UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

BERNARDO RECKZIEGEL

MEAN-VARIANCE OPTIMIZATION WITH DYNAMIC MODEL AVERAGING (DMA): AN APPLICATION TO FIXED-INCOME MARKET IN BRAZIL

Porto Alegre
2018

BERNARDO REXKZIEGEL

MEAN-VARIANCE OPTIMIZATION WITH DYNAMIC MODEL AVERAGING (DMA): AN APPLICATION TO FIXED-INCOME MARKET IN BRAZIL

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. João Frois Caldeira

Porto Alegre

2018

CIP - Catalogação na Publicação

Reckziegel, Bernardo
Mean-Variance Optimization with Dynamic Model Averaging (DMA): an application to fixed-income market in Brazil / Bernardo Reckziegel. -- 2018.
27 f.
Orientador: João Fróis Caldeira.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, BR-RS, 2018.

1. Renda-fixa. 2. Otimização de portfólios. 3. DMA. I. Caldeira, João Fróis, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

BERNARDO RECKZIEGEL

MEAN-VARIANCE OPTIMIZATION WITH DYNAMIC MODEL AVERAGING (DMA): AN APPLICATION TO FIXED-INCOME MARKET IN BRAZIL

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia.

Aprovada em: Porto Alegre, 31 de agosto de 2018.

BANCA EXAMINADORA:

Prof. Dr. João Fróis Caldeira – Orientador
Universidade Federal do Rio Grande do Sul – UFRGS

Prof. Dr. Eduardo de Oliveira Horta
Universidade Federal do Rio Grande do Sul – UFRGS

Prof. Dr. Tiago Pascoal Filomena
Universidade Federal do Rio Grande do Sul – UFRGS

Prof. Igor Clemente de Morais
University of California, Riverside - UCR

ABSTRACT

This work uses data from the Brazilian yield curve of interbank deposits to employ the Mean-Variance framework of Markowitz (1952). Expected returns and covariances are calculated according to Koop and Korobilis (2013) that relies on a Time-Varying Parameter VAR (TVP-VAR) with Dynamic-Model Averaging (DMA). It's found that despite having several desired properties and more accurate out-of-sample forecasts, in terms of Mean Squared Forecast Error (MSFE) and Mean Absolute Forecast Error (MAFE), the TVP-VAR-DMA performance it's not able to beat the random-walk on a risk-adjusted basis (Sharpe-ratio), yielding a near zero cumulative excess of return between the period of 2010-jan until 2017-jun. The cumulative underperformance over the selected fixed-income benchmarks, IRF-M and IMA-B 5, under the same period, are 14.25% and 5.48%, respectively.

Keywords: Fixed-Income. Portfolio Optimization. DMA

RESUMO

Este trabalho utiliza dados da curva de juros brasileira de depósitos interbancários para empregar o modelo de Média-Variância de Markowitz (1952). Os retornos esperados e as covariâncias são calculados de acordo com Koop and Korobilis (2013) e se baseiam em um modelo com Vetores Autoregressivos com Parâmetros Variáveis no Tempo (TVP-VAR) e Dynamic Model Averaging (DMA). Descobriu-se que apesar de conterem várias propriedades desejadas e previsões fora da amostra mais precisas, em termos de erro de previsão quadrático médio (MSFE) e erro de previsão absoluto médio (MAFE), o desempenho dos modelos TVP-VAR-DMA não foi capaz de superar Random-Walk em uma base ajustada para o risco (Sharpe-Ratio), produzindo um excesso de retorno sobre o CDI próximo de zero entre o período de 2010-jan até 2017-jun. A performance acumulada abaixo dos índices de renda fixa selecionados, IRF-M e IMA-B 5, no mesmo período, é de 14,25% e 5,48%, respectivamente.

Palavras-chave: Renda-fixa. Otimização de portfólios. DMA

SUMMARY

1 INTRODUCTION

Interest-rates are the main channel in which economic agents can use to make money go back and forth in time. Given this extraordinary feature, interest-rates dynamics have been extensively studied during the last 30 years. Litterman and Scheinkman (1991), for example, found that yield curve evolution is historically well described by a combination of three independent movements, interpreted as level, steepness and curvature. Understanding how these factors affect the shape of the term structure is important for predicting, allocating and hedging risk.

By the same token, Markowitz (1952) showed four decades earlier that, for many assets, risk and return commovements must be considered as a group in order to maximize return per unit of risk. That is, assets should be optimally chosen, and non-systematic risk diversified, whenever is possible. Given the solid foundation on which rests the portfolio optimization theory and the relevance of correctly managing the yield curve, this paper dives into the literature of fixed-income optimization, by applying Mean-Variance as a frame of reference.

A report from ANBIMA (2016) draws attention on the positive impact that a well-timed strategy could have in the Brazilian funds industry. The share of fixed-income assets held by retail funds corresponds to 47% locally, which is more than double typically seen worldwide. The same report also points out that hedge funds and retirement plans account for nearly 39% of the total Assets Under Management (AUM). Whereas this capital is not all invested in fixed-income obligations, high interest-rates certainly restrain managers' prospects in different markets (see nefin.com.br/risk_factors, for an example). In the aggregate, the fund industry manages resources in order of R$ 2.8 trillion (almost 45% of total public debt).

Data from B3 (Brazilian stock exchange) reveals that during the period of 2005 until 2017, interest-rates were responsible for 67% of the total futures contracts traded, followed by 20% in exchange rates and 6% in other financial products. Only in 2017, almost 2 billion were traded daily.

Given the magnitude of those numbers, it is surprising that so little attention has been dedicated to forecasting fixed-income in Brazil. Only recently has the specialized literature tried to fill in this gap. Vieira et al. (2017) for example, show that it is possible to improve random-walk forecasts for the Brazilian yield-curve up to 40%. Their model, however, requires too many explanatory variables which may prevent investors to make decisions on “real-time”. In a mean-variance context, Caldeira, Moura and Santos (2016), and Caldeira, Moura and Santos (2017) present evidence that forecast combination can pay off in a risk-adjusted base. The same authors also show that performance is improved when volatility forecasts are combined in such a way that “winning” models are calibrated to receive more weight in the portfolio optimization process Caldeira et al. (2017).

