Prediction and Allocation of Stocks, Bonds, and REITs in the US Market

Ana Sofia Monteiro[1] · Helder Sebastião[1] · Nuno Silva[1]

Abstract

This study employs dynamic model averaging and selection of Vector Autoregressive and Time-Varying Parameters Vector Autoregressive models to forecast out-of-sample monthly returns of US stocks, bonds, and Real Estate Investment Trusts (REITs) indexes from October 2006 to December 2021. The models were recursively estimated using 17 additional predictors chosen by a genetic algorithm applied to an initial list of 155 predictors. These forecasts were then used to dynamically choose portfolios formed by these assets and the riskless asset proxied by the 3-month US treasury bills. Although we did not find any predictability in the stock market, positive results were obtained for REITs and especially for bonds. The Bayesian-based approaches applied to just the returns of the three risky assets resulted in portfolios that remarkably outperform the portfolios based on the historical means and covariances and the equally weighted portfolio in terms of certainty equivalent return, Sharpe ratio, Sortino ratio and even Conditional Value-at-Risk at 5%. This study points out that Constant Relative Risk Averse investors should use Bayesian-based approaches to forecast and choose the investment portfolios, focusing their attention on different types of assets.

✉ Ana Sofia Monteiro
asmm@fe.uc.pt

1   CeBER, Faculty of Economics, University Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal

🖄 Springer

1 Introduction

Return predictability and portfolio selection are two of the most relevant topics in finance. The ability to forecast returns has important implications for portfolio asset allocation, as highlighted by Campbell et al. (2003) and McMillan (2021). The possibility of return predictability motivates the use of more robust estimation techniques and predictor selection to detect and use all the pertinent information available in the data. Most notably because, as it is well known, financial markets are subject to permanent shocks and structural breaks, and experience volatility clustering. Hence, well-defined investment strategies should use flexible forecasting models that accommodate those features of financial data.

Recent research has analyzed the predictability of stock and bond returns by considering several macro and microeconomic variables, such as inflation rates (Ludvigson & Ng, 2009), interest rates (Bandi et al., 2019), macroeconomic attention indices (Ma et al., 2022a), geopolitical risk (Ma et al., 2022b), climate change news risk (Huynh & Xia, 2020), financial stress index (Xu et al., 2023), valuation ratios, and dividend-price ratio (Cochrane, 2007; Golez & Koudijs, 2018). Typically, these studies rely on specific sets of predictors to forecast multiple-asset returns (Gao & Nardari, 2018). For instance, Welch and Goyal (2008) and Rapach et al. (2010) consider fifteen predictors, most of them financial variables. Hjalmarsson (2010) use four financial variables (stock price ratios and interest rates) to predict excess stock returns in 14 countries, Golez and Koudijs (2018) only use two valuation ratios as predictors of stock returns in four countries, and Zhang et al. (2019) only apply the short interest index and the aligned investor sentiment to forecast the returns of the S&P500. There is no consensus on the predictor space, but for sure its definition is of utmost importance since inadequate predictors reduce the predictive ability and, consequently, the performance of asset allocation strategies devised upon those models.

Traditionally, research on return predictability has mainly been building up in-sample empirical evidence. However, more recent literature has highlighted the power and robustness of out-of-sample analyses (Fisher et al., 2020; Welch & Goyal, 2008). The debate between the advantages and disadvantages of in-sample versus out-of-sample analyses has focused on different aspects, such as data snooping, data mining, spurious regressions, and instability in return predictability (Wu et al., 2013; Dichtl et al., 2021). This paper adopts an out-sample analysis, as we believe that it provides more reliable insights.

There is an extensive body of empirical literature comparing the predictive power of different models. Typically, asset returns are forecasted within a Vector Autoregressive (VAR) framework (Guidolin & Hyde, 2012). Despite its popularity, VAR models entail the danger of over-parameterization, leading to unreliable predictions. Nowadays, the toolbox of applied econometrics includes numerous modeling and forecasting tools to prevent the proliferation of parameters and reduce the number of parameters and model uncertainty. Amongst these tools, stand out time-varying parameters models, forecast combinations,

model averaging, and model selection techniques, which have been fueled by the noticeable advances in computational power. Although the existing literature proposes numerous extensions to VAR models, still there is no consensus on which framework is the best setup to forecast multiple assets. According to Fisher et al. (2020), not a single feature alone, but an ensemble of them, is required to handle the uncertainty and instability of financial markets, aiming at making good predictions. Hence, in line with the existing literature, this study analyses different model specifications and features, such as time-varying parameters, model/forecast combinations, and dynamic model selection/averaging to jointly obtain dynamic forecasts of three risky asset classes out-of-sample and use these forecasts to lead the investment.

Typically, the literature has concentrated on stocks and bonds. However, other assets, such as Real Estate Investment Trusts (REITs), have increasingly gained the interest of academics and practitioners. Recently, REITs have been seen as an alternative investment vehicle because they provide diversification benefits, improve the risk-return trade-off, and offer a non-negligible dividend income (see, for instance, Ling et al., 2020; Zhu & Lizieri, 2022). In this study, we focus on three classes of assets: Stocks, Bonds, and REITs. More precisely, the main goal of this study is to obtain the best portfolio of Stocks, Bonds, and REITs in the US from January 1976 to December 2021. The US is the world's largest economy, accounting for a quarter of the global Gross Domestic Product (GDP) and has the largest stock market capitalization in the world.[1] Besides that, we choose the US due to the availability of long and up-to-date time series and to have a higher level of comparability with most of the literature.

The applied forecasting schemes used in this study are in line with several papers, namely Banbura et al. (2010), Rapach et al. (2010), and Koop and Korobilis (2013). The methodological framework includes combinations of forecasts from conventional VARs, namely means and forecasts weighted by the Mean Squared Forecast Error (WMSFE) of the best models, and Bayesian Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA) with different sets of predictors. But this paper contributes to the existing literature in several aspects.

First, it includes REITs in the asset space. US REITs are firms that own or finance income-producing real estate across 13 property sectors. Investment into REITs can be performed through the purchase of individual company stocks, mutual funds, or exchange-traded funds (ETF). According to the NAREIT website (https://www.reit.com), currently "REITs of all types collectively own more than \$4.5 trillion in gross assets across the US, with public REITs owning around \$3 trillion in assets. Approximately 150 million (roughly 45%) American households invested in REITs." By September 2022, US-listed REITs had an equity market capitalization of more than \$1.4 trillion, almost 4% of the total capitalization of the S&P 500.

Second, the study begins with a large set of predictors that surpasses most of the ones used in related studies. Besides the lag returns of the assets, we additionally consider 155 predictors (19 for Stocks, 122 for Bonds, and 14 for

---

[1] World Bank Open Data Website (http://data.worldbank.org). Accessed on December 26, 2022.

REITs). A Genetic Algorithm (GA) is used to select the predictor space besides returns. Usually, the choice of the variables is made through ad hoc methods, which can potentially exclude pertinent variables. The GA search technique has been applied successfully in diverse optimization problems commonly associated with Machine Learning models to forecast returns and develop trading strategies (e.g., Bauer, 1994; Karathanasopoulos et al., 2016; Leigh et al., 2002; Ozcalici & Bumin, 2022). However, to the best of our knowledge, this is one of the first papers to apply GA to select the predictor space.

Third, the use of up-to-date data allows the study of the recent years (the out-of-sample period begins in October 2006), when the financial markets were characterized by abnormal turbulence due to the global financial crisis of 2008–2009 triggered by the US subprime and subsequent sovereign crises and to the COVID-19 pandemic (the World Health Organization's (WHO)) identified the outbreak as a Public Health Emergency of International Concern (PHEIC) on January 30, 2020, and a pandemic on March 11, 2020, (see, for instance, Ganie et al., 2022; Szczygielski et al., 2023).

Based on the literature, we believe that predictability is time-varying. Both the best predictor space and coefficients change over time, and therefore there is a need to update the forecasting models. Combining or selecting forecasts from different models may increase the forecasting accuracy and strengthen the forecasts against misspecification bias and measurement errors in the data (Timmermann, 2006). Thus, we expect from the beginning that the Bayesian frameworks will produce better results.

There are two main results of this study to highlight. First, although the forecasting accuracy is low, especially for the Stocks, the lagged returns of the S&P 500 present some ability to forecast the other markets. Second, the best portfolio strategy is the DMA/DMS with only the lagged returns in the predictor space, regardless of the level of risk aversion of a Constant Relative Risk Averse (CRRA) investor. Our results support the use of methodologies that incorporate dynamic modeling and parameter uncertainty and make use of combinations or selections of several models. Most importantly, highlight that the inclusion of other assets increases the overall predictability and improves the performance of the portfolios.

In a nutshell, we intend to provide additional evidence that may help investor and practitioners in devising reliable and robust investment strategies. The paper focus different forecasting models, provides a method for selecting the predictor space, and highlights the need to consider different types of assets.

The remaining of this study is structured into five sections. Section 2 presents a brief literature review. Section 3 describes and provides a preliminary analysis of the data. Section 4 outlines the basic theoretical concepts, the specifications of the models and presents the methodology to measure the forecast accuracy, construct the portfolios, and the metrics used to assess their performance. Section 5 shows the statistical and economic results obtained from different models. Finally, Sect. 6 highlights the main conclusions.

2 Literature Review

Throughout the years, several studies have analyzed the predictability of financial returns using different predictor spaces (e.g., Kothari & Shanken, 1997; Campbell & Shiller, 1988; Pontiff & Schall, 1998; Baker & Wurgler, 2000; Goetzmann et al., 2001; Lettau & Ludvigson, 2001; Guo, 2006; van Binsbergen & Koijen, 2010; Ferreira & Santa-Clara, 2011; Rapach et al., 2013; Neely et al., 2014; Maio & Santa-Clara, 2015; Golez & Koudijs, 2018; Jagannathan & Liu, 2019; le Bris et al., 2019; Bandi et al., 2019; Piatti & Trojani, 2019; Dai et al., 2021).

Goetzmann et al. (2001) analyze the US aggregate stock market and found little evidence of stock return predictability during most of the XIX century until 1925. Ferreira and Santa Clara (2011) address the predictability of international stock returns using dividend-price ratios, earnings growths, and price-earnings ratio growths from 1927 to 2007 finding substantial predictability, hence concluding that it would have been possible to profitably "time the market". Golez and Koudijs (2018) show that dividend-price ratios did not predict US stock returns from 1871 to 1945. However, there was some forecasting ability afterwards, until 2015. Piatti and Trojani (2019) also reach a similar conclusion to the previous study. Dai et al. (2021) find evidence of stock return predictability using technical indicators from 1989 to 2018.

Although mainstream research has focused solely on stocks and bonds, more recently, other types of assets, such as Real Estate Investment Trusts (REITs), have attracted the attention of researchers. According to Habbab et al. (2022), one of the most significant advantages of investing in REITs is to benefit from the real estate sector without having to pay a substantial amount or manage the underlying assets. Bhuyan et al. (2014) reinforce that REITs have been an alternative investment vehicle since the 1980s. Historically, REITs have been a desirable financial asset by providing diversification benefits, improving the risk-return trade-off, and supplying a non-negligible dividend income (see, for instance, Ling et al., 2020). Since REITs tend to adjust quickly to the cost of living, they also provide a hedge against inflation, turning their real return relatively stable. Furthermore, REITs returns show high predictability since their income comes from the underlying commercial real estate with long-term lease periods (Bhuyan et al., 2014; Fugazza et al., 2015). Beracha et al. (2019) point out that REITs had a high market valuation of 4% annually in the last ten years before the study was made.

There is extensive research on which variables are more suitable for predicting returns of stocks and bonds. In the early 1960s, several studies examined the forecast power of several technical indicators, such as moving averages, filter rules, and momentum oscillators. This line of research was recently recovered by some authors, such as Neely et al. (2014), Gao et al. (2018), Zhang et al. (2019), and Dai et al. (2021). Besides these indicators, the literature has provided a broad list of predictors, such as the dividend-price ratio (Campbell & Shiller, 1988; Cochrane, 2007), earnings-price ratio (Campbell & Shiller, 1988), book-to-market ratio (Kothari & Shanken, 1997), accruals (Hirshleifer et al., 2009),

nominal interest rate and interest rate spread (Fama, 1990; Rapach et al., 2016), volatility and downside risk (Bollerslev et al., 2014; Guo, 2006; Kilic & Shaliastovich, 2019), lagged industry returns (Hong et al., 2007), oil prices and oil-relative variables (Driesprong et al., 2008; Nonejad, 2018), investor sentiment (Huang et al., 2015), manager sentiment (Jiang et al., 2019), expected business conditions (Campbell & Diebold, 2009), labor income (Santos & Veronesi, 2006), aggregate output (Rangvid, 2006), output gap (Cooper & Priestly, 2009), inflation rate (Ludvigson & Ng, 2009), and main macroeconomic indicators (Wang et al., 2018). Welch and Goyal (2008) summarize a list of several variables that have been used in the literature with positive results. The present study considers not only this list but also other variables that have also been used for predicting returns in a multi-asset framework.

