently, Choueifaty and Coignard [2008] introduce “Most-Diversified Portfolios” whose weights maximize a “diversification ratio” defined as the ratio of the weighted average of asset volatilities divided by the portfolio volatility. They show that these portfolios have the strongest out performance potential relative to a market-cap benchmark. It should be mentioned that maximizing the diversification ratio is related to maximizing portfolio excess growth rate.
single infinitely divisible share per stock, the value process of the market portfolio is
\[ Z_\mu(t) = X(t)^\top 1_n, \]
and we define the market capitalization weight of a stock as follows
\[ \mu_i(t) = \frac{X_i(t)}{Z_\mu(t)}. \]

**Assumption 3:** There exists a positive constant \( \kappa \) such that
\[
\sup_{\tau > 0} \left| \int_0^\tau \frac{\partial}{\partial t} \mu(t)^\top \sigma(t) \mu(t) \right| dt < \kappa
\]

This assumption imposes that the covariance matrix of the market does not experience abrupt variations through time and is necessary for our central result.

### 2.2 Diversity

Following the definition introduced in Fernholz [2002] the financial market is diverse if there exist a constant \( \delta > 0 \) such that
\[
\max_{1 \leq i \leq n} \mu_i(t) \leq 1 - \delta.
\]

**Assumption 4:** The market coefficients \( a(t) \) 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 \( \gamma_i \) ends up dominating the market with unit probability.

It is nevertheless possible to explicitly construct diverse markets, as illustrated in Fernholz and Karatzas [2005] and reference therein.

### 2.3 Relative arbitrage

We recall here the notion of relative arbitrage introduced in Fernholz [2002].
**Definition 1:** A portfolio $\pi$ is an arbitrage relative to $\nu$ if

\[
P[Z^\pi_T] \geq P[Z^\nu_T] = 1
\]

\[
P[Z^\pi_T > Z^\nu_T] > 0.
\]

As shown for example in Fernholz and Karatzas [2005], the existence of a relative arbitrage necessarily implies that the process $H(t)$ is a strict local martingale. Notice, however, that the existence of relative arbitrage does not necessarily imply the existence of arbitrage in the classical sense as shown in Mijatovic and Urusov [2011]. It follows that pricing and utility maximization in a diverse market are still possible (see Fernholz and Karatzas [2005] for a discussion on this topic).

In order to analyze portfolio strategies that yield relative arbitrage, we first note that the logarithmic return of portfolio $\pi$ relative to a reference portfolio $\nu$ is given by

\[
d \log \left( \frac{Z^\pi(t)}{Z^\nu(t)} \right) = \sum_{i=1}^{n} \pi_i(t) d \log \left( \frac{X_i(t)}{Z^\nu(t)} \right) + \gamma^\pi(t) dt
\]

Remarkably, the portfolio excess growth rate $\gamma^\pi(t)$ can be written as in equation (1) using the covariance of stock return, $\sigma(t)$, or using the covariance of relative stock returns, namely

\[
\gamma^\pi(t) = \frac{1}{2} \left[ \pi(t)^\top \text{diag}(\sigma(t)) - \pi(t)^\top \sigma(t) \pi(t) \right]
\]

\[
= \frac{1}{2} \left[ \pi(t)^\top \text{diag}(\tau(t)) - \pi(t)^\top \tau^\nu(t) \pi(t) \right]
\]

where $\tau^\nu$ denotes the relative covariance matrix with elements

\[
\tau^\nu_{ij}(t) = \frac{1}{dt} d \log \left( \frac{X_i(t)}{Z^\nu(t)} \right) \log \frac{X_j(t)}{Z^\nu(t)}
\]

\[
= (\sigma_{ij}(t) - \sigma_{\nu ij}(t) - \sigma_{j \nu}(t) + \sigma_{\nu \nu}(t))
\]

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

We now turn to a specific reference portfolio, namely 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. This next result is central to our derivation of the beta-arbitrage and is summarized in the following proposition.

**Proposition 1:** The return of a portfolio $\pi$ relative to the market portfolio $\mu$ can be written as the sum of two terms: the portfolio-weighted sum of relative changes in holdings’ market capitalization weights minus a drift term proportional to the portfolio tracking variance. It is explicitly given by

$$
\frac{d \log \left( \frac{Z_{\pi}(t)}{Z_{\mu}(t)} \right)}{dt} = \sum_{i=1}^{n} \pi_i(t) \frac{d\mu_i(t)}{\mu_i(t)} - \frac{1}{2} \frac{1}{\mu_i(t)} \tau(t) \pi(t) dt.
$$

(3)

The second term on the RHS of equation (3) is the instantaneous squared ex-ante tracking error risk. It measures locally how portfolio returns are likely to deviate from market returns.

### 2.4 Beta Arbitrage

In this Section we turn to the central result of this paper and show that a portfolio that overweights high beta stocks and underweights low beta stocks is dominated by the market portfolio.

In the context of our model the beta of stock $i$ relative to the market portfolio is an instantaneous measure and is expressed as follows

$$
\beta_i(t) = \frac{\sigma_{i\mu}(t)}{\sigma^2(\mu(t), t)}
$$

(4)

where $\sigma^2(\mu(t), t) = \mu(t)^T \sigma(t) \mu(t)$ is the market variance and $\sigma_{i\mu}(t)$ is the instantaneous covariance between stock $i$ and the market portfolio.

Let us consider the beta-weighted portfolio $\hat{\pi}(t)$ defined according to the rule

$$
\hat{\pi}_i(t) = \mu_i(t) \left[ (1 + \alpha) \beta_i(t) - \alpha \right].
$$

(5)

This portfolio rule, for $\alpha > -1$, maps the market portfolio into a beta weighted portfolio, where holdings with beta higher (respectively, lower) than one have their weights...
inflated (respectively, deflated) relative to market weights. Conversely, for \( \alpha < -1 \), the portfolio rule maps the market portfolio into a portfolio, where holdings with beta higher (respectively, lower) than one have their weights deflated (respectively, inflated) relative to market weights. For \( \alpha = -1 \) portfolio and market weights coincide. Notice that, for very large values of \( \alpha \), positive or negative, the long-only market portfolio is mapped into a long-short portfolio. For instance, for \( \alpha = 9 \), a low beta security, with \( \beta_i(t) = 0.8 \), would have a portfolio weight equal to \( \hat{\pi}_i(t) = -\mu_i(t) \). A high beta security, with \( \beta_i(t) = 1.2 \) would have a portfolio weight equal to \( \hat{\pi}_i(t) = 3\mu_i(t) \). For \( \alpha = -11 \), a low beta security, with \( \beta_i(t) = 0.8 \) would have a portfolio weight equal to \( \hat{\pi}_i(t) = 3\mu_i(t) \). A high beta security, with \( \beta_i(t) = 1.2 \) would have a portfolio weight equal to \( \hat{\pi}_i(t) = -\mu_i(t) \).