To explore the potential benefits of diversification, this paper relies on Dynamic Model Averaging (DMA). DMA is a Bayesian strategy that consists of many time-varying regression models formed from combinations that the practioner considers relevant. It’s a flexible method that can be applied to univariate or multivariate series. Moreover, DMA also allows the set of reasonable models to change with time using what the literature has called “forgetting factors”. Through this mechanism, past model performance receives relatively less weight than current models and the estimation procedure can continuously adapt as new information arises. This unique feature is in harmony with a large body of literature that demonstrates that combinations of different models can be more accurate than the individual models themselves (Aiolfi, Timmermann, 2006; Hendry, Clements, 2004; Timmermann, 2006). Positive implications for volatility combinations schemes appears to exist as well (Christiansen, Schmeling and Schrimpf , 2012; Opschoor et al., 2014). This is one of the main theorical reasons of why DMA could be a powerful device at bond manager’s disposal.

From a practical point of view, however, conceiving an efficient DMA algorithm remains a challenging issue - see Catania and Nonejad (2016) and Chan et al. (2012). In attempt to solve this problem, Koop and Korobilis (2013) developed a method to estimate large TVP-VAR models without the need to perform MCMC algorithms. This mechanism can greatly speed up calculations. Secondly, the authors allow prior hyper-parameters to expand in a predefined range. Since the final estimates are derived from the famous Bayes-Rule, different values for the priors can also be thought as a way to define different models. Thirdly, “model uncertainty” is taken into account by allowing different VAR dimensions to be estimated simultaneously. This leads to Dynamic Size Selection (DSS), in which the algorithm automatically selects the VAR size with the highest posterior probability density. Finally, their model only requires three parameters to be estimated, which can, ultimately, wipe-out the “dimensionality curse”. This last contribution is of great value, especially for portfolio optimization problems, in which the manager routinely has dozens (or hundreds) of assets that have to be considered jointly.

In short, the problem can be summarized as follows: combination of different forecasting schemes seems to improve input estimates. In the same way, the quality of input estimates plays a pivotal role in mean-variance optimization. Why not, then, apply a method that can extend those ideas to combine both, returns and covariances, whenever is possible? Luckily, DMA is designed to do exactly that!

The text continues as follows: next section reports information about the dataset; section 3 details the key features of the models used; section 4 shows the results; and section 5 contains the final remarks.

2 DATA

The dataset used is based on the term structure of Brazilian interbank deposits (DI) starting on 2005-01-31 going up to 2017-06-30. The data is computed monthly and relies on futures contracts effectively traded at the B3 Exchange. Each contract has a nominal value of R$ 100.000,00 and a minimum of 5 contracts must be traded to open (or close) a new position. There is a total of 14 maturities that goes from a minimum of 3 months up to 5 years distributed in the following maturity set $\{3,6,9,12,15,18,21,24,27,30,36,42,48,60\}$.

As prices and rates often go on the opposite direction in fixed-income markets, to open a long position the investor must sell the interbank deposit rate, and to go short, the interest rate must be bought. A daily adjustment with the net position of each counterpart follows daily, as in most marketplaces, to minimize insolvency risk. With more than 2 billion traded daily, slippage can usually be managed properly, especially for strategies that cannot be labeled as high-frequency.

An important feature of fixed income markets is that asset prices can be recovered from the yield curve and transformed into a time series of zero coupon bonds. Simply use the formula:

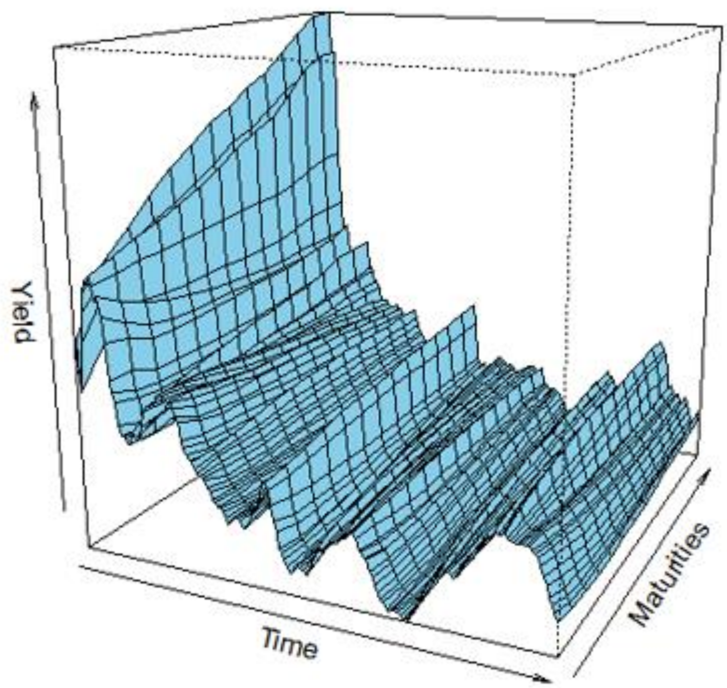
$$P_t = \frac{1}{(1 + y_t)^m} \tag{1}$$

in which P_t is the price of the zero-coupon bond at time t; y_t is the current yield; m is the bonds' maturity expressed in years; a final assumption implies that each asset has a value of R$1 when issue is due. To see why this makes sense, we need to remind ourselves that the yield of a bond is a measure of the average rate of return that will be earned, if the bond is bought at time t and held until maturity. Thus, bringing the traded price to the present, through the discount function, $(1 + y_t)^{-m}$, should result in a zero-coupon bond.

