Bond portfolio optimization: a dynamic heteroskedastic factor model approach

João F. Caldeira\textsuperscript{a}, Guilherme V. Moura\textsuperscript{b, *}, André A. P. Santos\textsuperscript{b}

\textsuperscript{a}Department of Economics  
Universidade Federal do Rio Grande do Sul  
\textsuperscript{b}Department of Economics  
Universidade Federal de Santa Catarina

Abstract

In this paper we use Markowitz’s approach to optimize bond portfolios. We derive closed form expressions for the vector of expected bond returns and for their conditional covariance matrix based on a general class of dynamic heteroskedastic factor models. These estimators are then used as inputs to obtain mean-variance and minimum variance optimal bond portfolios. An empirical application using the dynamic version of the Nelson-Siegel and Svensson yield curve factor models is carried out involving a data set of 14 contracts with different maturities. The results indicate that the optimized bond portfolios deliver an improved risk-adjusted performance in comparison to several benchmarks.

Keywords: fixed-income, portfolio optimization, yield curve, dynamic conditional correlation (DCC), forecast.

JEL C53, E43, G17

*Corresponding author. Tel. +55 48 3721-6652; fax +55 48 3721-9585. E-mail: guilherme.moura@ufsc.br
1. Introduction

The portfolio optimization approach proposed by Markowitz (1952) based on the mean-variance relationship among assets is one of the milestones of modern finance theory. In this framework, individuals choose their allocations in risky assets based on the trade-off between expected return and risk. This approach is widely used to support managers in portfolio construction and to develop quantitative asset allocation strategies (Cornuejols and Tütüncü, 2007). However, in the vast majority of the cases these applications are restricted to the construction of equity portfolios; see, for instance, Brandt (2009), DeMiguel and Nogales (2009) and DeMiguel et al. (2009a) for recent applications. Thus, much is known about the strengths, weaknesses and performance of optimized equity portfolios using the mean-variance approach, but little is known about optimal portfolios composed of fixed-income securities.

The literature shows few references suggesting the use of mean-variance approach to bond portfolio selection (see, for instance, Wilhelm, 1992; Korn and Koziol, 2006; Puhle, 2008). In practice, fixed-income portfolios tend to be selected in order to approximate the duration of a benchmark or to replicate the performance of this benchmark in terms of return and volatility (Fabozzi and Fong, 1994). In this sense, managers’ opinions about the evolution of the term structure of interest rates will determine to a large extent the composition of the fixed-income portfolio. Quantitative models for the term structure of interest rates, such as those proposed by Vasicek (1977), Cox et al. (1985), Nelson and Siegel (1987), Svensson (1994), Hull and White (1990), Heath et al. (1992), tend to be applied only to risk management or to price fixed-income derivatives, but not directly to portfolio selection.

At least two reasons can be pointed out in order to justify the lack of mean-variance optimization applied to fixed-income portfolios. The first reason is the relative stability and low historical volatility of this asset class, which discouraged the use of sophisticated methods to exploit the risk-return trade-off in fixed-income assets. However, this situation has been changing rapidly in recent years, even in markets where these assets have low default probability (Korn and Koziol, 2006). The recurrence of turbulent episodes in global markets usually brings high volatility to bond prices, which increases the importance of adopting portfolio selection approaches that, on one hand, take into account the risk-return trade-off in bond returns and, on the other hand, allow for the possibility of risk diversification across different maturities.

The second reason refers to the difficulties in modeling bond returns and the covariance matrix of bond returns (Wilhelm, 1992; Korn and Koziol, 2006; Puhle, 2008). Fabozzi and
Fong (1994) argue that if a covariance matrix of bond returns is available, the process of portfolio optimization using fixed-income securities is similar to that of equity portfolios. However, one should bear in mind that fixed-income securities have finite maturities and promise to pay face value at maturity. In this sense, the end of the year price of a bond with two years to maturity is indeed a random variable. But the price of that same bond in two years is a deterministic quantity (disregarding the risk of default) given by its face value. This implies that the statistical properties of price and return of a fixed-income security depends on their maturity. Thus, both bond price and bond return are non-ergodic processes, and therefore traditional statistical techniques cannot be used to directly model the expected return and volatility of these assets (Meucci, 2009).

In this paper, we put forward a novel approach to obtain closed-form estimators for the vector of expected bond returns and for the covariance matrix of bond returns. These estimators are then used as inputs to obtain optimal bond portfolios based on the mean-variance framework. The proposed approach is based on heteroskedastic dynamic factor models, as proposed by Santos and Moura (2011) for the case of equity portfolios, but now applied to the term structure of interest rates. According to Korn and Koziol (2006), the great advantage of using factor models for the term structure of interest rates is the possibility to model constant (or fixed) maturity yields. This allows the estimation of the conditional distribution of bond yields without relying on bond maturities (Meucci, 2009). To accomplish this, we use well established dynamic factor models for the term structure such as the dynamic version of the Nelson-Siegel model proposed by Diebold and Li (2006), and the four factor model proposed by Svensson (1994).

It is also worth noting that our approach to bond portfolio optimization differs in several respects to the few existing ones. Wilhelm (1992), Korn and Koziol (2006) and Puhle (2008), for instance, propose the use of the mean-variance paradigm to the selection of bond portfolios using the Vasicek (1977) model for the yield curve. However, it is well known that the single factor Vasicek model has limited forecasting power (see, for example, Duffee, 2002). In contrast, the approach proposed here is based on dynamic factor models which, besides allowing the derivation of closed-form moments for bond returns, are widely used to successfully forecast the yield curve (see, for example, BIS, 2005).

Furthermore, Wilhelm (1992), Korn and Koziol (2006) and Puhle (2008) consider only

---

1 See, for instance, De Pooter (2007), Diebold and Rudebusch (2011), Almeida et al. (2008), Caldeira et al. (2010b), and Rezende and Ferreira (2011) for an analysis of the predictive performance of factor models for the term structure.
homoscedastic factor models, thus ignoring the persistence in the volatility of bonds returns. In this paper, however, the conditional heteroscedasticity is explicitly modeled, and the mean-variance optimization are based on estimates not only for the vector of expected bond returns, but also on estimates of the conditional covariance matrix of bonds returns. To this end, we consider a parsimonious multivariate GARCH specification that allows for the estimation and forecast of conditional covariance matrices for high dimensional problems.

We provide an empirical application involving a data set of daily closing prices of constant-maturity future contracts of the Brazilian Inter Bank Deposit Future Contract (ID-future). More specifically, we use ID-future contracts traded on the Brazilian Mercantile and Futures Exchange (BM&F) with fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 42 and 48 months. Based on the estimates for the vector of expected returns of these assets and their conditional covariance matrix, we obtain optimal mean-variance and minimum variance portfolios and evaluate their out-of-sample performance in comparison to benchmarks used in the Brazilian fixed-income market. The results indicate that the proposed approach generates optimal mean-variance and minimum variance portfolios with improved risk-adjusted performance when compared to the benchmarks. Moreover, the results are shown to be robust against i) alternative econometric specifications used to model the yield curve, ii) alternative econometric specifications for the factor dynamics and iii) alternative re-balancing frequencies of the optimized portfolios.

The paper is organized as follows. Section 2 describes the factor models used for modeling the term structure, as well as the econometric specification for the conditional heteroscedasticity of the factors. Section 3 discusses an estimation procedure in multi-steps for the proposed model. Section 4 discusses the Markowitz mean-variance optimization for fixed-income portfolio, and bring an empirical application. Finally, Section 5 brings the final considerations.