Using proposition 1 and the definition of the beta-weighted portfolio, we can write the relative logarithmic return dynamics of \( \hat{\pi}(t) \) with respect to the market portfolio as follows

\[
d\log \left( \frac{Z_{\hat{\pi}}(t)}{Z_{\mu}(t)} \right) = (1 + \alpha) \sum_{i=1}^{n} \beta_i(t)d\mu_i(t) - \frac{(1 + \alpha)^2}{2} \sum_{i,j=1}^{n} \beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt. \tag{6}
\]

This result is derived in the Appendix.

The equation governing the instantaneous variation of a beta-weighted portfolio relative to the market portfolio contains a beta transfer term, equal to the sum of each holding’s ex-ante beta, times its change in market capitalization weight, and a drift term, which is proportional to the portfolio tracking risk relative to the market. The beta transfer term is stochastic: it is unpredictable, since it depends on stochastic changes in market capitalization weights. The negative drift term depends on beta dispersion in the market and can be assessed ex-ante. The next proposition shows how the return of the beta-weighted portfolio can be decomposed and constitutes the main theoretical contribution of this paper.

**Proposition 2:** The return of a beta-weighted portfolio \( \hat{\pi} \) relative to the market portfolio \( \mu \) is explicitly given by

\[
d\log \left( \frac{Z_{\hat{\pi}}(t)}{Z_{\mu}(t)} \right) = \frac{1 + \alpha}{2} \left[ d \log \left( \hat{\sigma}^2(\mu(t), t) \right) - \frac{\partial}{\partial t} \log \left( \hat{\sigma}^2(\mu(t), t) \right) dt \right] + d\Theta_t. \tag{7}
\]

It is the sum of three terms
1. The instantaneous change in the log market portfolio variance

2. The negative of the time differential of the market portfolio variance, which captures the variation of the market portfolio variance due to changes in assets covariances

3. A drift term

\[ d\Theta_t := -\frac{1 + \alpha}{2\hat{\sigma}^2(\mu(t), t)} \sum_{i,j=1}^{n} \mu_i(t)\mu_j(t)\tau_{ij}(t)\sigma_{ij}(t)dt + \frac{(1 - \alpha^2)}{2} \sum_{i,j=1}^{n} \beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt. \] (8)

The first two terms are bounded and the third term has always a strictly defined sign, which depends on the choice of the parameter \( \alpha \).

Proposition 2 shows that, at a sufficiently long time horizon, the behavior of the beta-weighted portfolio relative to the market is determined by the sign of the drift term, which eventually dominates the first two bounded terms. For instance, in the case \( \alpha \geq 0 \), which corresponds to a portfolio whose holdings with beta higher (respectively, lower) than one have their weights inflated (respectively, deflated) relative to market weights, the beta-weighted strategy actually underperforms the market in the long run. This mechanism also implies that a strategy overweighing low beta stocks and underweighing high beta stock dominates the market in long run and constitutes as such a relative arbitrage.

**Remark 1:** For \( \alpha = -1 \), Eq.(5) shows that portfolio weights and market weights become identical. Accordingly, the right hand side of Eq.(7) vanishes. For \( \alpha \geq 0 \) we prove in the Appendix that the drift term is strictly negatively defined and that the beta-weighted portfolio is dominated by the market in the long term. It is worth to notice, both in Eq.(6) and in the definition of the drift term \( d\Theta_t \) in Proposition 2, that for high values of \( |\alpha| \) (no matter the sign of \( \alpha \)) the drift term scales like \(-\alpha^2\), while noise scales like \( \alpha \). We saw before that in this situation the long-only market is mapped into a long-short portfolio.

**Remark 2:** From the dynamics of the relative return of the beta-weighted portfolio

\[ d\log \left( \frac{Z_{\hat{\pi}}(t)}{Z_\mu(t)} \right) = (1 + \alpha) \sum_{i=1}^{n} \beta_i(t)d\mu_i(t) - \frac{(1 + \alpha)^2}{2} \sum_{i,j=1}^{n} \beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt, \]
we can see that if the market is not diverse, the second term on the RHS vanishes as the market capitalization concentrates into one stock. In this case there are no beta-arbitrage opportunities. Diversity is not required for relative arbitrage in general, but it is necessary for beta-arbitrage.

3 Maximum Drift Portfolios

Looking at the definition of $d\Theta_t$, Eq.(8), we can find the range of $\alpha$ which makes the drift term positive, in such a way that the portfolio can beat the market over a long enough time horizon. We realize that the drift term, Eq.(8), can be written as a quadratic form in $\alpha$:

$$d\Theta_t := -\frac{1}{2\hat{\sigma}^2(\mu(t), t)} \left\{ (1 + \alpha) \sum_{i,j=1}^n \mu_i(t)\mu_j(t)\tau_{ij}(t)\sigma_{ij}(t)dt + 
(1 - \alpha^2) \sum_{i,j=1}^n \hat{a}^2(\mu(t), t)\beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt \right\}.$$ 

Both quadratic forms in the square bracket on the right hand side of the previous equation are strictly positively defined. Furthermore, the second quadratic form is always lower than the first one, since it is built on the Hadamard product of the relative return covariance matrix $\tau$, and the first element in the spectral decomposition of the covariance matrix $\sigma$.

