Systematic and multifactor risk models revisited
Michel Fliess, Cédric Join

To cite this version:

HAL Id: hal-00920175
https://hal-polytechnique.archives-ouvertes.fr/hal-00920175
Submitted on 17 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Systematic and multifactor risk models revisited

Michel Fliess, Cédric Join

Abstract Systematic and multifactor risk models are revisited via methods which were already successfully developed in signal processing and in automatic control. The results, which bypass the usual criticisms on those risk modeling, are illustrated by several successful computer experiments.
1 Introduction

Systematic, or market, risk is one of the most studied risk models not only in financial engineering, but also in actuarial sciences, in business and corporate management, and in several other domains. It is associated to the beta ($\beta$) coefficient, which is familiar in the investment industry since Sharpe’s capital asset pricing model (CAPM) [30]. The pitfalls and shortcomings of $\beta$ have been detailed by a number of excellent authors.\footnote{The literature questioning the validity of the beta coefficient is huge and well summarized in several textbooks (see, e.g., [4]). A recent and remarkable paper by Tofallis [32] has been most helpful in this study.} Replacing moreover time-invariant linear regressions by time-varying and/or nonlinear ones does not seem to improve this situation.\footnote{See, e.g., [1, 31], and the references therein.} The model-free standpoint advocated in [11] and [14] alleviates several of the known deficiencies but unfortunately cannot be extended to multifactor risk models which became also popular after Ross’ arbitrage pricing theory (APT) [29]. In order to encompass the univariate and multivariate cases, we propose here a unified definition, with the same advantages, namely a clear-cut mathematical foundation, which

- bypasses clumsy statistical and/or financial assumptions,
- leads to efficient computations.

Our approach is based on the following ingredients:

- As in our previous works [10, 11, 14, 15] we utilize the Cartier-Perrin theorem [5]. It shows that under a mild integrability condition any time series may decomposed as a sum of a mean, or trend, and of quick fluctuations.
- Classic mathematical tools like the Wronski determinants [24].
- We employ recent estimation and identification techniques\footnote{The use of advanced theory stemming from signal analysis is not new in finance. See, e.g., [23].} [20, 21], which are stemming from control theory and signal processing where they have been utilized quite successfully.\footnote{See, e.g., [8, 9, 16, 17, 18, 19, 27, 28, 33, 34, 35, 36], and the references therein.}

From a more practical standpoint, our main result is the derivation of two independent $\beta$ coefficients, the first one for the comparison between returns and the second one for the comparison between volatilities. It implies among other consequences that the importance of the popular $\alpha$ coefficient might vanish.

Our paper is organized as follows. After a short review of the Cartier-Perrin theorem, Section 2 details the new mathematical definitions of the coefficients $\alpha$ and $\beta$, and of $\beta$ alone. Section 3 develops the comparison with the classical settings. Computer illustrations are provided in Section 4.

Future publications will be exploiting the above advances in at least three directions:
1. The extension of Section 3.3 to skewness and kurtosis should be straightforward. Our understanding, which would not rely exclusively any more on “Gaussianism”, of the respective behaviors of various assets might therefore be quite enhanced.

2. According to the methods sketched in [15] and in [12], dynamic portfolio management and option pricing may be achieved by tracking quite independent performances with respect to returns and volatilities.

3. We will relate some instances of systemic risk to the abrupt changes [16] of some quantities like our new beta coefficients (see [10, 11, 14] for preliminary results).

2 Theoretical background

2.1 A short review on time series via nonstandard analysis

Take the time interval \([0, 1] \subset \mathbb{R}\) and introduce as often in nonstandard analysis the infinitesimal sampling

\[\mathcal{T} = \{0 = t_0 < t_1 < \cdots < t_N = 1\}\]

where \(t_{i+1} - t_i, 0 \leq i < N,\) is infinitesimal, i.e., “very small”.

A time series \(X(t)\) is a function \(X : \mathcal{T} \to \mathbb{R}\).

The Lebesgue measure on \(\mathcal{T}\) is the function \(\ell\) defined on \(\mathcal{T} \setminus \{1\}\) by \(\ell(t_i) = t_{i+1} - t_i\). The measure of any interval \([c, d] \subset \mathcal{T},\ c \leq d,\) is its length \(d - c\).