It's important to emphasize that the broad acceptance of the term structure as a stochastic system, instead of a deterministic process, it's a recent twist. This new rationale follows from the fact that the yield to maturity can generate a poor estimate if (1) expected returns of interest rates are volatile; (2) the yield curve is steeply sloped (either upper or downward); (3) there is a significant risk of default. In all of those cases the relevant return is the realized return.

The interest rate dynamics can be visualized on the image that follows.

Figure 1- Brazilian Interest-Rate Dynamics



Source: Bloomberg. Written by the author.

First, there is a downward trend in interest rates for all maturities. Second, yields and maturities have, most often than not, a positive correlation. This could be due to increasing risk-premium for long-term landing, as stated by the "Liquidity Preference Theory". Third, as the maturities go longer and longer, interest-rates become more volatile. For long-term maturities, 48 and 60 months, the annualized standard-deviation goes up to nearly 10%. Given their annualized sample return of 13%, uncertainty can be enormous.

The following table shows some descriptive statistics of bonds returns.

Table 1- Descriptive Statistics of Interest Rate Returns

Maturity (months)	Minimum	Median	Maximum	Standard Deviation	Skewness	Kurtosis
3	-0.0021	0.0000	0.0028	0.0009	0.3042	0.2544
6	-0.0045	0.0002	0.0066	0.0019	0.2657	0.5434
9	-0.0068	0.0007	0.0132	0.0031	0.3997	1.3328
12	-0.0108	0.0009	0.0203	0.0045	0.3408	1.8214
15	-0.0141	0.0011	0.0275	0.0061	0.2931	2.0876
18	-0.0196	0.0013	0.0369	0.0077	0.3331	2.9495
21	-0.0257	0.0016	0.0446	0.0093	0.2944	2.8759
24	-0.0346	0.0018	0.0544	0.0110	0.3027	3.6515
27	-0.0447	0.0020	0.0612	0.0127	0.0955	3.6569
30	-0.0518	0.0022	0.0680	0.0144	0.0899	3.3699
36	-0.0699	0.0027	0.0815	0.0177	-0.0547	3.7386
42	-0.0876	0.0031	0.0951	0.0209	-0.1246	3.9252
48	-0.1138	0.0036	0.1120	0.0247	-0.2475	4.7213
60	-0.1422	0.0067	0.1442	0.0314	-0.1361	4.5816

Source: written by the author.

Fixed-income returns seems to have a symmetric distribution, but as maturities increase, the empirical distribution smoothly moves from the right (positive skewed) to the left (negative skewed). There is also an increase in excess of kurtosis as the maturities expand.

2.1 BENCHMARKS

ANBIMA is the leading Brazilian institution for fixed-income appraisals. Most of its the records started only recently but different portfolios have been designed to fulfill investor's needs. From their data warehouse, two indexes seem appropriate for comparisons with the interest rates traded at B3 Exchange: IRF-M and IMA-B 5. IFR-M is an index based on a typical fixed-income instrument, where you have a pre-determined amount that is paid by the issuer and earned by an investor. IMA-B 5, on the other hand, is an index composed of Treasury Inflation Protected Securities (TIPS). This is a common instrument for investors that want to manage the risks of surprise in inflation. TIPS' principal adjusts upward along with consumer price inflation (IPCA), which provides investors with a guaranteed "real return" (or return net of inflation). In IMA-B 5, no asset can have more than 5 years until it is due.

Table 2 collapses five simple statistics of IMA-B 5 and IRF-M that are often used by practioners in the finance industry.

Table 2- Descriptive Statistics from the chosen Benchmarks

	IMA-B-5	IRF-M
Annualized Excess of Return	0.0168	0.0067
Annualized Standard Deviation	0.0346	0.0290
Annualized Sharpe-Ratio	0.4860	0.2304
Skewness	0.5573	-0.3415
Kurtosis	4.7906	3.2638

Source: written by the author.

3 THE OPTIMIZATION MODEL

The mean-variance framework allows us to build portfolios at each point in time based on returns and the risk appetite of each individual. Markowitz (2010) emphasizes that returns and variances must be estimated and obtained based on expected values, rather than on their historical performance. These estimates, however, are surrounded by uncertainty and do not always carry desirable properties Jorion (1986). Therefore, forecasting errors tend to be significant if the assumptions embedded in the estimation process are mispecified. According to Chopra et al. (2011), the "optimal" solution can be very sensitive to slight changes in inputs. For this reason, the section 3.1 briefly explores the theorical foundation supporting the usage of the quadratic utility function.

3.1 RISK AVERSION FORMULATION

The utility maximization problem can be solved by the quadratic utility formulation, usually written as

$$\begin{aligned} max \quad & w'\boldsymbol{\mu} - \zeta(w'\boldsymbol{\Sigma}w) \\ s.t. \quad & \sum_{i=1}^{N} w_i = 1 \\ & w_i \geq 0 \end{aligned} \tag{2}$$

in which $w = (w_1, w_2, w_3, \ldots, w_n)$ is the vector of weights invested in each asset; $\boldsymbol{\mu}$ is a NxM vector of expected returns for each maturity; $\boldsymbol{\Sigma}$ is a covariance matrix of dimension MxM; and ζ is the investor's (subjective) coefficient of risk aversion. When ζ is low, the penalty incurred by the additional amount of risk is also low. On the other hand, when ζ is high, the portfolio is highly penalized for an increased exposure to risk. A common practice is to calibrate the value of ζ in order to fit the portfolio into the investor desired level of risk (Fabozzi et al., 2007). At the present work, the value of $\zeta = 1$ is settled.

The quadratic utility function is important for theoretical and practical reasons. The theoretical perspective was well explored by Levy and Markowitz (1979) and Michaud, and Michaud (2008). As they point out, the quadratic utility is (locally) a good approximation for investors utility. Markowitz (2010) also shows that this formulation is well suited for asset returns that lie in the range of R= {-30%, +40%}. Moreover, investment companies routinely estimate potential returns for assets they cover. It would be naive to assume that this information is not brought into account in the decision-making process.