We define:

$$\phi(t) := \sum_{i,j=1}^n \mu_i(t)\mu_j(t)\tau_{ij}(t)\sigma_{ij}(t)dt,$$

and

$$\chi(t) := \sum_{i,j=1}^n \hat{a}^2(\mu(t), t)\beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt.$$ 

Then the drift term can be written as:

$$d\Theta_t := -\frac{1}{2\hat{\sigma}^2(\mu(t), t)} \left\{ (1 + \alpha) \phi - (1 - \alpha^2)\chi \right\}.$$
In order for the drift term to be positive, we must impose that the square bracket in the previous equation be strictly negatively defined. The two solutions of the previous quadratic equation are \(\alpha_1 = -1\) and \(\alpha_2(t) = 1 - \frac{\phi(t)}{\chi(t)}\). Since at each time \(t\) the condition \(\phi(t) \geq \chi(t)\) holds, it follows that \(d\Theta_t > 0\) for values of \(\alpha(t)\) between the two solutions of the quadratic equation, which are always both negative.

For values of \(\alpha(t) \geq 0\), the drift term is strictly negatively defined. We also see that it is possible to build a strategy where the parameter \(\alpha(t)\) is chosen at each step \(t\) in such a way that the drift is maximized. In fact \(d\Theta_t\) is maximized by choosing

\[
\alpha(t) = -\frac{\phi(t)}{2\chi(t)}.
\]

From Eq.(9) and from the definition of the variables \(\phi(t)\) and \(\chi(t)\), we see that the value of \(\alpha(t)\) becomes more and more negative when idiosyncratic risk becomes dominant in the market. However, investors should be careful at exploiting the optimal portfolio parametrization, when idiosyncratic risk grows significantly. In fact, for high values of \(|\alpha|\) the drift term becomes negative, since the long-only market is mapped into a long-short portfolio. One should complement the optimal portfolio parametrization, maximizing drift, with a constraint on tracking error risk, as follows:

\[
\pi(t)^\top \tau(t)\pi(t) = [1 + \alpha(t)]^2 \sum_{i,j=1}^{n} \beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t) \leq \zeta^2,
\]

where \(\zeta\) is a level of ex-ante tracking error risk not to be overcome. This condition is satisfied when \(\alpha(t)\) is in the range:

\[
-1 - \frac{\zeta \hat{\sigma}(\mu(t),t)}{\chi(t)^2} \leq \alpha(t) \leq -1 + \frac{\zeta \hat{\sigma}(\mu(t),t)}{\chi(t)^2}.
\]

When the optimal value of the parameter alpha is outside of the region compatible with the tracking risk constraint, the most reasonable choice is to take the alpha bound compatible with the tracking risk constraint and closest to the optimal specification.

1. If the optimal alpha specification is lower than the lower alpha bound respecting the
tracking risk constraint,

\[-\frac{\phi(t)}{2\chi(t)} < -1 - \frac{\zeta\hat{\sigma}(\mu(t), t)}{\chi(t)^{\frac{1}{2}}},\]

then take \(\alpha(t) = -1 - \frac{\zeta\hat{\sigma}(\mu(t), t)}{\chi(t)^{\frac{1}{2}}}\).

2. If the optimal alpha specification is higher than the upper alpha bound respecting the tracking risk constraint,

\[-\frac{\phi(t)}{2\chi(t)} > -1 + \frac{\zeta\hat{\sigma}(\mu(t), t)}{\chi(t)^{\frac{1}{2}}},\]

then take \(\alpha(t) = -1 + \frac{\zeta\hat{\sigma}(\mu(t), t)}{\chi(t)^{\frac{1}{2}}}\).

3. In all the other cases take \(\alpha(t) = -\frac{\phi(t)}{2\chi(t)}\).

We now turn to the empirical analysis of the performance of the beta-arbitrage strategies.

4 Empirical Tests

In this section we construct beta-arbitrage portfolios on different security types. We first focus on individual stocks and then apply our methodology to country and sector indices. In order to apply the results of the previous sections, we make the following assumption

**Assumption 5:** The relative price variation in the model \(\frac{dX_i}{X_i}\) equals the observed total return in the data \(\frac{dS_i + D_i dt}{S_i}\), where \(S_i\) and \(D_i\) are the observed stock price and dividend rate respectively.

Under this assumption, the reference market portfolio with respect to which we compute the individual stock’s beta corresponds to a dividend reinvested value weighted index.

4.1 Individual stocks on the US stock market

We consider an investment universe consisting of all available stocks in the CRSP database for the sample period 07/1925 to 12/2011. At the end of each month, we estimate the asset covariance matrix by using logarithmic returns calculated over a rolling window of 60
Based only on asset covariances and on market portfolio weights, we estimate at the end of each month the relative return covariance matrix $\tau_{ij}$ and the asset ex-ante beta’s, according to Eq. (4).

In this section, we define the relative performance of a strategy as the difference in log return between the beta-weighted strategy and the benchmark portfolio. We choose this performance criteria as it corresponds exactly to the theoretical results developed in the previous sections. As the rebalancing frequency increases, the relative performance corresponds to the return on a portfolio that is long in the beta-weighted portfolio and short in the benchmark portfolio.

### 4.1.1 Optimal $\alpha$ portfolio

Using the results of section 3 we construct a portfolio designed to maximize the expected drift term in the return decomposition of the beta arbitrage strategy. This amounts to choosing a value for the coefficient $\alpha$ that takes advantage of the heterogeneity of beta’s across assets and is explicitly given by Eq. (9). In this context $\alpha$ is adjusted every month following the variations in the beta dispersion.

The top panel of Figure 1 displays the cumulative log return of the $\alpha$-optimal portfolio against the benchmark. The performance is remarkable displaying a steady increase over the entire sample with a total cumulative result of 40%. The bottom panel of Figure 1 displays the evolution of the coefficient $\alpha$ which frequently reverts to a value of $-1$, when market conditions do not warrant a modification in the portfolio structure. When there is enough heterogeneity in the stock beta’s, the coefficient decreases significantly overweighing low beta stocks and underweighing high beta stocks.

We further test the performance of the long-short portfolio ($\alpha$-optimal against benchmark) by performing a regression analysis over non overlapping sub-samples of 10 years. In each sub-sample, we measure the ability of the Fama French factors to explain the

---

Our results are robust to alternative choices of the size of the rolling window.

Asset covariance are estimated using a one factor market model.
performance of the strategy. Figure 2 displays the value of the intercept of the regression over each sub-sample. The intercept oscillates around 0.5 % per year in the first half of the sample and increases to 2.5 % in the nineties and 1.5 % in the period 2001-2010. The coefficient values are statistically significant in 6 out of 8 sub-samples.