2. Bond portfolio optimization using yield curve models

In this section we consider the use of factor models for the yield curve to perform bond portfolio optimization according to the mean-variance approach proposed by Markowitz (1952). Factor models for the term structure of interest rates allow us to obtain closed form expressions for the expected yields, as well as for their conditional covariance matrix. From these moments, we shown how to compute the distribution of bond prices and bond returns, which will later be used as an input for portfolio optimization.
2.1. Dynamic factor models for the yield curve

Dynamic factor models play a major role in econometrics, allowing the explanation a large set of time series in terms of a small number of unobserved common factors (see, for instance, Fama and French, 1993; Stock and Watson, 2002). In general, the relationship between the time series and the common factors is assumed to be linear, and the weights of the individual factors are usually refereed to as factor loadings. In many applications in which dynamic factor models are considered, the individual series have a natural ordering in terms of one or more variables or indicators. For instance, if we consider a set of time series of bond yields, we can order the yields according to the time to maturity of each bond. Moreover, for this particular type of data set, it is usually assumed that the loads for a specific factor are a relatively smooth function of the variable that is used to order the time series. Many specifications for the yield curve can be viewed as dynamic factor models with a set of restrictions imposed on the factor loadings.

We consider a set of time series of bond yields with $N$ different maturities, $\tau_1, \ldots, \tau_N$. The yield at time $t$ of a security with maturity $\tau_i$ is denoted by $y_t(\tau_i)$ for $t = 1, \ldots, T$. The $N \times 1$ vector of all yields at time $t$ is given by

$$y_t = \begin{bmatrix} y_t(\tau_1) \\ \vdots \\ y_t(\tau_N) \end{bmatrix}, \quad t = 1, \ldots, T.$$  

The general specification of the dynamic factor model is given by

$$y_t = \Lambda(\lambda)f_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, \Sigma_t), \quad t = 1, \ldots, T, \quad (1)$$

where $\Lambda(\lambda)$ is the $N \times K$ matrix of factor loadings, $f_t$ is a $K$-dimensional stochastic process, $\varepsilon_t$ is the $N \times 1$ vector of disturbances and $\Sigma_t$ is an $N \times N$ conditional covariance matrix of the disturbances. We restrict the covariance matrix $\Sigma_t$ to be diagonal. This means that the covariance between the yields with different maturities is explained solely by the common latent factor $f_t$. The dynamic factors $f_t$ are modeled by the following stochastic process

$$f_t = \mu + \Upsilon f_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \Omega_t), \quad t = 1, \ldots, T, \quad (2)$$

where $\mu$ is a $K \times 1$ vector of constants, $\Upsilon$ is the $K \times K$ transition matrix, and $\Omega_t$ is the conditional covariance matrix of disturbance vector $\eta_t$, which are independent of the residuals $\varepsilon_t$ $\forall t$. 

5
The specification for \( f_t \) is general. Therefore, it is possible to model its dynamics using a variety of processes (see Jungbacker and Koopman, 2008). In modeling yield curves the usual specification for \( f_t \) is a vector autoregressive process of lag order 1 (Diebold et al., 2006; Caldeira et al., 2010b).

Next we present the factor models for the yield curve considered in this paper. The specifications considered are the two main variants of the original formulation of the Nelson and Siegel (1987) factor model, namely the dynamic Nelson-Siegel model proposed by Diebold and Li (2006), and the extension proposed by Svensson (1994). The alternative Nelson-Siegel specifications considered are all nested and can therefore be captured in the general formulation in (1) and (2) with different restrictions imposed on the loading matrix \( \Lambda(\lambda) \). For some specifications, additional restrictions on the factor dynamics and on the vector of conditional mean \( \mu \) are also required. The first specification considered for the term structure is based on the seminal paper of Nelson and Siegel (1987) in which the yield curve is approximated by a weighted sum of three smooth functions, and the functional form of these three functions depend on a single parameter. Diebold and Li (2006) use the Nelson-Siegel framework to develop a two-step procedure to forecast yields. The second specification considered is the four-factor model proposed by Svensson (1994) model to increase the flexibility and improve the performance of the Nelson-Siegel model by adding a second hump-shaped factor with a separate decay parameter. The model of Nelson and Siegel (1987) and the extension of Svensson (1994) are widely used by central banks and other market participants as an econometric model for the term structure of interest rates (Gimeno and Nave, 2009; BIS, 2005).

2.2. Dynamic Nelson-Siegel model

Nelson and Siegel (1987) show that the term structure can be surprisingly well fitted by a linear combination of three smooth functions. Although the Nelson-Siegel model is specified as a static model that does not account for the intertemporal evolution of the term structure, Diebold and Li (2006) show that the coefficients in \( f_t \) can be interpreted as three latent dynamic factors. The Nelson-Siegel yield curve, denoted by \( g_{NS}(\tau) \) is given by

\[
g_{NS}(\tau) = \xi_1 + \lambda^S \cdot \xi_2 + \lambda^C \cdot \xi_3, \tag{3}
\]

where

\[
\lambda^S(\tau) = \frac{1 - \exp(-\lambda \tau)}{\lambda \tau}, \quad \lambda^C(\tau) = \frac{1 - \exp(-\lambda \tau)}{\lambda \tau} - \exp(-\lambda \tau) \tag{4}
\]

and where \( \lambda, \xi_1, \xi_2 \) and \( \xi_3 \) are treated as parameters. The yield curve depends on these parameters which can be estimated via least squares based on the following nonlinear regression

6
model
\[ y_t(\tau_i) = g_{NS}(\tau_i) + u_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, n \]
where \( u_{it} \) has zero mean and possibly different variances for different maturities \( \tau_i \). One of the advantages of the Nelson-Siegel model is that the parameters \( \xi \) have a clear interpretation. The parameter \( \xi_1 \) controls for the level of the yield curve. The parameter \( \xi_2 \) can be associated with the slope of the yield curve since its loading \( \lambda^S(\tau) \) is higher for short maturities and lower for a long maturities. The loadings \( \lambda^C(\tau) \) for different time to maturities form a U-shaped function and thus \( \xi_3 \) can be interpreted as the curvature of the yield curve. The decomposition of the yield curve in to level, slope and curvature factors has been highlighted by Litterman and Scheinkman (1991).

The Nelson-Siegel yield curve can be also specified as a dynamic factor model by treating the \( \xi \) parameters as factors and letting them evolve as time-varying processes. In this case, we have the following specification:
\[ y_t = \Lambda_{NS}f_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, \Sigma_t), \quad (5) \]
where \( f_t \) is a \( 3 \times 1 \) vector \((K = 3)\) and \( \Sigma_t \) is a diagonal conditional covariances matrix. The loading matrix \( \Lambda_{NS} \) consists of the three columns \((1, \ldots, 1)'\), \( \lambda^S(\tau) \) and \( \lambda^C(\tau) \) respectively. This dynamic factor representation of the Nelson-Siegel model is proposed by Diebold et al. (2006).

### 2.2.1. Svensson model

Svensson (1994) extended the Nelson-Siegel specification by including an extra exponential term with a different decaying parameter. Formally, the model is an extension of the system described by equations (3) and (4) with the inclusion of a fourth component, \( \xi_4 \), with an independent decay parameter, \( \lambda_2 \). The fourth component, \( \left[ \lambda_2^C = \frac{1-\exp(-\lambda_2 \tau)}{\lambda_2 \tau} - \exp(-\lambda_2 \tau) \right] \), introduces a second medium-term component to the model. Similar to the Nelson-Siegel model, the Svensson model can be represented as a dynamic factor model:
\[ y_t = \Lambda_{SV}f_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, \Sigma_t), \quad (6) \]
where the dynamics of the \( 4 \times 1 \) vector \( f_t \) follows (2) while \( \Sigma_t \) is a diagonal conditional variance matrix. The \( i \)-th row of the loading matrix \( \Lambda_{SV} \) is given by \([1, \lambda^S_i(\tau_i), \lambda_1^C_i(\tau_i), \lambda_2^C_i(\tau_i)]\).
2.3. **Conditional covariance of the factor models for the yield curve**

Forecasting volatility of interest rates remains an important challenge in financial econometrics. A rich body of literature has shown that the volatility of the yield curve is, at least to some extent, related to the shape of the yield curve. For instance, the volatility of interest rates is usually high when interest rates are high and when the yield curve exhibits more curvature (see Cox et al. (1985), Litterman et al. (1991), and Longstaff and Schwartz (1992), among others). This suggests that the shape of the yield curve is a potentially useful instrument for forecasting volatility.

