Immunization in an affine term structure framework

Mario Alessandro Maggi
(Università di Pavia)

# 16 (02-02)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Febbraio 2002
Immunization in an affine term structure framework

Mario Alessandro Maggi*

January 2002

Abstract

In this work I deal with the affine term structure models (ATSM hereafter), namely the models where the valuation function \( P \) of a default-free (nonrisky) zero-coupon bond (ZCB in what follows) is exponentially affine in a vector of stochastic state variables (or risk factors). After a short introduction to the classical immunization theory, my work analyzes how to apply it to an ATSM. Namely, I will deal with ZCBs’ portfolio multifactor immunization. Non trivial problems arise. No special difficulty appears in finding a suitable immunization strategy in one factor models, but all becomes hard to deal with as soon as the dimension of the model grows. The problem may admit no solution even with two factors only.

1 Introduction

In this work I deal with the affine term structure models (ATSM hereafter), namely the models where the valuation function \( P \) of a default-free (nonrisky) zero-coupon bond (ZCB in what follows) is

\[
P(t, \tau) = \exp \left\{ A(\tau) - B(\tau) X(t) \right\},
\]

being: \( t \) the current valuation date, \( \tau \) the time to maturity, \( A = A(\tau) \) a real function, \( B = B(\tau) \) an \( \mathbb{R}^n \) vector function and \( X = X(t) \) the \( \mathbb{R}^n \) vector collecting the \( n \) stochastic risk factors.

In the last years the most part of the interest rates term structure theory has exhibited a relevant effort towards a progressive generalization and settling (e.g., see [5], [12], [14], [15], [19], [20], [21], and [22]). Moreover, several empirical researches have tested how much those models are adequate for the market (e.g., see [6], [8] and [25]).

Some papers in this field are widely recognized as cornerstones. Among them we find the 1996 paper by Duffie and Kan [21] and the 2000 paper by

*Università degli Studi di Pavia, Dipartimento di Economia Politica e Metodi Quantitativi, via S. Felice, 5, 27040 Pavia (Italy), e-mail: mario.maggi@unipv.it.
Dai and Singleton [14]. The former contains an ultimate and strictly formal characterization of ATSMs. Its main results are the valuation formula (1.1), together with an essential equivalence between it and the linearity, with respect to $X$, of the drift and of the conditional variance functions of the process leading the dynamics of the risk factors. The latter affords a detailed analysis of the affine models in a generalized formulation and, moreover, some deep reasoning on their structure, supported by an empirical comparison grounded of the U.S. interest rates of the last decades.

As [14] clearly shows, affine models constitute a wide family that encompasses the majority of the most known standard models as special cases (also the Vasicek and Cox, Ingersoll and Ross models, among the most celebrated ones).

The interest towards these models stems mainly from their tractability. Indeed, the functions $A$ and $B$ which are employed for valuation purposes may be numerically obtained in a satisfactory manner even in the worst cases. However, the use of ATSMs must be justified not only by their relative simplicity, but also by their flexibility. In fact, especially when several risk factors are involved, those models give a good representation of the market behavior.

Section 2 fixes some notation. Section 3 briefly introduces affine term structure models. Sections 4 contains a short introduction to the classical immunization theory and section 5 analyzes how to apply it to an ATSM. Section 6 applies the result of section 5 to ZCBs’ portfolios. Non trivial problems arise. No special difficulty appears in finding a suitable immunization strategy in one factor models, whereas all becomes hard to deal with as soon as the dimension of the model grows. The problem may admit no solution even with two factors. Section 7 concludes.

2 Notation and conventions

Let me fix some notation and conventions used in this work.

The scalar $a_{i}$ is the $i$-th entry of the vector $a$ and $a_{ij}$ is the entry in the $i$-th row and $j$-th column of the matrix $A = [a_{ij}]$. The vectors $A$ and $A'$ define the $i$-th row and the $j$-th column of the matrix $A$. A prime ($'$) denotes transposition. The matrix $I$ is the identity matrix. The ordinary inverse of a square matrix $A$ is denoted by $A^{-1}$. For a symmetric matrix $A$ the notation $A > 0$ ($A \geq 0$) means that $A$ is positive definite (semipositive definite). Vectors are treated as matrices with one row or one column. The vectors $[0]$, $[1]$ and $[1]^t$ are the null vector or matrix, the vector of all ones and the $i$-th basis vector respectively. Their dimension is the suitable one depending on the context. The product between matrix is defined in the standard way, so, with a row vector $x$ and a column vector $y$, both in $\mathbb{R}^n$, $xy$ defines the scalar product $\sum_{i=1}^{n} x_i y_i$, and $yx$ the outer product $A = [a_{ij}] = [y_i x_j]$.

The componentwise comparisons between both row- (or both column-) vectors $a$, $b \in \mathbb{R}^n$ follows this convention: $a \geq b$ and $a > b$ respectively mean $a_i \geq b_i$ and $a_i > b_i$, $\forall i$, whereas $a \geq b$ stands for $a \geq b$ but $a \neq b$. With $b = [0]$ I
respectively get a nonnegative, positive and semipositive vector $a$. In the same way I define the notation like $a \leq b$ and I use those signs also for matrices of the same order. Hence, when comparing scalars the “less (more) or equal then” sign is $\leq (\geq)$. The rank of $A$ and its trace $(\sum_{i=1}^{n} a_{ii})$ are denoted by $\text{rk}(A)$ and $\text{tr}(A)$ respectively.

For a (regular enough) real scalar function $f(t,x)$ with $t \in \mathbb{R}$ and $x \in \mathbb{R}^{n}$ the notation $\nabla_{x}(f)$ and $H_{x}(f)$ define the gradient (a row vector) and the Hessian matrix of $f(x)$ with respect to the components of the vector $x$ only.

## 3 Affine term structure models

In this section I introduce continuous time dynamic models for the term structure of interest rates, focusing on affine multifactor models. The models describing and analyzing yield curve and interest rate products have three decades of history. Since in seventies Black [10] and Vasicek [32] placed fixed income valuation into a continuous time arbitrage-free framework, many models have been developed and generalized to best fit the real world features of interest rates dynamics.

For simplicity’s sake I treat unit face value zCBs, whose positive price is $P$:

$$P = P(t, \tau), \quad \text{with } t \geq 0, \ \tau \geq 0,$$

where $t$ is the current date and $\tau$ is the time to maturity. The fixed income continuous time market is assumed to be frictionless: no transaction costs, bid-ask spread or any type of frictions are in place. Moreover, the function $P$ is assumed to be differentiable in $\tau$. All this means that it is possible to trade in continuous time over uncountably many zCBs, one for each possible time to maturity up to a finite horizon $T$.

In financial markets the dynamics of prices and interest rates are driven by macroeconomic variables, such as decisions of political economy and expectations, risk attitude and feelings of economic agents. The first models dealt with a single factor source of randomness in yield curve dynamics. For an introduction to these models see, e.g., [11]. In the following I will introduce the general case of an economy driven by $n$ sources of randomness and consider the single factor models as special cases.

Let the random factors or state variables are stacked in a column vector $X = X(t) \in \mathbb{R}^{n}$, whose dynamics is described by a vector SDE

$$dX(t) = \mu(X) \, dt + \sigma(X) \, dW(t), \quad (3.1)$$

where the column vector $W(t)$ is an $n$-dimensional independent Wiener process under the historical probability measure, whereas the column vector $\mu = \mu(X) : \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$ and $\sigma = \sigma(X) : \mathbb{R}^{n} \rightarrow \mathbb{R}^{n} \times \mathbb{R}^{n}$ maps $X \in \mathbb{R}^{n}$ into an $n \times n$ matrix.

---

1 Often further regularity conditions are needed in order to assure the existence of a solution of the SDE (see [28, chapter 2]).
Let $F(t, X)$ be the value of any fixed income security (i.e., a ZCB, a cap or a more complex one) whose value is assumed to depend on time and factors only. By the continuous time arbitrage theory, the term structure equation is derived

$$F_t + \nabla_X (F) (\mu - \sigma \Lambda(t)) + \frac{1}{2} \text{tr} (\sigma' H_X (F) \sigma) - r(t) F = 0,$$  \hfill (3.2)

where the $\mathbb{R}^n$ column vector $\Lambda(t)$ is the risk premia function\(^2\). Equation (3.2) is a (partial derivatives) equation to be satisfied by the prices of a fixed income securities in order to prevent arbitrage opportunities\(^3\), (see [11] or [28] for a detailed derivation of equation (3.2)).

In order to get a unique solution, equation (3.2) needs a boundary condition. For a $(t + \tau)$-maturity ZCB the obvious one is

$$F(t + \tau, X) = 1, \quad \text{or} \quad P(t, 0, X) = 1; \quad (3.4)$$

in general for all non path-dependent interest rate derivatives the suitable boundary condition is

$$F(t + \tau, X) = \Phi(X),$$

$\Phi$ being the pay-off conditions of the contract.

The dynamics of factors (3.1), hence the corresponding pricing function of ZCBs, may be of various kinds. The most tractable class, that however preserve some flexibility in representing the real world, is the ATSM class. A model is said to be affine if the ZCB’s pricing function is exponentially affine in the factors:

$$P(t, \tau) = \exp \{A(t, \tau) - B(t, \tau) X(t)\}, \quad (3.5)$$

being the real scalar $A$ and the row vector $B \in \mathbb{R}^n$ deterministic (and smooth) real functions, whereas $X$ defines the risk factor vector.

If a model is affine, then the yields to maturity, the forward rates and the short rate are affine as well.

Duffie and Kan [21, page 386] find the (substantial) equivalence between (3.5) and the fact that factors follow an affine diffusion under the risk neutral probability measure $Q$, i.e. the drift and volatility functions in (3.1) are affine

\(^2\)The $i$-th element $\Lambda_i$ of $\Lambda$ represents the risk premium relative to the $i$-th factor $X_i$. In order to guarantee the existence and the uniqueness of the stochastic processes involved the function $\Lambda(t)$ is assumed to verify

$$\int_0^T \|\Lambda(s)\|^2 \, ds < \infty, \quad \text{almost surely.}$$

\(^3\)The term structure equation is often put forward in the equivalent form (e.g., see [11, pages 44-45])

$$F_t + \nabla_X (F) (\mu - \sigma \Lambda(t)) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 F}{\partial X_i \partial X_j} \sigma_i \sigma_j - r(t) F = 0. \quad (3.3)$$
in the factors: this means that each entry of $\mu^Q$ and $\sigma'\sigma$ is affine in $X$.\textsuperscript{4} They prove that if $\mu^Q$, $\sigma'\sigma$ and $r(X)$ are affine in $X$, then \textit{i)} (3.5) holds, \textit{ii)} if (3.5) holds, then $\mu^Q$, $\sigma'\sigma$ and $r(X)$ are affine in $X$. The second implication holds generically, i.e. except on a closed set of null measure.