The benefits of incorporating constraints in the mean-variance optimization, on the other hand, is a topic that was initially explored by Frost and Savarino (1988). But it was in the work of Jagannathan and Ma (2003) that we found a robust support for its use. The authors argue that constraining portfolio weights to be positive is equivalent to reducing assets covariance by specific amount (that is, it has a "shrinkage" effect). By the same token, upper and lower bounds act to adjust small covariances upwards meanwhile reducing those that are considered (relatively) too high. In an attempt to benefit of those insights a second specification will also be used by adding the constraint of $w_i \leq 0.30$ into equation (2).

3.2 MOMENTS ESTIMATION

There is a good evidence showing that it's difficult to consistently overcome the Random-Walk forecasts, Diebold and Li (2006), Timmermann and Granger (2004), Demiguel and Uppal (2007). Therefore, simple models with few parameters to estimate arise as a natural starting point for predicting asset returns. Those will be called "Competitive Models" and are going to be used as a benchmark against 25 TVP-VAR specifications.

3.2.1 Random-Walk Model

In the simplest random walk process each successive change in y_t is drawn independently from a probability distribution with 0 mean. Thus, y_t is determined by

$$y_t = y_{t-1} + \epsilon_t \tag{3}$$

with $E(\epsilon_t) = 0$ and $E(\epsilon_t, \epsilon_{t-1}) = 0$ for all $t \neq (t-1)$. Such a process could be generated by successive flips of a coin, where the heads receive a value of 1 and a tail receives a value of -1.

The forecast for such a model is given by

$$y_{t+1} = y_t \tag{4}$$

Similarly, the forecast $t + T$ steps ahead is also y_t.

For the variance the sample estimator is used

$$\Sigma = \frac{1}{1 - N} y_t y_t{}' \tag{5}$$

3.2.2 Autoregressive Model (AR)

This is the famous AR (1) model. It has the main feature of allowing for mean-reverting and is described by the following format

$$y_t = \phi y_{t-1} + \epsilon_t \tag{6}$$

In which ϕ is the short-term impact of y_{t-1} in y_t, while the long-term average is given by $1/(1 - \phi)$. The forecast one step ahead is

$$y_{t+1} = \phi y_t \tag{7}$$

For the variance the sample estimator is used

$$\Sigma = \frac{1}{1 - N} y_t y_t{}' \tag{8}$$

3.2.3 Vector Autoregressive with Time-Varying Parameters (TVP-VAR)

This model can be written in the state-space model as

$$\begin{aligned} y_t &= Z_t \beta_t + \epsilon_t \\ \beta_t &= \beta_{t-1} + u_t \end{aligned} \tag{9}$$

in which y_t is a vector with the variables we want to explain; Z_t is a diagonal M x K matrix with p lags. Z_t contains M variables plus the intercept, so $K = (1 + pM)$; β is a K x K state vector that contains the time-varying parameters; ϵ_t is i.i.d. $N \sim (0, \Sigma_t)$, u_t is $N \sim (0, Q_t)$ and $cov(\epsilon, u) = 0, \forall t$.

To avoid burdening computations problems Koop and Korobilis (2013) suggests a new method to approximate Q_t and Σ_t values. To track closely the behavior of Q_t, a forgetting factor is added into the Kalman's filter as it was previously suggested by Raftery and Ettler (2010). The predicted covariance estimate of the Kalman Filter has a closed format and is normally written as

$$V_{t-1} = Z_t V_{t-1} Z_t{}' + Q_t \tag{10}$$

Koop and Korobilis (2013) suggests that Q_t can be substituted by

$$V_t = \frac{1}{\lambda} V_{t-1} \tag{11}$$

and does not need to be estimated any longer. A simple algebra shows that $Q_t = (\lambda^{-1} - 1)V_{t-1}$. Clearly, we can control the magnitude of the shocks that impact Z_t by adjusting λ instead of directly estimating Q_t. If $\lambda = 1$ the constant VAR model emerge, otherwise parameters will be time-varying. For instance, when $\lambda = 0.99$, in the context monthly data, observations five years ago will receive approximately 55% as much weight as last period's observation, which corresponds to smooth time-variation in Z_t. When $\lambda = 0.96$, observations 5 years into the past receive only about 8% as much weight as last period's observations, suggesting that relatively larger shocks hit the VAR coefficients. Evidently, while this procedure can be used to "train" the model, the increased variability in Z_t also results in higher variance prediction. As a consequence, estimating the main equation depends not only on the choice of the predictors in Z_t, but also the choice of λ.

At time $t - 1$, β have a distribution equal to

$$\beta_{t-1} \sim N(\beta_{t-1}, V_{t-1}) \tag{12}$$

in which β_{t-1} is the filtered state at time $t - 1$ and V_{t-1} the variance at the same point in time.

The matrix Σ is an approximation made through the Exponentially Weighted Moving Average model (EWMA), widely used in finance. EWMA is expressed as

$$\hat{\Sigma}_t = \kappa \hat{\Sigma}_{t-1} + (1 - \kappa) \hat{\epsilon}_t \hat{\epsilon}_t{}' \tag{13}$$

The residual of the measurement equation is $\hat{\epsilon}_t = y_t - \beta_t Z_t$ and can be obtained through the Kalman-Filter. Popularized by Morgan (1996), this is a multivariate GARCH that has a constant decay factor equal to κ. The decay parameter *kappa* could be interpreted as volatility persistence: the higher the κ, more weight is set into recent observations. On the other hand, $(1 - \kappa)$ is usually viewed as a metric of reaction to market events. The lower the κ, the higher is $(1 - \kappa)$ and more pressure is put into impredictable shocks.