Despite the large amount of studies dealing with fitting and forecasting of the yield curve, only recently attention has been turned to the presence of conditional heteroskedasticity in the term structure of interest rates. In most cases, the models for the yield curve adopt the assumption of constant volatility for all maturities. This issue is particularly important since the assumption of constant interest rate volatility has important practical implications for risk management policies, as it neglects the time-varying characteristic of interest rate risk. Furthermore, interest rate hedging and arbitrage operations are also influenced by the presence of time-varying volatility as, in these operations, it is often necessary to compensate for the market price of interest rate risk. Another important implication is that in the presence of conditional volatilities the confidence intervals for the forecasts obtained from these models will be possibly miscalculated in finite samples. Some recent approaches designed to overcome these limitations have been proposed by Bianchi et al. (2009), Haustsch and Ou (2010), Koopman et al. (2010) and Caldeira et al. (2010a).

In this paper, the effects of time-varying volatility are incorporated using a multivariate GARCH specification proposed by Santos and Moura (2011). To model $\Omega_t$, the conditional covariance matrix of the factors in (2), alternative specifications can be considered, including not only multivariate GARCH models but also multivariate stochastic volatility models (see Harvey et al., 1994; Aguilar and West, 2000; Chib et al., 2009). In this paper, we consider the dynamic conditional correlation model (DCC) proposed by Engle (2002), which is given by:

$$\Omega_t = D_t \Psi_t D_t$$

---

2See Poon and Granger (2003) and Andersen and Benzoni (2010) for recent surveys on volatility forecasting.

3See, for instance, Filipovic (2009), for a review on interest rate modeling.

4See Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009) for a detailed review of multivariate GARCH models.
where $D_t$ is a $K \times K$ diagonal matrix with diagonal elements given by $h_{fkt}$, where $h_{fkt}$ is the conditional variance of the $k$-th factor, and $\Psi_t$ is a symmetric correlation matrix with elements $\rho_{ij,t}$, where $\rho_{ii,t} = 1$, $i, j = 1, \ldots, K$. In the DCC model, the conditional correlation $\rho_{ij,t}$ is given by:

$$
\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}
$$

where $q_{ij,t}$, $i, j = 1, \ldots, K$, are the elements of the $K \times K$ matrix $Q_t$, which follows a GARCH-type dynamics:

$$
Q_t = (1 - \alpha - \beta) \bar{Q} + \alpha z_{t-1}'z_{t-1}' + \beta Q_{t-1}
$$

where $z_{ft} = (z_{f1t}, \ldots, z_{fKt})$ is the standardized vector of returns of the factors, whose elements are $z_{fkt} = f_{it}/\sqrt{h_{fkt}}$, $\bar{Q}$ is the unconditional covariance matrix $z_t$, $\alpha$ and $\beta$ are non negative scalar parameters satisfying $\alpha + \beta < 1$.

To model the conditional variance of the measurement errors $\varepsilon_t$ in (1), we assume that $\Sigma_t$ is a diagonal matrix with diagonal elements given by $h_{tei}$, where $h_{tei}$ is the conditional variance of $\varepsilon_i$. Moreover, a procedure similar to Cappiello et al. (2006) is applied and alternative specifications of the univariate GARCH type are used to model $h_{tei}$. In particular, we consider the GARCH model of Bollerslev (1986), the asymmetric GJR-GARCH model Glosten et al. (1993), the exponential GARCH (EGARCH) model of Nelson (1991), the threshold GARCH (TGARCH) model of Zakoian (1994), the asymmetric exponent GARCH (APARCH) model of Ding et al. (1993), asymmetric GARCH (AGARCH) model of Engle (1990), and the non-linear asymmetric GARCH (NAGARCH) model of Engle and Ng (1993). In all models, their simplest form is adopted in which the conditional variance depends on one lag of both past returns and conditional variances. Appendix 2 lists the specifications of each of these models. The same procedure is applied to the choice of the GARCH specification for the conditional variance of the factors in (7). In all cases, the choice of the specification used is based on Akaike Information Criterion (AIC).

2.4. Mapping the moments of yields into the moments of returns

The Markowitz approach to portfolio optimization requires estimates of the expected return of each bond, as well as the covariance matrix of bond returns. However, the factor models for the term structure of interest rates discussed above are designed to model only bond yields. Nevertheless, it is possible to obtain expressions for the expected bond return and for the conditional covariance matrix of bond returns based on the distribution of the
expected yields. The following proposition defines this distribution.\(^5\)

**Proposition 1.** Given the system of equations in (1) and (2), the distribution of expected yields \(y_{t|t-1}\) is \(N(\mu_{yt}, \Sigma_{yt})\) with \(\mu_{yt} = \Lambda f_{t|t-1}\) and \(\Sigma_{yt} = \Lambda \Omega_{t|t-1} \Lambda' + \Sigma_{t|t-1}\), where \(f_{t|t-1}\) is a one-step-ahead forecast of the factors and \(\Omega_{t|t-1}\) and \(\Sigma_{t|t-1}\) are one-step-ahead forecasts of the conditional covariance matrices in (1) and (2), respectively.

Using the results of Proposition 1, we next show that it is possible to derive the distribution of expected fixed-maturity bond prices. Taking into account that the price of a bond at time \(t\), \(P_t(\tau)\), is the present value at time \(t\) of $1 receivable \(\tau\) periods ahead, and letting \(y_{t|t-1}\) denote the one step ahead forecast of its continuously compounded zero-coupon nominal yield to maturity, we obtain the vector of expected bond prices \(P_{t|t-1}\) for all maturities:

\[
P_{t|t-1} = \exp\left(-\tau \otimes y_{t|t-1}\right),
\]

where \(\otimes\) is the Hadamard (elementwise) multiplication and \(\tau\) is the vector of maturities. Since \(y_{t|t-1}\) follows a Normal distribution, \(P_{t|t-1}\) has a log-normal distribution with mean given by:

\[
\mu_{pt} = \exp\left\{\left[-\tau \otimes \mu_{yt} + \frac{\tau^2}{2} \otimes \text{diag}(\Sigma_{yt})\right]\right\},
\]

where \(\text{diag}(\Sigma_{yt})\) is a vector containing the diagonal elements of \(\Sigma_{yt}\). The covariance of bond prices, \(\Sigma_{pt}\), has elements given by:

\[
\sigma_{pi,j}^2 = \exp\left\{\left[-\tau^i \mu_{yt}^i - \tau^j \mu_{yt}^j + 0.5 \left(\tau^2 \sigma_{yi,i}^2 + \tau^2 \sigma_{yj,j}^2\right)\right]\right\} \cdot \left[\exp\left(\tau^i \tau^j \sigma_{yi,j}^2\right) - 1\right]
\]

Using the formula for log-returns, one can find a closed form expression for the vector of expected returns of bonds as well as for their conditional covariance matrix. Proposition 2 defines these expressions.

**Proposition 2.** Given the system of equations in (1) and (2) and Proposition 1, the vector of expected bond returns, \(\mu_{r_{t|t-1}}\), and their conditional covariance matrix \(\Sigma_{r_{t|t-1}}\), which is

---

\(^5\)Appendix 1 presents proofs of all propositions.
positive-definite $\forall t$, are given by:

$$\mu_{r_{t|t-1}} = -\tau \otimes \mu_{yt} + \tau \otimes y_{t-1},$$

(13)

$$\Sigma_{r_{t|t-1}} = \tau' \tau \otimes \begin{pmatrix} \Lambda \Omega_{t|t-1} \Lambda' + \Sigma_{\tau|t-1} \\ \Sigma_{\eta} \end{pmatrix}.$$  

(14)

The results in Proposition 2 show that it is possible to obtain closed form expressions for the expected bond returns and their covariance matrix based on yield curve models such as the ones by Nelson and Siegel (1987) and Svensson (1994). These estimates are key ingredients to the problem of portfolio selection based on mean-variance paradigm proposed by Markowitz, as discussed in section 4.