From now onwards I assume that the drift and volatility functions are time independent and satisfy, under the risk-neutral measure the equations

\[
\mu^Q(X) = K^Q(\Theta^Q - X(t)), \\
\sigma(X) = \Sigma\sqrt{S},
\]

where:

- the $Q$ superscript distinguishes parameters under the risk neutral measure from the corresponding ones under the historical one;
- $\Theta^Q$ is a column vector of $R^n$;
- $K^Q$, $\Sigma$ and $S$ are ($n \times n$) nonsingular matrix;
- $S$ is diagonal and its diagonal $i$-th entry is

\[
S_i(X(t)) = \alpha_i + \beta_iX(t), \quad i = 1,2,\ldots,n,
\]

with $\alpha_i \in R$ and $\beta_i$ a row vector of $R^n$;
- $\sqrt{S}$ is diagonal matrix solving $\sqrt{S}\left(\sqrt{S}\right)' = S$.

In this case the zcbs’ pricing function is

\[
P(t,\tau) = \exp\{A(\tau) - B(\tau)X(t)\},
\]

thus for an ATSM the term structure equation (3.2) becomes

\[
\begin{cases}
B_r' = -(K^Q)'B'(\tau) - \frac{1}{2} \sum_{i=1}^{n} [\Sigma' B'(\tau)]^2_i \beta'_i + \delta'_1 \\
A_r = -(\Theta^Q)'(K^Q)'B'(\tau) + \frac{1}{2} \sum_{i=1}^{n} [\Sigma' B'(\tau)]^2_i \alpha_i - \delta_0.
\end{cases}
\]

This system of linear quadratic (of the Riccati class) ordinary differential equations is to be completed with the initial conditions. The boundary condition (3.4) for a zcb is

\[
P(t,0) = 1 = \exp\{A(t,0) - B(t,0)X(t)\}, \quad \forall t, \forall X(t),
\]

therefore $A(t,0) = 0$ and $B(t,0) = [0]$:

\[
A(0) = 0, \quad B(0) = [0].
\]

\textsuperscript{4}The entry $\sigma_{ij}$ of the conditional volatility matrix $\sigma'\sigma$ of the factors $X$ equals the inner product between the $i$-th and the $j$-th row of matrix $\sigma$, just as in (3.3).
In order to solve this system it is necessary to know the values of the parameters $\Theta^Q, K^Q, \Sigma, \alpha, B$ of the factors’ risk-neutral diffusion, and the coefficients $\delta_0, \delta_1$ of the short rate as well. The system (3.7)–(3.8) admits analytical solution in special cases only, hence in general it is necessary to use an approximation algorithm, i.e. the fourth-order Runge-Kutta method. This fact allows to say that, although complex, ATSMs are numerically tractable. In fact, given a set of parameters, finding the (approximate) pricing function is a quite easy matter.

Summing up, in ATSM the SDE driving the factors may be written as

$$dX(t) = K(\Theta - X(t))\ dt + \Sigma \sqrt{\gamma} \, dW(t),$$  \hspace{1cm} (3.9)

under the historical probability measure, or

$$dX(t) = K^Q(\Theta^Q - X(t))\ dt + \Sigma \sqrt{\gamma} \, dW(t),$$

under the risk neutral one.

4 Classical immunization

This section contains a short introduction to the classical immunization theory. In the fixed income market the main risk is represented by the changes in prices of assets due to unexpected movements in interest rates. Such movements can be the effects of monetary policy, changes in investor’s risk aversion, economic growth, inflation, or other macro- or micro-economic reasons that here I consider as exogenous. These price movements produce an interest rate risk. For example, in general, an overall rise in interest rates depreciates long positions in fixed income products\(^5\). Sometimes the potential loss in long positions, roughly due to a rise in interest rates, is called net value risk, whereas the potential loss in short positions, caused by fall in interest rates, is called reinvestment risk.

There are many works that look for strategies to make bond portfolios insensitive to changes in interest rates or, better, to compose them in a way that gives some gain from interest rates movements within a certain class of possible shocks. This subject is called immunization and the various kinds of immunization differ in the assumptions that are adopted about the set of possible shocks\(^6\). For instance, the so called global immunization does not constrain the size of perturbations in interest rates, whereas local immunization allows sufficiently small shocks only. The fundamental assumption in all kinds of immunization theory is that the prices of fixed income instruments are affected by interest rates only.

\(^5\)This effect is present in equity market too, but it is not the main source of volatility of stock prices.

\(^6\)Indeed there are other definitions of immunization. For example in [29], the Authors call immunization a portfolio strategy that keeps the portfolio “value as close as possible to the value of another asset: the target”. This definition does not completely agree with the fact that many immunization techniques look for a good (i.e. never negative) response of the portfolio value to the shocks in interest rates.
The classical technique of immunization is based on the Hicks-Macaulay-
Redington theory (see [27] and [30]). Its objective is to set up a fixed income
assets portfolio whose total value can not fall down after an interest rates
move- ment. In other words, the today (present at time $t$) value of the portfolio is
required to be a local minimum of the value of the portfolio with respect to
the term structure of interest rates. The admissible set of possible shocks in term
structure characterizes each model of immunization. For instance, Redington
[30] deals with parallel shifts of a flat term structure of interest rates, whereas
Fisher and Weil [23] extend the model by allowing that those shifts are applied
to any kind of initial term structure. A survey of classical immunization theory
as well as other refinements may be found, e.g. in [16].

In order to summarize classical immunization theory, consider a stream of
cash flows $a$ that pays out $a_i \neq 0$ at time $t + \tau_i$, for $i = 1, \ldots, n$. Define its market
(no arbitrage) value $V(t)$ at time $t$ as the present value of $a$ at the current term
structure of interest rates implied in the ZCB’s prices term structure:

$$V(t) = \sum_{i=1}^{n} a_i P(t, \tau_i) = \sum_{i=1}^{n} a_i e^{-Y(t, \tau_i) \tau_i}.$$  

The partial derivative of $V$ with respect to a parallel shift $\varepsilon$ in yields, when
valued at the current term structure (i.e. at the null shift $\varepsilon = 0$), is

$$\left. \frac{\partial}{\partial \varepsilon} V(t) \right|_{\varepsilon=0} = -V(t) D(a), \quad (4.1)$$

where

$$D(a) = \frac{\sum_{i=1}^{n} a_i e^{-Y(t, \tau_i) \tau_i \tau_i}}{V(t)},$$

is the duration of the asset or of the financial project that produces the cash
flows composing $a$.

In a similar way, the second derivative of $V$ with respect to a parallel shift
$\varepsilon$ in yields, still valued at the current term structure, is

$$\left. \frac{\partial^2}{\partial \varepsilon^2} V(t) \right|_{\varepsilon=0} = V(t) D^2(a), \quad (4.2)$$

where

$$D^2(a) = \frac{\sum_{i=1}^{n} (\tau_i)^2 a_i e^{-Y(t, \tau_i) \tau_i \tau_i}}{V(t)},$$

is the convexity (or Redington’s dispersion) of $a$.

The vector $a$ of the cash flows may be split into the $\mathbb{R}^n$ subvectors $a^1$ and
$a^2$, containing different cash flows of $a$ and zero entries sufficient to get two $\mathbb{R}^n$

vectors:

$$a = a^1 + a^2 \quad (4.3)$$

7
and $V(t)$ may be written accordingly as

$$V(t) = V^1(t) + V^2(t) = \sum_{a_i \in a^1} a_i P(t, \tau_i) + \sum_{a_i \in a^2} a_i P(t, \tau_i), \quad (4.4)$$

hence I get

$$D(a) = \frac{V^1(t) D(a^1) + V^2(t) D(a^2)}{V(t)},$$

$$D^2(a) = \frac{V^1(t) D^2(a^1) + V^2(t) D^2(a^2)}{V(t)}.$$

If $a^1$ contains only the positive cash flows of $a$ and $a^2$ the negative ones, then the Redington theorem follows:

**Theorem 4.1 (Redington [30])** Split the vector $a$ of cash flows as follows:

$$a = a^1 + a^2, \quad \text{with: } a^1 \geq [0], a^2 \leq [0], \quad (4.5)$$

and assume that the relations

$$V^1(t) = -V^2(t), \quad (4.6)$$

$$D(a^1) = D(a^2), \quad (4.7)$$

$$D^2(a^1) > D^2(a^2), \quad (4.8)$$

hold. Then $a$ is locally immunized against parallel shifts of a flat term structure of interest rates.

In fact, if $a$ fulfils (4.6), then (4.4) shows that the current value $V(t)$ is null, in which case, thanks to (4.1) and (4.2), properties (4.7) and (4.8) respectively give

$$\frac{\partial}{\partial \varepsilon} V(t) \bigg|_{\varepsilon=0} = 0, \quad \frac{\partial^2}{\partial \varepsilon^2} V(t) \bigg|_{\varepsilon=0} > 0, \quad (4.9)$$

so $V(t)$ exhibits an internal local minimum with respect to an additive shift $\varepsilon$ in the current flat term structure. Owing to the use of condition (4.7), the use of this kind of immunization is often called duration matching strategy.

Let me remark that condition (4.6) (often called wealth constraint) on the null total value $V(t)$ and (4.5) on the signs of $a^1$ and $a^2$, are not strictly necessary in order to verify (4.9), as in general it is sufficient that the relations

$$V^1(t) D(a^1) + V^2(t) D(a^2) = 0$$

$$V^1(t) D^2(a^1) + V^2(t) D^2(a^2) > 0,$$

hold for any partition $(a^1, a^2)$ of vector $a$, as it is pointed out in [18, chapter 15].
If, instead of inflows and outflows, \( a^1 \) and \( a^2 \) represent the cash flows of two financial projects composing the portfolio \( a \), it is sensible to look for the weights \( w_1 \) and \( w_2 \) (with \( w_1 + w_2 = 1 \)) that immunize it, i.e. for the relative weights such that

\[
\begin{align*}
w_1 V^1(t) D(a^1) + w_2 V^2(t) D(a^2) &= 0, \\
w_1 V^1(t) D^2(a^1) + w_2 V^2(t) D^2(a^2) &= 0.
\end{align*}
\]

A portfolio whose value is locally insensitive to the shock may be called a critical portfolio.