We can assert that the book-to-market or size factor exposures of the optimal $\alpha$—strategy can not explain the observed risk adjusted performance.

4.1.2 Tracking error

The value of the coefficient $\alpha$ that maximizes the drift of the Beta weighted portfolio is unbounded. Under particular market conditions, it may imply taking on significant short position and thus, moving far away from the benchmark portfolio.

A portfolio manager may wish to limit the distance between the beta weighted portfolio and the benchmark by controlling the tracking error. The results of section 3 allow us to attain this objective. The top panel of Figure 3 displays the evolution of two alternative portfolios with annualized tracking errors of 1 and 4 percent respectively, on the US stock sample over the period 07/1925 to 12/2011. The bottom panel displays the evolution of the tracking error bounds for the two portfolios. We can see that the low tracking error bounds are attained very frequently over the entire sample. This translates into a very stable portfolio evolution, albeit with a smaller return.

4.2 Indices

In this section we analyze the performance of the beta-arbitrage strategies on equity country indices and international equity sectors. In this case we do not need to impose assumption 5 as we take the indices structure as given. We first consider the performance
of strategies with constant positive $\alpha$ and show how close the theoretical return decomposition is from the empirical results. For the strategies that overweight high beta securities, we can identify a significant negative drift in both samples, in accordance with the result of proposition 2.

4.2.1 Underperforming strategies on country and sector indices

MSCI Country Indices

We consider a universe of equity country indices from Morgan Stanley Capital International (MSCI). Among all the developed countries within the MSCI World index, we take those which have been part of the index since its inception. Our sample includes four Pacific countries (Japan, Australia, Singapore and Hong Kong), two North American countries (United States and Canada), and twelve European Countries (Germany, France, United Kingdom, Italy, Switzerland, Austria, Spain, Denmark, Netherlands, Sweden, Norway and Belgium). From monthly market capitalizations we extract market portfolio weights, $\mu_i(t_k)$, at the end of each month $k$.

At the end of each month, we estimate the asset covariance matrix by using logarithmic returns calculated over a rolling window of 60 months. Based only on asset covariances and on market portfolio weights, we estimate at the end of each month:

- the relative return covariance matrix $\tau_{ij}$. For its evaluation, we can use Eq.(2), with the reference portfolio $\nu$ replaced by the market portfolio $\mu$;
- the asset ex-ante beta’s, according to Eq.(4);
- the weights of the ex-ante beta-weighted portfolio, according to Eq.(5);
- the drift term, according to proposition 2;
- the change in logarithm of market ex-ante variance, minus the explicit time differential of market ex-ante variance, according to Eq.(17) in the Appendix. We refer to this variable as Ito noise, and we interpret it as the time variation of market ex-ante variance.

4 The use of monthly returns is justified by our assumption that asset logarithmic returns obey Brownian motion processes and by the fact that at short time scales (e.g. daily), the distribution of logarithmic asset returns is known to deviate from normality.

5 Our results are robust to alternative choices of the size of the rolling window.
variance due solely to changes in market portfolio weights. For its evaluation, we can use either the left or the right hand side of Eq.(17) in the Appendix.

The results are shown in Figure 4 for portfolio weights defined according to Eq.(5) with $\alpha = 1$. The red line shows the time evolution of cumulative relative logarithmic return between the ex-ante beta weighted country portfolio and the market cap weighted country universe. Its increments correspond to the left hand side of Eq.(7). The dashed blue line shows the time evolution of the cumulative drift term, $\Theta(t)$. The dash-dotted black line shows the time evolution of the cumulative Ito noise term, which corresponds to the integral of Eq.(17) in the Appendix multiplied a factor $\frac{1+\alpha}{2}$.

We observe that cumulative Ito noise does remain bounded, oscillating around zero. The continuous red line in Figure 4 shows the sum of cumulative Ito noise, cumulative drift term, and differential dividend rate between the portfolio and the market cap weighted index. Its increments correspond to the right hand side of Eq.(7), with the drift term now properly accounting for dividends.

If our assumption that monthly logarithmic returns obey Brownian motion processes is confirmed, we expect the continuous red line and the continuous green line in the chart (respectively, observation and model prediction) to remain all the time very close to each other. Our expectation is nicely verified over the country universe considered. We also visually remark that the pattern of cumulative relative logarithmic return is driven by the cumulative drift term, as our model predicts.

Table 2 reports the detail of the results for different choices of the parameter $\alpha$ in the definition of portfolio weights, Eq.(5). The empirical results confirm the increasingly negative drift between portfolio and benchmark for high values of $\alpha$, predicted in Eq.(8).
MSCI Sectors Indices

We apply our result to the universe of MSCI World GICS (Global Industry Classification Standards) equity sectors. We use monthly data over the sample period January 1995 to November 2010. The results are shown in Figure 5 for portfolio weights defined according to Eq. (5) with $\alpha = 1$. Table 3 reports the detail of the results for different choices of the parameter $\alpha$ in the definition of portfolio weights, Eq. (5).

We observe that the monthly contribution of the drift term is of similar magnitude to the one observed for the country universe. From Figure 5 we can see that the pattern of cumulative relative logarithmic return is driven by the cumulative drift term, as predicted by the model.

We can also observe the large negative Ito noise contribution coming from the beta transfer effect that occurred in the period 2000-2002, when the burst of the tech bubble determined a huge shift from higher beta sectors into lower beta sectors.

To gain further insight into the determinants of the behavior of the beta-portfolio, we compute a Gini index of the concentration of betas across the 10 MSCI equity sectors and compare it to the expected drift term every month. The results are displayed in Figure 6. We can see that the expected drift term takes large negative values when the dispersion is large and is close to zero when beta distribution is highly concentrated. Beta arbitrage strategies have a significant impact when Beta dispersion within the investment universe is important.
4.2.2 Optimal $\alpha$–strategy on country and sector indices

In this section we use the results of section 3 to construct portfolio strategies that aim at maximizing the expected drift term on both MSCI country and sector indices. In both cases we document significant over-performance of the proposed strategy against the market benchmark.

Country indices

We first consider the universe of equity country indices from Morgan Stanley Capital International (MSCI) described in section 4.2.1. We follow the methodology described in section 3 and in particular we use Eq. (9) to define the time varying $\alpha$ which maximizes the expected drift term in the return decomposition of the Beta arbitrage strategy.