As pointed out by Litterman and Scheinkman (1991), the return on a fixed maturity zero-coupon bond can be decomposed into two parts. The first part is a result of the capitalization received due to ageing of the bond and the second part is attributed to the change in market prices of constant maturity bonds. Furthermore, Litterman and Scheinkman (1991) point out that the first part is deterministic, while the second part is subject to uncertainty regarding the changes in prices. Clearly, portfolio optimization is only concerned with the second part.

However, for comparison with other benchmarks, it is also necessary to compute the deterministic part of the return. The total return will be given by the income generated by the capitalization based on the interest rate on the bond, plus capital appreciation given by the variation in market prices. Following Jones et al. (1998) and de Goeij and Marquering (2006), the total return (between $t$ and $t + p$) on a bond with fixed maturity $\tau$ is given by:

$$R_{t,t+p}(\tau) = \frac{P_t(\tau)}{P_{t-p}(\tau)} - 1 + \frac{p}{252}y_{t-p}(\tau) = \exp(r_{t,t+p}) - 1 + \frac{p}{252}y_{t-p},$$

(15)

where $p$ is given on weekdays and $r_{t,t+p}$ is the part of the return generated by changes in yields of fixed maturities from period $t$ to $t + p$.

3. Estimation procedure

In this section, we propose an estimation procedure for the parameters of the yield curve and of volatility models. The estimation is performed in a multi-step procedure in which

\cite{footnote}{See equation (19) in Appendix 1 for details concerning the calculation of $r_{t,t+p}$.}
the parameters of the factor model are first estimated, and resulting residuals are used to estimate the volatility models discussed in Section 2.3.

3.1. Estimation of the yield curve models

The most straightforward approach to estimate the factors and parameters of the system (1) and (2) consists of a two-step procedure proposed by Diebold and Li (2006). In the first step, the measurement equation is treated as a cross section for each period of time, and ordinary least squares (OLS) is employed to estimate the factors for all time periods individually. In the second step, the dynamic specification of the factors are estimated. To simplify the estimation procedure, Diebold and Li (2006) suggest reducing the parameter vector by calibrating the value of $\lambda$, rather than treating it as an unknown parameter.

The first step produces time series for the $K$ factors, $\{\{\beta_{k,t}\}_{t=1}^{T}\}_{k=1}^{K}$. The next step is to estimate the factor dynamics of the state equations. We estimate separate AR(1) models for each factor, thus assuming that $\Upsilon$ in (2) is diagonal, as well as a VAR(1) by assuming that $\Upsilon$ in (2) is a full matrix instead.

The choice of the decay parameters for the Nelson-Siegel and Svensson models are restricted to the interval between 0.04 and 0.5, since these values correspond to the maximum value of the curvature loadings at 48 and 6 months, respectively, which are the highest and lowest maturities of the data set considered in the empirical application in Section 4. Given these boundaries, we construct the grid $\Phi = \{0.04 + 0.001j\}_{j=1}^{501}$. Given $\lambda_j \in \Phi$ and the correspondent matrix of factor loadings $\Lambda(\lambda_j)$, the vector of factors $f_t$ is estimated by OLS for each period $t$, and the decay parameter $\hat{\lambda} \in \Phi$ is chosen to minimize the root mean squared error (RMSE). More specifically, $\hat{\lambda}$ is chosen to minimize the difference between the expected yield, $y_{t|t-1}$, and the observed yield, $y_t$. The calibration problem can be represented as:

$$\hat{\lambda} = \arg\min_{\lambda \in \Phi} \sqrt{\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} (y_t(\tau_i) - y_{t|t-1}(\tau_i, \lambda, f_{t|t-1}))^2}$$

where $T$ is the number of yield curves in the sample.

In the case of Svensson model, the problem is similar except that in this case it is necessary to estimate two decay parameters ($\hat{\lambda}_1, \hat{\lambda}_2$). Then, $(\lambda_1, \lambda_2)$ are the solution to the following problem:

$$\left(\hat{\lambda}_1, \hat{\lambda}_2\right) = \arg\min_{(\lambda_1, \lambda_2) \in \Theta} \sqrt{\frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} (y_t(\tau_i) - y_{t|t-1}(\tau_i, \lambda_1, \lambda_2, f_{t|t-1}))^2}.$$
A multicolinearity problem may arise when the decay parameters $\lambda_1$ and $\lambda_2$ have similar values. When this is the case, the Svensson model boils down to the three-factor base model but with a curvature factor equal to the sum of $\beta_3t$ and $\beta_4t$. In this case, $\beta_3t$ and $\beta_4t$ will have the same factor loading, yielding two perfectly collinear regressors. This extreme situation has been reported before, for instance in Xiao (2001) and De Pooter (2007). To avoid this problem, we adopt the modification proposed by De Pooter (2007) which consists of replacing the last term of $\lambda_2^C_2$, i.e. $-\exp\left(\frac{-\tau}{\lambda_2t}\right)$, for $-\exp\left(\frac{-2\tau}{\lambda_2t}\right)$.

3.2. Estimation of the covariance matrix of bond yields

To obtain the conditional covariance matrix of the factors, $\Omega_{t|t-1}$, a DCC specification in (7) is used. The estimation of the DCC model can be conveniently divided into volatility part and correlation part. The volatility part refers to estimating the univariate conditional volatility models of the factors using a GARCH-type specification. The parameters of univariate volatility models are estimated by quasi maximum likelihood (QML) assuming Gaussian innovations.\(^7\) The correlation part refers to the estimation of the conditional correlation matrix in (8) and (9). To estimate the parameters of the correlation matrix, we employ the composite likelihood (CL) method proposed by Engle et al. (2008). As pointed out by Engle et al. (2008), the CL estimator provides more accurate parameter estimates in comparison to the two-step procedure proposed by Engle and Sheppard (2001) and Sheppard (2003), especially in large problems.

4. Application to portfolio optimization of fixed-income securities

To illustrate the applicability of the proposed estimators of expected bond returns and conditional covariance matrix of bond returns defined in Propositions 1 and 2, we consider the mean-variance optimization problem of fixed-income portfolios (Markowitz, 1952; Korn and Koziol, 2006; Puhle, 2008). The formulation of mean-variance portfolio is given by

$$\min_w w_t \sum_{r_t|t-1} w_t - \frac{1}{\delta} w_t' \mu_{r_t|t-1}$$

subject to

$$w_t' = 1$$

\(^7\)A review of issues related to the estimation of univariate GARCH models, such as the choice of initial values, numerical algorithms, accuracy, and asymptotic properties are given by Berkes et al. (2003), Robinson and Zaffaroni (2006), Francq and Zakoian (2009) and Zivot (2009). It is important to note that even when the normality assumption is inappropriate, the QML estimator of univariate GARCH models based on maximizing the Gaussian likelihood is consistent and asymptotically normal, provided that the conditional mean and variance of the GARCH model are correctly specified, see Bollerslev and Wooldridge (1992).
where $\mu_{r|t-1}$ is a one-step-ahead forecast of the expected bond returns, $\Sigma_{r|t-1}$ is a one-step-ahead prediction of the conditional covariance matrix of bond returns, $w_t$ is the vector of optimal weights, $\iota$ is a vector of ones with dimension $N \times 1$, and $\delta$ is the coefficient of risk aversion. In the case where short-sales are restricted, a constraint to avoid negative weights is added to (16), i.e. $w_t \geq 0$. Previous works show that adding such a restriction can substantially improve performance, especially reducing the turnover of the portfolio, see Jagannathan and Ma (2003), among others.

A variation of the mean-variance portfolio optimization problem is the minimum variance portfolio problem. In this formulation, individuals are assumed to be extremely risk averse, such that $\delta \to \infty$. The formulation of the optimal minimum variance portfolio is given by

$$
\min_{w_t} w_t \Sigma_{r|t-1} w_t
$$

subject to

$$
w_t' \iota = 1.
$$

(17)

Similar as before, a restriction to avoid negative weights is added to (17), or $w_t \geq 0$. In both cases, the optimal weights with short-selling restrictions are obtained by using numerical optimization methods.