Critiques and generalizations about this framework come from a large number of works (too many to cite all, see for example [4], [18, Chapter 15], [31] and references therein; for a recent review also see [7]). Although this theory is unrealistic and for someone it does not seem consistent with an arbitrage-free treasury market, its application gives quite good results.

It is unrealistic because almost always the actual movement in interest rates are not parallel: different maturity rates, although correlated, move in different ways. It is inconsistent with the absence of arbitrage because if it were possible to set up a risk-free portfolio with zero value and that appreciates in some future state of nature and do not lose value in anyone (as theorem 4.1 says), an arbitrage opportunity is found (see [9]).

This last point of view is rather controversial. In fact immunization is not actually a way to find a free lunch, but a technique that tries to eliminate from portfolios the interest rate risk due to instantaneous shifts, of stochastic magnitude and timing, in interest rates.

At least the following main issues are to be evaluated in order to clarify whether immunization does give shape to a free lunch (later in section 5.2 I will come back on this matter):

- The wealth constraint (4.6) is the first one that a free lunch usually requires, so free lunch is ruled out when that constraint does not hold. But this is a poor argument, as the addition to the cash flows \( a \) of an asset or liability whose value \( -V(t) \) is insensitive to interest rates movements leads to a null total value.

- As I will point out later in section 5.2 with respect to factor immunization, in immunization theory it is common to “freeze” the time; in other words, the aim is to look at the interest rate shock effects only, without letting values to be affected from the passage of time. All this may be viewed as a refinement of the classical what if tool, which regards the effects on the portfolio value of instantaneous changes in interest rates, rather than what happens when time goes by.

- Another difference pertains to the nature of the strategy. When local immunization is possible, then the resulting portfolio appreciates only locally, therefore also some interest rate movements may exist large enough to depreciate the portfolio. So this is not a free lunch, even though I do
not consider the model risk. In order to see in an immunized portfolio a free lunch strategy, it is necessary to assign zero probability to large movements in factors. Even though this would agree with the affine setting (and indeed it would not), in this case an old question arises: dealing with local results, when a movement may be considered small enough or too large?

- The question “is immunization a free lunch?” may be significant, at the most, only when immunization is global in a strict sense, i.e. when current rates lead to a global minimum of $V(t)$ with respect to all possible interest rate movements leading to still positive and finite rates.

Although parallel shifts in term structure represent an abstraction, empirical works have found that they explain 80% to 90% of the actual shocks (see for example [26]), so duration matching can practically work quite well. This explains why duration immunization is one of the most widely used method for hedging bond portfolios: the model is a good first approximation of what really happens in treasury market.

4.1 Recent developments

In recent years the research on interest rate management has more and more used stochastic models borrowed from stock market theory in order to value and use interest rate derivative securities for hedging and speculation. The theory introduced in section 3 is clearly derived from the Black-Scholes setting and its core is the absence of arbitrage opportunities. This framework has produced some hedging schemes for fixed income products. In fact, the strategy put forward in [6], [12] and [26] in order to eliminate interest rate risk is an extension of the Black-Scholes delta-hedging technique; in other words, the objective is to compose a portfolio whose value has all first derivatives with respect to the underlying factors null, so as to eliminate its volatility. It seems therefore that the tools employed in those models are the same as in stock market analysis.

On the other side various works (for example [4], [7], [18, Chapter 15] and [31]) continue in the classical field of immunization, analyzing sophisticated shapes of shocks and generalizing theory. In this field some Authors (for example [24]) try to hedge a treasury portfolio by minimizing a measure of risk and obtaining a reduction in variability of the portfolio value.

In what follows I will try to apply the classical approach to a risk neutral model of term structure. This trial may be viewed as a refinement of the classical what if tool, which regards the effects on the portfolio value of changes in interest rates, rather than the effects of the passage of time. My objective is to look for the existence of immunization strategies, if any, and for conditions under which it is possible to eliminate interest rate risk by choosing a suitable composition of a portfolio. In other words, my aim is to look for a bridge between the two approaches.
5 Factor immunization in an ATSM framework

Given the framework outlined in section 3, it seems sensible to extend immunization to an affine factor model

\[ P(t, \tau) = \exp \left\{ A(\tau) - B(\tau) X(t) \right\}, \]

where the risk factor vector \( X(t) \) can take values in a convex set \( D \subseteq \mathbb{R}^n \).

My aim is to try to find an immunized portfolio, whose nonzero value I call \( \Pi \), composed by a proportion \( w_i \) of asset \( i \), with \( i = 1, \ldots, k \). The price of one unit of asset \( i \) is \( G_i \). The (relative) weights \( w_i \) are gathered in the row vector \( w \in \mathbb{R}^k \) and sum up to unity. The portfolio composed by a number \( \gamma_i \) of asset \( i \) units has the value

\[ \Pi = \sum_{i=1}^{k} \gamma_i G_i \]  

(5.1)

and the relative increment in this value is

\[ \frac{d\Pi}{\Pi} = \sum_{i=1}^{k} \frac{\gamma_i}{\Pi} dG_i = \sum_{i=1}^{k} \frac{w_i}{G_i} dG_i, \quad \text{with: } w_i = \frac{\gamma_i G_i}{\Pi}. \]  

(5.2)

Here the non zero value \( \Pi \) of the portfolio is obviously understood: a usual assumption as soon as relative weights are introduced (e.g., see [11, section 5.3]).

In the case of a ZCB, (5.2) becomes

\[ \frac{d\Pi}{\Pi} = \sum_{i=1}^{k} \frac{w_i}{\Pi} dP(t, \tau_i) \]

whereas it looks as

\[ \frac{d\Pi}{\Pi} = \sum_{i=1}^{k} w_i \sum_{j=1}^{s} c_{ij} \frac{dP(t, \tau_j)}{P(t, \tau_j)}, \]

for a set of \( k \) coupon bonds priced \( G_i(t, \tau_i) \) and with coupon \( c_{ij} \) paid-out at the date \( t + \tau_j \), \( j = 1, \ldots, s \).

5.1 Convexity of zcbs and of coupon bonds

In classical immunization theory the values of simple instruments such as ZCBs and coupon bonds are convex functions of interest rates shifts. In fact \( P(t, \tau) = e^{-(Y(t, \tau) + \varepsilon)\tau} \) is convex in \( \varepsilon \) and

\[ G = \sum_{j=1}^{s} c_j e^{-(Y(t, \tau_j) + \varepsilon)\tau_j}, \]

being a linear combination of convex functions with positive weights, is convex in \( \varepsilon \) as well. There exist more general conditions on the shape of the shocks
that guarantee the convexity of these assets (see [16]), but for the extension put forward in section 5.2 this setting is suitable.

In this multifactor setting it is possible to extend the following preliminary results.

- The gradient and the Hessian matrix of \( P(t, \tau) \) with respect to the factor risks \( X_1 \) to \( X_n \) are respectively,

\[
\nabla_X (P) = -P(t, \tau) B(\tau), \\
H_X (P) = P(t, \tau) B'(\tau) B(\tau),
\]

and, being \( B'(\tau) B(\tau) \) an outer product, the Hessian is a positive semi-definite matrix with rank 1.

- The value of a coupon bond with maturity \( \tau_s \) may be expressed as the sum of a portfolio of zCBs in the form

\[
G = G(t, c) = \sum_{j=1}^{s} c_j P(t, \tau_j)
\]

as soon as \( c_j \) defines the coupon to be cashed at time \( (t + \tau_j) \) and the face value is already added to the last coupon \( c_s \). The gradient and the Hessian of \( G \) are respectively

\[
\nabla_X (G) = - \sum_{j=1}^{s} c_j P(t, \tau_j) B(\tau_j), \\
H_X (G) = \sum_{j=1}^{s} c_j P(t, \tau_j) B'(\tau_j) B(\tau_j), \quad (5.3)
\]

and the Hessian, being a linear combination with positive weights of positive semi-definite matrix, is at least positive semi-definite\(^7\).

Being \( P(t, \tau) \) a positive valued function, both zCBs and coupon bonds have an at least positive semi-definite Hessian \( \forall X \in D \), the subset of \( \mathbb{R}^n \) where the state vector \( X \) can take values, so in both cases \( G(t, c) \) is (weakly) convex with respect to the factors.

In order to simplify the notation:

- for a zCB with maturity \( \tau_i \) and price \( G_i = P(t, \tau_i) \) I define

\[
B_{G_i} = B(\tau_i), \quad H_{G_i} = B'(\tau_i) B(\tau_i); \quad (5.4)
\]

\(^7\)This is a convention: at least means that this matrix is semi-definite or definite. In other words I define the classes of positive definite and positive semi-definite matrices as two disjoint classes, as in most textbooks. The symmetry of \( H \) obviously comes from the symmetry of each outer product in (5.3).
• for a coupon bond with value \( G_i = G_i(t, c) \) and maturity \( \tau_i \) I write

\[
B_{G_i} = \sum_{j=1}^{k} \frac{c_{ij} P(t, \tau_j)}{G_i} B(\tau_j), \quad H_{G_i} = \sum_{j=1}^{k} \frac{c_{ij} P(t, \tau_j)}{G_i} B'(\tau_j) B(\tau_j);
\]

• for a portfolio composed by \( k \) assets (ZCBs and/or coupon bonds) with relative weights \( w_j \) and total value \( \Pi \) I put

\[
B_\Pi = \sum_{i=1}^{k} w_i B_{G_i}, \quad H_\Pi = \sum_{i=1}^{k} w_i H_{G_i}.
\]

5.2 \( \mathcal{B} \)-matching

As in the case of classical duration theory, the first step is to find a critical portfolio locally insensitive to shocks in factors, i.e. whose value \( \Pi \) defined in (5.1) has a zero gradient with respect to the factors:

\[
\nabla_X (\Pi) = \sum_{i=1}^{k} \gamma_i G_i B_{G_i} = \Pi \sum_{i=1}^{k} w_i B_{G_i} = \Pi B_\Pi = [0],
\]

in other words \( B_\Pi = [0] \). This simply means to find the relative weights vector \( w \) solving the system

\[
\begin{cases} 
    wB = [0] \\
    w1 = 1,
\end{cases} \tag{5.5}
\]

where \( B \) is the \( k \times n \) matrix that stacks the vectors \( B_{G_i} \), \( i = 1, \ldots, k \). This explains the title of the current section\(^8\).

Let me remark that a portfolio that fulfills condition (5.5) is delta-hedged as well. In fact the sde driving the portfolio value is

\[
\frac{d\Pi}{\Pi} = \mu_\Pi dt + wB \Sigma \sqrt{\Sigma} dW(t), \tag{5.6}
\]

as it can easily be found by applying Itô’s lemma to the affine diffusion of factors (3.9). Obviously, in a no-arbitrage environment, the instantaneous return of this portfolio must equal the risk free rate \( r(t) \).