Earlier studies mostly report in-sample evidence on return predictability. The predominance of in-sample studies could be justified by using all available data, which increases the power of econometric tests (Neely et al., 2014). As argued by the authors, in-sample estimations tend to produce efficient and precise estimates of the parameters. However, in-sample tests may be biased if the predictor and return innovations are correlated and the predictor is highly persistent (Stambaugh, 1999). That bias potentially leads to substantial size distortions in the usual t-tests on the significance of the variables.

The focus on in-sample predictability has been gradually shifting to out-of-sample predictability (Feunou et al., 2018; Dai et al., 2020). For instance, Welch and Goyal (2008) and Thornton and Valente (2012) show that although some variables successfully predicted returns in-sample, they were not significant out-of-sample. Predictions based on these variables failed to consistently outperform the simple historical average benchmark forecast in terms of Mean Squared Forecast Error (MSFE).

Whether returns are predictable out-of-sample or not is still an ongoing debate. According to Wu et al. (2013), the conflicting empirical results presented in the literature may be related to problems such as data mining, spurious regressions, and instability of return predictability. Hence, recent studies have provided adaptive methods that improve forecasting in a dynamic setup. Just to name a few: time-varying parameters or time-varying volatility (Dangl & Halling, 2012), diffusion indexes (Ludvigson & Ng, 2009), combinations of many potential return predictors (Rapach et al., 2013; Fisher et al., 2020; Zhang et al., 2018; Bahrami et al., 2019, Gargano et al., 2019), inclusion of regime shifts (Hammerschmid & Lohre, 2018). Nevertheless, some recent studies have argued in favor of traditional predictive regressions, showing that these methods, updated with schemes to resolve parameter uncertainty and instability, outperform the historical average forecast in out-of-sample experiments (see, for instance, Rapach et al., 2013; Koop & Korobilis, 2013; Fisher et al., 2020).

Most commonly, asset returns are forecasted in a Vector Autoregressive (VAR) framework (Guidolin & Hyde, 2012). VAR models provide a coherent way to generate internally consistent multiperiod forecasts that account for concurrent and dynamic correlations across the variables (Elliott & Timmermann, 2008). VAR models are a valuable tool for a small number of assets and additional predictors.

However, the amplification of the asset or predictor space implies an increase in the number of parameters that may lead to an enlargement of the estimation error. Different methodologies have been developed to deal with this issue, such as Bayesian methods that make use of the high computational power that is now available to researchers. Koop and Korobilis (2013) and Dangl and Halling (2012) are two examples of applications of such methodologies with positive results.

According to Parslow et al. (2013), the main advantage of Bayesian methods is their potential to systematically incorporate previous knowledge of models and parameters. Additionally, Bayesian frameworks allow to moderate prior information (see, for instance, Barberis, 2000; Fugazza et al., 2015). Bayesian models may also include numerous efficient shrinkage tools to prevent the proliferation of parameters and eliminate parameter/model uncertainty. Two examples are the Bayesian model averaging and selection applied to VAR and time-varying VAR (Dangl & Halling, 2012; Elliott et al., 2013; Koop & Korobilis, 2013).

The Bayesian-based dynamic models have four main advantages. First, they do not require Markov Chain Monte Carlo simulations. Instead, they rely on estimated discount factors that characterize the degree of variation of the VAR coefficients. Second, they allow Dynamic Model Switching (DMS), which mitigates over-parameterization. This method selects a model over a set of different dimensions based on the past predictive likelihoods of the dependent variables. Third, they allow for time-varying parameters. Usually, forecasting models assume that coefficients are constant over time, although there is plenty of evidence of instability in the relationship between asset returns and predictors. Fourth, Bayesian setups may average forecasts from different models over time (Dynamic Model Averaging). Several studies highlighted the substantial benefits of combining forecasts across different models, such as the improvement in the predictive performance (Tian et al., 2021), the strengthening of the forecast against misspecification bias and measurement errors in the data (Timmermann, 2006), the handling of uncertainty and instability of financial markets (Fisher et al., 2020), and the providing of diversification benefits (Atiya, 2020).

Most academic studies regarding asset allocation focus on myopic portfolio optimization problems (see, for instance, DeMiguel et al., 2009; Daskalaki & Skiadopoulos, 2011; Cenesizoglu & Timmermann, 2012). However, recent literature has shown substantial utility benefits when investors incorporate return predictability and model uncertainty into their investment decisions (Diris et al., 2015; Rapach & Zhou, 2013). Hence, several studies, such as Johannes et al. (2014), Gargano et al. (2019), Gao and Nardari (2018), and Fisher et al. (2020), have implemented Bayesian dynamic approaches to portfolio strategies with positive results.

3 Data Description and Preliminary Analysis

3.1 Asset Classes

This study considers three US-based asset classes: Stocks, Bonds, and REITs. The period under scrutiny spans from January 1976 to December 2021 (553 monthly
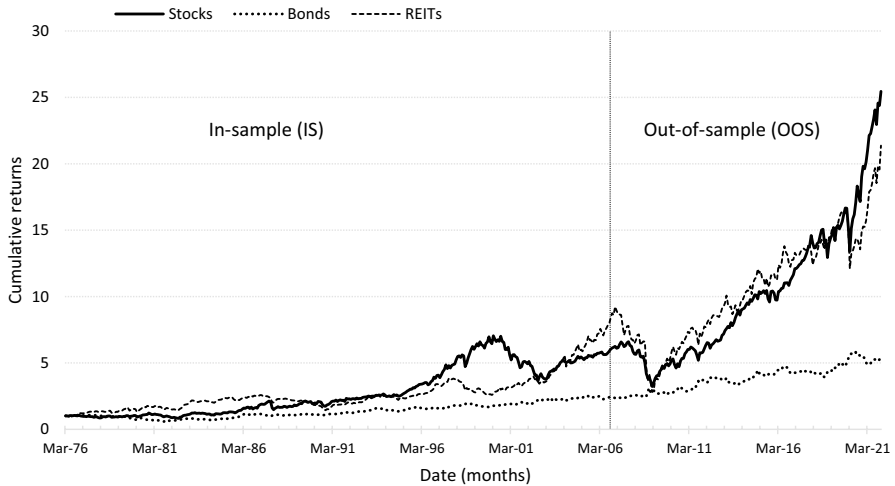
Fig. 1 Cumulative log-returns of the total return indexes of the Stocks, Bonds, and REITs, computed as $I_t = \exp\left(\sum_{\tau=1}^{t} r_{i,\tau}\right)$ from March 1976 to December 2021 (550 monthly observations). Stocks are proxied by the S&P 500 Total Return Index, Bonds are proxied by the Barclays Capital US Aggregate Bond Index, both collected from the Refinitiv Eikon database, and REITs are proxied by the FTSE Nareit US Real Estate All Equity REITs Index, collected from the NAREIT website (https://www.reit.com/data-research). The vertical line separates the in-sample (IS) period (2/3 of the sample, corresponding to 367 months) and the out-of-sample (OOS) period (1/3 of the sample, corresponding to 183 months)

observations). The three initial months were used to compute some predictors. Hence the sample was reduced to 550 observations, from March 1976 to December 2021. Stocks are proxied by the S&P 500 Total Return Index, and Bonds are proxied by the Barclays Capital US Aggregate Bond Index, both collected from the Refinitiv Eikon database. REITs are proxied by the FTSE Nareit US Real Estate All Equity REITs Index, collected from the NAREIT website (https://www.reit.com/data-research). Hence, we only consider publicly available stocks of REITs. The risk-free rate is proxied by the monthly yield-to-maturity of 3-month US Treasury Bills.

Figure 1 plots the cumulative returns of the total return indexes of the Stocks, Bonds, and REITs. All series are notably more volatile after 2000, and Stocks and REITs are more sensitive to the business cycle than Bonds. However, the cumulative returns of Stocks and REITs dominate those of Bonds throughout the overall period, especially after March 2009. In the previous 19 months, from September 2007 until March 2009, the Stocks and REITs indexes fell -58.53% and -93.26% while the Bonds increased by 20.08% due to the subprime crisis. Nevertheless, a Buy-and-Hold investment during the overall period would earn an annual rate of return of approximately 7.32% and 6.91% for REITs and Stocks, respectively, while this figure for Bonds is only 3.66%.

To analyze the predictability of the returns of these three classes of assets, we divided the sample into an in-sample (IS) period (2/3 of the sample, corresponding to 367 months) and an out-of-sample (OOS) period (1/3 of the sample, corresponding to 183 months). The analysis is conducted out-of-sample recursively. More specifically, for each month out-of-sample, beginning at $t_0 + 1 = 368$ until $T = 550$, the returns are forecasted using all the information until the previous month, i.e., the

Table 1 Descriptive statistics of log-returns

Variable	Sample	Stocks	Bonds	REITs
Mean (%)	Total R_t R_t	0.59***	0.30**	0.56***
	IS	0.48**	0.24	0.56***
	OOS	0.80**	0.42*	0.54
Median (%)	Total R_t R_t	0.96	0.41	0.70
	IS	0.82	0.43	0.63
	OOS	1.37	0.22	1.08
StD (%)	Total R_t R_t	4.33	3.12	4.89
	IS	4.28	3.07	3.78
	OOS	4.42	3.22	6.50
Max (%)	Total R_t R_t	12.23	13.46	24.65
	IS	12.23	13.08	9.77
	OOS	12.11	13.46	24.65
Min (%)	Total R_t R_t	− 24.82	− 11.93	− 36.05
	IS	− 24.82	− 10.41	− 16.64
	OOS	− 18.33	− 11.93	− 36.05
Skew	Total R_t R_t	− 0.83	0.13	− 1.49
	IS	− 0.83	0.05	− 0.80
	OOS	− 0.98	0.25	− 1.46
Kurt	Total R_t R_t	5.91	4.76	12.80
	IS	6.46	4.64	5.73
	OOS	5.54	5.43	10.34
Jarque–Bera	Total R_t R_t	257.29***	72.77***	2404***
	IS	224.73***	41.24***	153.58***
	OOS	51.56***	30.16***	526.36***

This table reports some summary statistics of Stocks, Bonds, and REITs monthly log-returns over the period from March 1976 to December 2021 (550 observations). *Full* refers to the overall sample, IS is the in-sample period (2/3 of the sample, corresponding to 367 months), and OOS is the out-of-sample period (1/3 of the sample, corresponding to 183 months). The statistics are the mean, standard deviation, *StD*, maximum and minimum values, *Max* and *Min*, respectively, skewness, *Skew*, kurtosis, *Kurt*, and the Jarque–Bera normality test. The values are in percentages, except *Skew*, *Kurt*, and *Jarque–Bera*. The table also presents the results of the tests on the significance of the mean. Asterisks *, **, ***, indicate significance at the 10%, 5%, and 1% levels, respectively

forecast at $t_0 + 1$ is obtained with the previous 367 months, the forecast at $t_0 + 2$ is obtained with the previous 368 months, and so forth until month T, to which the forecast is obtained using the previous 549 months. The recursive (expanding) estimation window to generate out-of-sample forecasts is usually used in the literature on return predictability (see, e.g., Rapach et al., 2010, 2016; Neely et al., 2014; Zhang et al., 2019).

The forecasts are then used to allocate the investment into portfolios formed by Stocks, Bonds, REITs, and the riskless asset (3-month US Treasury Bills).

Accordingly, the portfolios are rebalanced monthly. Almadi et al. (2014) analyzed the performance of dynamic portfolios rebalanced at the same frequency as the forecast horizon showing that monthly rebalancing provides the best performance in the presence of transaction costs of 0.5%.