The integral over \([c, d]\) of the time series \(X(t)\) is the sum

\[\int_{[c,d]} X \, d\tau = \sum_{t \in [c,d]} X(t)\ell(t)\]

\(X\) is said to be \(S\)-integrable if, and only if, for any interval \([c, d]\) the integral \(\int_{[c,d]} |X| \, d\tau\) is limited, i.e. not infinitely large, and, if \(d - c\) is infinitesimal, also infinitesimal.

\(X\) is \(S\)-continuous at \(t_i \in \mathcal{T}\) if, and only if, \(f(t_i) \simeq f(\tau)\) when \(t_i \simeq \tau\). \(X\) is said to be almost continuous if, and only if, it is \(S\)-continuous on \(\mathcal{T} \setminus R\), where \(R\) is a rare subset. \(X\) is Lebesgue integrable if, and only if, it is \(S\)-integrable and almost continuous.

---

5 See, e.g., [6, 7] for basics in nonstandard analysis.

6 \(a \simeq b\) means that \(a - b\) is infinitesimal.

7 The set \(R\) is said to be rare [5] if, for any standard real number \(\alpha > 0\), there exists an internal set \(B \supset A\) such that \(m(B) \leq \alpha\).
A time series $X : \mathbb{T} \to \mathbb{R}$ is said to be quickly fluctuating, or oscillating, if, and only if, it is $S$-integrable and $\int_A X d\tau$ is infinitesimal for any quadrable subset.$^8$

Let $X : \mathbb{T} \to \mathbb{R}$ be a $S$-integrable time series. Then, according to the Cartier-Perrin theorem [5], the additive decomposition

$$X(t) = E(X)(t) + X_{\text{fluctuat}}(t)$$

holds where

- the mean $E(X)(t)$ is Lebesgue integrable,
- $X_{\text{fluctuat}}(t)$ is quickly fluctuating.

The decomposition (1) is unique up to an infinitesimal.

Remark 1. Decomposition (1), where $E(X)(t)$ is “smoother” than $X(t)$, provides, to the best of our knowledge, the first complete theoretical justification (see [10]) of the trends in technical analysis (see, e.g., [2, 25]).

### 2.2 Multivariate factors

#### 2.2.1 Arithmetical average

Assume that $X : \mathbb{T} \to \mathbb{R}$ is $S$-integrable. Take a quadrable set $A \subseteq \mathbb{T}$ such that $\int_A d\tau$ is appreciable, i.e., non-infinitesimal. The arithmetical average of $X$ on $A$, which is written $AV_A(X)$, is defined by

$$AV_A(X) = \frac{\int_A X d\tau}{\int_A d\tau}$$

It follows at once from Equation (1) that the difference between $AV_A(X)$ and $AV_A(E(X))$ is infinitesimal, i.e.,

$$AV_A(X) \simeq AV_A(E(X))$$

In practice, $A$ is a time interval $[t - L, t]$, with an appreciable length $L$. Set, if $t \geq L$,

$$X(L, t) = AV_{[t-L,t]}(X) = \frac{\int_{t-L}^t X d\tau}{L} \simeq \frac{\int_{t-L}^t E(X) d\tau}{L}$$

Introduce

$^8$ A set is quadrable [5] if its boundary is rare.

$^9$ See [26] for a more down to earth exposition.
\[ X^{[\nu]}(L, t) = \frac{1}{(\nu - 1)!} \int_{-L}^{t} (t - \tau)^{\nu - 1} X d\tau \quad (3) \]

It corresponds via the classic Cauchy formula to an iterated integral of order \( \nu \) (see, e.g., [22]). Note that \( X^{[1]}(L, t) = X(L, t) \).

### 2.3 Alpha and betas

Take \( n + 1 \) \( S \)-integrable time series \( Y, X_1, \ldots, X_n : \mathfrak{T} \rightarrow \mathbb{R} \). Assume, without any loss of generality, that their values at any \( t_i \in \mathfrak{T} \) is bounded by a given limited number. Set

\[ \mathcal{Y}(L, t) = \alpha(L, t) + \beta_1(L, t) X_1(L, t) + \cdots + \beta_n(L, t) X_n(L, t) \]

\( \alpha(L, t), \beta_i(L, t) \in \mathbb{R}, i = 1, \ldots, n, \) are not yet uniquely determined.