This point may appear inconsistent with immunization theory, that looks for a portfolio that never loses value after a shock in interest rates. In immunization literature a common basic assumption is that the changes in interest rates are

---

\(^8\)The same result also holds should I consider the gradient of the relative change (5.2) in the portfolio value, as in that case, being \( \frac{1}{\Pi} \) and \( \Pi > 0 \), I get

\[
\nabla_X \left( \frac{d\Pi}{\Pi} \right) = \nabla_X \left( \sum_{i=1}^{k} w_i \frac{dG_i}{G_i} \right) = \sum_{i=1}^{k} w_i B_{G_i} = [0].
\]
instantaneous. Although unrealistic, this assumption is commonly made, as it allows to examine the effect on an interest rate change separately from the effect of the passage of time on the value of a fixed income instrument. In fact, the increment considered in (5.2) is with respect to changes in factor values, whereas in (5.6) the increment is over time. So the aim of immunization is to protect the portfolio against changes that may happen with a stopped time, i.e., against jumps. The affine diffusion (3.9) however does not allow discontinuity in the factors’ path, so, when it is possible to immunize a portfolio (in a sense I will define in section 5.3), the specification risk (or model risk) can be reduced.

5.3 Convexity

After the matching of B’s, the second step is to look for a suitable second order condition. Should it be possible to find a portfolio B-matched with a positive definite Hessian matrix

$$\sum_{i=1}^{k} \gamma_i G_i H_{G_i} = \prod_{i=1}^{k} \frac{\gamma_i G_i}{\Pi} = \prod_{i=1}^{k} w_i H_{G_i} = \Pi H_{\Pi},$$

then a local immunization would be in place.

Therefore my problem is now to find a row vector w of relative weights wᵢ solving the system

$$\begin{cases} wB = [0] \\ w1 = 1 \\ \sum_{i=1}^{k} w_i H_{G_i} \succ 0, \end{cases}$$

(5.7)

where the sign “$\succ 0$” is a shorthand for “positive definite”.

There is no guarantee that this problem admits solution and, when it does, there could be many solutions. Thus, after finding the set of all vectors $w \in \mathbb{R}^k$ that verify

$$\begin{cases} wB = [0] \\ w1 = 1, \end{cases}$$

(5.8)

it seems sensible to look for a weights vector that fulfils the positive definiteness property.⁹

6 Dealing with zcbs only

Owing to the proverb “first things first”, in the following sections I deal with zcbs only, as I hope to identify some first preliminary results, if any, that may be the starting points for generalizations. I recall that, dealing with zcbs, the

⁹ Obviously, should problem (5.7) admit many solutions, a further problem would arise: the choice of a somehow optimal w, a problem that I do not deal with here.
$i$-th row of the matrix $B$ is simply $B(\tau_i)$, just as in relation (5.4), so I may define

$$H_{G_i} = B'(\tau_i) B(\tau_i)$$

(6.1)

$$H_w = \sum_{i=1}^{k} w_i [H_{G_i}] = \sum_{i=1}^{k} w_i [B'(\tau_i) B(\tau_i)],$$

(6.2)

and problems (5.7) and (5.8) admit the special appearance

$$\begin{cases}
  wB = [0] \\
  w1 = 1 \\
  H_w = 0,
\end{cases}$$

(6.3)

$$\begin{cases}
  wB = [0] \\
  w1 = 1.
\end{cases}$$

(6.4)

### 6.1 On the column rank of $B$

Although in general false, the implication

$$\begin{cases} 
  M = M' \\
  x'Mx = 0 \\
  x \neq [0]
\end{cases} \Rightarrow \begin{cases} 
  Mx = [0] \\
  x \neq [0]
\end{cases}$$

holds for $M = H_{G_i}$, thanks to the definition of $H_{G_i}$ given by (6.1), so that it is

$$x'H_{G_i}x = x' [B'(\tau_i) B(\tau_i)]x = (x'B'(\tau_i))(B(\tau_i)x) = \|B(\tau_i)x\|^2.$$  

It is now plain that a vector $x \neq [0]$ exists which makes each $x'H_{G_i}x = 0$ if and only if each $(B(\tau_i)x)$ is null\footnote{The set of vectors making $x'H_{G_i}x = 0$ is a hyperplane $(x = [0]$ excepted). Now I am simply trying to exclude that all such sets collapse into one set.}:

$$\{x'H_{G_i}x = 0, \forall i\} \Leftrightarrow \{B(\tau_i)x = 0, \forall i\},$$

in other words if and only if the system

$$\begin{cases}
  Bx = [0] \\
  x \neq [0].
\end{cases}$$

admits a solution $x$. This is tantamount to state that the $n$ columns of $B$ are linearly dependent, i.e. that the rank of $B$ is less than $n$.

From now on I assume that $B$ has a full column rank, simply because I want to get rid not only of the case of a factor which does not affect any bond (as it happens when $B$ has a zero column), but also of the presence of any redundant factor, i.e. of a factor whose effects on bonds are a linear combination of the
effects of other factors. I also remark that $B$ must contain more ZCBs than factors (i.e. $k > n$), as with \( \text{rk}(B) = n = k \) the unique solution of \( wB = [0] \) is the zero vector, that obviously can not fulfill the sum constraint \( w1 = 1 \). Summing up, I assume

\[
\text{rk}(B) = n < k. \tag{6.5}
\]

### 6.2 The case of semipositive weights

Suppose that a vector \( w \geq [0] \) solving system (6.4) exists. Then \( H_w \) is at least positive semidefinite:

\[
\begin{align*}
wB &= [0] \\
w1 &= 1 \\
w &\geq [0]
\end{align*}
\]

(\( \geq 0 \) means “positive semidefinite”). Moreover, I get a positive definite \( H_w \) as soon as I can exclude the existence of vectors \( x \neq [0] \) making \( x'H_wx = 0 \), i.e. solving the system

\[
\begin{align*}
\sum_{i=1}^{k} w_i (x'H_{G_i}x) &= 0 \\
x &\neq [0].
\end{align*}
\]

Owing to signs of \( w \), the unsolvability of this system means that, although each \( H_{G_i} \) (being positive semidefinite only) admits some vector \( x \neq [0] \) leading to \( x'H_{G_i}x = 0 \), no vector \( x \neq [0] \) exists which makes each \( x'H_{G_i}x = 0 \). But in the last section I already dealt with this risk by simply assuming that \( B \) has a full column rank. I sum up all this stuff in the following theorem.

**Theorem 6.1** Assume that the matrix $B$ of order $k \times n$ has a full column rank and that the system

\[
\begin{align*}
wB &= [0] \\
w &\geq [0] \\
w1 &= 1,
\end{align*}
\]

admits a solution $w$. Then the matrix

\[
H_w = \sum_{i=1}^{k} w_i [B' (\tau_i) B (\tau_i)]
\]

is positive definite.

**Proof.** Each matrix $[B' (\tau_i) B (\tau_i)]$ is positive semidefinite by construction and $w$ is a semipositive vector, therefore $H_w \geq 0$. The assumption on the linear independence of the columns of $B$ ensures that no $x \neq [0]$ exists that makes null all the quadratic forms $x' (B' (\tau_i) B (\tau_i)) x$, $i = 1, \ldots, k$. Hence $H_w > 0$. \( \blacksquare \)
6.3 A global immunization result

The two simple conditions pointed out in the theorem above (there exists a weights vector \( w \) solving system (6.4) and \( \text{rk}(B) = n < k \)) afford a local immunization result for a portfolio of ZCBs. This is the contents of the following theorem:

**Theorem 6.2** Consider a portfolio of \( k \) ZCBs and \( n \) risk factors. Suppose that \( B \) has a full column rank and that the system

\[
\begin{align*}
& wB = [0] \\
& w \geq [0] \\
& w1 = 1
\end{align*}
\]

admits a solution \( w \). Then the portfolio is locally immunized.

**Proof.** Theorem 6.1 already states that the portfolio of \( k \) ZCBs is locally immunized, as each vector \( w \) solving the system above makes: i) the portfolio locally insensitive to the movements in risk factors, and ii) the Hessian \( H_w \) positive definite.

As the result that theorem 6.2 affords is rather nice, it is sensible to ask whether and how much its assumptions are heavy. The first condition (\( B \) has a full column rank) seems artless, as it simply avoids that the portfolio contains a redundant ZCB, i.e. whose reaction to risk factors is some linear combination of the reactions of the other \( (k - 1) \) ZCBs.

The second condition regards the solvability of system (6.7) and deserves a suitable treatment. On this matter the following theorem may be useful.

**Theorem 6.3** The system (6.7) is solvable if and only no vector \( z \in \mathbb{R}^n \) solves the system

\[
\begin{align*}
& Bz > [0] \quad \text{or} \quad Bz < [0] \\
& z \text{ sign unrestricted.}
\end{align*}
\]

**Proof.** System (6.7) may be written in the equivalent form

\[
\begin{align*}
& w [B1] = [0] [1] \\
& w \geq [0]
\end{align*}
\]

and the set of rules contained in [17] to generate alternative theorems for linear systems ensures that this last system is solvable if and only if no pair \((v, \alpha)\), with \( v \in \mathbb{R}^n \) and \( \alpha \in \mathbb{R} \), solves the system

\[
\begin{align*}
& Bv + \alpha1 \geq [0] \\
& \alpha \leq 0 \\
& v \text{ sign unrestricted} \\
& Bv + \alpha1 > [0], \text{ or } \alpha < 0,
\end{align*}
\]
or, which is the same, solves at least one of the systems

\[
\begin{cases}
  Bz > [0] \\
  z \text{ sign unrestricted,}
\end{cases}
\quad
\begin{cases}
  Bz \geq 1 \\
  z \text{ sign unrestricted.}
\end{cases}
\]

As the unsolvability of the first system entails the same property for the second one, system (6.7) admits a solution \( w \) if and only if system (6.8), version \( Bz > [0] \), admits no solution \( z \), whereas the version \( Bz < [0] \) directly follows from the absence of sign restrictions on \( z \).

Should the weights vector \( w \) be required (strictly) positive, then, starting from the solvability of the system

\[
\begin{cases}
  wB = [0] \\
  w > [0] \\
  w1 = 1,
\end{cases}
\]

instead of (6.7), the same proof path leads to the unsolvability of the system

\[
\begin{cases}
  Bz \geq [0] \quad (\text{or } Bz \leq [0]) \\
  z \text{ sign unrestricted.}
\end{cases}
\]

The results above, derived with linear algebra analysis, afford a local immunization strategy. Exceptionally, as the portfolio weights are constrained to be nonnegative, the same results become global, thanks to the fact that a \( w \geq [0] \) leads to a convex portfolio value. Indeed, this nice outcome may be obtained by means of the following lemma 2.1 by Balbás and Ibáñez [4] that I present below with some minor adjustments.