Table 1 reports the summary statistics of Stocks, Bonds, and REITs returns for the overall sample, IS, and OOS periods. In the overall sample, the mean and median monthly stock returns are the highest (0.59% and 0.96%, respectively). Most of the means are significantly different from zero. The variability of these returns, measured by the standard deviation and range, is high but is surpassed by REITs. The means of Stocks and Bonds are higher OOS, while the median is higher for Stocks and REITs. The variability measured by the standard deviation and range is also, loosely speaking, higher OOS, especially for REITS, for which the standard deviation almost doubled. Bonds are the asset class that has the lowest mean, median, and variability. Bonds have positive skewness, whilst Stocks and especially REITs are negatively skewed. The three asset classes present mild excess kurtosis, and REITs stand out as the one with the most leptokurtic distribution, especially in the OOS period. All series are non-normal, as indicated by the Jarque–Bera test.

3.2 Additional Predictors

The initial database compiles a comprehensive set of predictive variables documented in the literature on asset return predictability. The Appendix lists these variables (19 for stocks, 122 for Bonds, and 14 for REITs), presenting their abbreviations, summary descriptions, online sources, and transformations made to obtain the variables used in the literature and to correct non-stationarity.

It is well known that the profusion of predictive variables in asset returns regressions leads to overfitting and a poor performance out-of-sample. Thus, to avoid this problem, a Genetic Algorithm (GA) is used for selecting the most relevant predictors. The Darwinian process of natural selection that drives the evolution of species inspired this method, which may solve a wide array of optimization problems. It has proved valuable in variable selection for multivariate regressions, as it achieves good performance and its execution time is lower than that of alternative algorithms (Leardi & Gonzáles, 1998; Leardi et al., 1992). The GA begins by randomly creating an initial population of candidate solutions (chromosomes). Then, the performance of each chromosome is evaluated. In the next stage, the chromosomes are selected based on their performance and are combined using crossover operations. Also, some chromosomes are mutated according to the mutation probability. These operations lead to the creation of a new population. The procedure is repeated until the termination condition is reached (for further details, see, for example, Leardi et al., 1992).

In this research, the GA is implemented using the Matlab Regression toolbox (see Consonni et al., 2021, for further details). The initial population has 50 individuals, and the GA is run for 200 generations. The crossover and mutation probabilities are set to 0.5 and 0.01, respectively. The performance of each chromosome is assessed

Table 2 Selected predictors for each asset class

Stocks	Bonds	REITs
Dividend price ratio	Manufacturing and trade sales	Term spread
Dividend payout ratio	New order for non-defense capital goods	Change in adult population
Price-earnings ratio	1-year federal fund spread (1YFFS)	Change in employment (Δ Emp)
Inflation	CPI durables	
Dividend yield	Average hours of earnings in construction sector	
Cyclically adjusted PE ratio	Crude oil	
Robust minus weak factor		
Conservative minus aggressive factor (CMA)		

This table reports, for each asset class, the predictors chosen by the Genetic Algorithm. The initial predictor space has 155 variables, 19 for stocks, 122 for Bonds, and 14 for REITs (see the Appendix). The algorithm uses an initial population of 50 individuals, and it is run for 200 generations. The crossover and mutation probabilities are 0.5 and 0.01, respectively, and the performance of each chromosome is assessed using the fivefold cross-validated RMSE. A maximum number of predictive variables of 10 is superimposed into the algorithm. The predictors in bold represent the best ones for each asset class

using a fivefold cross-validated Root Mean Squared Error (RMSE). A maximum number of 10 predictors is superimposed into the algorithm.

Table 2 displays the predictors selected by the GA and highlights, in bold, the best predictor for each asset class.

Table 3 presents the correlations of asset returns and the three most important additional predictors. In the full sample, only the lagged predictor of bonds is significantly correlated with that asset class. Also, there is no significant predictor for Stocks, but Bonds and REITs are significantly correlated with three predictors implying that these two asset classes probably present some degree of predictability. The additional predictors are not significantly cross-correlated, which is a desirable feature, as Zhang et al. (2019) highlights.

The absence of significant correlations for Stocks remains OOS. In this period, the number and significance of correlations for Bonds and especially for REITs decreases. In OOS, the correlations between Bonds and lagged Stocks and lagged REITs returns are significant at the 1% and 5% levels, respectively. Lagged stock returns are also correlated with REITs (at the 5% level). It seems that OOS, the conservative minus aggressive factor (CMA) may help predict REITs.

In sum, this analysis points out that probably Stocks are not forecastable, but the information on the stock market may help predict REITs and especially Bonds.

Table 3 Autocorrelations and cross-correlations

	Stocks$_t$	Bonds$_t$	REITS$_t$	CMA$_{t-1}$	1YFFS$_{t-1}$
Total sample					
Stocks$_{t-1}$	0.03	−0.16***	0.13***		
Bonds$_{t-1}$	0.07	0.05	0.16 ***		
REITS$_{t-1}$	0.06	−0.17***	0.09**		
CMA$_{t-1}$	−0.05	0.04	0.02		
1YFFS$_{t-1}$	0.01	0.14***	0.04	−0.01	
Δ Emp$_{t-1}$	−0.02	−0.00	−0.01	0.01	0.05
IS					
Stocks$_{t-1}$	−0.00	−0.10**	0.10*		
Bonds$_{t-1}$	0.06	0.07	0.23***		
REITS$_{t-1}$	0.08	−0.07	0.10*		
CMA$_{t-1}$	−0.11**	0.02	−0.06		
1YFFS$_{t-1}$	0.02	0.17***	0.08	−0.00	
Δ Emp$_{t-1}$	0.00	−0.10**	−0.10*	−0.01	0.10*
OOS					
Stocks$_{t-1}$	0.08	−0.25***	0.17**		
Bonds$_{t-1}$	0.07	0.03	0.10		
REITS$_{t-1}$	0.04	−0.30**	0.08		
CMA$_{t-1}$	0.11	0.08	0.14*		
1YFFS$_{t-1}$	0.01	0.09	0.11	0.07	
Δ Emp$_{t-1}$	−0.04	0.03	0.01	0.02	−0.02

This table reports the correlations between asset returns and the three most important additional predictors in the total sample, IS and OOS. The table presents the autocorrelations of order one between returns, the correlations between returns and lagged values of the additional predictors and the contemporaneous cross-correlations between the additional predictors. The significance test of these correlations uses Bartlett's standard errors Asterisks *, **, ***, indicate significance at the 10%, 5%, and 1% levels, respectively

4 Methodology

This section presents the basic theoretical concepts and model specifications. It begins by presenting the Vector Autoregressive model (VAR) and the procedures used to obtain forecasts based on the combinations of VAR estimated with different predictor spaces. Next, it presents the Time-Varying Parameter Vector Autoregressive model (TVP-VAR), the procedures used to estimate these models with forgetting factors, and the methods to combine or select these models. Finally, it shows the metrics of forecasting accuracy and portfolio economic performance from the perspective of a CRRA investor.

4.1 Vector Autoregressive Models (VAR) and Forecast Combinations

Since its introduction by Sims (1980), VAR has been widely used for forecasting purposes. These models are a straightforward multivariate generalization of univariate autoregressions and can generate dynamic forecasts that ensure consistency across different endogenous variables and forecasting horizons. Many researchers have used large VAR with tens of dependent variables (see, among many others, Banbura et al., 2010; Carriero et al., 2009, and Koop & Korobilis, 2013).

As in Campbell et al. (2003), this study implements first-order vector autoregressive models, VAR(1), to capture the linear dynamics of asset returns and other predictors,[2] such that:

$$\mathbf{y}_t = \mathbf{X}_t \mathbf{B} + \boldsymbol{\epsilon}_t, \tag{1}$$

where $\mathbf{y}_t$ is a column vector containing the observations on $m = m_1 + m_2$ time series. This vector includes a $(m_1 \times 1)$ vector of excess log-returns at time t, $\mathbf{r}_t$, and a $(m_2 \times 1)$ vector of other endogenous variables, $\mathbf{u}_t$. In our empirical application, $\mathbf{r}_t = \begin{bmatrix} r_{1,t} r_{2,t} r_{3,t} \end{bmatrix}\prime$, where $r_{1,t}$, $r_{2,t}$, and $r_{3,t}$ are the excess log-returns over the risk-free rate of Stocks, Bonds, and REITs, respectively. $\mathbf{X}_t$ is a $(m \times k)$ such that:

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{x}\prime_t & 0 & \cdots & 0 \\ 0 & \mathbf{x}\prime_t & \cdots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & \cdots & 0 & \mathbf{x}\prime_t \end{bmatrix}, \tag{2}$$

where $\mathbf{x}_t$ is a column vector containing an intercept and one lag of each of the m variables, and therefore $k = m(1 + m)$. $\mathbf{B}$ is the coefficient matrix, and $\boldsymbol{\epsilon}_t$ is a vector of shocks *i.i.d.* $\mathcal{N}(0, \boldsymbol{\Sigma})$. Notice that shocks can be cross-sectionally correlated.

In the present study, the VAR(1) includes 3 asset classes and 3 predictors (one predictor for each asset class). Thus, there is a total of $8 \times 6 \times 3 = 144$ different model specifications corresponding to all possible combinations of predictors in $\mathbf{u}_t$.

The forecasts of the excess log-returns vector are obtained using several methods for combining the individual forecasts from models $\mathcal{M}_j$, with $j = 1, 2, \ldots, n$. More specifically, the individual model forecasts, $\mathbb{E}\big(\mathbf{r}_t | \mathcal{M}_j, \mathcal{F}_{t-1}\big)$, and covariances matrices forecasts, $\widehat{Cov}\big(\mathbf{r}_t | \mathcal{M}_j, \mathcal{F}_{t-1}\big)$, which use information up to time $t - 1$, $\mathcal{F}_{t-1}$, are then used to compute the Mean, i.e., the average of the forecasts of individual models, and the Weighted Mean Squared Forecasting Errors (WMSFE) forecasts of $\mathbf{r}_t$. For each model j, the WMSFE is computed as

---

[2] The VAR(1) avoids additional lags that would require a larger state vector with many parameters. Nevertheless, this representation is not restrictive, since a VAR of any order can be rewritten as a VAR(1) by increasing the number of variables in the state vector.

$$WMSFE_j = \frac{1}{T - t_0} \sum_{t=t_0+1}^{T} \mathbf{e}'_{j,t} [\widehat{Cov}(\mathbf{r}_t)]^{-1} \mathbf{e}_{j,t}, \tag{3}$$

where $\mathbf{e}_{j,t}$ is the column vector of forecast errors of model j at time t in the out-of-sample period, with $t = t_0 + 1, t_0 + 2, \ldots, T$. The forecast error of asset i in model j, is the difference between the realized excess log-return, $r_{i,t}$, and the one-step-ahead forecast, $\widehat{r}_{i,j,t}$, i.e., $e_{i,j,t} = r_{i,t} - \widehat{r}_{j,i,t}$. $\widehat{Cov}(\mathbf{r}_t)$ is the sample estimate of the asset excess log-returns unconditional covariance matrix (see Rapach et al., 2010).

The combinations based on the WMSFE heighten certain individual forecasts due to the covariance matrix. This procedure attributes higher penalties to forecast errors with lower variability, that is, those on which the investor is highly confident, and penalizes more lightly diffuse forecasts. Also, penalties are lower when forecast errors have the same (opposite) sign for positively (negatively) correlated assets than when the reverse pattern holds. Furthermore, combinations based on the WMSFE allow the weights on individual forecasting models to reflect their past predictive accuracy. More specifically, the weight of model j at t is given by its WMSFE in the period before, $t - 1$, that is:

$$\omega_{j,t} = \frac{\varphi_{j,t}^{-1}}{\sum_{l=1}^{n} \varphi_{l,t}^{-1}}, \tag{4}$$

where $\varphi_{j,t}$ is the sorted WMSFE for the j model at time t according to the WMSFE. The weight $\omega_{j,t}$ may be computed using all models or just a subset of these models. In this study we consider the 10%, 20%, 30%, 40%, and 50% best models.

Regardless of the method used, there are several advantages to using forecast combinations. They can be seen as a diversification strategy in asset allocation, may capture different aspects of business conditions and provide information signals to models and predictive power variations through time (Bates & Granger, 1969; Rapach et al., 2010). For instance, if the correlation between individual forecasts is weak, their combination may produce less unstable models, rendering more stable forecasts, reducing forecast risk, and improving forecast performance under model instability and uncertainty (Rapach & Zhou, 2013).