Define the time series \( 1 : \mathfrak{T} \rightarrow \mathbb{R}, t_i \mapsto 1 \). Its arithmetical average is always 1. Equation (3) yields

\[ 1^{[\nu]}(L, t) = \frac{L^{\nu - 1}}{\nu!} \]

Introduce the Wronskian-like determinant (see, e.g., [24])

\[ W_{1,X_1,\ldots,X_n}(L, t) = \begin{vmatrix} 1 & X_1^{[1]}(L, t) & \cdots & X_n^{[1]}(L, t) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{L^n}{(n+1)!} & X_1^{[n+1]}(L, t) & \cdots & X_n^{[n+1]}(L, t) \end{vmatrix} \quad (4) \]

\( X_1, \ldots, X_n \) are said to be \( \alpha \)-W-independent on \([t - L, t]\) if, and only if, \( W_{1,X_1,\ldots,X_n}(L, t) \) is appreciable.

Introduce the \((n + 1) \times (n + 2)\) matrix

\[ \mathcal{M}_{1,X_1,\ldots,X_n}(L, t) = \begin{pmatrix} \mathcal{Y}^{[1]}(L, t) & 1 & X_1^{[1]}(L, t) & \cdots & X_n^{[1]}(L, t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{L^n}{(n+1)!} & X_1^{[n+1]}(L, t) & \cdots & X_n^{[n+1]}(L, t) \end{pmatrix} \quad (5) \]

Assume that \( X_1, \ldots, X_n \) are \( \alpha \)-W-independent on \([t - L, t]\). Then the matrix (5) is of rank \( n + 1 \). The Cramer rule yields limited values for \( \alpha(L, t), \beta_1(L, t), \ldots, \beta_n(L, t) \) in Equation (2.3):

\[ \alpha(L, t) = \frac{\mathcal{Y}^{[1]}(L, t) X_1^{[1]}(L, t) \cdots X_n^{[1]}(L, t)}{W_{1,X_1,\ldots,X_n}(L, t)} \]

\[ \beta_1(L, t) = \frac{\cdots \frac{L^n}{(n+1)!} X_1^{[n+1]}(L, t) \cdots X_n^{[n+1]}(L, t)}{W_{1,X_1,\ldots,X_n}(L, t)} \]
\[
\beta_n(L, t) = \begin{vmatrix}
1 & X_1^{[1]}(L, t) & \cdots & X_{n-1}^{[1]}(L, t) & Y^{[1]}(L, t) \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\frac{\varepsilon}{(n+1)!} & X_1^{[n+1]}(L, t) & \cdots & X_{n-1}^{[n+1]}(L, t) & Y^{[n+1]}(L, t)
\end{vmatrix}
\]  

\]

\begin{align*}
W_{X_1, \ldots, X_n}(L, t) &= X_1^{[1]}(L, t) \cdots X_{n}^{[1]}(L, t) \\
& \quad \cdots \cdots \\
& \quad X_1^{[n]}(L, t) \cdots X_{n}^{[n]}(L, t)
\end{align*}

Remark 2. Replacing in Equation (2.3) the arithmetic averages by the original time series yields

\[
Y = \alpha(L, t) + \beta_1(L, t)X_1 + \cdots + \beta_n(L, t)X_n + e_{[t-L, t]}
\]

where \( \int_{t-L}^{t} e_{[t-L, t]} d\tau \) is infinitesimal.

### 2.4 Betas alone

Let us drop \( \alpha \). Equation (2.3) becomes

\[
\overline{Y}(L, t) = \sum_{i=1}^{n} \beta_i(L, t)X_i(L, t)
\]  

(6)

Determinant (4) is replaced by

\[
W_{X_1, \ldots, X_n}(L, t) = X_1^{[1]}(L, t) \cdots X_{n}^{[1]}(L, t) \\
& \quad \cdots \cdots \\
& \quad X_1^{[n]}(L, t) \cdots X_{n}^{[n]}(L, t)
\]

\( X_1, \ldots, X_n \) are said to be \( W \)-independent on \([t-L, t]\) if, and only if, \( W_{X_1, \ldots, X_n}(L, t) \) is appreciable. Matrix (5) is replaced by the \( n \times (n+1) \) matrix

\[
\mathcal{M}_{Y, X_1, \ldots, X_n}(L, t) = \begin{pmatrix}
Y^{[1]}(L, t) & X_1^{[1]}(L, t) & \cdots & X_{n}^{[1]}(L, t) \\
\vdots & \ddots & \ddots & \ddots \\
Y^{[n]}(L, t) & X_1^{[n]}(L, t) & \cdots & X_{n}^{[n]}(L, t)
\end{pmatrix}
\]

Assume that \( X_1, \ldots, X_n \) are \( W \)-independent on \([t-L, t]\). Then \( \mathcal{M}_{Y, X_1, \ldots, X_n}(L, t) \) is of rank \( n \). Limited values for \( \beta_i(L, t), i = 1, \ldots, n \) are again given by the Cramer rule. Although we do not give again the formulae, it goes without saying that these numerical values are in general different from those derived in Sect. 2.3. We do not repeat also Remark 2.