First of all let me remark that in [4] the set of possible shocks is represented by a subset \( K \) of the real function space defined on the time to maturity interval \([0, T]\). These functions represent the shocks that cause the perturbations in the term structure function. The convexity of \( K \) is the standard convexity of a set.

**Lemma 6.4 (Lemma 2.1 of [4])** Let the set \( K \) of possible shocks be convex and the asset pricing functions be positive and convex in the shocks. Let \( \mu_0 \geq 0 \) and indicate with \( \Pi_w \) the infimum value taken after a shock in interest rates by a portfolio whose relative weights are \( w \). Then there exists a portfolio \( w^* \geq [0] \) such that

\[
\Pi_{w^*} \geq \Pi_{w^*} \mu_0
\]

if and only if, for every possible shock, there is at least an asset \( i \) (this index depends on the shock) such that the corresponding ratio between the value \( G^+_i \) after and the value \( G_i \) before a shock satisfies the relation

\[
G^+_i \geq G_i \mu_0.
\]
When $\mu_0 = 1$, lemma 6.4 states that there exists a semipositive weight portfolio that does not lose value if and only if, for each possible shock, there exists an asset that does not lose value. In the framework outlined in this section, lemma 6.4 applies with $\mu_0 = 1$. Moreover, for ATSMs dealing with zcbs only, provided that the set $D$ is convex, then the assumptions of lemma 6.4 are fulfilled. In fact, for an ATSM the term structure function is (here the dependence on $X$ is emphasized)

$$Y (t, \tau, X) = \left( -A (\tau) \frac{1}{\tau} \right) + \left( \frac{1}{\tau} B (\tau) \right) X (t).$$

So a shock $\Delta Y$ may be written as

$$\Delta Y = Y (t, \tau, X (t) + \Delta X) - Y (t, \tau, X (t)) = -\frac{1}{\tau} B (\tau) \Delta X.$$

Therefore the set $K$ is defined as the set of functions

$$K = \left\{ \Delta Y = -\frac{1}{\tau} B (\tau) (X_1 - X_2), \; X_1, X_2 \in D \right\}$$

and, provided that $D$ is convex, $K$ is convex as well. In fact, for $0 \leq z \leq 1$, if $\Delta Y_1, \Delta Y_2 \in K$, there exist four possible realizations of the factor vector $X_1, X_2, X_3, X_4 \in D$ such that

$$k = z \Delta Y_1 + (1 - z) \Delta Y_2 = -\frac{1}{\tau} B (\tau) [z (X_1 - X_2) + (1 - z) (X_3 - X_4)] =$$

$$= -\frac{1}{\tau} B (\tau) [z X_1 + (1 - z) X_3 - (z X_2 + (1 - z) X_4)],$$

and, being $(z X_1 + (1 - z) X_3), (z X_2 + (1 - z) X_4) \in D$, the shock $k$ belongs to $K$, that is therefore convex, as lemma 6.4 requires. Moreover, the price function $P$ of a zcb is always positive and its convexity in the factors (and thus in their variations $\Delta X (t)$) has been already shown in section 5.1. The following corollary derived from lemma 6.4 says what is relevant for an ATSM dealing with zcbs only.

**Corollary 6.5** *In an ATSM a zcb portfolio with a semipositive weights vector can be globally immunized if and only if the system*

$$\left\{ \begin{array}{l}
\mathbf{B} z > [0] \\
z \text{ sign unrestricted.}
\end{array} \right. \quad (6.9)$$

*is unsolvable.*

**Proof.** I already know that the value of a zcb is convex in $\Delta X (t)$ (recall section 5.1). From lemma 6.4 I also know that a semipositive weights immunized portfolio exists if and only if every admissible shock $\Delta X (t)$ there is at least a zcb of price $P (t, \tau_i)$ (the index $i$ depends on the shock) such that

$$\exp \{ A (\tau_i) - B (\tau_i) [X (t) + \Delta X (t)] \} \geq \exp \{ A (\tau_i) - B (\tau_i) X (t) \},$$

19
\[ \forall \Delta X(t), \ \exists i : B(\tau_i) \Delta X(t) \leq 0, \]

or, equivalently,

\[ \exists \Delta X(t) : B(\tau_i) \Delta X(t) > 0, \ \forall i. \]

Writing this last property in matrix notation and substituting \( \Delta X(t) \) with a generic \( \mathbb{R}^n \) vector \( z \), I get the result: a ZCBs portfolio with a semipositive weights vector can be globally immunized if and only if the system (6.9) has to be unsolvable.

Corollary 6.5 shows that the results of theorems 6.2 and 6.3 are indeed global. This fact may arouse again some arbitrage flavour hidden in the immunization strategies. Indeed, in this case the local argument does not apply. A possible way out is proposed in [4]. The Authors interpret their result in this light: because the market is arbitrage free, there must exist a shock such that all assets lose value, i.e. that makes the system

\[
\left\{ \begin{array}{l}
Bz < [0] \\
z \text{ sign unrestricted}
\end{array} \right.
\]

solvable. But I remark that, owing to the freedom in signs of \( z \), this is equivalent to ask the system (6.8) to be solvable, i.e. that no immunization with semipositive weights is allowed. This is why I can not agree with the way out put forward in [4].

6.4 Its meaning

Now I deal with the two following problems:

- Is there any sensible meaning in the fact that no \( z \) makes \( Bz > [0] \)?
- What does the existence of a weights vector \( w \geq [0] \) making \( wB = [0] \) and \( w1 = 1 \) mean?

Although the theorem 6.3 ensures that these are two equivalent questions, their answers may have a different contents.

The unsolvability of system (6.8) is exactly the price to be paid in order to get an immunization result (such as theorem 6.2) by simply working on a first-order conditions only. The vector \( z \) in \( Bz \) may be viewed as a vector of increments \( \Delta X(t) \) in the risk factors values \( X(t) \), so \( Bz \) measures the total response of the portfolio to those factors\(^{11} \). Hence the unsolvability of system (6.8) means that no new determination of the risk factors exists that leads to a total response to risk factors with the same non zero sign for each bond. This conclusion is rather obvious: I can not hope to set up a portfolio insensitive to

\(^{11}\)Indeed, I could refer to the levels of the risk factors as well.
all changes in risk factors when all its elements react in the same direction. Remark that this trivial conclusion allows to rule out that a vector \( z \) exists which makes \( Bz > [0] \) as soon as \( B \) has a positive or a negative column.

The sign restriction on \( w \) has a plain meaning: no short sales are allowed, in other words the portfolio may contain assets only and no liability, as already remarked in [4]. But a portfolio without short sales (no liability) is indeed a too strong assumption, hence it is sensible to drop it off.

### 6.5 The case of unrestricted weights

Now I drop the sign restriction on the weights vector and I come back to the old problem of finding a solution \( w \) to the system (6.3), that here I rewrite for my case

\[
\begin{cases}
  wB = [0] \\
  w1 = 1 \\
  w \text{ sign unrestricted}
\end{cases}
\]

Following a special convention that will appear useful later, here I explicitly state that vector \( w \) is sign unrestricted. I have already justified in section 6.1 the assumptions (6.5) on the full column rank of \( B \) and on the inequality \( k > n \). As they do not depend on any sign restriction on the weights vector, here I still confirm them.

Maybe the first thing to do in order to manage problem (6.10) is to ensure the solvability of the system

\[
\begin{cases}
  wB = [0] \\
  w1 = 1 \\
  w \text{ sign unrestricted.}
\end{cases}
\]

Here is a theorem which directly faces this matter.

**Theorem 6.6** The system (6.11) is solvable if and only no sign unrestricted vector \( z \) gives a vector \( Bz \) whose entries have a common non zero value.

**Proof.** I write system (6.11) in the equivalent form

\[
\begin{cases}
  w \begin{bmatrix} B & 1 \end{bmatrix} = \begin{bmatrix} [0] & 1 \end{bmatrix} \\
  w \text{ sign unrestricted},
\end{cases}
\]

and I apply to it the same set of rules contained in [17] to generate alternative theorems for linear systems. They ensures that this last system is solvable if

\[\text{For an equivalent point of view see [4].}\]
and only if no pair \((z, \alpha)\), with \(z \in \mathbb{R}^n\) and \(\alpha \in \mathbb{R}\), solves the system

\[
\begin{aligned}
Bz &= \alpha 1 \\
z &\text{ sign unrestricted} \\
\alpha &> 0.
\end{aligned}
\]

Being \(z\) sign unrestricted, I can drop the sign requirement on \(\alpha\), so I get the system

\[
\begin{aligned}
Bz &= \alpha 1 \\
(z, \alpha) &\text{ sign unrestricted.}
\end{aligned}
\]

Moreover, thanks to the assumption on \(\text{rk}(B)\), I can rule out the case \(\alpha = 0\) so to state that system (6.11) is unsolvable if and only if a vector \(z\) exists such that all the entries of \(Bz\) have a common non zero value.

\[\blacksquare\]

### 6.6 The meaning of a necessary first order condition

Now I come back to the same questions already discussed in section 6.4. As vector \(w\) is now sign unrestricted, the constraint on no short sales drops and it only remains to find any meaning for the contents of theorem 6.6. As \(Bz\) measures the total response of the portfolio to a shift \(\Delta X (t)\) in the risk factors, the solvability of system (6.11) means that no shift of the risk factors exists that leads to exactly the same (non zero) common response for each bond: an obviously less binding condition than the one pointed out in theorem 6.3, where a response of a common sign for each bond was required.

Although theorem 6.6 settles the solvability of system (6.11) in a clear and quite meaningful way, the solvability of system (6.10) still remains an open problem, as the two examples below show.

The following example shows that problem (6.10) may have no solution in any way, even when system (6.11) is solvable.

**Example 6.7** Let

\[
B = \begin{bmatrix}
-1 & 1 \\
-2 & \frac{1}{2} \\
1 & 1
\end{bmatrix}.
\]

Then the unique vector that makes \(wB = 0\) and \(w1 = 1\) is

\[
w = \begin{bmatrix}
-\frac{5}{2} & 2 & \frac{3}{2}
\end{bmatrix},
\]

but the corresponding Hessian

\[
H_w = \begin{bmatrix}
7 & 2 \\
2 & \frac{1}{2}
\end{bmatrix}
\]

is indefinite, hence problem (6.10) admits no solution with this \(B\).
In the following example, instead, all runs smoothly.