4.2 Time-Varying Parameter Vector Autoregressive (TVP-VAR)

A Time-Varying Parameter-Vector Autoregressive model of order 1, TVP-VAR(1), may be represented as follows:

$$\mathbf{y}_t = \mathbf{X}_t \mathbf{B}_t + \boldsymbol{\varepsilon}_t, \text{ and } \mathbf{B}_t = \mathbf{B}_{t-1} + \boldsymbol{\nu}_t, \tag{5}$$

where $\boldsymbol{\varepsilon}_t$ is *i.i.d.* $\mathcal{N}(0, \boldsymbol{\Sigma}_t)$ and $\boldsymbol{\nu}_t$ is *i.i.d.* $\mathcal{N}(0, \mathbf{Q}_t)$. $\boldsymbol{\varepsilon}_\tau$ and $\boldsymbol{\nu}_t$ are independent for all τ and t.

Traditionally TVP-VAR were usually estimated using forgetting factors (also known as discount factors). This method is still used in recent applications due to

its simplicity and fast-tracking (see, for instance, Dangl & Halling, 2012; Koop & Korobilis, 2013).

Let $\mathbf{Y}_{t-1} = (y_1, \ldots, y_{t-1})'$ be all the observations until $t-1$. Using the Kalman filter, the Bayesian inference of $\mathbf{B}_t$ is

$$\mathbf{B}_{t-1}|\mathbf{Y}_{t-1} \sim \mathcal{N}(\mathbf{B}_{t-1|t-1}, \mathbf{V}_{t-1|t-1}). \tag{6}$$

The distribution of the state vector in the next period, using the same information set, is:

$$\mathbf{B}_t|\mathbf{Y}_{t-1} \sim \mathcal{N}(\mathbf{B}_{t|t-1}, \mathbf{V}_{t|t-1}), \text{ with } \mathbf{V}_{t|t-1} = \mathbf{V}_{t-1|t-1} + \mathbf{Q}_t \tag{7}$$

By replacing $\mathbf{Q}_t = (\lambda^{-1} - 1)\mathbf{V}_{t-1|t-1}$ into the previous equation one obtains

$$\mathbf{V}_{t|t-1} = \frac{1}{\lambda}\mathbf{V}_{t-1|t-1}, \tag{8}$$

where λ denotes the forgetting factor, with $0 < \lambda \leq 1$. This implies that observations h-periods in the past have a weight λ^h in the filtered estimate of $\mathbf{B}_t$ and that the variance of the coefficient vector increases by a factor of $1/\lambda$ per period. Dangl and Halling (2012) apply two granularity choices for λ: $\lambda \in \{0.96, 0.98, 1.00\}$ and $\lambda \in \{0.96, 0.97, 0.98, 0.99, 1.00\}$ and conclude that λ should have a lower bound of 0.98 to mitigate the variability of the coefficients. Alternatively, Koop and Korobilis (2013) implement a more robust technique. Instead of simply setting λ equal to a fixed value, they estimate:

$$\lambda_t = \lambda_{min} + (1 - \lambda_{min})L^{\theta_t}, \tag{9}$$

where $\theta^t = -NINT(\widetilde{\boldsymbol{\varepsilon}}'_{t-1}\widetilde{\boldsymbol{\varepsilon}}_{t-1})$, $NINT$ rounds to the nearest integer, and $\widetilde{\boldsymbol{\varepsilon}}'_{t-1} = \mathbf{y}_t - \mathbf{X}_t\mathbf{B}_{t|t-1}$ is the one-step-ahead prediction error produced by the Kalman filter. Following Koop and Korobilis (2013) we set $\lambda_{min} = 0.96$ and $L = 1.1$ to obtain values between 0.96 and 1 for the forgetting factor.

To eliminate the need to simulate the multivariate stochastic volatility in the measurement equation we use the Exponentially Weighted Moving Average (EWMA) estimator for the error covariance matrix:

$$\widehat{\boldsymbol{\Sigma}}_t = \kappa\widehat{\boldsymbol{\Sigma}}_{t-1} + (1 - \kappa)\widetilde{\boldsymbol{\varepsilon}}_t\widetilde{\boldsymbol{\varepsilon}}'_t, \tag{10}$$

Our choice of the decay factor draws on the Riskmetrics in JP Morgan/Reuters (1996) technical note, which computes the value of κ that minimizes the root mean squared prediction error for the variance using more than 480 series. Based on their estimates, they suggest using a decay factor in the interval (0.94, 0.98) for monthly data. In this research, we choose the value in the middle of the proposed range (0.96). The computation of $\widehat{\boldsymbol{\Sigma}}_t$ also requires the choice of an initial condition for $\boldsymbol{\Sigma}_0$, which is set equal to the sample covariance matrix of $\mathbf{y}_{t_0}$.

Models such as TVP-VAR are designed to accommodate gradual changes in the coefficients and are unable to adjust to abrupt changes, which reduces their performance. A way to deal with the possibility of significant changes is to allow the

switching between different models. Thus, we enable the TVP-VAR(1) to change dimensions over time by using a Dynamic Model Selection (DMS) procedure. This procedure requires the estimation of the forgetting factors, the definition of the priors, and the definition of various dimensions of the combinations of the TVP-VAR models. We use the DMS to select the optimal values for the VAR shrinkage parameter in a time-varying manner. In this way, the DMS is a recursive algorithm where the important recursions are similar to the forecasting and updating equations of the Kalman filter. Following Koop and Korobilis (2013), the model prediction and updating equations using a forgetting factor α are derived from:

$$\pi_{j,t|t-1} = \frac{\pi_{j,t-1|t-1}^{\alpha}}{\sum_{l=1}^{n} \pi_{l,t-1|t-1}^{\alpha}},$$ (11)

which is the probability that model j will be chosen, given the information up to $t-1$, and,

$$\pi_{j,t|t} = \frac{\pi_{j,t|t-1} p_j(\mathbf{y}_t|\mathbf{y}_{t-1})}{\sum_{l=1}^{n} \pi_{l,t|t-1} p_l(\mathbf{y}_t|\mathbf{y}_{t-1})},$$ (12)

where $p_j(\mathbf{y}_t|\mathbf{y}_{t-1})$ is the predictive likelihood (the predictive density of model j calculated at $\mathbf{y}_t$). The probability used to select models can be written as:

$$\pi_{j,t|t-1} \propto \prod_{k=1}^{t-1} \left[p_j\left(\mathbf{y}_{t-k}|\mathbf{y}_{t-k-1}\right) \right]^{\alpha^k}.$$ (13)

The symbol $\propto$ means that the probability is proportional to the expression on the right. Model j receives more weight at time t according to the accuracy of its forecasts in the recent past. The weight of past predictive densities is controlled by the forgetting factor, α, which has similar features to the forgetting factor defined before, λ.

In our study, we set $\alpha = 0.99$ as in Koop and Korobilis (2013), which implies that the forecast performance 60 periods ago receives 55% as much weight as the forecast performance in the last period. We do not define any other values for α. As Koop and Korobilis (2013) showed, choosing α between 0.95 and 1 has a minor impact on the results. Furthermore, Dangl and Halling (2012) and Hill and Rodrigues (2022) show that models with α close to 1 tend to outperform models that forget past information faster. In sum, we consider $\alpha = 0.99$, and $\kappa = 0.96$.

Our approach does not require the estimation of $\mathbf{Q}_t$, and, as we referred before, it uses an EWMA estimator of $\mathbf{\Sigma}_t$ that requires prior information on $\mathbf{B}_0$. In the literature, it is usual to use training sample priors to produce hyperparameters that monitor the degree of shrinkage when working with large VAR or TVP-VAR (Banbura et al., 2010). However, Koop and Korobilis (2013) used a different approach that allows for the estimation of the shrinkage hyperparameter in a time-varying manner. To do so, they applied an automatic updating procedure, which is less demanding

computationally since it does not require the re-estimation of the shrinkage priors or the model at each point in time.

This study uses a normal prior for $\mathbf{B}_0$ like a Minnesota prior. The prior mean is set as $\mathbb{E}(\mathbf{B}_0) = 0$, and the prior covariance matrix of $\mathbf{B}_0$ is diagonal, such that the i-th diagonal element is equal to $\underline{a}$ for the intercepts and δ for the lag coefficients. Hence, δ controls for the degree of shrinkage on the VAR coefficients,[3] and $\underline{a} = 10^3$, for the intercepts to be uninformative.

A large degree of shrinkage is needed to obtain good forecasting performances in large VAR and TVP-VAR. To do so, we estimate δ at each point in time using a method similar to the one used for the forgetting and decay factors. As in Koop and Korobilis (2013), we use a grid for δ, such that $\delta \in \{10^{-10}, 10^{-5}, 0.001, 0.005, 0.01, 0.05, 0.1\}$.

In sum, the DMS implies choosing the model with the highest value of $\pi_{j,t|t-1}$ to obtain the forecast at time t. Since $\pi_{j,t|t-1}$ varies over time, the forecasting model may change, allowing the model switching feature. Besides the DMS, we also consider Dynamic Model Averaging (DMA), which uses $\pi_{j,t|t-1}$ as the weighting scheme.

We also consider different dimensions (i.e., different state space sets) when implementing the DMS and DMA procedures. These models are denominated: (1) "Small", which only considers the first lags of the excess returns, (2) "Medium", which includes three additional predictors (the best one for each asset chosen by the GA), (3). "Large", which includes the 17 predictors chosen by the GA, and (4) "Full", which is selected or averaged, for the DMS and DMA, respectively, across the Small, Medium, and Large models.

A crucial point of TVP-VAR selection and averaging is the calculation of $\pi_{j,t|t-1}$. When forecasting at time t, this probability for each model j is evaluated and the value of δ and the dimension of TVP-VAR(1) that maximizes it is used. This is done recursively using Eqs. (11) and (12) and setting the initial probability of selecting each model equal to $\pi_{j,0|0} = 1/n$ for all models. However, when dealing with TVP-VAR with different dimensions we have different predictive densities, $p_j(\mathbf{y}_t|\mathbf{y}_{t-1})$, since $\mathbf{y}_t$ has different dimensions, rendering them incomparable. A possible solution is to use the same predictive densities for all dimensions. So, we use the predictive density of the Small model since the variables in this model are common to all models. In other words, the DMS is determined by the joint predictive likelihood of the three asset excess returns.

4.3 Forecasting Accuracy, Asset Allocation, and Portfolio Performance

The accuracy of the point forecasts of each forecasting scheme is measured using the Mean-Squared Forecast Errors (MSFE). These schemes are compared against the recursive historical average, which is the average of past asset returns up to the date on which the prediction is made. The forecasting schemes and the resulting

---

[3] Note that this differs from the Minnesota prior in that the latter contains two shrinkage parameters (corresponding to own lags and other lags) and these are set to be constant. For ease of computation, we only use one shrinkage parameter, as in Banbura et al. (2010).

are hereafter denoted by s. Following Campbell and Thompson (2008), the additional predictive power of asset i in scheme s can be measured by the pseudo-R^2 out-of-sample:

$$\text{pseudo} - R_{i,s}^2 = 1 - \frac{MSFE_{i,s}}{MSFE_{i,ha}} = 1 - \frac{\sum_{t=t_o+1}^{T} e_{i,s,t}^2}{\sum_{t=t_o+1}^{T} e_{i,ha,t}^2}. \tag{14}$$

$e_{i,s,t} = r_{i,t} - \hat{r}_{i,s,t}$ is the forecast error of s, and $e_{i,ha,t} = r_{i,t} - \bar{r}_{i,t}$ is the forecast error in relation to $\bar{r}_{i,t}$ (the recursive historical average) of asset i at time t. A scheme s produces better predictions than the historical average if the pseudo-$R_{i,s}^2$ is positive.

This analysis is complemented by the adj-MSFE test of Clark and West (2007). This test uses an approximately normal modified version of the MSFE statistic, which the authors show to be undersized. The null hypothesis is that the MSFE of s and the historical average are equal, whereas the alternative hypothesis is that the s predictions are more accurate. The most convenient way to implement this one-side test is to compute:

$$\hat{f}_{i,s,t} = \left(r_{i,t} - \bar{r}_{i,t-1}\right)^2 - \left[\left(r_{i,t} - \hat{r}_{i,s,t-1}\right)^2 - \left(\bar{r}_{i,t} - \hat{r}_{i,s,t-1}\right)^2\right], \tag{15}$$

and then regress $\hat{f}_{i,s,t}$ on a constant and use the resulting t-statistics.