The top panel of Figure 7 displays the cumulative log return against the country index benchmark of the unconstrained $\alpha$-optimal portfolio and of an $\alpha$-optimal portfolio with an annualized tracking error bound of 8%. The bottom panel displays the optimal $\alpha$ along with the $\alpha$-bounds consistent with the 8% annualized tracking error limit.

In Figure 8, we decompose the long short portfolio performance to identify the contribution of the expected drift term. We focus on the portfolio with 8% tracking error bound. We can see that the expected drift term (in red), which we seek to maximize by adequately choosing the coefficient $\alpha$, accounts for most of the performance of the strategy.

Sector indices

We perform a similar test on sector indices. The top panel of Figure 9 displays the cumulative log return against the sector index benchmark of the unconstrained $\alpha$-optimal portfolio and of an $\alpha$-optimal portfolio with an annualized tracking error bound of 4%.
The bottom panel displays the optimal $\alpha$ value along with the $\alpha$-bounds consistent with the 4% annualized tracking error limit.

This test confirm our previous result. The $\alpha$—optimal portfolio significantly outperforms the benchmark when there is enough dispersion in the beta distribution. Given the limited number of sector indices, it is not surprising to observe rather long period of inactivity with low absolute values of $\alpha$. For instance, in the period 2003-2008 the $\alpha$—optimal portfolio remains very close to the benchmark.

\begin{center}
Insert Figure 9 here
\end{center}

5 Conclusion

Diversity in financial markets is a sufficient condition to allow for beta-arbitrage opportunities. Beta-weighted portfolios, which over-invest in high beta stocks and under-invest in low beta stocks, are dominated by the market portfolio in the long run. This result is not obtained by assuming a particular set of portfolio constraints, nor by assuming restrictions on financial securities. It is simply due to the continuous rebalancing of beta-weighted portfolio and the boundedness of market variance.

Our results complement the existing literature on beta-arbitrage by showing how at least part of the explanation can be attributed to the very structure of beta-weighted portfolios. We show that returns of these portfolios obey a particular decomposition formula and we analyze empirically whether this decomposition finds support in the data. We find strong support for our explanation at various levels, on equity country indices, equity sectors and individual stocks.

Finally, we consider a form of optimal beta-arbitrage which maximizes the expected return relative to a market portfolio. This is achieved by considering a time varying over-weighting of low beta stocks. We provide an explicit characterization of this optimal portfolio by relying on the return decomposition formula and show that it generates significant performance on equity country indices, equity sectors and individual stocks.
References


Appendix

Proof of proposition 1

Using Proposition 1.2.5 in Fernholz [2002], we obtain

\[ d \log \left( \frac{Z_\pi(t)}{Z_\mu(t)} \right) = \sum_{i=1}^{n} \pi_i(t) d \log \mu_i(t) + \gamma^*_\pi(t). \] (11)

Using the equivalence result in Eq.(2), we may rewrite Eq.(11) as

\[ d \log \left( \frac{Z_\pi(t)}{Z_\mu(t)} \right) = \sum_{i=1}^{n} \pi_i(t) d \log \mu_i(t) + \frac{1}{2} \left[ \pi(t)^\top \text{diag}(\tau(t)) - \pi(t)^\top \pi(t) \right]. \] (12)

Using Ito’s lemma, we obtain the dynamics of the logarithm of the market capitalization

\[ d \log \mu_i(t) = \frac{d\mu_i(t)}{\mu_i(t)} - \frac{\tau_{ii}(t)}{2} dt \]

which together with equation (12) implies the result in Proposition 1

\[ d \log \left( \frac{Z_\pi(t)}{Z_\mu(t)} \right) = \sum_{i=1}^{n} \pi_i(t) \frac{d\mu_i(t)}{\mu_i(t)} - \frac{1}{2} \pi(t)^\top \pi(t) dt. \]

A derivation of equation (6)

Using equation (3) and the portfolio weights as in Eq.(5), we obtain

\[ d \log \left( \frac{Z_\eta(t)}{Z_\mu(t)} \right) = \sum_{i=1}^{n} (1 + \alpha) \beta_i(t) - \alpha d\mu_i(t) - \frac{1}{2} \sum_{i,j=1}^{n} \mu_i(t)\mu_j(t)((1 + \alpha)\beta_i(t) - \alpha)((1 + \alpha)\beta_j(t) - \alpha)\tau_{ij}(t) dt \] (13)

Knowing that the sum of market portfolio weights is an invariant, always equal to one, we must have:

\[ \sum_{i=1}^{n} d\mu_i(t) = 0 \]

The matrix \( \tau(t) \) is positive semi-definite and of rank \( n - 1 \). In addition, for reference portfolio \( \eta \), the vector of portfolio weights \( \eta(t) \) belongs to kernel of the relative covariance...
matrix \( \tau^\eta(t) \). For each element \( i \) of the vector \( \tau^\eta(t)\eta(t) \), we have

\[
\left[ \tau^\eta(t)\eta(t) \right]_i = \sum_{j=1}^{n} \left( \sigma_{ij}(t) - \sigma_{in}(t) - \sigma_{jn}(t) + \sigma_{\eta\eta}(t) \right) \eta_j(t) = \\
= \sigma_{in}(t) - \sigma_{in}(t) - \sigma_{\eta\eta}(t) + \sigma_{\eta\eta}(t) = 0.
\]

It follows that the kernel of \( \tau(t) \) is generated by \( \mu(t) \) and we can rewrite equation (13) as

\[
d \log \left( \frac{Z_{\eta}(t)}{Z_{\mu}(t)} \right) = (1 + \alpha) \sum_{i=1}^{n} \beta_i(t) \mu_i(t) - \frac{(1 + \alpha)^2}{2} \sum_{i,j=1}^{n} \beta_i(t) \mu_i(t) \beta_j(t) \mu_j(t) \tau_{ij}(t) dt,
\]

which is equation (6).