4.1. Data and implementation details

We use time series of yields of Inter Bank Deposit Future Contract (ID-future), which is one of the largest fixed-income markets among emerging economies, collected on a daily basis. The ID-future contract with maturity $\tau$ is a zero-coupon future contract in which the underlying asset is the ID-future interest rate accrued on a daily basis, capitalized between trading period $t$ and $\tau$. The contract value is set by its value at maturity, $R$100,000.00, discounted according to the accrued interest rate negotiated between the seller and the buyer. A similar data set is also used by Almeida and Vicente (2009).

The Brazilian Mercantile and Futures Exchange (BM&F) is the entity that offers the ID-future contract and determines the number of maturities with authorized contracts. In general, there are around 20 maturities with authorized contracts every day, but not all of them have liquidity. Approximately 10 maturities have contracts with high liquidity. In

---

8In this paper, we follow DeMiguel and Nogales (2009) and assume that $\delta = 1$.

9The ID-future rate is the average daily rate of Brazilian interbank deposits (borrowing/lending), calculated by the Clearinghouse for Custody and Settlements (CETIP) for all business days. The ID-future rate, which is published on a daily basis, is expressed in annually compounded terms, based on 252 business days.
2010 the ID-future market traded a total of 293 million contracts corresponding to US$ 15 billion. The ID-future contract is very similar to the zero-coupon bond, except for the daily payment of marginal adjustments. Every day the cash flow is the difference between the adjustment price of the current day and the adjustment price of the previous day, indexed by the ID-future rate of the previous day.

When buying a future ID-future contract for the price at time \( t \) and keeping it until maturity \( \tau \), the gain or loss is given by:

\[
100,000 \left( \prod_{i=1}^{\zeta(t,\tau)} \left( 1 + y_i \right) \right)^{1/252} \left( 1 + ID^* \right) \zeta(t,\tau)_{252} - 1 ,
\]

where \( y_i \) denotes the ID-future rate, \((i-1)\) days after the trading day. The function \( \zeta(t,\tau) \) represents the number of working days between \( t \) and \( \tau \).

We use time series of daily closing yields of the ID-future contracts. In practice, contracts with all maturities are not observed on a daily basis. Therefore, based on the observed rates for the available maturities, the data were converted into fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 42 and 48 months, using the cubic splines interpolation method originally proposed by McCulloch (1971, 1975).\(^{10}\) It is worth noting that the data set contains maturities with highest liquidity for January 2006 to December 2010 (\( T = 986 \) observations). The assessment of the performance of the model is made by splitting the full sample into two parts. The first 500 observations are used to estimate the parameters of all models according to the procedures discussed in Section 3. The remaining 486 observations are used to analyze the out-of-sample performance of optimal bond portfolios.

Table 1 reports descriptive statistics for the Brazilian interest rate yield curve based on the ID-future market. For each time series we report the mean, standard deviation, minimum, maximum and the sample autocorrelations at lags of one day, one week, and one month. The summary statistics confirm the presence of stylized facts common to yield curve data: the average curve is upward sloping and concave, volatility is decreasing with maturity, autocorrelations are very high specially for short maturity.

Figure 4.1 displays a three-dimensional plot of the data set and illustrates how yield levels and spreads vary substantially throughout the sample. The plot also suggests the presence of an underlying factor structure. Although the yield series vary heavily over time for each

\(^{10}\)For further details and applications of this method, see Hagan and West (2006) and Hayden and Ferstl (2010).
Table 1: Descriptive statistics

The Table reports descriptive statistics of daily yields for different maturities. The last three columns contain the autocorrelations with a lag of one day, one week and one month, respectively. The sample period runs from Jan. 2006 to Dec. 2010.

<table>
<thead>
<tr>
<th>Maturity (Months)</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>$\hat{\rho}(1)$</th>
<th>$\hat{\rho}(5)$</th>
<th>$\hat{\rho}(21)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>10.82</td>
<td>1.65</td>
<td>8.58</td>
<td>14.34</td>
<td>0.220</td>
<td>2.006</td>
<td>0.999</td>
<td>0.997</td>
<td>0.969</td>
</tr>
<tr>
<td>6</td>
<td>10.88</td>
<td>1.67</td>
<td>8.59</td>
<td>14.52</td>
<td>0.264</td>
<td>2.071</td>
<td>0.999</td>
<td>0.997</td>
<td>0.968</td>
</tr>
<tr>
<td>9</td>
<td>10.94</td>
<td>1.69</td>
<td>8.58</td>
<td>14.69</td>
<td>0.306</td>
<td>2.132</td>
<td>0.999</td>
<td>0.996</td>
<td>0.967</td>
</tr>
<tr>
<td>12</td>
<td>11.09</td>
<td>1.72</td>
<td>8.61</td>
<td>15.32</td>
<td>0.386</td>
<td>2.241</td>
<td>0.999</td>
<td>0.995</td>
<td>0.961</td>
</tr>
<tr>
<td>15</td>
<td>11.34</td>
<td>1.73</td>
<td>8.73</td>
<td>16.04</td>
<td>0.495</td>
<td>2.373</td>
<td>0.998</td>
<td>0.992</td>
<td>0.950</td>
</tr>
<tr>
<td>18</td>
<td>11.60</td>
<td>1.72</td>
<td>8.99</td>
<td>16.40</td>
<td>0.572</td>
<td>2.461</td>
<td>0.998</td>
<td>0.989</td>
<td>0.938</td>
</tr>
<tr>
<td>21</td>
<td>11.85</td>
<td>1.68</td>
<td>9.35</td>
<td>16.92</td>
<td>0.655</td>
<td>2.565</td>
<td>0.997</td>
<td>0.986</td>
<td>0.925</td>
</tr>
<tr>
<td>24</td>
<td>12.04</td>
<td>1.61</td>
<td>9.55</td>
<td>17.12</td>
<td>0.718</td>
<td>2.659</td>
<td>0.996</td>
<td>0.982</td>
<td>0.911</td>
</tr>
<tr>
<td>27</td>
<td>12.21</td>
<td>1.55</td>
<td>9.79</td>
<td>17.26</td>
<td>0.805</td>
<td>2.815</td>
<td>0.995</td>
<td>0.979</td>
<td>0.894</td>
</tr>
<tr>
<td>30</td>
<td>12.33</td>
<td>1.49</td>
<td>10.06</td>
<td>17.44</td>
<td>0.912</td>
<td>3.026</td>
<td>0.995</td>
<td>0.975</td>
<td>0.877</td>
</tr>
<tr>
<td>33</td>
<td>12.43</td>
<td>1.45</td>
<td>10.27</td>
<td>17.62</td>
<td>1.005</td>
<td>3.290</td>
<td>0.994</td>
<td>0.972</td>
<td>0.859</td>
</tr>
<tr>
<td>36</td>
<td>12.50</td>
<td>1.41</td>
<td>10.42</td>
<td>17.78</td>
<td>1.085</td>
<td>3.568</td>
<td>0.993</td>
<td>0.968</td>
<td>0.843</td>
</tr>
<tr>
<td>42</td>
<td>12.60</td>
<td>1.32</td>
<td>10.71</td>
<td>17.83</td>
<td>1.281</td>
<td>4.180</td>
<td>0.992</td>
<td>0.961</td>
<td>0.814</td>
</tr>
<tr>
<td>48</td>
<td>12.68</td>
<td>1.24</td>
<td>11.09</td>
<td>17.93</td>
<td>1.465</td>
<td>4.910</td>
<td>0.990</td>
<td>0.955</td>
<td>0.788</td>
</tr>
</tbody>
</table>

Figure 1: Evolution of the yield curve

The figure plots the evolution of term structure of interest rates (based on ID-future contracts) for the time horizon of 2006:01-2010:12. The sample consisted of the daily yields for the maturities of 1, 3, 4, 6, 9, 12, 15, 18, 24, 27, 30, 36, 42 and 48 months.
of the maturities, a strong common pattern in the 15 series over time is apparent. For most months, the yield curve is an upward sloping function of time to maturity. For example, last year of the sample is characterized by rising interest rates, especially for the shorter maturities, which respond faster to the contractionary monetary policy implemented by the Brazilian Central Bank in the first half of 2010. It is clear from Figure 4.1 that not only the level of the term structure fluctuates over time but also its slope and curvature. The curve takes on various forms ranging from nearly flat to (inverted) $S$-type shapes.