To analyze the asset allocation amongst the three risky asset classes and the riskless asset, it is assumed that the investor has a CRRA utility function $U(W_t) = \frac{W_t^{1-\gamma}}{1-\gamma}$, where $W_t = \exp\{r_{p,t}\}$ denotes the investor's wealth at time t, and γ, with $\gamma > 1$, is the relative risk aversion coefficient. At each point in time, the investor chooses the allocation amongst these assets that maximizes the 1-period-ahead expected utility $\mathbb{E}_{t-1}[U(W_t)]$. The optimal weights of s are given by the solution of the following constrained maximization problem, where the investor maximizes the difference between the portfolio's simple expected excess return and the portfolio's variance multiplied by half the coefficient of relative risk aversion (Fisher et al., 2020):

$$\arg\max_{\mathbf{w}_{s,t-1}} \mathbf{w}'_{s,t-1}\left(\hat{\boldsymbol{\mu}}_{s,t|t-1} + \frac{1}{2}diag\hat{\boldsymbol{\Sigma}}_{s,t|t-1}\right) - \frac{\gamma}{2}\mathbf{w}'_{s,t-1}\hat{\boldsymbol{\Sigma}}_{s,t|t-1}\mathbf{w}_{s,t-1}, s.t. : (1)\mathbf{w}'_{s,t-1}\iota \leq 1 \text{and} (2)w_{s,t-1} \geq 0, \forall s, t. \tag{16}$$

The vector $\mathbf{w}_{s,t-1}$ denotes the weights of the risky assets in the portfolio given strategy s, $\hat{\boldsymbol{\mu}}_{s,t|t-1} = \mathbb{E}(\mathbf{r}_t|s, \mathcal{F}_{t-1})$ and $\hat{\boldsymbol{\Sigma}}_{s,t|t-1} = \widehat{cov}(\mathbf{r}_t|s, \mathcal{F}_{t-1})$ are the mean and covariance of the predictive density of the vector of risky asset $\mathbf{r}_t$, computed using the information available at time $t-1$ according to s, and ι is a conformable vector of ones · Following much of the asset allocation literature, short-selling (i.e., negative portfolio weights) is ruled out. The investment portfolio is rebalanced each month. We have opted for a monthly rebalancing frequency as a strategic choice aimed at mitigating the impact of transaction costs. Also, data has a monthly periodicity; hence it seems adequate to rebalance the portfolio once the information set is updated. This idea is present in Almadi et al. (2014) and Maeso and Lionel (2020).

Assuming that the excess returns of the risky assets are log-normally distributed, following (Campbell et al., 2003), the portfolio log-return resulting from s at time t is:

$$r_{s,t} = r_{F,t} + \widehat{\mathbf{w}}'_{s,t-1}\left(\mathbf{r}_t - r_{F,t}\mathbf{\iota}\right) + \frac{1}{2}\widehat{\mathbf{w}}'_{s,t-1}diag(\widehat{\boldsymbol{\Sigma}}_{s,t|t-1}) - \frac{1}{2}\widehat{\mathbf{w}}'_{s,t-1}\widehat{\boldsymbol{\Sigma}}_{s,t|t-1}\widehat{\mathbf{w}}_{s,t-1}$$

(17)

where $r_{F,t}$ represents the continuously compounded risk-free rate.

Besides presenting the descriptive statistics of the discrete returns of the portfolios (mean, standard deviation, skewness, and kurtosis), the performance of the portfolios is also studied using three metrics. The discrete returns are computed as $R = \exp(r) - 1$.

The Certainty Equivalent Return (CER) is the risk-free rate that the investor is willing to accept rather than adopting the risky portfolio. Considering monthly data, the annualized CER can be expressed as follows:

$$\widehat{CER}_s = \left(\frac{1}{T - t_0}\sum_{t=t_0+1}^{T}\widehat{W}_{s,t}^{1-\gamma}\right)^{\frac{12}{1-\gamma}} - 1$$

(18)

where $\widehat{W}_{s,t} = 1 + \widehat{R}_{s,t}$ is the realized wealth at time t resulting from s, T is the total number of periods, $t_0 + 1$ is the initial observation out-of-sample.

The annualized Sharpe ratio (SR) measures the desirability of a risky investment strategy, by dividing the portfolio average excess discrete return, by the standard deviation of the excess discrete return. In other words, the SR measures the reward per unit of variability:

$$\widehat{SR}_s = \frac{\sqrt{12}\left(\widehat{\mu}_{R_s} - \mu_{R_F}\right)}{\sqrt{\frac{1}{T-t_o}\sum_{t=t_o+1}^{T}\left\{\left(\widehat{R}_{s,t} - R_{F,t}\right) - \left(\widehat{\mu}_{R_s} - \mu_{R_F}\right)\right\}^2}},$$

(19)

The annualized Sortino ratio (SOR) considers the downside risk, that is the negative deviations from a certain target point B that constitutes the minimum acceptable rate of return.

$$\widehat{SOR}_s = \frac{\sqrt{12}\left(\widehat{\mu}_{R_s} - B\right)}{\sqrt{\frac{1}{Q}\sum_{\widehat{R}_{s<B}}^{Q}\left\{min\left(\widehat{R}_s - B\right), 0\right\}^2}},$$

(20)

In the computation of SOR, it is considered that $B = 0$, and therefore the denominator is exogenous from the sample mean. Finally, we also present the CVaR at 5%.

These analyses are conducted without and with transaction costs. The transaction costs are incorporated considering a proportional cost rate of 0.5%. This figure is usually used in the literature (see, for instance, DeMiguel et al., 2009, and Almadi et al., 2014). One may argue that in this framework a transaction cost of 0.5%

Table 4 - Out-of-sample pseudo-R^2

Forecasting schemes	Model	Stocks	Bonds	REITs
Mean		−2.80	3.57**	−1.31
WMSFE	10%	−5.55	3.00**	−3.37
	20%	−4.91	3.73**	−2.33
	30%	−4.34	3.52**	−1.99
	40%	−3.71	3.61**	−1.73
	50%	−3.36	3.68**	−1.41
DMS and DMA	Small	−0.98	2.11	0.44
	Medium	0.28	−0.04	3.76
	Large	−10.23	−5.07	−2.65
DMS	Full	−3.74	0.07	−1.25
DMA	Full	−1.28	0.07	0.47

This table reports the out-of-sample pseudo-R^2 (Campbell & Thompson, 2008), in percentage, of the different forecasting schemes based on the Mean, WMSFE (with the best 10%, 20%, 30%, 40%, and 50% models) and DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1) models. The DMS and DMA statistics are the same, except for the Full model. Associated with these statistics, the table also reports the significance of the adj-MSFE of Clark and West (2007). Two asterisks, **, indicate significance at the 5% level

understates the real transaction costs, but one should notice that there are publicly available ETFs on the three indexes under scrutiny. Notice that at time t the weights are different from the ones estimated at time $t-1$ due to the changes in prices. The new weight for each risky asset i, after the price change but before rebalancing, is given by $w_{i,s,t}^+ = \frac{\widehat{w}_{i,s,t-1} exp\{r_{i,t-1}\}}{\sum_k \widehat{w}_{k,s,t-1} exp\{r_{j,t-1}\}+(1-\sum_k \widehat{w}_{k,s,t-1})}$, where $\left(1-\sum_k \widehat{w}_{k,s,t-1}\right)$ represents the weight of the risk-free asset before the price change. Therefore, the rebalancing at time t has a transaction cost of $tc_{s,t} = 0.5\% \sum_k |\widehat{w}_{k,s,t} - w_{k,s,t}^+|$. The portfolio discrete return after transaction costs at time t is given by $R_{s,t}^{tc} = \left(1-tc_{s,t}\right)R_{s,t}$.

These metrics are compared with those of two benchmark portfolios: The portfolio based on the returns of the portfolio based on recursive historical means and unconditional covariances (called hereafter Historical) and the equally weighted portfolio that only includes the risky assets (denoted hereafter by 1/N). According to DeMiguel et al. (2009), this is a portfolio very hard to outperform, especially if transaction costs are taken into account.

5 Empirical Results

This section presents the results of the forecasting accuracy of the various model combination schemes for each asset class and the performance metrics for the portfolios created and rebalanced using the different strategies.

Table 5 Statistics of out-of-sample portfolio discrete returns for different risk aversion coefficients

Strategy γ		Mean (%)			StD (%)			Skewness			Kurtosis			Jarque–Bera		
		3	5	10	3	5	10	3	5	10	3	5	10	3	5	10
Without transaction costs																
Mean		8.76	7.92	6.96	14.06	13.61	11.54	−0.09	−0.12	−0.11	0.42	0.44	**0.41**	33.09	43.44	30.19
WMSFE	10%	9.24	8.52	7.20	**13.09**	12.30	10.39	−0.05	−0.07	−0.12	0.48	0.57	0.90	55.42	104.66	446.29
	20%	11.28	9.48	7.68	14.10	12.78	10.57	0.12	−0.03	−0.12	0.63	0.55	0.88	150.82	89.81	409.21
	30%	11.76	9.84	7.92	14.55	12.96	10.67	0.24	0.00	−0.11	0.84	0.57	0.86	382.16	103.43	385.30
	40%	12.24	10.44	8.16	14.93	13.30	10.74	**0.32**	0.06	−0.10	1.02	0.61	0.85	649.51	134.33	369.09
	50%	12.24	11.16	8.52	15.03	14.48	11.12	0.31	**0.37**	−0.02	0.99	1.15	0.85	606.89	885.94	365.76
DMS and DMA	Small	**23.04**	**23.64**	**20.52**	14.34	13.37	11.57	−0.07	0.06	0.18	0.55	0.47	0.53	91.27	50.71	91.71
	Medium	19.32	19.32	16.92	16.04	15.07	12.71	0.26	0.35	0.41	0.79	0.90	0.92	321.79	478.77	518.66
	Large	8.28	6.60	4.68	13.79	**9.08**	**5.72**	−0.32	−0.17	0.45	0.56	0.58	1.15	134.64	119.22	917.30
DMS	Full	12.36	10.56	8.04	13.93	10.12	7.24	−0.12	0.20	**0.88**	0.56	0.72	1.76	104.96	238.47	2643.41
DMA	Full	17.40	20.04	19.20	16.87	15.42	12.71	−0.08	0.03	0.17	**0.37**	**0.33**	0.43	15.83	5.77	43.54
Historical		9.24	8.88	7.80	19.57	19.50	18.22	−0.39	−0.39	−0.46	0.77	0.78	0.93	337.57	348.44	553.93
1/N			8.64			11.74			−0.38			0.72			273.64	
With proportional transaction costs of 0.5%																
Mean		3.60	2.76	1.92	14.03	13.58	11.50	−0.10	−0.13	−0.13	0.44	0.46	**0.41**	40.18	51.37	31.74
WMSFE	10%	2.40	2.04	1.92	13.20	12.37	10.46	−0.05	−0.07	−0.13	0.49	0.59	0.96	59.90	118.57	529.50
	20%	4.56	3.12	2.64	14.13	12.82	10.60	0.10	−0.04	−0.13	0.61	0.56	0.93	138.10	98.60	483.38
	30%	5.04	3.48	2.76	14.55	12.99	10.70	0.21	−0.01	−0.13	0.81	0.58	0.91	337.05	109.05	452.82
	40%	5.64	4.08	3.00	14.90	13.30	10.81	**0.29**	0.04	−0.12	0.97	0.61	0.89	557.67	132.98	432.0
	50%	5.76	4.80	3.36	15.00	14.45	11.12	0.28	**0.33**	−0.05	0.94	1.09	0.88	518.02	766.81	414.17
DMS and DMA	Small	**18.96**	**19.92**	**17.64**	14.13	13.13	11.29	−0.11	0.01	0.13	0.55	0.47	0.54	96.11	48.84	93.81
	Medium	16.08	16.20	14.52	16.04	15.07	12.61	0.26	0.35	0.37	0.80	0.92	0.92	332.84	499.98	506.50
	Large	7.92	6.12	4.44	13.75	**9.01**	**5.65**	−0.33	−0.19	0.40	0.56	0.57	1.10	137.54	118.86	800.91

Table 5 (continued)

Strategy	γ	Mean (%)			StD (%)			Skewness			Kurtosis			Jarque–Bera		
		3	5	10	3	5	10	3	5	10	3	5	10	3	5	10
DMS	Full	11.28	9.72	7.44	13.72	9.87	7.00	−0.16	0.15	**0.81**	0.55	0.67	1.61	101.96	184.17	2161.76
DMA	Full	15.84	17.52	16.68	16.77	15.31	12.54	−0.10	0.00	0.13	**0.37**	**0.32**	0.43	17.02	4.56	37.57
Historical			0.96	−0.24	19.43	19.36	18.12	−0.38	−0.39	−0.46	0.77	0.78	0.93	330.77	342.54	550.30
1/N			8.52			11.74			−0.38			0.72			276.49	

This table reports the annualized mean, standard deviation, skewness, and kurtosis of the discrete returns of the portfolios based on the following strategies: Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). Also, the last three columns report the Jarque–Bera normality test, which are all significant at the 1% level, except the one for the DMA Full with γ=5, without and with transaction costs (these values are in italic). The portfolios were selected considering three risk aversion coefficients (γ=3, 5 and 10). "Historical" refers to the portfolio based on the historical means and covariances and 1/N is the equally weighted portfolio of the three risky assets. The mean and standard deviation are in percentage. In bold are the best values (highest mean and skewness and lowest standard deviation and kurtosis)

Table 4 reports the pseudo-R^2 and the significance of the adj-MSFE test. Although there is only one model for Small, Medium, and Large sets, the selection/average of δ is still made, among the 7 possible values reported in Sect. 4. The results of the DMS and DMS are the same for the Small, Medium, and Large models (the same situation occurs in Koop & Korobilis, 2013).