Remark 3. If \( n = 1 \) and \( \int_{t-L}^{t} Y d\tau \neq 0 \),

\[
\beta_1(L, t) = \frac{X_1^{[1]}(L, t)}{Y(L, t)} = \frac{\int_{t-L}^{t} X_1 d\tau}{\int_{t-L}^{t} Y d\tau}
\]  

(7)
3 Comments

3.1 The model-free standpoint

The length $L$ of the time window $[t - L, t]$ may be chosen quite short, i.e., of a size compatible with what is needed for calculating the trends in [10]. Updating the various factors is achieved by letting slide this time window. Let us emphasize that the linearity of the local models (2.3) and (6), which are valid only during a short time interval, does not imply therefore a global time-invariant linearity as assumed in the CAPM and APT settings. This model-free standpoint has already been proved to be quite efficient in control theory.\(^\text{10}\)

Remark 4. Equations (2.3) and (6) should not be viewed as time-varying linear relations, since the values of their coefficients depend nonlinearly on $X_1, \ldots, X_n$.

3.2 Reverse formula

Take for simplicity’s sake $n = 1$ in Equation (2.3). Then

$$\overline{Y}(L, t) = \alpha + \beta_1(L, t)X_1(L, t)$$

yields, if $\beta_1(L, t) \neq 0$,

$$X_1(L, t) = -\frac{1}{\alpha(L, t)} + \frac{1}{\beta_1(L, t)}Y(L, t)$$

The same reverse formula would have also been derived from the linear algebra of Section 2.4.

Now, we restrict ourselves for simplicity’s sake to a CAPM-like equation

$$r(t) = \alpha + \beta R(t) + \epsilon(t)$$

(8)

where

- $r(t)$ and $R(t)$ are the values at time $t$ of some returns,
- $\epsilon(t)$ is a zero-mean stochastic processes,
- $\alpha$ and $\beta$ are constant.

As pointed out in [32], the classic least square techniques utilized with Equation (8) do not lead to the most natural reverse formula.

\(^{10}\) See [13]. Many successful concrete engineering applications may be found in the references.
3.3 Volatility

3.3.1 Today’s situation

Consider again Equation (8). This global linear time-invariant equation leads to usual systematic, or market, risk calculation, i.e., to

\[ \text{var}(r) = \beta^2 \text{var}(R) + \text{var}(\epsilon) \] (9)

where \( \text{var}(\epsilon) \) should be “small” if there is a “good” diversification. It explains

1. why increasing \( \beta \) also increases the risk,
2. the importance of generating a “good” \( \alpha \) in Equation (8).

If, as emphasized in [32], Equation (8) does not hold, i.e., there is no global linear time-invariant relationship, Equation (9) is then erroneous. The whole “philosophy” which was built in order to justify the utilization of the CAPM and of its extensions like the APT (see, e.g., [4]) might therefore break down.\footnote{See also the harsh quotations and comments in [3].}

Remark 5. Equation (2) shows that the quick fluctuations do not appear in Equations (2.3) and (6). Those equations are therefore useful for comparing the time evolution of means, i.e., trends, and certainly not for the comparison of the corresponding volatilities.

3.3.2 A remedy

We start by reviewing the definitions of (co)variances and volatility given in [14, 15]. Take two \( S \)-integrable time series \( X, Y \) such that their squares and the squares of \( E(X) \) and \( E(Y) \) are also \( S \)-integrable. Then the following property is obvious: \( XY, E(X)E(Y), E(X)Y_{\text{fluctuat}}, X_{\text{fluctuat}}E(Y), X_{\text{fluctuat}}Y_{\text{fluctuat}} \) are all \( S \)-integrable. Assume moreover that \( E(X) \) and \( E(Y) \) are differentiable in the following sense: there exist two Lebesgue integrable time series \( f, g : \mathbb{T} \to \mathbb{R} \), such that, \( \forall \ t \in \mathbb{T} \), with the possible exception of a limited number of values of \( t \),

\[
E(X) = E(X)(0) + \int_0^t f(\tau)d\tau,
\]

\[
E(Y) = E(Y)(0) + \int_0^t g(\tau)d\tau.
\]

Integrating by parts shows that the products \( E(X)Y_{\text{fluctuat}} \) and \( X_{\text{fluctuat}}E(Y) \) are quickly fluctuating.