4.2. Performance evaluation

The performance of optimal mean-variance portfolios and of minimum variance portfolios is evaluated in terms of average return ($\hat{\mu}$), average excess return relative to the risk-free rate$^{11}$ ($\hat{\mu}_{ex}$), standard deviation (volatility) of returns ($\hat{\sigma}$), Sharpe Ratio (SR), and turnover. These statistics are calculated as follows:

$$\hat{\mu} = \frac{1}{T-1} \sum_{t=1}^{T-1} w_t' R_{t+1}$$
$$\hat{\mu}_{ex} = \frac{1}{T-1} \sum_{t=1}^{T-1} (w_t' R_{t+1} - CDI_{t+1})$$
$$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T-1} (w_t' R_{t+1} - \hat{\mu})^2}$$
$$SR = \frac{\hat{\mu}_{ex}}{\hat{\sigma}}$$
$$\text{Turnover} = \frac{1}{T-1} \sum_{t=1}^{T-1} \sum_{j=1}^{N} (|w_{j,t+1} - w_{j,t}|),$$

where $w_{j,t}$ is the weight of the asset $j$ in the portfolio in period $t$, but before the re-balancing and $w_{j,t+1}$ is the desired weight of the asset $j$ in period $t + 1$. As pointed out by DeMiguel et al. (2009b), the turnover, as defined above, can be interpreted as the average fraction of wealth traded in each period.

Similar to DeMiguel et al. (2009a) and Santos et al. (2012), the stationary bootstrap of Politis and Romano (1994) with $B=1,000$ resamples and block size $b = 5^{12}$ was used to test

---

$^{11}$We consider the risk free rate to be the one-month DI-future contract (CDI).

$^{12}$We performed extensive robustness tests to define the block size, using values for $b = 5$ to $b = 250$. 

17
the statistical significance of differences between the volatilities and Sharpe ratios of optimal portfolios relative to a benchmark. *p*-values were obtained using the methodology suggested in Ledoit and Wolf (2008, Note 3.2). The benchmark used is the fixed-income index IRF-M discussed in Section 4.3.

4.3. Benchmark indices

The Brazilian Association of Financial and Capital Markets (ANBIMA) releases several market indices composed of fixed-income securities, which are widespread among market participants. These indices are used in the industry as benchmarks for fixed-income funds and managers in Brazil. In order to meet the needs of different types of investors, there is a set of indices that represent the evolution of the different types of bonds traded. Among the indices available, some can be highlighted:

**IRF-M:** Composed of non-indexed zero-coupon government bonds (NTN-F e LTN);

**IMA-B:** Composed of inflation-indexed government bonds, where the principal is indexed to IPCA price index (NTN-B);

**IMA-C:** Composed of inflation-indexed government bonds, where the principal is indexed to IGP-M price index (NTN-C);

**IMA-S:** Composed of government bonds, where the principal is indexed to the SELIC rate (LFT).

Taking into account that the database used is composed of zero-coupon DI-future contracts for various maturities (less than and greater than one year), we chose to use the fixed-income index IRF-M as a benchmark, since this index is based on zero-coupon bonds with several maturities, including maturities lower than 1 year.

4.4. Results

In this section, we present the out-of-sample results of optimal mean-variance and minimum variance portfolios of ID-future contracts. To assess the robustness of the results, we consider several econometric specifications to model the vector of expected returns and covariance matrix of the returns. More specifically, we consider two specifications for the factor model (Nelson-Siegel and Svensson) and two specifications for the dynamic factors: VAR(1)

Regardess of the size of the block, the test results for the variance and Sharpe ratio are similar to those reported here.
and AR(1). The covariance matrices of returns are obtained using the DCC specification presented in section 2.3.

Optimal portfolio compositions are recalculated (re-balanced) on a daily basis. However, the transaction costs involved in this re-balancing frequency can degrade the performance of the portfolios and hinder its implementation in practice. Thus, the performance of optimized portfolios is also evaluated in the case of weekly, monthly and quarterly re-balancing. A potentially negative effect of adopting a lower frequency of re-balancing is that the optimal compositions may become outdated.

Tables 2 and 3 report the performance results of mean-variance and minimum variance portfolios, respectively. In addition, Table 4 reports the performance evaluation of the benchmark indexes discussed in Section 4.3. The statistics of average return, average excess return, standard deviation and Sharpe ratio are annualized. The results in Table 2 show that all specifications outperform the benchmark index (IRF-M) in terms of average return and average excess return on all re-balancing frequencies. For example, the specification Nelson-Siegel/AR/DCC with quarterly re-balancing generates an average return of 21.2% and average excess return of 11.8%, while the benchmark index generates 2.11% and 1.8% for these figures. However, we observe that the standard deviation (risk) of the mean-variance portfolios is statistically higher than that obtained by the benchmark in the majority of specifications, since the standard deviation of the optimized portfolio ranges from 1.66% to 5.15%, while the standard deviation of the benchmark index is 1.65%. Nevertheless, when examining the Sharpe ratio, the mean-variance portfolios show better risk-adjusted performance than that obtained by the benchmark in majority of the specifications. For instance, the specification Nelson-Siegel/AR/DCC generates a Sharpe ratio of 2.3 at quarterly re-balancing, while the reference index generates a Sharpe ratio of 1.1. Finally, we find that, as expected, the quarterly re-balancing generates turnover substantially lower when compared to the daily re-balancing.

Table 3 reports the results of the portfolios obtained by the criterion of minimum variance. We observe that in terms of average return and average excess return, the results are inferior in comparison to the reference index IRF-M. However, the standard deviation of minimum variance portfolios is substantially (and statistically) lower than that obtained by the benchmark. Throughout all econometric specifications and over all re-balancing frequencies, the standard deviation of minimum variance portfolios ranges from 0.20% to 0.24%, while the benchmark index has a standard deviation of 1.6%. Consequently, the risk-adjusted return of minimum variance portfolios measured by Sharpe ratio is statistically higher than that of the reference index in all cases, since it varies from 2.7 to 3.4 while the Sharpe ratio of the
Table 2: Out-of-sample performance of optimal mean-variance portfolios

The Table reports performance statistics for mean-variance portfolios using ID-future contracts with maturities equal to 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 42 and 48 months. The optimal portfolios are re-balanced with daily, weekly, monthly and quarterly frequencies. The statistics of returns, standard deviation and Sharpe ratio are annualized. The excess return is calculated using the one-month ID-future as a risk-free asset. Asterisks indicate that the coefficient is statistically different to that obtained by the benchmark index (IRF-M) at a significance level of 10%.