Finally, it worth mention a word about the prior behavior. Its assumed that $E(\beta_0) = 0$, which is not too far from reality, since asset returns are usually seen as stationary. The prior construction is motivated on the work of Litterman and Sims (1984), with minor changes to let online estimation and forecasting to be carried forward. The variance has a dimension of K x M, which contains, in the principal diagonal, elements of the hyper-parameter γ that controls the degree of shrinkage on the VAR coefficients. Koop and Korobilis (2013) set $\gamma = \{10^{-10}, 10^{-5}, 0.001, 0.005, 0.01, 0.05, 0.1\}$. Given the differences between short and long-term maturities this set is expanded to $\gamma = \{10^{-10}, 10^{-5}, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 3, 5, 10\}$.

In short, only three parameters are unknown: λ, κ, β. If they are designed to stay within a pre-specified range, however, the recursive nature of the Kalman-Filter takes care of matrix Z_t and no parameter need to be estimated, which whips out the "dimensionality-curse". A common practice is to initially set $\beta_0 = 1$ to emulate a random-walk behavior. For κ and λ the following values are used: $\kappa = \{0.96, 0.97, 0.98, 0.99, 1\}$ and $\lambda = \{0.96, 0.97, 0.98, 0.99, 1\}$. There is one option for β_0, five for κ and five for λ, totaling 25 possible combinations for the TVP-VAR specifications (1x5x5).

3.3 DYNAMIC MODEL AVERAGING (DMA)

The objective of Dynamic Model Averaging is to calculate the probability that a given model should be used in forecasting, given set of explanatory variables. In simple terms, DMA uses the information available at $t - 1$ to predict a set of selected variables each point in time. The predictive density is then used to calculate the probability that each model has of been the best. The optimal forecast is constructed weighting the different models by their past performance. Mathematically:

$$\pi_t \propto \prod_{i=1}^{t-1} [\, p_j * y_{t-i}]^{\eta} \tag{14}$$

Thus, model j will receive more weight at time t if its predictive density, p_j, is high, and the opposite is also true. Following recommendations of Koop and Korobilis (2013) and Raftery and Ettler (2010) the value of η is set at 0.99.

4 EMPIRICAL RESULTS

To evaluate model's accuracy its necessary first to define appropriate tools for forecast comparison. Traditionally, the manager is interested not only in the point forecast, but also in the magnitude of its expected variance. Some tension will always exist between these two, since its not always possible to reduce one without relaxing the other. Therefore, the Mean Squared Forecasting Error (MSFE) and Mean Absolute Forecasting Error (MAFE) stands as natural evaluation metrics in most papers in applied finance. MSFE is defined as

$$RMFE = (\sum_{i=1}^{N} \frac{(y_i - \hat{y}_i)^2}{N})^{1/2} \tag{15}$$

in which y_i is the observed time-series; $\hat{y}_i$ is the value forecasted by the model at hand; and N is the number of out-of-sample evaluations. For MAFE, on its turn, is labeled as

$$MAFE = \sum_{i=1}^{N} \frac{|(y_i - \hat{y}_i)|}{N} \tag{16}$$

For portfolio comparisons the out-of-sample Sharpe-Ratio (SR) is computed. Sharpe-Ratio is characterized by the equation

$$SR = \frac{\frac{1}{T-1}[\sum_{i=1}^{N-1}(w_t R_{t+1}) - CDI_{t+1}]}{\frac{1}{T-1}[\sum_{i=1}^{N-1}(w_t R_{t+1} - \hat{\mu})]} \tag{17}$$

which can be clearly simplified to $SR = \frac{\hat{\mu}_{t+1}}{\hat{\sigma}_{t+1}}$.

The first 60 periods of data are used to train the model and forecasting comes after recursively. The out-of-sample forecasts are constructed thought direct forecasting, as described in Marcellino et al. (2006), starting in jan-2010. Portfolios are rebalanced monthly. In each period, 2/3 of the sample (that is, 42 of the 60 realizations) are used to adjust the Minnesota-prior, and only then the recursive nature of the Kalman-Filter starts. As informed earlier, the parameters are set to obey the following sets $\kappa = \{0.96, 0.97, 0.98, 0.99, 1\}$ and $\lambda = \{0.96, 0.97, 0.98, 0.99, 1\}$. This yields a total of 25 TVP-VAR models to be tested.

Table 3 shows the accuracy ratios of each TVP-VAR-DMA against the Random-Walk, for different values of λ and κ.

Table 3 - Out-of-Sample Accuracy Measures

Model	λ	κ	RMFE	MAFE
Random-Walk	-	-	1.0000	1.0000
TVP-VAR-DMA	0.96	1.00	1.2180	1.1501
TVP-VAR-DMA	0.97	1.00	1.2208	1.1506
TVP-VAR-DMA	1.00	0.96	1.2212	1.1376
TVP-VAR-DMA	0.98	1.00	1.2262	1.1520
TVP-VAR-DMA	0.99	0.96	1.2289	1.1437
TVP-VAR-DMA	0.99	1.00	1.2307	1.1529
TVP-VAR-DMA	1.00	1.00	1.2328	1.1530
TVP-VAR-DMA	1.00	0.97	1.2328	1.1458
TVP-VAR-DMA	0.98	0.96	1.2334	1.1468
TVP-VAR-DMA	0.99	0.97	1.2350	1.1476
TVP-VAR-DMA	1.00	0.98	1.2358	1.1490
TVP-VAR-DMA	0.99	0.98	1.2388	1.1516
TVP-VAR-DMA	0.97	0.97	1.2393	1.1530
TVP-VAR-DMA	0.99	0.99	1.2394	1.1531
TVP-VAR-DMA	1.00	0.99	1.2396	1.1524
TVP-VAR-DMA	0.98	0.97	1.2397	1.1519
TVP-VAR-DMA	0.97	0.96	1.2397	1.1520
TVP-VAR-DMA	0.96	0.96	1.2416	1.1538
TVP-VAR-DMA	0.97	0.98	1.2417	1.1553
TVP-VAR-DMA	0.98	0.98	1.2419	1.1543
TVP-VAR-DMA	0.96	0.97	1.2426	1.1556
TVP-VAR-DMA	0.98	0.99	1.2426	1.1560
TVP-VAR-DMA	0.96	0.98	1.2429	1.1572
TVP-VAR-DMA	0.97	0.99	1.2435	1.1572
TVP-VAR-DMA	0.96	0.99	1.2436	1.1584
AR (1)	-	-	1.3162	1.1737

Source: written by the author.