**Proof of proposition 2**

The first term in equation (6) can be written as follows

\[
(1 + \alpha) \sum_{i=1}^{n} \beta_i(t) \mu_i(t) = \frac{1 + \alpha}{2} \sum_{i=1}^{n} \frac{\partial}{\partial \mu_i(t)} \log \left( \sigma^2(\mu(t), t) \right) d\mu_i(t)
\]

(15)

The Ito expansion of the logarithm of ex-ante variance, which depends on stochastic market portfolio weights, gives:

\[
d \log \left( \sigma^2(\mu(t), t) \right) = \frac{\partial}{\partial t} \log \left( \sigma^2(\mu(t), t) \right) dt + \sum_{i=1}^{n} \frac{\partial}{\partial \mu_i(t)} \log \left( \sigma^2(\mu(t), t) \right) d\mu_i(t) + \\
+ \frac{1}{2} \sum_{i,j=1}^{n} \mu_i(t) \mu_j(t) \tau_{ij}(t) \frac{\partial^2}{\partial \mu_i(t) \partial \mu_j(t)} \log \left( \sigma^2(\mu(t), t) \right) dt
\]

(16)

The second derivative of the logarithm of ex-ante variance in market portfolio weights can be written as:

\[
\frac{\partial^2}{\partial \mu_i(t) \partial \mu_j(t)} \log \left( \sigma^2(\mu(t), t) \right) = 2 \frac{\sigma_{ij}(t)}{\sigma^2(\mu(t), t)} - 4 \beta_i(t) \beta_j(t)
\]
The stochastic behavior of the instantaneous change in the logarithm of market variance, minus its explicit time differential, can therefore be written as:

\[
d\log(\sigma^2(\mu(t), t)) - \frac{\partial}{\partial t} \log(\hat{\sigma}^2(\mu(t), t)) dt = 2 \sum_{i=1}^{n} \beta_i(t) d\mu_i(t) + \\
+ \frac{1}{\sigma^2(\mu(t), t)} \sum_{i=1}^{n} \mu_i(t) \mu_j(t) \tau_{ij}(t) \sigma_{ij}(t) dt - 2 \sum_{i,j=1}^{n} \beta_i(t) \mu_i(t) \beta_j(t) \mu_j(t) \tau_{ij}(t) dt
\]

(17)

Combining Eq.(15) and Eq.(16), we obtain:

\[
(1 + \alpha) \sum_{i=1}^{n} \beta_i(t) d\mu_i(t) = \frac{(1 + \alpha)}{2} \left[ d\log(\sigma^2(\mu(t), t)) - \frac{\partial}{\partial t} \log(\hat{\sigma}^2(\mu(t), t)) dt - \\
- \frac{1}{\sigma^2(\mu(t), t)} \sum_{i,j=1}^{n} \mu_i(t) \mu_j(t) \tau_{ij}(t) \sigma_{ij}(t) dt + 2 \sum_{i,j=1}^{n} \beta_i(t) \mu_i(t) \beta_j(t) \mu_j(t) \tau_{ij}(t) dt \right]
\]

(18)

Using Eq.(18), Eq.(6) can be rewritten as:

\[
d\log \left( \frac{Z^*(t)}{Z^*(\mu(t))} \right) = \frac{1 + \alpha}{2} \left[ d\log(\hat{\sigma}^2(\mu(t), t)) - \frac{\partial}{\partial t} \log(\hat{\sigma}^2(\mu(t), t)) dt - \\
- \frac{1}{\sigma^2(\mu(t), t)} \sum_{i,j=1}^{n} \mu_i(t) \mu_j(t) \tau_{ij}(t) \sigma_{ij}(t) dt - \\
- (1 - \alpha) \sum_{i,j=1}^{n} \beta_i(t) \mu_i(t) \beta_j(t) \mu_j(t) \tau_{ij}(t) dt \right].
\]

(19)

The first term on the RHS equals the instantaneous change in the logarithm of market variance. This term needs to be bounded, since market variance is bounded. Assumption 3 makes the second term bounded as well. It remains to study the third term. We must assess the sign of the double sum:

\[
\sum_{i,j=1}^{n} \mu_i(t) \mu_j(t) \tau_{ij}(t) \sigma_{ij}(t),
\]

(20)

which is nothing but the quadratic form built on the Hadamard product of the covariance matrix of stock relative returns, measured with respect to the market portfolio, and of the covariance matrix of stock absolute returns, \(\sigma_{ij}(t)\).

Let \(\{\lambda_1(t), \ldots, \lambda_n(t)\}\) be the eigenvalues of the covariance matrix of stock absolute returns,
For each \( k \in \{1, \ldots, n\} \) we define \( e_k(t) = [e_{k1}(t), \ldots, e_{kn}(t)]^T \) as the normalized eigenvector of \( \sigma_{ij}(t) \) corresponding to the eigenvalue \( \lambda_k(t) \). Then we can rewrite the quadratic form Eq.(20) as

\[
\sum_{k=1}^{n} \lambda_k(t) \sum_{i,j=1}^{n} \mu_i(t)e_{ki}(t)\tau_{ij}(t)\mu_j(t)e_{kj}(t).
\]

Defining for each \( k \in \{1, \ldots, n\} \) the vector \( v_k(t) = [\mu_1(t)e_{k1}(t), \ldots, \mu_n(t)e_{kn}(t)]^T \), we can rewrite Eq.(21) as

\[
\sum_{k=1}^{n} \lambda_k(t) \left[ v_k(t)^T \tau(t)v_k(t) \right],
\]

so that we obtain:

\[
-\frac{1}{\hat{\sigma}^2(\mu(t),t)} \sum_{i,j=1}^{n} \mu_i(t)\mu_j(t)\tau_{ij}(t)\sigma_{ij}(t)dt = -\frac{\sum_{k=1}^{n} \lambda_k(t) \left[ v_k(t)^T \tau(t)v_k(t) \right]}{\hat{\sigma}^2(\mu(t))} dt.
\]

The matrix \( \tau(t) \) is positive semi-definite. It is also of rank \( n-1 \) and its kernel is generated by \( \mu(t) \), as the analog of Eq.(14), when the reference portfolio \( \eta(t) \) coincides with the market portfolio, shows.

Hence \( v_k(t)^T \tau(t)v_k(t) = 0 \) if and only if \( v_k(t) \) is proportional to \( \mu(t) = (\mu_1(t), \ldots, \mu_n(t)) \).

Otherwise we must have \( v_k(t)^T \tau(t)v_k(t) > 0 \).