Generally, Table 4 reports low predictability of the forecasting schemes when compared with the historical mean, except for the Mean and WMSFE combinations for Bonds, which have a pseudo-R^2 statistically significant at the 5% level, ranging from 3.00 to 3.73%. For Stocks and REITs, the DMS and DMA of Medium models present the best results, with pseudo-R^2 of 0.28 and 3.76, respectively, but none of them are significant. These mixed results highlight the problem of model uncertainty faced when dealing with forecasting returns (Fugazza et al., 2015). The good results for Bonds are in line with Fisher et al. (2020).

We continue the analysis by studying the portfolios of CRRA investors with three risk aversion coefficients ($\gamma = 3, 5,$ and 10). We have not discarded any forecasting scheme or any asset. Table 5 reports the mean, standard deviation, skewness, kurtosis, and Jarque–Bera normality test of the discrete returns of the portfolio strategies. The portfolio returns are analyzed without transaction costs and with proportional transaction costs of 0.5%.

Several patterns emerge from Table 5 that we describe hereafter. First, the consideration of transaction costs does not alter the order of the statistics, most notably for the strategies with the best results. Transaction costs obviously decrease the mean returns for all strategies, but also decrease the standard deviation and the skewness and increase the kurtosis for most strategies. Second, the returns of the portfolios are non-normal, except for the portfolios resulting from the DMA Full with $\gamma = 5$. Third, the best results in terms of standard deviation, skewness, and kurtosis are scattered among the pairs strategy/γ, but most of the differences are small. Fourth, the increase of the risk aversion coefficient decreases the mean (with the exceptions of the DMS and DMA Small and Medium and DMA Full strategies, when this coefficient increases from 3 to 5), naturally decreases the standard deviation for all strategies and the results are mixed for the skewness and kurtosis, as in Diris et al. (2015). Fifth, the inclusion of more models in the WMSFE strategy improves the portfolios in terms of mean returns and skewness at a cost of an increase in the standard deviation, and in some cases increase in the kurtosis, but the strategy based on the mean forecast presents poor results. Sixth, most strategies outperform the Historical strategy without transaction costs regarding the mean, all have lower standard deviation and higher skewness, and most have lower kurtosis than the Historical strategy. This trend is more marked when transaction costs are considered, which turns all strategies superior in terms of mean, standard deviation, and skewness. Seventh, the 1/N strategy is better than the Historical (the only exception is the mean return of the Historical for $\gamma = 5$ without transaction costs), and its superiority clearly increases with transaction costs, with a visible impact on the mean returns. This implies that the 1/N strategy is more difficult to outperform, especially in the presence of transaction costs. Eight, although the majority of pairs strategy/γ have higher means than the 1/N portfolio without transaction costs, even
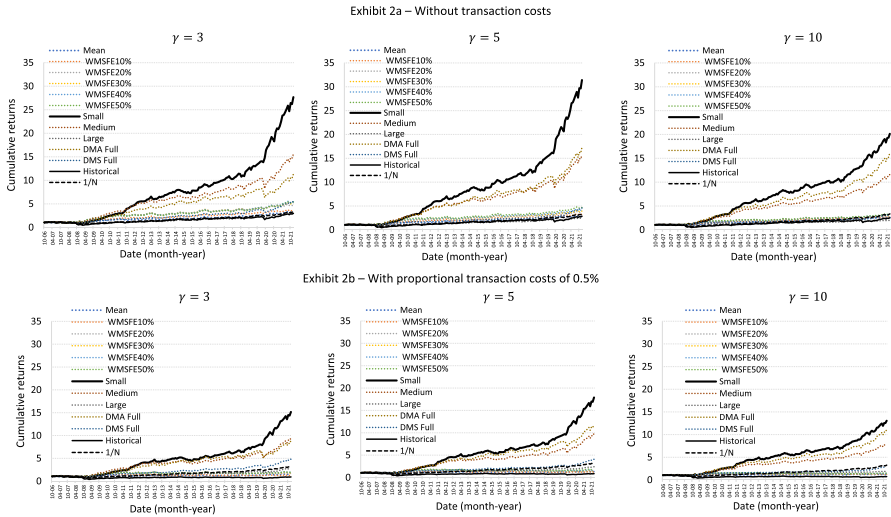
Fig. 2 Cumulative returns of the different portfolio strategies

at the expense of a higher standard deviation, this number reduces significantly when transaction costs are considered. In this case, only the DMS and DMA Small, Medium, and Full present better results in terms of mean returns. Finally, and most strikingly, a pattern appears from the comparison of the best results, which suggests that given the small differences in terms of standard deviation, skewness, and kurtosis, the best strategy is the Small, due to the evident high mean returns, especially for $\gamma = 5$. On average, across the three risk aversion coefficients, the Small strategy has a mean return of approximately 21% higher than the second-best model, the Medium strategy, with and without transaction costs, 159% higher than the Historical and 1/N strategies without transaction costs and 2,041%, and 121% higher than the Historical and 1/N strategies with transaction costs.

The superiority of the Small strategy is quite remarkable. This can be visualized in Fig. 2 which shows the cumulative returns of all strategies with and without transaction costs for $\gamma = 3$, 5 and 10.

The fact that the Small strategy, which only considers the first lags of the excess returns, outperforms the other strategies is in line with Welch and Goyal (2008), which conclude that none of the used additional variables has worked in forecasting equity excess returns. This is reinforced by Goyal et al. (2023). However, one the one hand, these good results also may imply that traditional predictive regressions, updated with schemes to resolve parameter uncertainty and instability, work out-of-sample (see, for instance, Rapach et al., 2013; Koop & Korobilis, 2013; Fisher et al., 2020) and, on the other hand, that the inclusion of REITs improve the risk-return trade-off, as suggested by Ling et al. (2020) and Zhu and Lizieri (2022).

Table 6 shows the out-of-sample performance metrics of the different portfolio strategies for the three different risk aversion coefficients. For each strategy/γ, we conduct a robust comparison with the Historical and 1/N portfolios using bootstrap p-values. The results presented in this table reinforce what has been discussed

Table 6 Portfolio performance out-of-sample for different risk aversion coefficients

Strategy\risk coef:γ		CER (%)			SR (%)			SOR (%)			CVaR at 5% (%)		
		3	5	10	3	5	10	3	5	10	3	5	10
Without transaction costs													
Mean		5.91	3.12	−0.13**	56.12	51.61	52.87	153.23	129.03	124.48	8.41**	8.65**	7.52***
WMSFE	**10%**	6.87	4.71	1.00**	64.05	62.08	60.54*	164.33	151.73	140.86	7.56**	7.45**	6.51***
	20%	8.65	5.47	1.38*	73.57*	67.45*	64.70*	203.15	170.07	155.63	7.54*	7.46**	6.59***
	30%	9.10	5.77	1.44**	75.01*	69.39*	65.60*	213.02*	182.10	157.73	7.54*	7.46**	6.64***
	40%	9.48	6.15*	1.61**	76.29*	71.55*	67.46*	217.02*	188.11	162.45	7.57*	7.45**	6.67**
	50%	9.40	6.46*	1.70*	75.55*	71.22*	68.47*	220.96*	204.61	171.51	7.68**	7.44**	6.70**
DMS and DMA	Small	21.81***,+++	21.11***,+++	15.11***,+++	153.57***,+++	169.61***,+++	168.68***,+++	338.88***,+++	412.56***,+++	450.56***,+++	7.27***	6.16***,++	4.74***,+++
	Medium	16.90***,+++	15.02***,+++	10.26***,+++	115.04***	121.90***	126.26***,++	320.92***,+++	354.38***,+++	377.30***,+++	7.95***	7.09***	5.45***,++
	Large	5.43	4.52*	3.25**	54.01**	62.54***	67.19***,+++	126.50	160.74**	213.82***,++	9.48***	5.78***,+++	2.91***,+++
DMS	Full	9.74*,++	8.37**,++	6.01***,++	81.91***	95.24***,++	98.09***,+++	199.76**,+	276.77***,+++	400.16***,+++	8.02***	4.84***,+++	2.38***,+++
DMA	Full	13.94***,+++	15.12***,+++	12.24***,+++	97.87***	123.64***	143.51***,+++	249.21***,++	318.63***,+++	404.19***,+++	9.93**	7.89***	5.25***,++
Historical		2.67	−3.05	−20.86	42.52	41.15	38.13	99.42	96.26	88.70	13.90	13.88	13.15
I/N		6.64			66.07			142.29			7.84		
With proportional transaction costs of 0.5%													
Mean		0.64	−2.05*	−5.16**	19.87	14.13	8.69	68.89	52.69	41.10	8.87***	9.11**	8.03***
WMSFE	**10%**	−0.26	−1.90*	−4.31**	11.63	9.56	10.66	52.11	45.00	49.66	8.21**	8.02***	7.05***
	20%	1.56	−1.07*	−3.82**	25.80	17.81	16.74	96.63	68.68	63.82	8.18**	8.02**	7.10**
	30%	2.03	−0.79*	−3.80**	29.06	20.38	17.81	108.08	76.63	67.37	8.16***	8.00***	7.14**
	40%	2.49*	−0.42*	−3.65**	32.16	23.92	19.96	118.35	86.08	71.44	8.16***	8.02**	7.17**
	50%	2.54*	−0.10*	−3.54**	32.55	27.61	22.52	118.14	100.24	78.50	8.23**	8.02**	7.20**
DMS and DMA	Small	17.13***,+++	16.80***,+++	11.96***,+++	127.57***,+++	144.40***,+++	147.26***,+++	289.59***,+++	342.67***,+++	386.42***,+++	7.53***	6.45***,+	5.00***,+++
	Medium	13.11***,+++	11.49***,+++	7.54***,+++	94.39***	101.35***	107.85***,+	268.97***,++	289.06***,+++	321.44***,+++	8.14***	7.31***	5.73***,++
	Large	4.97**	4.14***	3.00***	50.92***	58.75***	63.36***,+++	119.03***	150.29***	201.50***,+++	9.53***	5.83***,+++	2.96***,+++
DMS	Full	8.73***	7.61***,+	5.47***,+	75.80***	89.37***,++	92.87***,+++	187.06***	259.77***,+++	363.59***,+++	8.12***	4.93***,+++	2.47***,+++
DMA	Full	12.17***,++	12.38***,+++	9.51***,+++	88.89***	108.59***	125.58***,++	223.55***,+	280.47***,+++	370.69***,+++	10.10**	8.18***	5.50***,++

Table 5 (continued)

Strategy\risk coef:γ	CER (%)			SR (%)			SOR (%)			CVaR at 5% (%)		
	3	5	10	3	5	10	3	5	10	3	5	10
Historical	−4.95	−10.40	−26.86	3.01	0.48	−6.06	16.77	11.29	−3.06	14.47	14.47	13.79
1/N		6.46			64.64			139.10			7.86	