**Example 6.8** Let

\[
B = \begin{bmatrix}
0.1 & 0.9 \\
0.5 & 0.19 \\
0.41 & 0.1 \\
0.7 & 0.059
\end{bmatrix}.
\]

Then, among the vectors making \( wB = [0] \) and \( w1 = 1 \), I may choose

\[
w = \frac{1}{15} \begin{bmatrix}
395.1 \\
-3079.1 \\
1574 \\
1221
\end{bmatrix}
\]

in order to get the corresponding Hessian

\[
H_w = \frac{1}{15} \begin{bmatrix}
32.3518 & -47.3314 \\
-47.3314 & 76.288597
\end{bmatrix}
\]

positive definite, hence a solution \( w \) of problem (6.10).

### 6.7 One factor immunization for a portfolio of zcbs

Now I will try to detect some immunization opportunities given an ATSM and a set of \( k \) zcbs, obviously with more bonds than factors, as condition (6.5) requires. I will work on the case \( n \leq 3 \) only, as the majority of ATSMs have up to 3 factors. The one factor ATSM models were the first studied in seventies since the fundamental work of Vasiček in 1977 [32]. The one factor ATSM model under the martingale measure \( Q \) is

\[
P(t, \tau) = \exp \{ A(\tau) - B(\tau) X(t) \},
\]

\[
dX(t) = K^Q (\Theta^Q - X(t)) \, dt + \Sigma \sqrt{\alpha + \beta X(t)} \, dW(t),
\]

where all parameter and functions are scalars.

### 6.8 One factor immunization with 2 zcbs

Let me begin with 2 zcbs whose prices are \( P_1 = P(t, \tau_1) \) and \( P_2 = P(t, \tau_2) \). This case of two assets and one risky factor is similar to the classical delta-hedging scheme. The portfolio value is

\[
\Pi = \gamma_1 P_1 + \gamma_2 P_2,
\]

and the matrix \( B \) is simply a column vector of \( \mathbb{R}^2 \), say

\[
B = \begin{bmatrix}
B(\tau_1) \\
B(\tau_2)
\end{bmatrix} = \begin{bmatrix}
b_1 \\
b_2
\end{bmatrix}.
\]
So, defining, as usual, \( w_i = \frac{2}{n} \), and employing the standard test for positive definiteness (all north-west minors are positive), the problem have this look:

\[
\begin{align*}
& w_1 b_1 + w_2 b_2 = 0 \\
& w_1 + w_2 = 1 \\
& w_1 (b_1)^2 + w_2 (b_2)^2 > 0.
\end{align*}
\] (6.12)

This problem is trivial and a bit of algebra is enough to fix it as in the following theorem.

**Theorem 6.9** A two ZCBs portfolio in one factor ATSM can be globally immunized if and only if the matrix

\[
B = \begin{bmatrix}
  b_1 \\
  b_2
\end{bmatrix}
\]

fulfills the condition

\[ b_1 b_2 < 0. \] (6.13)

**Proof.** With \( b_1 = b_2 \) the subsystem containing the first two equations in (6.12) leads to \( B = [0] \) (then the assumption \( \text{rk}(B) = n \) is violated) and the problem is unsolvable. With \( b_1 \neq b_2 \) the same equations gives

\[ w = \frac{1}{b_2 - b_1} \begin{bmatrix}
  b_2 \\
  -b_1
\end{bmatrix} \] (6.14)

and inserting this result into the last inequality in (6.12) directly gives the condition (6.13), that obviously entails the properties

\[ b_1 \neq 0, \ b_2 \neq 0, \ b_1 \neq b_2. \] (6.15)

Thanks to (6.13) and (6.14), the entries of vector \( w \) above have the same sign, that is positive in order to fulfill the sum constraint in (6.12). So corollary 6.5 ensures a global result.

The cases ruled out by relations (6.15) have no real interest, as a \( b_1 = 0 \) states that the \( i \)-th ZCB does not react to the risk factor, whereas the case \( b_1 = b_2 \) shows that the two bonds have exactly the same reaction, just as it happens with two ZCBs with the same maturity, hence with the same price.

Remark that the condition (6.13) can not be imposed a priori, simply because \( B \) stacks some values of the function \( B(\tau) \) that describes the link among interest rates. Generally its actual form is the output of an estimation or calibration process starting from data, therefore nothing ensures that \( B \) obeys that condition. Moreover, thanks to the continuity of the price of a ZCB (and hence of the function \( B(\tau) \)), it is always possible to choose two ZCBs whose times to maturity are close enough so as to not verify the assumption \( b_1 b_2 = B(\tau_1) B(\tau_2) < 0 \) of theorem 6.9. This means that, given a model, it is always possible to find two ZCBs that do not allow immunization. The conclusion is that, in general, immunization is not always possible with one risk factor and two ZCBs.
6.9 One factor immunization with 3 zcbs

In order to give portfolio immunization a better chance, I try to get a degree of freedom by considering, still with one risk factor, a three zcbs portfolio. Its value is

$$\Pi = \gamma_1 P_1 + \gamma_2 P_2 + \gamma_3 P_3;$$

now the matrix $B$ is simply a column vector of $\mathbb{R}^3$, say

$$B = \begin{bmatrix} B(\tau_1) \\ B(\tau_2) \\ B(\tau_3) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

and the assumption (6.5) requires that each $b_i$ is non zero and the entries have distinct values. Conditions (6.3) become

$$\begin{cases} w_1 b_1 + w_2 b_2 + w_3 b_3 = 0 \\ w_1 + w_2 + w_3 = 1 \\ w_1 (b_1)^2 + w_2 (b_2)^2 + w_3 (b_3)^2 > 0, \end{cases}$$

and they may be managed in the following theorem.

**Theorem 6.10** Consider a three zcbs portfolio in one factor ATSM with

$$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad b_i \neq b_j \text{ if } i \neq j,$$

$$b_i \neq 0, \forall i, \quad b_i \neq b_j \text{ if } i \neq j.$$ (6.17)

Then the portfolio can be locally immunized. All the weights vectors $w$ leading to this result are of the form

$$w_2 = \frac{w_1 (b_1 - b_3) + b_3}{b_3 - b_2}, \quad w_3 = \frac{w_1 (b_2 - b_1) - b_2}{b_3 - b_2},$$

with

$$w_1 \leq \frac{b_2 b_3}{(b_1 - b_3)(b_1 - b_2)},$$ (6.18)

according to the sign ($\geq 0$) of the quantity

$$\frac{b_2 b_3}{(b_1 - b_3)(b_1 - b_2)}.$$ (6.19)

**Proof.** Following the same way already employed for theorem 6.9, I write system (6.16) in the equivalent form

$$\begin{cases} w_2 = \frac{w_1 (b_1 - b_3) + b_3}{b_3 - b_2}, \quad w_3 = \frac{w_1 (b_2 - b_1) - b_2}{b_3 - b_2} \\ w_1 (b_1)^2 + \frac{w_1 (b_1 - b_3) + b_3}{b_3 - b_2} (b_2)^2 + \frac{w_1 (b_2 - b_1) - b_2}{b_3 - b_2} (b_3)^2 > 0. \end{cases}$$

25
The last inequality is equivalent to

\[ w_1 \text{sgn} [(b_1 - b_2) (b_3 - b_1)] < \frac{b_2 b_3 \text{sgn} [(b_1 - b_2) (b_3 - b_1)]}{(b_1 - b_3) (b_1 - b_2)}, \]

where the function \( \text{sgn} (x) \) means the sign of \( x \), i.e. \( \text{sgn} (x) = 1, 0, -1 \) according to \( x \geq 0 \). This solution is meaningful, as the assumption (6.17) ensures that \( \text{sgn} [(b_1 - b_2) (b_3 - b_1)] \neq 0 \), hence \( w_1 \) is defined by any value given by (6.18)–(6.19). So there are as many solutions of system (6.16) as many choices of \( w_1 \) values.

Maybe some remarks on the assumptions (6.17) may be useful. It is quite easy to show that my problem is unsolvable when the vector \( \mathbf{B} \) has all entries equal or has more than one zero entry. Although the cases of a null \( b_i \) and of all equal entries have no have real interest (see section 6.9), for completeness’ sake I may explore the following complementary cases (proofs are trivial, hence I omit them):

- if two \( b_i \) are equal, say \( b_1 = b_2 \), then a solution exists if and only if the inequality \( b_1 b_3 < 0 \) holds (it obviously implies \( b_1 \neq b_3 \)), just as in the case with two bonds, and all solutions are of the kind

\[
    w = \begin{bmatrix}
        (b_1 - b_1) w_2 + b_3 \\
        b_3 - b_1
    \end{bmatrix}
    \begin{bmatrix}
        w_2 \\
        b_1 - b_3
    \end{bmatrix};
\]

- if exactly one \( b_i \) is null, say \( b_1 = 0 \), then a solution exists if and only if \( b_2 \neq b_3 \), in which case all solutions may be set up as follows: choose any \( w_3 \) such that

\[
    \text{sgn} (w_3) = \text{sgn} [b_3 (b_4 - b_2)],
\]

then take

\[
    w = \begin{bmatrix}
        w_3 (b_3 - b_2) + b_2 \\
        b_3 - b_2
    \end{bmatrix}
    \begin{bmatrix}
        -w_3 b_2 \\
        b_3
    \end{bmatrix}.
\]

Theorem 6.10 shows that three \( \mathbf{ZCBs} \) are enough to immunize a portfolio against one risk factor, no matter what values the matrix \( \mathbf{B} \) takes (i.e. no matter what form the function \( B (\tau) \) takes and what maturity are selected for the three \( \mathbf{ZCBs} \), obviously provided that conditions (6.17) hold. This fact, together with theorem 6.9, shows that 3 is exactly the minimum number of \( \mathbf{ZCBs} \) to obtain this result. Moreover, it is easy to describe the whole set of the relative weights \( w_1 \) to be chosen in order to get immunized portfolios.