<table>
<thead>
<tr>
<th>Factor Model</th>
<th>Factor Dynamics</th>
<th>Covariance Matrix</th>
<th>Average Return (%)</th>
<th>Average Excess Return (%)</th>
<th>Standard Deviation (%)</th>
<th>Sharpe Ratio</th>
<th>Turnover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily re-balancing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>AR</td>
<td>DCC</td>
<td>17.763</td>
<td>8.384</td>
<td>3.785*</td>
<td>2.214*</td>
<td>0.990</td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>VAR</td>
<td>DCC</td>
<td>16.772</td>
<td>7.393</td>
<td>3.065*</td>
<td>2.411*</td>
<td>0.480</td>
</tr>
<tr>
<td>Svensson</td>
<td>AR</td>
<td>DCC</td>
<td>15.340</td>
<td>5.960</td>
<td>3.128*</td>
<td>1.905*</td>
<td>1.277</td>
</tr>
<tr>
<td>Svensson</td>
<td>VAR</td>
<td>DCC</td>
<td>13.632</td>
<td>4.253</td>
<td>2.063</td>
<td>2.061</td>
<td>0.651</td>
</tr>
<tr>
<td>Weekly re-balancing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>AR</td>
<td>DCC</td>
<td>17.736</td>
<td>8.357</td>
<td>3.602*</td>
<td>2.319*</td>
<td>0.238</td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>VAR</td>
<td>DCC</td>
<td>15.233</td>
<td>5.853</td>
<td>2.921*</td>
<td>2.003*</td>
<td>0.150</td>
</tr>
<tr>
<td>Svensson</td>
<td>AR</td>
<td>DCC</td>
<td>15.186</td>
<td>5.807</td>
<td>2.698*</td>
<td>2.151</td>
<td>0.298</td>
</tr>
<tr>
<td>Svensson</td>
<td>VAR</td>
<td>DCC</td>
<td>12.627</td>
<td>3.247</td>
<td>1.951</td>
<td>1.663</td>
<td>0.219</td>
</tr>
<tr>
<td>Monthly re-balancing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>AR</td>
<td>DCC</td>
<td>16.568</td>
<td>7.189</td>
<td>3.967*</td>
<td>1.811*</td>
<td>0.054</td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>VAR</td>
<td>DCC</td>
<td>16.317</td>
<td>6.938</td>
<td>3.003*</td>
<td>2.310*</td>
<td>0.046</td>
</tr>
<tr>
<td>Svensson</td>
<td>AR</td>
<td>DCC</td>
<td>15.206</td>
<td>5.827</td>
<td>2.633*</td>
<td>2.212*</td>
<td>0.073</td>
</tr>
<tr>
<td>Svensson</td>
<td>VAR</td>
<td>DCC</td>
<td>12.752</td>
<td>3.373</td>
<td>1.830</td>
<td>1.842</td>
<td>0.074</td>
</tr>
<tr>
<td>Quarterly re-balancing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>AR</td>
<td>DCC</td>
<td>21.231</td>
<td>11.851</td>
<td>5.148*</td>
<td>2.301*</td>
<td>0.023</td>
</tr>
<tr>
<td>Nelson-Siegel</td>
<td>VAR</td>
<td>DCC</td>
<td>15.610</td>
<td>6.230</td>
<td>2.281*</td>
<td>2.730*</td>
<td>0.020</td>
</tr>
<tr>
<td>Svensson</td>
<td>AR</td>
<td>DCC</td>
<td>16.304</td>
<td>6.924</td>
<td>2.928*</td>
<td>2.364*</td>
<td>0.028</td>
</tr>
<tr>
<td>Svensson</td>
<td>VAR</td>
<td>DCC</td>
<td>12.648</td>
<td>3.269</td>
<td>1.912</td>
<td>1.708</td>
<td>0.028</td>
</tr>
</tbody>
</table>

IRF-M is 1.1. Similar to the results obtained with the mean-variance portfolios, we observe that the turnover of mean-variance portfolios and minimum variance reduce substantially as the re-balancing frequency decreases. Finally, a comparative analysis of the performance of mean-variance and minimum variance portfolios shows that the latter generates higher Sharpe ratios, and a substantially lower standard deviation. Thus, the results suggest that, in fact, the minimum variance portfolios serve their purpose of generating optimal compositions that are less risky relative to the benchmark and also with respect to other optimized portfolios.

Figure 2 plots the cumulative returns of mean-variance (upper graph) and minimum variance (lower graph) portfolios based on specifications Nelson-Siegel/AR/DCC and Svens-
Table 3: Out-of-sample performance of optimal minimum variance portfolios

The Table reports performance statistics for minimum-variance portfolios using ID-future contracts with maturities equal to 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 42 and 48 months. The optimal portfolios are re-balanced with daily, weekly, monthly and quarterly frequencies. The statistics of returns, standard deviation and Sharpe ratio are annualized. The excess return is calculated using the one-month ID-future as a risk-free asset. Asterisks indicate that the coefficient is statistically different to that obtained by the benchmark index (IRF-M) at a significance level of 10%.

<table>
<thead>
<tr>
<th>Factor Model</th>
<th>Factor Dynamics</th>
<th>Covariance Matrix</th>
<th>Average Return (%)</th>
<th>Average Excess Return (%)</th>
<th>Standard Deviation (%)</th>
<th>Sharpe Ratio</th>
<th>Turnover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nelson-Siegel AR DCC</td>
<td>10.078</td>
<td>0.699</td>
<td>0.205</td>
<td>3.393</td>
<td>0.073</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel VAR DCC</td>
<td>10.071</td>
<td>0.691</td>
<td>0.206</td>
<td>3.350</td>
<td>0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson AR DCC</td>
<td>10.067</td>
<td>0.687</td>
<td>0.223</td>
<td>3.076</td>
<td>0.098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson VAR DCC</td>
<td>10.048</td>
<td>0.669</td>
<td>0.222</td>
<td>3.002</td>
<td>0.098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel AR DCC</td>
<td>10.076</td>
<td>0.696</td>
<td>0.206</td>
<td>3.379</td>
<td>0.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel VAR DCC</td>
<td>10.082</td>
<td>0.702</td>
<td>0.207</td>
<td>3.389</td>
<td>0.039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson AR DCC</td>
<td>10.062</td>
<td>0.683</td>
<td>0.223</td>
<td>3.051</td>
<td>0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson VAR DCC</td>
<td>10.049</td>
<td>0.669</td>
<td>0.222</td>
<td>3.011</td>
<td>0.051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel AR DCC</td>
<td>10.087</td>
<td>0.708</td>
<td>0.204</td>
<td>3.457</td>
<td>0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel VAR DCC</td>
<td>10.088</td>
<td>0.708</td>
<td>0.205</td>
<td>3.448</td>
<td>0.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson AR DCC</td>
<td>10.036</td>
<td>0.657</td>
<td>0.221</td>
<td>2.962</td>
<td>0.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson VAR DCC</td>
<td>10.013</td>
<td>0.633</td>
<td>0.221</td>
<td>2.861</td>
<td>0.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel AR DCC</td>
<td>10.075</td>
<td>0.696</td>
<td>0.206</td>
<td>3.376</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nelson-Siegel VAR DCC</td>
<td>10.074</td>
<td>0.695</td>
<td>0.209</td>
<td>3.308</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson AR DCC</td>
<td>10.021</td>
<td>0.642</td>
<td>0.232</td>
<td>2.759</td>
<td>0.014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svensson VAR DCC</td>
<td>10.011</td>
<td>0.632</td>
<td>0.230</td>
<td>2.739</td>
<td>0.014</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Performance of benchmark indices

The Table reports performance statistics for several benchmark indices. The IRF-M1, IRF-M1+ and IRF-M indices consist of non-indexed government bonds, whereas the indices IMA-B 5, B-5 + IMA and IMA-B are composed of indexed bonds. The statistics of returns, standard deviation and Sharpe ratio are annualized. The excess return is calculated using the one-month DI-future contract as a risk-free asset.

<table>
<thead>
<tr>
<th></th>
<th>Average Return (%)</th>
<th>Average Excess Return (%)</th>
<th>Standard Deviation (%)</th>
<th>Sharpe Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF-M 1</td>
<td>10.288</td>
<td>0.909</td>
<td>0.391</td>
<td>2.324</td>
</tr>
<tr>
<td>IRF-M 1+</td>
<td>11.971</td>
<td>2.592</td>
<td>2.835</td>
<td>0.914</td>
</tr>
<tr>
<td>IRF-M</td>
<td>11.205</td>
<td>1.826</td>
<td>1.655</td>
<td>1.103</td>
</tr>
<tr>
<td>IMA-B 5</td>
<td>12.705</td>
<td>3.326</td>
<td>1.603</td>
<td>2.075</td>
</tr>
<tr>
<td>IMA-B 5+</td>
<td>18.925</td>
<td>9.546</td>
<td>5.386</td>
<td>1.772</td>
</tr>
<tr>
<td>IMA-B</td>
<td>15.658</td>
<td>6.278</td>
<td>3.269</td>
<td>1.920</td>
</tr>
</tbody>
</table>

son/AR/DCC considering a quarterly re-balancing frequency. We observe that, in the case of mean-variance portfolios, the difference in returns relative to the benchmark is consistently positive throughout the out-of-sample period. For the minimum variance portfolios, we observe that the specifications achieve a similar performance to the benchmark in terms of cumulative returns. However, the visual inspection shows that the optimal portfolios have a substantially lower variance than that of the reference index.