In table 3 above, each $MSFE_{TVP-VAR}$ and $MAFE_{TVP-VAR}$ was divided by the MSFE and MAFE of the Random-Walk model. That is

$$RMSFE_i = \frac{MSFE_{TVP-VAR}}{MSFE_{RW}} \tag{18}$$

and

$$RMAFE_i = \frac{MAFE_{TVP-VAR}}{MAFE_{RW}} \tag{19}$$

In this case, if either the $RMSFE_i$ or the $RMAFE_i$ are higher than 1, the respective model displays a more erratic and unstable out-of-sample forecasts compared to the Random-Walk. As we can see, there were an average dilution of 24%, in terms of MSFE, and 15% in terms of MAFE. Its interesting to note that the average forecast errors do not seems to be too sensitive to changes in κ. Possibly, the high level of colinearity among maturities makes it harder to extract relevant information from return and covariance matrices. For λ, the same pattern arises.

The portfolio returns are evaluated and compared to the Brazilian risk-free rate (CDI), which is the most widely used benchmark interest-rate in fixed-income and hedge-fund industries. The statistics that follow on table 4 were constructed (and annualized) based on portfolios excess of return.

Table 4 - Mean-Variance Optimization Statistics

Model	λ	K	Return	Standard Deviation	Sharpe	Average Maturity
Random-Walk	-	-	0.0019	0.0220	0.0876	5.3692
TVP-VAR-DMA	0.98	0.99	-0.0043	0.0094	-0.4591	7.7216
TVP-VAR-DMA	1.00	0.96	-0.0043	0.0094	-0.4608	4.7312
TVP-VAR-DMA	1.00	1.00	-0.0040	0.0090	-0.4409	6.1217
TVP-VAR-DMA	1.00	0.98	-0.0039	0.0090	-0.4394	4.7302
TVP-VAR-DMA	0.96	0.99	-0.0045	0.0086	-0.5286	4.5410
TVP-VAR-DMA	0.98	0.97	-0.0045	0.0086	-0.5253	5.7948
TVP-VAR-DMA	0.96	0.97	-0.0049	0.0085	-0.5763	7.9669
TVP-VAR-DMA	1.00	0.99	-0.0044	0.0084	-0.5253	5.1827
TVP-VAR-DMA	0.97	0.96	-0.0048	0.0084	-0.5714	3.7818
TVP-VAR-DMA	0.97	0.99	-0.0047	0.0081	-0.5810	5.1699
TVP-VAR-DMA	0.98	0.98	-0.0047	0.0081	-0.5799	5.1958
TVP-VAR-DMA	0.96	0.98	-0.0049	0.0081	-0.6072	5.1269
TVP-VAR-DMA	0.99	0.97	-0.0047	0.0081	-0.5867	7.2608
TVP-VAR-DMA	0.99	1.00	-0.0049	0.0080	-0.6050	9.0805
TVP-VAR-DMA	0.97	1.00	-0.0047	0.0080	-0.5851	6.2235
TVP-VAR-DMA	0.96	1.00	-0.0048	0.0080	-0.6030	5.2346
TVP-VAR-DMA	0.99	0.99	-0.0047	0.0080	-0.5907	4.5176
AR (1)	-	-	-0.0045	0.0079	-0.5683	7.3117
TVP-VAR-DMA	0.97	0.98	-0.0049	0.0079	-0.6254	5.4120
TVP-VAR-DMA	0.98	0.96	-0.0048	0.0079	-0.6052	4.0277
TVP-VAR-DMA	0.99	0.98	-0.0051	0.0076	-0.6717	4.5274
TVP-VAR-DMA	1.00	0.97	-0.0051	0.0076	-0.6747	4.0727
TVP-VAR-DMA	0.97	0.97	-0.0050	0.0076	-0.6632	6.3558
TVP-VAR-DMA	0.98	1.00	-0.0050	0.0075	-0.6660	7.2257
TVP-VAR-DMA	0.96	0.96	-0.0052	0.0074	-0.6986	9.1665
TVP-VAR-DMA	0.99	0.96	-0.0054	0.0071	-0.7585	7.2391

Source: written by the author.

Table 4 is sorted from the highest to the smallest Sharpe ratio. Stands out that most portfolios have a similar shape (exception made to the Random-Walk portfolio, that showed higher returns and variances). The end up results rested upon very short maturities, which can be seen by the last column. Indeed, returns from the spot rates increase only marginally as we go from short to the long maturities. The average annualized sample return of the 3-month rate diverged from the 5 year's rate by only 2.20%, while the annualized sample standard deviation increased by more than 9%. This can partially explain why the Mean-Variance is soo insensitive to the TVP-VAR specifications and oriented toward short-term rates.

The table 5 shows the results for the portfolio optimization when the restriction $w_i \leq 0.30$ is incorporated into the utility function presented in equation (2). As expected, there is an increase in average maturities holdings. But most important, Sharpe ratios, although still negative, have a significant improvement. The bulk of the difference comes from standard deviations, that for most portfolios, had a symbolic increase.