Assumption 1 is equivalent to saying that the covariance matrix \( \sigma(t) \) must be positively definite at all times. Under these hypotheses, the rank of the covariance matrix is at least two, and its eigenvalues are all positive and bounded away from zero. This means that at least one of the \( v_k \) corresponding to a non zero eigenvalue is not parallel to \( \mu \), which results into a strictly negatively defined third term. For the same reasons, the sign of the double sum:

\[
\sum_{i,j=1}^{n} \beta_i(t)\mu_i(t)\beta_j(t)\mu_j(t)\tau_{ij}(t)dt
\]

is strictly positively defined, unless all the asset beta’s are equal (and identically equal to one). However, notice that in Eq.(19) this double sum is multiplied by a term equal to \( \frac{(1-\alpha^2)}{2} \), resulting into a strictly negatively defined contribution for \( \alpha > 1 \). For \( \alpha = 1 \) the
contribution is exactly equal to zero. For $0 \leq \alpha < 1$ the contribution is positive, but it is
more than offset by the contribution coming from the third term in Eq.(19). In fact, the
combination of third and fourth term in Eq.(19) can be written as:

$$-rac{1 + \alpha}{2\hat{\sigma}^2(\mu(t), t)} \left[ \sum_{i,j=1}^{n} \mu_i(t)\mu_j(t)\tau_{ij}(t) \left( \sigma_{ij}(t) - (1 - \alpha)\beta_i(t)\beta_j(t)\hat{\sigma}^2(\mu(t), t) \right) \right] dt.$$ 

Each covariance matrix element $\sigma_{ij}(t)$ can be written as its systematic, beta-dependent
component, plus a residual component $\sigma_{\epsilon,ij}(t)$

$$\sigma_{ij}(t) = \beta_i(t)\beta_j(t)\hat{\sigma}^2(\mu(t), t) + \sigma_{\epsilon,ij}(t). \quad (22)$$

The covariance matrix is positive definite by assumption. This means that, for any given
weight vector at any given time, $\pi(t)$, the quadratic form

$$\hat{\sigma}^2(\pi(t), t) = \sum_{i,j=1}^{n} \pi_i(t)\pi_j(t)\sigma_{ij}(t)$$

is positive. We can therefore write:

$$0 \leq \hat{\sigma}^2(\pi(t), t) = \beta_i^2(\pi(t))\hat{\sigma}^2(\mu(t), t) + \sum_{i,j=1}^{n} \pi_i(t)\pi_j(t)\sigma_{\epsilon,ij}(t).$$

The right hand side of the previous equation is just the decomposition of total portfolio
variance into its systematic and residual components. We observe that the market weight
vector $\mu(t)$ belongs to the kernel of the residual covariance matrix $\sigma_{\epsilon,ij}(t)$, which is therefore
positive semi-definite.

By replacing Eq.(22) into Eq.(19), we get:

$$-rac{1 + \alpha}{2\hat{\sigma}^2(\mu(t), t)} \left[ \sum_{i,j=1}^{n} \mu_i(t)\mu_j(t)\tau_{ij}(t) \left( \sigma_{\epsilon,ij}(t) + \alpha\beta_i(t)\beta_j(t)\hat{\sigma}^2(\mu(t), t) \right) \right] dt,$$

which is certainly negatively defined for $0 \leq \alpha < 1$, since the residual covariance matrix is
positive definite and since kernel of the matrix $\tau(t)$ is spanned by the sole vector of market
weights.
Tables

<table>
<thead>
<tr>
<th></th>
<th>constant</th>
<th>Market</th>
<th>SMB</th>
<th>HML</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong></td>
<td>0.0004</td>
<td>-0.0196</td>
<td>-0.0008</td>
<td>0.0260</td>
</tr>
<tr>
<td><strong>t-stat</strong></td>
<td>2.2146</td>
<td>-2.7236</td>
<td>-0.1094</td>
<td>3.8774</td>
</tr>
</tbody>
</table>

**Table 1:** Regression results for the CRSP stock universe optimal alpha portfolio. The dependent variable is the monthly relative log return of the optimal alpha portfolio against the benchmark portfolio, the independent variables are the Market, Size and Book-to-Market factors. The sample period is August 1932 to December 2011. t-statistics are computed using Newey-West standard errors.

<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 1$</th>
<th>$\alpha = 5$</th>
<th>$\alpha = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average monthly relative log return (observed)</td>
<td>$-0.057%$</td>
<td>$-0.205%$</td>
<td>$-0.458%$</td>
</tr>
<tr>
<td>Monthly relative log return volatility (observed)</td>
<td>1.076%</td>
<td>3.242%</td>
<td>6.018%</td>
</tr>
<tr>
<td>Average monthly turnover</td>
<td>5.084%</td>
<td>11.329%</td>
<td>19.881%</td>
</tr>
<tr>
<td>Average monthly relative drift contribution (predicted)</td>
<td>$-0.038%$</td>
<td>$-0.184%$</td>
<td>$-0.496%$</td>
</tr>
<tr>
<td>Monthly relative drift contribution volatility (predicted)</td>
<td>0.029%</td>
<td>0.169%</td>
<td>0.519%</td>
</tr>
<tr>
<td>Average monthly Ito noise contribution</td>
<td>$-0.014%$</td>
<td>$-0.042%$</td>
<td>$-0.077%$</td>
</tr>
<tr>
<td>Monthly relative Ito noise volatility</td>
<td>1.081%</td>
<td>3.244%</td>
<td>5.948%</td>
</tr>
<tr>
<td>Average monthly differential dividend rate contribution</td>
<td>$-0.005%$</td>
<td>$-0.015%$</td>
<td>$-0.027%$</td>
</tr>
<tr>
<td>Monthly differential dividend rate contribution volatility</td>
<td>0.035%</td>
<td>0.106%</td>
<td>0.193%</td>
</tr>
</tbody>
</table>