In summary, the results reported in Tables 2 and 3 indicate that the optimal mean-variance and minimum variance portfolios based on the proposed estimates for the vector of expected returns and the covariance matrix of the returns have superior out-of-sample performance with respect to the benchmark index in several respects. First, average and cumulative returns of the mean-variance portfolio exceeds those obtained by the benchmark in all specifications considered. Second, the standard deviation of portfolios of minimum variance is substantially lower than that obtained by the reference index, so that the risk adjusted return of the optimal portfolio is statistically higher. Moreover, the results were robust with respect to: i) the factor model used to model the yield curve, ii) the econometric specification for the transition equation of the factors, and iii) the portfolio re-balancing frequency.
Figure 2: Cumulative returns
5. Concluding remarks

The mean-variance approach introduced by Markowitz (1952) to obtain optimal portfo-
lios has been widely used by market participants and largely documented in the academic
literature. However, the use of this methodology for the optimization of fixed-income portfo-
lios has received little attention in the literature. In order to address this shortcoming, this
paper adopts the mean-variance approach to bond portfolio optimization based on models for
the term structure of interest rate, and on recent specifications to model large-dimensional
conditional covariance matrix of bond yields.

We show that factor models for the yield curve simplify the process of bond portfolio
optimization, since it allows the computation of expected bond returns and their conditional
covariance matrix. We show how to obtain closed-form expressions for the vector of expected
returns and for their conditional covariance matrix based on a general class of heteroskedas-
tic dynamic factor models, and use these estimators to obtain optimal mean-variance and
minimum variance bond portfolios. In particular, we consider the dynamic version of the
Nelson-Siegel model proposed by Diebold and Li (2006) and the four-factor model of Svens-
son (1994).

An application involving a data set of actively traded fixed-income future contracts on the
Brazilian interbank rate is used to illustrate the applicability of the proposed approach. The
results show that the optimized bond portfolios exhibit attractive risk-return profiles. For
example, the portfolios obtained by the mean-variance criterion deliver annualized Sharpe
ratio between 1.66 and 2.42, versus 1.10 for the benchmark index IRF-M. In practice, an in-
vester bears transaction costs when changing the composition of its portfolio over time, and
these costs are a function of frequency and magnitude of changes in the portfolio. Although
transaction costs are not taken into account directly, we compute the turnover of portfolios,
which shows that there is a stability of optimal compositions specially for lower re-balancing
frequencies. In addition, portfolios obtained by the minimum variance criterion show stan-
dard deviation between 0.20 and 0.2, while the reference index has a standard deviation of
1.66, resulting in Sharpe ratios significantly superior to the benchmark in all specifications
considered. Finally, all results were robust with respect to the factor specification used to
model the yield curve, to the dynamics of factors, and to the portfolio re-balancing frequency.
References


Appendix 1

Proof of Proposition 1
Taking expectation of the factor model for the yields in (1), we have

\[ \mu_t = E_{t-1} [y_t] = \Lambda (\lambda) E_{t-1} [f_t] = \Lambda (\lambda) f_{t|t-1} \]  

where \( f_{t|t-1} \) are one step ahead predictions of the factors. The corresponding conditional covariance matrix is given by:

\[
\Sigma_{y_t} = E_{t-1} [(y_t - E_{t-1} [y_t]) (y_t - E_{t-1} [y_t])] \\
= E_{t-1} [(\Lambda f_t + \varepsilon_t - \Lambda E_{t-1} [f_t]) (\Lambda f_t + \varepsilon_t - \Lambda E_{t-1} [f_t])'] \\
= E_{t-1} [(\Lambda (f_t - E_{t-1} [f_t]) + \varepsilon_t) (\Lambda (f_t - E_{t-1} [f_t]) + \varepsilon_t)'] \\
= E_{t-1} [(\Lambda (\mu_t + \gamma f_{t-1} + \eta_t - \mu - \gamma f_{t-1}) + \varepsilon_t) (\Lambda (\mu_t + \gamma f_{t-1} + \eta_t - \mu - \gamma f_{t-1}) + \varepsilon_t)'] \\
= E_{t-1} [(\Lambda \eta_t + \varepsilon_t) (\Lambda \eta_t + \varepsilon_t)'] \\
= E_{t-1} [\Lambda \eta_t \eta_t' + \varepsilon_t \varepsilon_t'] \\
= \Lambda \Omega_{t|t-1} \Lambda' + \Sigma_{y_t|t-1}.
\]

since cross-product between \( \eta_t \) and \( \varepsilon_t \) vanishes because of independence.

\[ \square \]

Proof of Proposition 2
Using the log-return expression, we get:

\[ r_t = \log \left( \frac{P_t}{P_{t-1}} \right) = \log P_t - \log P_{t-1} = -\tau \otimes (y_t - y_{t-1}). \]  

(19)

Since \( y_{t|t-1} \sim N (\mu_t, \Sigma_{y_t}) \) where \( \mu_t \) e \( \Sigma_{y_t} \) are defined in Proposition 1, it is known that the expected returns \( r_{t|t-1} \) follow \( N (\mu_{r_t}, \Sigma_{r_t}) \) where

\[ \mu_{r_{t|t-1}} = -\tau \otimes (E_{t-1} [y_t] - E_{t-1} [y_{t-1}]) = -\tau \otimes \mu_t + \tau \otimes y_{t-1}, \]  

(20)
\[
\Sigma_{r_{|t-1}} = \tau' \tau \otimes \left[ \Lambda \Omega_{t|t-1} \Lambda' + \Sigma_{t|t-1} \right]_{\Sigma_{yt}}.
\]

The positivity of the matrix \( \Sigma_{rt} \) can be demonstrated as follows. The first term in brackets, \( \Lambda \Omega_{t} \Lambda' \), is positive-definite since \( \Omega_{t|t-1} \) is diagonal and possesses only positive elements on its diagonal. The second term, \( \Sigma_{t|t-1} \) is positive-definite for the same reason. Since \( \tau \) contains only positive elements, \( \tau' \tau \) is also a positive-definite matrix. Finally, Schur product theorem ensures that the Hadamard product between \( \Sigma_{yt} \) and \( \tau' \tau \) is positive definite. \qed
Appendix 2: Univariate GARCH models considered

In this appendix we describe the univariate GARCH specifications that were used to model the conditional variance of the factors and the conditional variance of the measurement errors.

GARCH:

\[ \sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \]

Glosten-Jagannathan-Runkle GARCH (GJR-GARCH):

\[ \sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \gamma I[\epsilon_{t-1} < 0] \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \]

Exponential GARCH (EGARCH):

\[ \ln(\sigma_t^2) = \omega + \alpha \frac{|\epsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} + \gamma \frac{\epsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \beta \sigma_{t-1}^2 \]

Threshold GARCH (TGARCH):

\[ \sigma_t = \omega + \alpha |\epsilon_{t-1}| + \gamma I[\epsilon_{t-1} < 0] |\epsilon_{t-1}| + \beta \sigma_{t-1} \]

Asymmetric power GARCH (APARCH):

\[ \sigma_t^\lambda = \omega + \alpha (|\epsilon_{t-1}| + \gamma \epsilon_{t-1})^\lambda + \beta \sigma_{t-1}^\lambda \]

Asymmetric GARCH (AGARCH):

\[ \sigma_t^2 = \omega + \alpha (\epsilon_{t-1} + \gamma)^2 + \beta \sigma_{t-1}^2 \]

Nonlinear asymmetric GARCH (NAGARCH):

\[ \sigma_t^2 = \omega + \alpha (\epsilon_{t-1} + \gamma \sqrt{\sigma_{t-1}^2})^2 + \beta \sigma_{t-1}^2 \]