Table 5 - Mean-Variance Optimization Statistics
(Weights constrained to be between 0 and 0.30)

Model	λ	κ	Return	Standard Deviation	Sharpe	Average Maturity
Random-Walk	-	-	0.0001	0.0208	0.0045	9.5294
TVP-VAR-DMA	1.00	0.98	-0.0008	0.0173	-0.0441	8.3131
TVP-VAR-DMA	0.98	0.99	-0.0012	0.0170	-0.0681	7.8031
TVP-VAR-DMA	0.96	0.96	-0.0010	0.0169	-0.0595	7.1656
TVP-VAR-DMA	0.98	1.00	-0.0010	0.0168	-0.0589	8.1051
TVP-VAR-DMA	0.97	0.97	-0.0012	0.0168	-0.0721	7.4121
TVP-VAR-DMA	0.99	0.99	-0.0011	0.0167	-0.0644	8.3156
TVP-VAR-DMA	0.97	0.98	-0.0010	0.0166	-0.0614	7.7259
TVP-VAR-DMA	1.00	1.00	-0.0013	0.0163	-0.0827	7.1611
TVP-VAR-DMA	0.96	1.00	-0.0005	0.0162	-0.0320	8.3566
TVP-VAR-DMA	0.96	0.99	-0.0012	0.0162	-0.0751	9.9417
TVP-VAR-DMA	0.98	0.97	-0.0013	0.0162	-0.0794	9.0277
TVP-VAR-DMA	1.00	0.96	-0.0014	0.0162	-0.0880	8.1087
TVP-VAR-DMA	0.99	0.96	-0.0010	0.0162	-0.0635	8.0369
TVP-VAR-DMA	0.96	0.98	-0.0013	0.0159	-0.0806	7.2437
TVP-VAR-DMA	0.99	0.97	-0.0017	0.0159	-0.1059	8.9592
TVP-VAR-DMA	0.96	0.97	-0.0014	0.0159	-0.0854	9.8886
TVP-VAR-DMA	0.97	0.99	-0.0014	0.0158	-0.0901	7.7316
TVP-VAR-DMA	1.00	0.97	-0.0014	0.0158	-0.0881	7.6648
TVP-VAR-DMA	0.98	0.96	-0.0014	0.0157	-0.0922	9.0277
TVP-VAR-DMA	0.99	1.00	-0.0016	0.0145	-0.1072	9.1144
TVP-VAR-DMA	0.99	0.98	-0.0018	0.0145	-0.1244	9.8176
TVP-VAR-DMA	0.97	1.00	-0.0019	0.0143	-0.1315	7.1088
TVP-VAR-DMA	1.00	0.99	-0.0020	0.0143	-0.1395	7.1081
TVP-VAR-DMA	0.98	0.98	-0.0021	0.0139	-0.1524	7.6165
AR (1)	-	-	-0.0020	0.0136	-0.1490	8.2865
TVP-VAR-DMA	0.97	0.96	-0.0025	0.0130	-0.1914	8.3638

Source: written by the author.

A curious reader might raise the question of whether or not the results would be distinct under a different parametrization set for ζ and w_i. It happens that, for this specific dataset, the answer could be "not much" and worth spending a few words on why this is the case. The parameter ζ that controls the amount of penalization included in the objective function was very insensitive to alternative specifications, $\zeta = \{2, 3, 4, 5\}$. This can be explained by the fact that

$\zeta = 1$ is already considered too loose, so increasing ζ did not caused any major changes in the end up result. Stated in a different way, the increase in ζ tended to push portfolios towards to a more restrictive point in the optimum hyperplane without an equivalent compensation in terms of Sharpe Ratio. Obviously, that is the opposite of what we want. The weights vector, on the other hand, were also tested for $w_i = \{0.20, 0.40, 0.50\}$ with results almost completely governed by the asymmetric relationship within risk and return. As mentioned above, the increase in return does not offset the volatility expansion as we go from short term maturities to more extended ones. As an immediate consequence, when restriction is set to be too tight, the weight vector is always bidding, which effectively defines the optimization. For obvious reasons, this is also an undesirable result.

From the period of 2010-jan until 2017-jun, when the out-of-sample portfolios were generated, the IFR-M and IMA-B 5 had a cumulative excess of return over the CDI of 14.25% and 5.48%, respectively. Given that the TVP-VAR-DMA portfolios delivered an annualized near zero cumulative excess of return, those two values can be viewed as the total underperformance of TVP-VAR-DMA under the tested period.

5 FINAL CONSIDERATIONS

Brazilian real interest-rates are among the highest in the world. It does not come as a surprise, then, that most brazilian asset managers tilt their wealth into fixed-income streams. Nevertheless, interest-rates have been falling during the last 15 years and no manager can neglect the challenges that may arise in a lower and more stable interest-rates environment. This is notably true for retirement funds and insurance companies that hold long-term liabilities covered by short-term assets.

For this reason, this work approached fixed-income management from a singular point of view by applying a TVP-VAR-DMA model in a Mean-Variance context. The TVP-VAR model designed by Koop and Korobilis (2013) carries interesting properties that can be valuable for quantitative asset managers. Among several features its possible to highlight (i) the possibility to knock-out the “Dimensionality Course” by the use of forgetting-factors; (ii) the possibility to lessen Model Uncertainty thought DMA.

The first allows the Kalman-filter to be estimated without the need to appeal to MCMC algorithms, which is computely efficient. The second is calibrated by the posterior probability density, but it can also be regulated by a grid of values in the hyper-parameter γ.

The end results show that, neither the that portfolios constructed using TVP-VAR-DMA, Autoregressive Model, or the Random-Walk, were able to beat the selected fixed-income benchmarks, IRF-M and IMA-B 5, on a risk-adjusted basis. Yet, the DMA out-of-sample forecasts indicate that the features designed by Koop and Korobilis (2013) may help investors to sucesfully extract information from economic data avaiable, as showed in Caldeira, Moura and Santos (2015) and Dangl and Halling (2012).

REFERENCES

AIOLFI, Marco; TIMMERMANN, Allan. Persistence in forecasting performance and conditional combination strategies. **Journal of Econometrics**, v. 135, n. 1, p. 31-53, 2006.

BEST, Michael J.; GRAUER, Robert R. The analytics of sensitivity analysis for mean-variance portfolio problems. **International Review of Financial Analysis**, v. 1, n. 1, p. 17-37, 1992.

CALDEIRA, João et al. Combining multivariate volatility forecasts: an economic-based approach. **Journal of Financial Econometrics**, v. 15, n. 2, p. 247-285, 2017.