**Table 2:** Numerical results for the comparison of the observed and predicted behavior of an ex-ante beta weighted portfolio of 18 MSCI equity country indices vs. its market cap weighted analog. Portfolio weights are defined according to Eq.(5) for different values of the parameter $\alpha$. The results for $\alpha = 1$ correspond to the graph in Figure 4.
<table>
<thead>
<tr>
<th></th>
<th>$\alpha = 1$</th>
<th>$\alpha = 5$</th>
<th>$\alpha = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average monthly relative log return (observed)</td>
<td>$-0.231%$</td>
<td>$-0.782%$</td>
<td>$-1.655%$</td>
</tr>
<tr>
<td>Monthly relative logarithmic return volatility (observed)</td>
<td>$1.684%$</td>
<td>$5.145%$</td>
<td>$9.744%$</td>
</tr>
<tr>
<td>Average monthly turnover</td>
<td>$5.815%$</td>
<td>$13.168%$</td>
<td>$23.113%$</td>
</tr>
<tr>
<td>Average monthly relative drift contribution (predicted)</td>
<td>$-0.035%$</td>
<td>$-0.223%$</td>
<td>$-0.681%$</td>
</tr>
<tr>
<td>Monthly relative drift contribution volatility (predicted)</td>
<td>$0.026%$</td>
<td>$0.178%$</td>
<td>$0.559%$</td>
</tr>
<tr>
<td>Average monthly Ito noise contribution</td>
<td>$-0.162%$</td>
<td>$-0.486%$</td>
<td>$-0.892%$</td>
</tr>
<tr>
<td>Monthly relative Ito noise volatility</td>
<td>$1.753%$</td>
<td>$5.259%$</td>
<td>$9.642%$</td>
</tr>
<tr>
<td>Average monthly differential dividend rate contribution</td>
<td>$-0.022%$</td>
<td>$-0.065%$</td>
<td>$-0.120%$</td>
</tr>
<tr>
<td>Monthly differential dividend rate contribution volatility</td>
<td>$0.017%$</td>
<td>$0.052%$</td>
<td>$0.096%$</td>
</tr>
</tbody>
</table>

**Table 3:** Numerical results for the comparison of the observed and predicted behavior of an ex-ante beta weighted portfolio of the 10 MSCI World GICS equity sectors vs. its market cap weighted analog. Portfolio weights are defined according to Eq.(5) for different values of the parameter $\alpha$. The results for $\alpha = 1$ correspond to the graph in Figure 5.
Figure 1: Performance of the optimal alpha strategy in the CRSP stock universe. Portfolio weights are defined according to Eq.(5) with $\alpha$ given by Eq. (9). We use monthly total returns over the period July 1925 to December 2011. Stock covariances are estimated on the basis of a one factor model. The top panel displays the cumulative relative log return of the optimal alpha portfolio against the CRSP market benchmark and the bottom panel displays the optimal $\alpha$ value.
Figure 2: Intercept value in yearly percentage point from sub sample regressions of the long short optimal alpha strategy (long in the optimal alpha strategy in the CRSP stock universe and short in the benchmark) against the 3 Fama French Factors.
Figure 3: Performance of the optimal alpha strategy with tracking error bounds in the CRSP stock universe. Portfolio weights are defined according to Eq. (5) with $\alpha$ given by Eq. (9) under the tracking error constraints given by Eq. (10). We use monthly total returns over the period July 1925 to December 2011. Stock covariances are estimated on the basis of a one factor model. The top panel displays the cumulative relative log return against the CRSP market benchmark of the optimal alpha portfolio with annualized tracking error of, respectively, 1% (dotted green line) and 4% (red line). The bottom panel displays the optimal $\alpha$ value along with the $\alpha$ bounds corresponding to an annualized tracking error of, respectively, 1% (dotted green line) and 4% (red line).
Figure 4: Graphical comparison between the observed and predicted behavior of an ex-ante beta weighted portfolio vs. its market cap weighted analog, defined over a universe of 18 MSCI equity country indices. Portfolio weights are defined according to Eq.(5) with $\alpha = 1$. We use monthly total returns and market capitalizations over the period January end 1970-November end 2010. The estimation of country covariances is based on rolling windows of 5 years (60 months). The trajectories shown correspond to the period January 1975-November 2010.
Figure 5: Graphical comparison between the observed and predicted behavior of an ex-ante beta weighted portfolio vs. its market cap weighted analog, defined over a universe of the 10 MSCI World GICS equity sectors. Portfolio weights are defined according to Eq. (5) with $\alpha = 1$. We use monthly total returns and market capitalizations over the period January end 1995-November end 2010. The estimation of sector covariances is based on rolling windows of 5 years (60 months). The trajectories shown correspond to the period January 2000-November 2010.
Figure 6: Scatter plot of the expected drift term of the beta portfolio against the concentration of Beta across sectors (Gini Index) defined over a universe of the 10 MSCI World GICS equity sectors. Portfolio weights are defined according to Eq. (5) with $\alpha = 1$. We use monthly total returns and market capitalizations over the period January end 1995-November end 2010. The estimation of sector covariances is based on rolling windows of 5 years (60 months). The trajectories shown correspond to the period January 2000-November 2010.
Figure 7: Performance of the optimal alpha strategy with tracking error bounds for the MSCI country indices. Portfolio weights are defined according to Eq.(5) with $\alpha$ given by Eq. (9) under the tracking error constraints given by Eq. (10). We use monthly total returns over the period January 1975-November 2010. The top panel displays the cumulative relative log return against the MSCI country index benchmark of the unconstrained optimal $\alpha$ portfolio (blue line) and of the optimal $\alpha$ portfolio subject to an annualized tracking error constraint of 8 % (red line). The bottom panel displays the optimal $\alpha$ value along with the $\alpha$-bounds consistent with the 8% annualized tracking error limit (red line).
Figure 8: Drift contribution of the optimal alpha strategy with an annualized tracking error bound of 8 % for the universe of MSCI country indices. Portfolio weights are defined according to Eq. (5) with \( \alpha \) given by Eq. (9) under the tracking error constraints given by Eq. (10). We use monthly total returns over the period January 1975-November 2010. The drift contribution in red corresponds to the process \( d\Theta(t) \) of Proposition 2.
Figure 9: Performance of the optimal alpha strategy with tracking error bounds for the universe of MSCI sector indices. Portfolio weights are defined according to Eq. (5) with $\alpha$ given by Eq. (9) under the tracking error constraints given by Eq. (10). We use monthly total returns over the period January 2000-November 2010. The top panel displays the cumulative relative log return against the MSCI sector index market benchmark of the unconstrained optimal alpha portfolio (blue line) and of the optimal alpha portfolio subject to an annualized tracking error limit of 4 % (red line). The bottom panel displays the $\alpha$ optimal value along with the $\alpha$ bounds consistent with an annualized tracking error limit of 4 % (red line).