Remark 6. Let us emphasize that the product \( X_{\text{fluctuat}}Y_{\text{fluctuat}} \) is not necessarily quickly fluctuating.

The following definitions are natural:

1. The covariance of two time series \( X \) and \( Y \) is
\[
\text{cov}(XY) = E((X - E(X))(Y - E(Y))) \\
\approx E(XY) - E(X(t)) \times E(Y)
\]

2. The variance of the time series \( X \) is
\[
\text{var}(X) = E\left((X - E(X))^2\right) \\
\approx E(X^2) - (E(X))^2
\]

3. The volatility of \( X \) is the corresponding standard deviation
\[
\text{vol}(X) = \sqrt{\text{var}(X)} \tag{10}
\]

The definition of volatility given by Equation (10) associates to a time series \( X \) another time series \( \text{vol}(X) \), which is called the volatility time series. Take now \( n + 1 \) time series \( Y, X_1, \ldots, X_n \), which satisfy the above assumptions on integrability and differentiability. We may repeat for the \( n + 1 \) time series \( \text{vol}(Y), \text{vol}(X_1), \ldots, \text{vol}(X_n) \) the same calculations as in Sections 2.3 and 2.4. It yields new relations between those volatilities.

Remark 7. Take \( n = 1 \) as in the CAPM setting. We now have two time-varying betas for comparing the two assets:

1. The first one, derived from Sections 2.3 or 2.4, compares an averaged time evolution of their values or returns.
2. The second one, derived from Section 3.3.2, compares an averaged time evolution of their corresponding volatilities.\(^{12}\)

4 Some computer experiments

4.1 Monovariate \( \beta \)

Figures 1-(a), 1-(b), 1-(c) exhibit the daily time series behaviors of the the S&P 500 and of the two following assets:

1. IBM from 1962-01-02 until 2009-07-21 (11776 days),
2. JPMORGAN CHASE (JPM) from 1983-12-30 until 2009-07-21 (6267 days).

The corresponding returns are given in Figures 2-(a), 2-(b), 2-(c) and their volatilities in Figures 3-(a), 3-(b), 3-(c). We took \( L = 500 \) for the length \( L \) of the sliding windows.

Compare, as in Section 2.4, \( i.e., \) without \( \alpha \), those various assets.

\(^{12}\) Let us stress that the “famous” \( \alpha \) coefficient, in Equation (8), of the CAPM might therefore become quite obsolete.
The comparison between IBM and S&P 500 utilizes Formula (7). Figures 4-(a), 4-(b) and 4-(c) show three \( \beta \)s corresponding respectively to the values, the returns and the volatilities.

### 4.2 Bivariate \( \beta \)

A bivariate extension is provided by a rather academic example where we want to “explain” the S&P 500 via IBM and JPM. Set therefore

\[
R_{\text{S&P 500}} = \beta_1 R_{\text{IBM}} + \beta_2 R_{\text{JPM}}
\]

where \( R_{\text{S&P 500}} \) (resp. \( R_{\text{IBM}}, R_{\text{JPM}} \)) is the return of S&P 500 (resp. IBM, JPM). According to Section 2.4 we have to invert the determinant of the \( 2 \times 2 \) matrix

\[
B = \begin{pmatrix}
\int_{t-L}^{t} R_{\text{IBM}}(\tau) d\tau & \int_{t-L}^{t} R_{\text{JPM}}(\tau) d\tau \\
\int_{t-L}^{t} \tau R_{\text{IBM}}(\tau) d\tau & \int_{t-L}^{t} \tau R_{\text{JPM}}(\tau) d\tau
\end{pmatrix}
\]

Several sizes \( L = 100, 300, 500 \) for the sliding windows are utilized in parallel in order, if \( \det(B) \approx 0 \), to pick up the size where \( |\det(B)| \) is the greatest. Figure 4-(d) exhibits quite convincing results.

### References


Fig. 1 Time values

Fig. 2 Returns

Fig. 3 Volatility of returns
Systematic and multifactor risk models revisited

Fig. 4 Betas