This table reports the annualized Certainty Equivalent Return (CER), Sharpe Ratio (SR), Sortino ratio (SOR), and CVaR at 5%, in percentage, of the portfolio returns based on the following strategies: Mean, WMSFE (10%, 20%, 30%, 40%, and 50% best models), DMS and DMA of the Small, Medium, Large and Full TVP-VAR(1). The portfolios were selected considering three relative risk aversion coefficients (γ=3, 5, and 10). These strategies are compared against two benchmarks. "Historical" refers to the portfolio based on historical means and covariances, and 1/N is the equally weighted portfolio of risky assets. The difference between the performance metrics of each strategy and the benchmark portfolios was tested using bootstrap p-values obtained from 100,000 bootstrap samples created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004). This p-values are computed considering the alternative hypothesis that this difference is positive (except for the CVaR, for which the alternative hypothesis is that the difference is negative). The significance of the tests is presented with asterisks (*, ** or ***) for the Historical portfolio and with plus signs (+,++ or +++) for the 1/N portfolio. One, two, and three asterisks or plus signs indicate significance at the 10%, 5%, and 1% levels, respectively. In bold are the best results (highest CER, SR, and SOR and lowest CVaR)

previously. The Mean and WMSFE strategies show poor performance for any of the risk aversion coefficients considered, but the picture is completely different for the DMS and DMA strategies. In these strategies, only the CER and SOR of the DMA Large for $\gamma = 3$, without transaction costs, is not significantly greater than the CER of the Historical strategy at the 10% level. The statistical significance of the difference of these metrics against the 1/N strategy is generally lower, namely for the Large, DMS Full and DMA Full, especially after transaction costs. The second-best strategy, the Medium portfolio, presents a better CER, SR, and SOR than the 1/N strategy at the 1% level, except the SR and CVaR without transaction costs. But undoubtably, the "winning horse" is the Small strategy (see, for instance, Koop & Korobilis, 2013). With and without transaction costs, this strategy has a higher CER, SR, and SOR than the Historical and 1/N strategies at any γ considered, all significant at the 1% level. The CVaR are lower than the 1/N portfolio, but it is not significant for $\gamma = 3$ and is significant at the 10% level for $\gamma = 5$. The differences in the SR and SOR, without transaction costs are remarkable and show an increasing trend with γ. Just to give a general idea, on average, across the three risk aversion coefficients, the Small strategy has an SR which is approximately 304% and 148%, and a SOR which is 323% and 182% higher than the Historical and 1/N strategies, respectively, without transaction costs. With transaction costs, the average SR of the Historical portfolios is even negative. The average SR of the Small strategy is 116% higher than the one of the 1/N portfolio. With transaction costs, the average SOR of the Small strategy is higher than the Historical and 1/N portfolios by 3,967% and 144%, respectively.

6 Conclusion

This research aims to study the predictability of monthly returns and portfolio allocation among three US-based assets, namely Stocks, proxied by the S&P 500 Total Return Index, Bonds proxied by the Barclays Capital US Aggregate Bond Index, and REITs, proxied by the FTSE Nareit US Real Estate All Equity REITs Index. The period spans from March 1976 to December 2021, however, the analysis is conducted out-of-sample, beginning in October 2006 (last 183 months), which corresponds to a period of high turbulence in the stock and REITs markets, due to the global financial crises and the impact of the Covid-19 pandemic.

We use several procedures to obtain the forecasts of the three classes of assets. The forecasts resulting from the mean and the Weighted Mean Squared Forecasting Errors (WMSFE) of different VAR(1) models, Dynamic Model Selection (DMS) and Dynamic Model Averaging (DMA) applied to TVP-VAR(1) models, which differ according to the predictor space used. We begin with a large set of predictors, a total of 155 (19 for stocks, 122 for Bonds, and 14 for REITs), and resort to a Genetic Algorithm to shrink the predictor space. The Genetic Algorithm renders a set of 17 predictors (8 for Stocks, 6 for Bonds, and 3 for REITs).

The results are negative for Stocks, but some positive results are obtained for REITs and especially for Bonds, for which the Mean and WMSFE present significant pseudo-R^2 at the 5% level. A possible explanation for these results is that

the lagged stock returns are correlated to Bonds and REITs, the lagged returns of REITs are correlated with Bonds, and an additional predictor from the stock market (Conservative minus aggressive factor) is correlated with REITs. In sum, although we did not find any predictability in Stocks, this class of assets helps predicting REITs and most notably Bonds.

Then we devised several portfolio strategies for CRRA investors considering three risk aversion parameters, 3, 5, and 10. These portfolios are rebalanced monthly, considering the forecasts obtained from recursive estimations.

Although the portfolios with the best results in terms of standard deviation, skewness, and kurtosis are scattered among the different portfolios strategies, it is remarkable the performance of the DMS and DMA portfolios that only uses as predictors lagged returns of the three asset classes, independently of the risk aversion coefficient. Without considering transaction costs, the annual mean return of this strategy is 23.04%, 23.64%, and 20.52%, while for the portfolio based on the historical mean and covariances are 9.24%, 8.88%, and 7.80% for the risk aversion parameters of 3, 5 and 10, respectively. The equally weighted portfolio only presents an annual mean return of 8.64% while the corresponding figure with proportional transaction costs of 0.5% is 8.52%. The inclusion of proportional trading costs of 0.5% greatly penalizes the mean returns of the historical portfolio, which presents the values of 1.44%, 0.96%, and -0.24% for the three risk aversion parameters. With transaction costs, the best portfolio strategy, although more penalized than the equally weighted portfolio, maintains its superiority, with mean returns of 18.96%, 19.92%, and 17.64%.

The aforementioned mean return discrepancies translate into the superiority of the best strategy in terms of Certainty Equivalent, Sharpe ratio, and Sortino ratio, which are all significantly higher at the level of 1% than the corresponding figures for the historical and equally weighted portfolios with and without transaction costs for all risk parameters considered (except the Sharpe ratio with transaction cost and a risk aversion parameter of 3, which is significant at the 5% level). Additionally, even the CVaR at 5% of the best strategy are lower than those of the benchmark portfolios, with most of them statistically significant. Just to put the above discussion in context, with transaction cost for a risk aversion parameter of 3, the best strategy shows a CER of 17.13%, a SR of 127.57%, a SOR of 289.59%, and a CVaR of 7.53% while the equally weighted portfolio shows a CER of 6.46%, a SR of 64.64%, a SOR of 139.10%, and a CVaR of 7.86%.

In a nutshell, our paper presents two main conclusions. First, it supports the literature that argues in favor of methodologies that incorporate dynamic modelling and parameter uncertainty and uses combinations or selection of several models. Second, predictability mainly comes from the information of the stock market to other markets, hence including into the portfolio other assets tends to improve its performance. Therefore, our research points out that a promising line of research is to include in the analysis other assets, for instance, commodities, industry indexes, or other international markets. In future work, we intend to explore this line of research.

Appendix

See Tables 7, 8 and 9.

Table 7 Description of Stocks predictors

Variable	Description	Source	Transformation
DP	Dividend price ratio	Welch and Goyal (2008)	–
Inf	Inflation	Welch and Goyal (2008)	–
Tbl	Treasury bills	Welch and Goyal (2008)	–
Lty	Long term yield	Welch and Goyal (2008)	–
Tms	Term spread	Welch and Goyal (2008)	–
Rf	Risk free rate	Welch and Goyal (2008)	–
b/m	Book-to-market value	Welch and Goyal (2008)	–
EP Ratio	Earning price ratio	Welch and Goyal (2008)	–
DP Ratio	Dividend payout ratio	Welch and Goyal (2008)	–
Svar	Stock variance	Welch and Goyal (2008)	–
D/y	Dividend yield	Welch and Goyal (2008)	–
Dfy	Default yield spread	Welch and Goyal (2008)	–
CAPE	Cyclically adjusted price-to-earnings ratio	Robert Shiller Database	–
Factors: SMB	Small minus big factor	Fama and French (2018)	–
Factors: HML	High minus low factor	Fama and French (2018)	–
Factors: RMW	Robust minus weak factor	Fama and French (2018)	–
Factors: CMA	Conservative minus aggressive factor	Fama and French (2018)	–
Ntis	Net equity expansion	Welch and Goyal (2008)	–
Ltr	Long term rate	Welch and Goyal (2008)	–

This table presents the 19 variables used to predict Stocks returns, proxied by the S&P 500 Total Return Index, and their online sources. We do not transform the data as it is common practice for the predictors considered. "–" indicates no transformations. Welch and Goyal (2008) database is available at the website https://sites.google.com/view/agoyal145, Robert Shiller database is available at http://www.econ.yale.edu/~shiller/data.htm, Fama and French (2018) database is available at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Table 8 Description of Bonds predictors

Variable	Description	Source	Transformation
PIy	Personal income	Ludvigson and Ng (2009)	Δln
PIless transfers	Personal income minus transfers	Ludvigson and Ng (2009)	Δln
RealConsumption expensures	Real consumption expensures	Ludvigson and Ng (2009)	Δln
M&Tsales	Manufacturing and trade sales	Ludvigson and Ng (2009)	Δln
Retail sales	retail sales	Ludvigson and Ng (2009)	Δln
IP:total	Industrial total production index	Ludvigson and Ng (2009)	Δln
IP:products	Industrial production index	Ludvigson and Ng (2009)	Δln
IP:final products	Industrial final production index	Ludvigson and Ng (2009)	Δln
IP:cons goods	Industrial consumption production index	Ludvigson and Ng (2009)	Δln
IP:cons durable goods	Industrial consumption durables production index	Ludvigson and Ng (2009)	Δln
IP:cons non-durable goods	Industrial consumption non-durables production index	Ludvigson and Ng (2009)	Δln
IP:bus equipament	Industrial bus equipment production index	Ludvigson and Ng (2009)	Δln
IP:matls	Industrial materials production index	Ludvigson and Ng (2009)	Δln
IP:dble material	Industrial durable materials production index	Ludvigson and Ng (2009)	Δln
IP:nondble materials	Industrial non-durable materials production index	Ludvigson and Ng (2009)	Δln
IP:mfg manufacturing	Industrial manufacturing production index	Ludvigson and Ng (2009)	Δln
IP:residental utilities	Industrial residential utilities production index	Ludvigson and Ng (2009)	Δln
IP:fuels	Industrial fuels productioni	Ludvigson and Ng (2009)	Δln
Caputil	Capacity utilization	Ludvigson and Ng (2009)	Δ
Helpwanted	Index of help-wanted advertising in newspapers	Ludvigson and Ng (2009)	Δ
Helpwanted/emp	Employment ratio	Ludvigson and Ng (2009)	Δln
EmpCPS total	Civilian labor force: employed, total	Ludvigson and Ng (2009)	Δln
EmpCPS	Civilian labor Force: employed	Ludvigson and Ng (2009)	Δln
wps1	PPI finish goods	Ludvigson and Ng (2009)	Δ2ln
wps2	PPI finish consumer goods	Ludvigson and Ng (2009)	Δ2ln
wps3	PPI intermediate goods	Ludvigson and Ng (2009)	Δ2ln
wps4	PPI crude price	Ludvigson and Ng (2009)	Δ2ln

Table 8 (continued)

Variable	Description	Source	Transformation
NAPMempl	Employment index	Ludvigson and Ng (2009)	–
inflation	Inflation	Welch and Goyal (2008)	–
Macro Factors	8 macro factors	Ludvigson and Ng (2009)	–
U:all	Unemployment rate	Ludvigson and Ng (2009)	Δ
U:mean	Mean unemployment duration	Ludvigson and Ng (2009)	Δ
U<5	Less than 5 years unemployment	Ludvigson and Ng (2009)	$\Delta\ln$
U5-4	5-4 years unemployment	Ludvigson and Ng (2009)	$\Delta\ln$
U15+	More than 15 years unemployment	Ludvigson and Ng (2009)	$\Delta\ln$
U15-26	15–26 years unemployment	Ludvigson and Ng (2009)	$\Delta\ln$
U27+	More than 27 years unemployment	Ludvigson and Ng (2009)	$\Delta\ln$
UIclaims	Average weekly initial claims	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:total	Total employees on non-farm payrolls	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:gds	Employees on goods sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:mining	Employees on mining sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:const	Employees on construction sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:mfg	Employees on manufacturing goods sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:dble	Employees on durables goods sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:nondbles	Employees on non-durables goods sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:services	Employees on services sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:TTU	Employees on transports sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:wholesale	Employees on wholesale sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:retail	Employees on retail sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:FIRE	Employees on financial sector	Ludvigson and Ng (2009)	$\Delta\ln$
Emp:Govt	Employees on government sector	Ludvigson and Ng (2009)	$\Delta\ln$
Avghrs	Average weekly hours of production	Ludvigson and Ng (2009)	–
Overtime:mfg	Overtime: average weekly hours of production on manufactory goods	Ludvigson and Ng (2009)	Δ