It is important to remark that the choice of the weights does not depend upon the factor actual value. Then a portfolio of \( \mathbf{ZCBs} \) with constant times to maturity \( \tau \) need to be readjusted to keep unchanged both those times and the relative weights of assets as time goes by. Obviously, whether and how to use the freedom degree afforded by theorem 6.10 in choosing \( w_1 \) is an open problem.
6.10 An example

Now consider an example taken from a one factor CIR model. Given the real world dynamics of the (short rate) factor (with \( \sigma = \Sigma \sqrt{\beta} \) and \( r(t) = X(t) \))
\[
dX(t) = \mathcal{K} (\Theta - X(t)) \, dt + \sigma \sqrt{X(t)} \, dW^P(t),
\]
and the price of risk \( \Lambda = \ell \sqrt{X(t)} \),\(^{13} \) the corresponding risk neutral dynamics is
\[
dX(t) = \mathcal{K}^Q (\Theta^Q - X(t)) \, dt + \sigma \sqrt{X(t)} \, dW^Q(t),
\]
with
\[
\mathcal{K}^Q = \mathcal{K} + \ell, \quad \Theta^Q = \frac{\mathcal{K} \Theta}{\mathcal{K} + \ell}
\]

In this case the pricing function is (see \([13]\))
\[
P(t, \tau) = \exp \{ A(\tau) - B(\tau) X(t) \}
\]
where\(^{14} \)
\[
A(\tau) = 2 \frac{\mathcal{K}^Q \Theta^Q}{\sigma^2} \ln \left[ \frac{2 \gamma \exp \left\{ \frac{(\mathcal{K}^Q + \gamma)^2}{2} \right\}}{(\mathcal{K}^Q + \gamma)(e^{\gamma \tau} - 1) + 2 \gamma} \right],
\]
\[
B(\tau) = \frac{2 (e^{\gamma \tau} - 1)}{(\mathcal{K}^Q + \gamma)(e^{\gamma \tau} - 1) + 2 \gamma}
\]
and
\[
\gamma = \sqrt{(\mathcal{K}^Q)^2 + 2 \sigma^2}.
\]

In order to build up an example of an immunized portfolio in this setting, I assign the parameters the values estimated in \([25]\) from U.S. Treasury bills during 1964–89:
\[
\Theta = 0.0154, \quad \mathcal{K} = 12.43, \quad \ell = -6.08, \quad \sigma = 0.49,
\]
and then
\[
\mathcal{K}^Q = 6.35, \quad \Theta^Q = 0.03014519685.
\]

\(^{13}\)This form of risk premium is equivalent to the completely affine parametrization. In fact \( \Lambda \) is proportional to the volatility
\[
\Lambda = \frac{\ell}{\sigma^2} \sqrt{X(t)}.
\]
Thus in the standard notation of ATSMs the coefficient should be \( \lambda = \frac{\ell}{\sigma^2} \). Here the parameter \( \ell \) is what is called \( \lambda \) by CIR \([13]\) and Gibbons and Ramaswamy \([25]\).

\(^{14}\)Sorry, here \( \gamma \) is not the quantity of an asset in a portfolio. Rather I use \( \gamma \) because it is standard to call it in such a way since the original work of CIR \([13]\).
The functions $A$ and $B$ become

\[
A(\tau) = \frac{1.594519 \ln 12.7753982}{12.7376991 \exp (6.3876991\tau) + 0.0376991} \exp (6.3688496\tau)
\]

\[
B(\tau) = 2 \frac{\exp (6.3876991\tau) - 1}{12.7376991 \exp (6.3876991\tau) + 0.0376991}
\]

Now I consider a portfolio composed by zcbs of maturities 1 month, 3 months and one year\textsuperscript{15}. The function $B(\tau)$ evaluated at these maturities takes values leading to the matrix

\[
\mathbf{B} = \begin{bmatrix}
    b_1 \\
    b_2 \\
    b_3 
\end{bmatrix} = \begin{bmatrix}
    B\left(\frac{1}{12}\right) \\
    B(0.25) \\
    B(1)
\end{bmatrix} = \begin{bmatrix}
    0.06469551 \\
    0.12514096 \\
    0.15674933
\end{bmatrix}
\]

From theorem 6.10 I check the sign of

\[
B^2\left(\frac{1}{12}\right) - B\left(\frac{1}{12}\right) B(0.25) - B(0.25) B(1) + B(0.25) B(1) = 0.0055642 > 0,
\]

then with a weight $w_1$ for 1 month zcb such that

\[
w_1 > 3.5253299
\]

holds, the portfolio is locally immunized.

Let me remark that this result, although theoretically correct, may clash with common sense. In fact the weights of an immunized portfolio could be, for example,

\[
w_1 = 3.6, \quad w_2 = -5.5252582, \quad w_3 = 2.9252582,
\]

that is, in order to immunize a 10 000 € net value portfolio it is necessary to take a very large short position, namely of 55 252.58 €, in the 6 months zcb.

The similar exercise considering 3 months, 6 months and one year zcbs produces a remarkably more disappointing result. In fact

\[
B(0.25) = 0.12514096, \quad B(0.5) = 0.1505562, \quad B(1) = 0.15674933.
\]

From theorem 6.10 I check the sign of

\[
B^2(0.25) - B(0.25) B(0.5) - B(0.25) B(1) + B(0.5) B(1) = 0.00080334 > 0,
\]

and with a weight of 3 months zcb such that

\[
w_1 > 29.377039
\]

\textsuperscript{15}I choose these maturities because in [25] only the zcbs up to 12 months are considered.
holds, the portfolio is immunized, as with the choice

\[ w_1 = 30, \quad w_2 = -127.80319, \quad w_3 = 98.80319, \]

that is, in order to immunize a 10'000€ net value portfolio one must take a short position of 1'278 031.90€.

A word of caution is worth about the nature of only local result I have got. Indeed, in the example above the matrix \( B \) makes system (6.8) solvable, so the corollary 6.5 can not be applied, hence a global result is not ensured. In fact, the figure 1, plotting the value of the portfolio value after a shock in the factor against the shock \( \Delta X \) (starting from \( X = 0.5 \)), shows that an only local result is obtained.

### 6.11 Multifactor immunization

The results put forward in section 6.9 (see theorem 6.10) allow to state that a portfolio of 3 zcbs with unrestricted weights always may be immunized against one risk factor as soon a special but artless condition is fulfilled (all bonds do react to that factor with different reaction patterns). So it is rather obvious to hope that somehow similar results may be found for the case of many factors, at least under some meaningful (i.e. non only formal) conditions on the matrix \( B \). Moreover, as the existing tractable models has commonly up to 3 factors, my conjecture may be refered to the cases of 2 and 3 factors only.

In the examples 6.7 and 6.8 I have already shown that things may go wrong even with two factors. Moreover, on one side it is quite easy to find a sensible example leading to no immunization result with 3 factors, but on the other I have not yet found a non trivial example leading to results of such a kind. The following section shows how to reach this disappoint outcome.
6.12 Two factors immunization with 3 zcbs

Let the treasury market be driven by 2 stochastic factors. The assumptions (6.5) requires \(\text{rk}(\mathbf{B}) = 2 < k\). I do not begin with 3 zcbs as this way there is no degree of freedom to be spent in order to make the portfolio immunized. Just like in the one factor case (see theorem 6.10), at least \((n + 2)\) assets are needed. So I deal with a portfolio composed by 4 zcbs. Their prices are \(P_1 = P(t, \tau_1)\), \(P_2 = P(t, \tau_2)\), \(P_3 = P(t, \tau_3)\) and \(P_4 = P(t, \tau_4)\), so the portfolio value is

\[
\Pi = \gamma_1 P_1 + \gamma_2 P_2 + \gamma_3 P_3 + \gamma_4 P_4,
\]

and conditions (6.3) become

\[
\begin{bmatrix}
  w_1 & w_2 & w_3 & w_4 \\
\end{bmatrix}
\begin{bmatrix}
  B_1(\tau_1) & B_2(\tau_1) & 1 \\
  B_1(\tau_2) & B_2(\tau_2) & 1 \\
  B_1(\tau_3) & B_2(\tau_3) & 1 \\
  B_1(\tau_4) & B_2(\tau_4) & 1 \\
\end{bmatrix}
= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}
\]

\[
H_w = \sum_{i=1}^4 w_i [B' (\tau_i) B (\tau_i)] > 0.
\]

I suppose that:

- no unconstrained vector \(z\) exists which leads a vector \(\mathbf{B}z\) whose all entries have a common non zero value;
- there exists an unconstrained vector \(z\) leading to a positive \(\mathbf{B}z\).

As already stated in theorems 6.3 and 6.6, the first assumption ensures the existence of at least one weights vector \(w\) fulfilling the first order condition for immunization, whereas the second one rules out the case of a semipositive \(w\). Now I solve the first system in (6.20). If one solution appears, then check whether the corresponding Hessian is positive definite, i.e. whether the 2 north-west principal minors of

\[
H_w = \sum_{i=1}^k w_i [B' (\tau_i) B (\tau_i)] = \begin{bmatrix}
(H_w)_{11} & (H_w)_{12} \\
(H_w)_{12} & (H_w)_{22} \\
\end{bmatrix}
\]

are positive:

\[
\begin{cases}
(H_w)_{11} > 0 \\
(H_w)_{11} (H_w)_{22} - [(H_w)_{12}]^2 > 0.
\end{cases}
\tag{6.20}
\]

If many solutions appear, I may give their formal description by writing, say, \(w_2, w_3, \text{ and } w_4\) as linear functions of \(w_1\). The first inequality in system (6.20) is linear (affine) in \(w_1\), the second one is quadratic (complete) in \(w_1\). But now there is no guarantee that a solution for the whole system exists. In fact the following example does not allow immunization.
Example 6.11 Let matrix \( B \) be as follows:

\[
\begin{bmatrix}
B_1(\tau_1) & B_2(\tau_1) \\
B_1(\tau_2) & B_2(\tau_2) \\
B_1(\tau_3) & B_2(\tau_3) \\
B_1(\tau_4) & B_2(\tau_4)
\end{bmatrix}
= \begin{bmatrix}
-0.07 & 0.15 \\
0.04 & 0.3 \\
0.59 & 0.55 \\
0.82 & 0.75
\end{bmatrix}.
\]

Then the test (6.20) means

\[
\begin{align*}
0.2605429w_1 - 0.5532571 & > 0 \\
-0.0028286w_1^2 + 0.002619w_1 + 0.0054206 & > 0,
\end{align*}
\]

but there is no \( w_1 \) verifying both inequalities.

I may continue this example by applying here the same idea as in one factor case: in order to gain a degree of freedom, I may add another asset, i.e. a fifth ZCB, so getting a portfolio that nests the one in the previous example. For instance, adding to \( B \) the corresponding row made with \( B_1(\tau_5) = 0.83 \) and \( B_2(\tau_5) = 0.77 \), then the test (6.20) becomes

\[
\begin{align*}
(H_w)_{11} &= 0.2605429w_1 - 0.5532571 - 0.0188457w_5 \\
(H_w)_{11}(H_w)_{22} - [(H_w)_{21}]^2 &= -0.0282857w_1^2 + 0.002619w_1 + \\
&-0.00011918w_1w_5 + 0.0054206 + 0.001568w_5 - 0.00003353w_5^2.
\end{align*}
\]

Again, \((H_w)_{11}\) is linear in the subvector \([w_1 \quad w_5]\) (the weights left free by the first order condition), and \(\det (H_w)\) is quadratic in the same. The first one is positive for \(w_1 > 2.1234779 + 0.072332426w_5\), but unfortunately in this region the determinant of \(H_w\) is not positive.