CALDEIRA João, F.; MOURA Guilherme V.; SANTOS, André A. P. Yield curve forecast combinations based on bond portfolio performance. **Journal of Forecasting**, 2017.

CALDEIRA João, F.; MOURA Guilherme V.; SANTOS, André A. P. Previsões macroeconômicas baseadas em modelos TVP-VAR: evidências para o Brasil. **Revista Brasileira de Economia**, v. 69, n. 4, p. 407-428, 2015.

CALDEIRA João, F.; MOURA Guilherme V.; SANTOS, André A. P. Bond portfolio optimization using dynamic factor models. **Journal of Empirical Finance**, v. 37, p. 128-158, 2016.

CATANIA, Leopoldo; NONEJAD, Nima. **Dynamic model averaging for practitioners in economics and finance: The eDMA package**. [S.l.]. 2016.

CHAN, Joshua et al. Time varying dimension models. **Journal of Business & Economic Statistics**, v. 30, n. 3, p. 358-367, 2012.

CHOPRA, Vijay et al. The effect of errors in means, variances, and covariances on optimal portfolio choice. **The Kelly Capital Growth Investment Criterion: Theory and Practice**, v. 3, p. 249, 2011.

CHRISTIANSEN, Charlotte; SCHMELING, Maik; SCHRIMPF, Andreas. A comprehensive look at financial volatility prediction by economic variables. **Journal of Applied Econometrics**, v. 27, n. 6, p. 956-977, 2012.

DANGL, Thomas; HALLING, Michael. Predictive regressions with time-varying coefficients. **Journal of Financial Economics**, v. 106, n. 1, p. 157-181, 2012.

DEMIGUEL, Victor; GARLAPPI, Lorenzo; UPPAL, Ramam. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? **The review of Financial studies**, v. 22, n. 5, p. 1915-1953, 2007.

DIEBOLD, Francis X.; LI, Canlin. Forecasting the term structure of government bond yields. **Journal of econometrics**, v. 130, n. 2, p. 337-364, 2006.

FABOZZI, Frank J. et al. Robust portfolio optimization. **The Journal of Portfolio Management**, v. 33, n. 3, p. 40-48, 2007.

FROST, Peter A.; SAVARINO, James E. For better performance: Constrain portfolio weights. **The Journal of Portfolio Management**, v. 15, n. 1, p. 29-34, 1988.

GRANGER, Clive W. J.; RAMANATHAN, Ramu. Improved methods of combining forecasts. **Journal of forecasting**, v. 3, n. 2, p. 197-204, 1984.

HENDRY, David F.; CLEMENTS, Michael P. Pooling of forecasts. **The Econometrics Journal**, v. 7, n. 1, p. 1-31, 2004.

JAGANNATHAN, Ravi; MA, Tongshu. Risk reduction in large portfolios: Why imposing the wrong constraints helps. **The Journal of Finance**, v. 58, n. 4, p. 1651-1683, 2003.

JOBSON, David J.; KORKIE, Robert M. Putting Markowitz theory to work. **The Journal of Portfolio Management**, v. 7, n. 4, p. 70-74, 1981.

JORION, Philippe. Bayes-Stein estimation for portfolio analysis. **Journal of Financial and Quantitative Analysis**, v. 21, n. 3, p. 279-292, 1986.

JORION, Philippe. Bayesian and CAPM estimators of the means: Implications for portfolio selection. **Journal of Banking & Finance**, v. 15, n. 3, p. 717-727, 1991.

KOOP, Gary; KOROBILIS, Dimitris. Large time-varying parameter VARs. **Journal of Econometrics**, v. 177, n. 2, p. 185-198, 2013.

LEVY, Haim; MARKOWITZ, Harry. Approximating expected utility by a function of mean and variance. **The American Economic Review**, v. 69, n. 3, p. 308-317, 1979.

LEVY, Moshe; LEVY, Haim. For Better Performance: Constrain Portfolio Weights Differentially and Globally, 2014.

LITTERMAN, Robert; SIMS, Christopher. Forecasting and conditional projections using a realistic prior distribution. **Econometric Reviews**, v. 3, p. 1-100, 1984.

LITTERMAN, Robert; SCHEINKMAN, Jose. Common factors affecting bond returns. **Journal of Fixed-Income**, v. 1, n. 1, p. 54-61, 1991.

MARCELLINO, Massimiliano et al. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. **Journal of econometrics**, v. 135, n. 2, p. 499-526, 2006.

MARKOWITZ, Harry. Portfolio selection. **The journal of finance**, v. 7, n. 1, p. 77-91, 1952.

MARKOWITZ, Harry. Portfolio theory: as I still see it. **Annu. Rev. Financ. Econ.**, v. 2, n. 1, p. 1-23, 2010.

MICHAUD, Richard. O. The Markowitz optimization enigma: Is 'optimized'optimal? **Financial Analysts Journal**, v. 45, n. 1, p. 31-42, 1989.

MICHAUD, Richard O.; MICHAUD, Robert O. **Efficient asset management:** a practical guide to stock portfolio optimization and asset allocation. [S.l.]: Press, Oxford University, 2008.

MORGAN, J. **Riskmetrics technical document**. JP Morgan. [S.l.]. 1996.

OPSCHOOR, Anne et al. Improving Density Forecasts and Value-at-Risk Estimates by Combining Densities. **Tinbergen Institute Discussion Paper**, 2014.

RAFTERY, Kárný, ETTLER, Pavel. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. **Technometrics**, v. 52, n. 1, p. 52-66, 2010.

TIMMERMANN, Allan. Forecast combinations. **Handbook of economic forecasting**, v. 1, p. 135-196, 2006.

TIMMERMANN, Allan; GRANGER, Clive W. J. Efficient market hypothesis and forecasting. **International Journal of forecasting**, v. 20, n. 1, p. 15-27, 2004.

VIEIRA, Fausto et al. Forecasting the Brazilian yield curve using forward-looking variables. **International Journal of Forecasting**, v. 33, n. 1, p. 121-131, 2017.