I conjecture that with the matrix \( B \) above it is not possible to immunize a ZCBs portfolio (at least up to 5 assets\(^{16}\)). To be frank, my suspicion is that the technique of getting additional freedom degrees by simply adding a sensible number of other assets (no matter of the form of the function \( B(\tau) \)) works no more with \( n \geq 2 \), i.e. when my problem loses the linearity it enjoyed with one factor. In fact, as soon as I am aware, this procedure do not guarantee the existence of a region of the space of portfolio weights \( R^k \) where the linear and the quadratic functions defining the northwest minors of the Hessian matrix are both positive. The fact that the coefficients of the two functions are linked (all they come from the \( B(\tau) \) function) does not seem to help much. Moreover, nothing changes when I write the relevant inequalities in terms of the minors of the matrix \( B \). The same conclusions obviously apply to 3 or more factor models.

In fact, the dimension of the state variable vector determines the degree of the polynomials giving the northwest minors of \( H_w \) as functions of the weights \( w_i \) that the first order condition leaves free.

The relevant point is that, although it is tedious but possible to test whether a particular matrix \( B \) of low order (i.e. with \( n \leq 3 \)) leads to immunization

\(^{16}\)I have tried even with 6 assets, but without success.
results, finding general and financially meaningful conditions ensuring such a result for a generic matrix $B$ is still an open and intrinsically difficult problem when a mixed sign weights vector $w$ is needed. Moreover, in one factor case there exists a (minimum) number of assets that assures the feasibility of immunization, whereas in the multifactor setting, I have not found such a matter.

7 Conclusions

In the previous section I tried to extend classical immunization results to an affine factor model with one or more risk factors, dealing, for simplicity's sake, in a world with ZCBs only with price defined as

$$P(t, \tau) = \exp \{ A(\tau) - B(\tau) X(t) \}.$$

Let me sum up the main outcomes I got:

1) There are not heavy problems in finding a critical portfolio weights vector $w$, i.e. that fulfills the necessary first order conditions in order to get immunization results. Under an artless assumption on the column rank of $B$ and on a positive spread between the number of ZCBs and the number of factors, namely inequalities (6.5), the theorem 6.6 identifies a financially meaningful condition on the matrix $B$ that is equivalent to the existence of such a portfolio.

2) Then two complementary cases may be pointed out. In the first one a semipositive critical vector $w$ exists, whereas in the second one only critical vectors with mixed signs exist. Which of the two cases occurs depends on the unsolvability of system (6.8): again a condition whose sensible financial meaning I found.

3) In the former case the relevant second order conditions are automatically fulfilled, whereas in the latter the positive definiteness of the relevant Hessian matrix is explicitly needed. This property can be managed in an easy and satisfactory way (see theorem 6.10) when the model regards one risk factor: three ZCBs are needed as soon as I rule out all redundant special cases (bonds that do not react and/or exactly react in the same way). Unfortunately, with more than one factor all may happen: some cases lead to immunization results and some other do not. Moreover, a set of formal second order conditions may be found in a trivial and standard way, but these rules keep strictly formal, i.e. they may be checked for a special choice of the matrix $B$, but I was not able to give them a substantial and clear meaning from the financial point of view.\footnote{This lack of explicit meaning, together with the tedious way to obtain those rules, are the reasons why I did not present them in this work.}

4) Accepting the idea that immunization is inconsistent with the absence of arbitrage opportunities (see section 4 for a discussion on this point) implies that one factor ATSMs are not a proper tool to analyze treasury market. So, in order to eliminate these opportunities, a multifactor setting is needed. This can be another (I think poor) supporting argument in favor of multifactor models.

5) Obviously, I should avoid to throw the baby out with the bathwater. Indeed, I can recover some things from the wreckage, as two results still stand: the
immunization result when a semipositive critical weights vector exists, and the fact that, even in the worst case, first order conditions (5.5) keep an autonomous relevant interest in hedging interest rate risk in order to reduce variability of portfolio returns by means of classical hedging techniques, as in [6], [12] and [26].

7.1 A promising way out

Here I try to focus an idea that seems a promising way out in order to overcome some disappointing results I met on the way for immunization outcomes.

Till now I was seeking for conditions ensuring a minimum value position for a portfolio of zcb’s by simply employing a standard unconstrained minimization technique. Instead, now I may work as follows:

- let the variations \( \Delta X = \Delta X(t) \) in risk factors stay in a closed convex subset, say \( \mathcal{X} \), of \( D \);
- let the change \( \Delta \Pi \) in the portfolio value \( \Pi \) be approximated by the second order Taylor expansion
  \[
  \Delta \Pi \approx \nabla_X (\Pi) (\Delta X) + (\Delta X)^T H_X (\Pi) (\Delta X), \quad \Delta X \in \mathcal{X};
  \]
- choose the portfolio weights vector \( w \) which leads to minimize the worst loss coming from a shift \( \Delta X \); in other words, solve in \( w \) the problem
  \[
  \max_{w_{n=1}} \left( \min_{\Delta X \in \mathcal{X}} (\Delta \Pi) \right).
  \]

Obviously, owing to the local approximation (7.1) and to the choice of \( \mathcal{X} \) for the variations in risk factors, also this framework can lead to local results only. Also note that similar special maximin strategies for bond portfolios are analyzed in Balbás and Ibáñez [4]. Although it contains no room neither for an integrative portfolio with bid-ask spread, nor for transactions costs, the problem (7.2) recalls the platinum rule introduced in Wilmott [33] for a derivatives portfolio, a rule employed both in Crash Metrics® as a central tool and in the Lagrangian uncertainty volatility model by Avellaneda et al. [1], [2] and [3]. This interesting (and rather nonorthodox) approach is well suited for this situation for at least two reasons:

- First, it is a technique born with the aim of hedging a portfolio against unexpected shocks of uncertain timing and magnitude.
- Second, the philosophy behind the application of this rule is rather pragmatic: here the approximation of the changes in the portfolio value is derived within a well defined and widely accepted model (i.e.: the affine pricing formula in these pages, and the Black-Scholes derivative formulae in Crash Metrics®), but nobody can blindly rely on it, as one always should expect a possible crash in factors.

33
Maybe this proposal deserves attention because this procedure, and the approach of this section as well, explicitly tries to account for model risk and mis specification.

As the reader can easily see, this dissertation has some lacks and leaves some questions without a fixed answer. Moreover, there are some subjects that are only sketched out.

Let me enumerate some items that need a deeper investigation and that could be the topics of further research.

1. The analysis dealt with in section 4 put forward some immunization strategies that should be empirically corroborated.

2. In section 4 I have introduced the immunization analysis for zcb's portfolios. This affords the instrument leading to extend the analysis to more complex fixed income assets.

3. Because of the difficulties in dealing with multifactor second order conditions (see section 6.11), an investigation of the immunization improvement discussed in section 7.1 should be an interesting subject.

References


List of the lately published Technical Reports
(available at the web site: "http://economia.unipv.it/Eco-Pol/quaderni.htm").

### Quaderni di Dipartimento

<table>
<thead>
<tr>
<th>#</th>
<th>Date</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>124</td>
<td>01-01</td>
<td>P.Bertoletti</td>
<td>Why regulate prices? Some notes on the price cap methods</td>
</tr>
<tr>
<td>125</td>
<td>02-01</td>
<td>C.Castagnetti</td>
<td>Estimating the risk premium for swap spreads Two econometric GARCH-based techniques</td>
</tr>
<tr>
<td>126</td>
<td>02-01</td>
<td>M.Colombo, M.Delmaestro</td>
<td>The Determinants of Structural Inertia: Technological and Organizational Factors</td>
</tr>
<tr>
<td>127</td>
<td>02-01</td>
<td>M.Colombo, M.Delmaestro</td>
<td>How effective are Technology Incubators? Evidence from Italy.</td>
</tr>
<tr>
<td>128</td>
<td>02-01</td>
<td>E.Regazzini, A.Lijoi, I.Pruenster</td>
<td>Distributional results for means of normalized random measures with independent increments</td>
</tr>
<tr>
<td>129</td>
<td>02-01</td>
<td>A.Lijoi</td>
<td>Approximating priors by finite mixtures of conjugate distributions for and exponential family</td>
</tr>
<tr>
<td>130</td>
<td>02-01</td>
<td>R.Lucchetti, E.Rossi</td>
<td>Artificial Regression Testing in the GARCH-in-mean model</td>
</tr>
<tr>
<td>131</td>
<td>02-01</td>
<td>D.Sonedda</td>
<td>Employment Effects of Progressive Taxation in a Unionised Economy</td>
</tr>
<tr>
<td>132</td>
<td>04-01</td>
<td>P.L.Conti, A.Lijoi, F. Ruggeri</td>
<td>A Bayesian approach to the analysis of telecommunications systems performance</td>
</tr>
<tr>
<td>133</td>
<td>04-01</td>
<td>C. Bianchi</td>
<td>A Reappraisal of Verdoon’s Law for the Italian Economy: 1951-1997</td>
</tr>
<tr>
<td>134</td>
<td>09-01</td>
<td>A.Roverato, G.Consonni</td>
<td>Compatible Prior Distributions for DAG Models</td>
</tr>
<tr>
<td>135</td>
<td>11-01</td>
<td>D.Sonedda</td>
<td>On the dynamics of unemployment and labour tax progression</td>
</tr>
<tr>
<td>136</td>
<td>11-01</td>
<td>F.Chelli, L.Rosti</td>
<td>Gender Discrimination, Entrepreneurial Talent and Self-Employment in Italy ‘If you think you’re so discriminated against, why don’t you set up on your own?’</td>
</tr>
<tr>
<td>137</td>
<td>12-01</td>
<td>L. E. Nieto-Barajas, I. Pruenster, S.G. Walker</td>
<td>Normalized Random Measures driven by Increasing Additive Processes</td>
</tr>
<tr>
<td>138</td>
<td>02-02</td>
<td>A.Lijoi, E.Regazzini</td>
<td>Means of a Dirichlet process and multiple hypergeometric functions</td>
</tr>
</tbody>
</table>