Table 8 (continued)

Variable	Description	Source	Transformation
Avghrs: mfg	Average weekly hours of production on manufactory goods	Ludvigson and Ng (2009)	–
1-year forward rate	1-year forward rate	US Department of the Treasury	–
2-year forward rate	2-year forward rate	US Department of the Treasury	–
5-year forward rate	5-year forward rate	US Department of the Treasury	–
10-year forward rate	10-year forward rate	US Department of the Treasury	–
T10Y2YM	10-year treasury constant maturity minus 2-month treasury constant maturity	FRED	–
T10Y3YM	10-year treasury constant maturity minus 3-month treasury constant maturity	FRED	–
Permit	New house permits	Ludvigson and Ng (2009)	ln
Permitne	New house permits, Northeast	Ludvigson and Ng (2009)	ln
permitmw	New house permits, Midwest	Ludvigson and Ng (2009)	ln
permits	New house permits, South	Ludvigson and Ng (2009)	ln
permitw	New house permits, West	Ludvigson and Ng (2009)	ln
amd	New order for consumer goods	Ludvigson and Ng (2009)	Δln
and	New order for durable goods	Ludvigson and Ng (2009)	Δln
amdmu	New order for non-defence capital goods	Ludvigson and Ng (2009)	Δln
Businv	Total business inventories	Ludvigson and Ng (2009)	Δln
isratio	Inventories-to-sales ratio	Ludvigson and Ng (2009)	Δ
m1	M_1 money stock	Ludvigson and Ng (2009)	Δ2ln
m2	M_2 money stock	Ludvigson and Ng (2009)	Δ2ln
m2real	M_2 real money stock	Ludvigson and Ng (2009)	Δ2ln
Amb	St. Louis adjusted monetary base	Ludvigson and Ng (2009)	Δ2ln
Totresns	Total reserves	Ludvigson and Ng (2009)	Δ2ln
nonborr	Reserves of depository institutions	Ludvigson and Ng (2009)	Δ
busloans	Commercial and industrial loans	Ludvigson and Ng (2009)	Δ2ln
Realln	Real estate loans	Ludvigson and Ng (2009)	Δln
nonrev	Nonrevolving credit	Ludvigson and Ng (2009)	Δ2ln

Table 8 (continued)

Variable	Description	Source	Transformation
conspi	Consumer credit to personal income	Ludvigson and Ng (2009)	ln
Ff	Federal fund rate	Ludvigson and Ng (2009)	ln
cp3m	Commercial paper rate	Ludvigson and Ng (2009)	ln
tb3m	3-month treasury Bill	Ludvigson and Ng (2009)	ln
tb6m	6-month treasury Bill	Ludvigson and Ng (2009)	ln
gs1	1-year treasury rate	Ludvigson and Ng (2009)	ln
gs5	5-years treasury rate	Ludvigson and Ng (2009)	ln
gs10	10-years treasury rate	Ludvigson and Ng (2009)	ln
Aaa	Aaa bond yield	Ludvigson and Ng (2009)	ln
Baa	Baa bond yield	Ludvigson and Ng (2009)	ln
Comp	Composite federal fund spread	Ludvigson and Ng (2009)	–
tb3	3-month federal fund spread	Ludvigson and Ng (2009)	–
tb6	6-month federal fund spread	Ludvigson and Ng (2009)	–
t1yffm	1-year federal fund spread	Ludvigson and Ng (2009)	–
t5yffm	5-year federal fund spread	Ludvigson and Ng (2009)	–
tw	Trade weight for Us Dollar	Ludvigson and Ng (2009)	Δln
Exs	Exchange rate Switzerland	Ludvigson and Ng (2009)	Δln
Exj	Exchange rate Japan	Ludvigson and Ng (2009)	Δln
Exu	Exchange rate UK	Ludvigson and Ng (2009)	Δln
Exc	Exchange rate Canada	Ludvigson and Ng (2009)	Δln
oil	Crude oil	Ludvigson and Ng (2009)	Δ2ln
Ppic	PPI metals	Ludvigson and Ng (2009)	Δ2ln
Cpiau	CPI all items	Ludvigson and Ng (2009)	Δ2ln
Cpiap	CPI apparel	Ludvigson and Ng (2009)	Δ2ln
Cpitr	CPI transports	Ludvigson and Ng (2009)	Δ2ln
Cpimed	CPI medical care	Ludvigson and Ng (2009)	Δ2ln

Table 8 (continued)

Variable	Description	Source	Transformation
cus1	CPI commodities	Ludvigson and Ng (2009)	Δ2ln
cus2	CPI durables	Ludvigson and Ng (2009)	Δ2ln
cus3	CPI services	Ludvigson and Ng (2009)	Δ2ln
Cpiul	CPI all but food	Ludvigson and Ng (2009)	Δ2ln
cus4	CPI all but shelter	Ludvigson and Ng (2009)	Δ2ln
cus5	CPI all but medical care	Ludvigson and Ng (2009)	Δ2ln
Pcepi	Personal consumption chain	Ludvigson and Ng (2009)	Δ2ln
ddur	Personal consumption durable	Ludvigson and Ng (2009)	Δ2ln
Dnd	Personal consumption non-durable	Ludvigson and Ng (2009)	Δ2ln
dse	Personal consumption Services	Ludvigson and Ng (2009)	Δ2ln
ces7	Average hours of earnings in goods sector	Ludvigson and Ng (2009)	Δ2ln
ces8	Average hours of earnings in construction sector	Ludvigson and Ng (2009)	Δ2ln
ces9	Average hours of earnings in Manufacturing sector	Ludvigson and Ng (2009)	Δ2ln
Mzms	Money stock	Ludvigson and Ng (2009)	Δ2ln
Invest	Securities in bank credit	Ludvigson and Ng (2009)	Δ2ln
vxocls	VXO's	Ludvigson and Ng (2009)	–
CP	Forward rate factor	Cochrane and Piazzesi (2009)	–

Notes: This table presents the 122 variables used to predict Bonds, proxied by the Barclays Capital US Aggregate Bond Index, the sources, and the transformations made to obtain the variables used in the literature and to correct non-stationarity ("Δ" is the first difference, "ln" is the log transformation, "Δln" is the first log difference, and "Δ2ln" is the second log difference. "–" indicates no transformations). Ludvigson and Ng (2009) database is available at https://www.sydneyludvigson.com/data-and-appendixes, Cochrane and Piazzesi (2009) database is available at https://www.aeaweb.org/articles?id=10.1257/0002828053828581, US Department of the Treasury database is available at https://home.treasury.gov/data/treasury-open-data, Federal Reserve Bank of St. Louis (FRED) database is available at http://research.stlouisfed.org/fred2/categories/22, Welch and Goyal (2008) database is available at https://sites.google.com/view/agoyal145

Table 9 Description of REITs predictors

Variable	Description	Source	Transformation
Dividend price ratio	Dividend price ratio	Welch and Goyal (2008)	–
Inflation	inflation	Welch and Goyal (2008)	–
Tms	Term spread	Welch and Goyal (2008)	–
Dfy	Default yield spread	Welch and Goyal (2008)	–
chng income	Change in income	FRED	–
chng adult population	Change in adult population	FRED	$\Delta\ln$
chng employment	Change in employment	FRED	–
housing starts	Rate of growth of housing units	Federal Housing Finance Agency	$\%\Delta$
Mortgage loan amount	Amount of mortgage loan	Federal Housing Finance Agency	$\%\Delta$
Purchase price	Houses purchase price	Federal Housing Finance Agency	$\%\Delta$
Loan to price ratio	loan to price ratio	Federal Housing Finance Agency	$\Delta\ln$
Rent vacancy rate	Rent vacancy rate	Federal Housing Finance Agency	$\ln$
Change in mortgage loan	Percentage change mortgage loan	Federal Housing Finance Agency	–
30-year conventional mortgage rate	30-Year conventional mortgage	Federal Housing Finance Agency	$\Delta\ln$

This table presents the 14 variables used to predict REITs returns, proxied by FTSE Nareit US Real Estate All Equity REITs Index, the data online sources, the transformations made to obtain the variables used in the literature and to correct non-stationarity ("Δ" is the first difference, "ln" is the log transformation, "Δln" is the first log difference, and "Δ2ln" is the second log difference. "–" indicates no transformations). Welch and Goyal (2008) database is available at https://sites.google. com/view/agoyal145 and Federal Reserve Bank of St. Louis (FRED) database is available at http://research.stlouisfed.org/fred2/categories/22, Federal Housing Finance Agency:

https://www.fhfa.gov/DataTools/Downloads/Pages/National-Mortgage-Database-Aggregate-Data.aspx, and http://www.freddiemac.com/pmms/pmms30.html

Data Availability The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

Almadi, H., Rapach, D. E., & Suri, A. (2014). Return predictability and dynamic asset allocation: How often should investors rebalance? *The Journal of Portfolio Management, 40*(4), 16–27. https://doi.org/10.3905/jpm.2014.40.4.016

Atiya, A. F. (2020). Why does forecast combination work so well? *International Journal of Forecasting, 36*(1), 197–200. https://doi.org/10.1016/j.ijforecast.2019.03.010

Bahrami, A., Shamsuddin, A., & Uylangco, K. (2019). Are advanced emerging market stock returns predictable? A regime-switching forecast combination approach. *Pacific-Basin Finance Journal, 55*, 142–160. https://doi.org/10.1016/j.pacfin.2019.02.003

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance, 55*(5), 2219–2257. https://doi.org/10.1111/0022-1082.00285

Banbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics, 25*, 71–92. https://doi.org/10.1002/jae.1137

Bandi, F. M., Perron, B., Tamoni, A., & Tebaldi, C. (2019). The scale of predictability. *Journal of Econometrics, 208*(1), 120–140. https://doi.org/10.1016/j.jeconom.2018.09.008

Barberis, N. (2000). Investing for the long run when returns are predictable. *The Journal of Finance, 55*(1), 225–264. https://doi.org/10.1111/0022-1082.00205

Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operational Research Quarterly, 20*, 451–468. https://doi.org/10.1057/jors.1969.103

Bauer, R. J., Jr. (1994). *Genetic algorithms and investment strategies*. UK: John Wiley & Sons.

Beracha, E., Feng, Z., & Hardin, W. G. (2019). REIT operational efficiency: Performance, risk, and return. *The Journal of Real Estate Finance and Economics, 58*(3), 408–437. https://doi.org/10.1007/s11146-018-9655-2

Bhuyan, R., Kuhle, J., Ikromov, N., & Chiemeke, C. (2014). Optimal portfolio allocation among REITs, stocks, and long-term bonds: An empirical analysis of US financial markets. *Journal of Mathematical Finance, 4*(2), 104–112. https://doi.org/10.4236/jmf.2014.42010

Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. *Journal of Financial and Quantitative Analysis, 49*, 633–661.

Campbell, J. Y., & Shiller, R. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*(3), 195–228. https://doi.org/10.1093/rfs/1.3.195

Campbell, J. Y., Chan, Y. L., & Viceira, L. M. (2003). A multivariate model of strategic asset allocation. *Journal of Financial Economics, 67*(1), 41–80. https://doi.org/10.1016/s0304-405x(02)00231-3

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies, 21*, 1509–1531. https://doi.org/10.1093/rfs/hhm055

Campbell, S. D., & Diebold, F. X. (2009). Stock returns and expected business conditions: A half century of direct evidence. *Journal of Business and Economic Statistics, 27*, 266–278. https://doi.org/10.1198/jbes.2009.0025

Carriero, A., Kapetanios, G., & Marcellino, M. (2009). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting, 25*, 400–417. https://doi.org/10.1016/j.ijforecast.2009.01.007

Cenesizoglu, T., & Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance, 36*(11), 2974–2987. https://doi.org/10.1016/j.jbankfin.2012.06.008

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics, 138*, 291–311. https://doi.org/10.1016/j.jeconom.2006.05.023

Cochrane, J. H. (2007). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies, 21*(4), 1533–1575. https://doi.org/10.1093/rfs/hhm046

Cochrane, J. H., & Piazzesi, M. (2009). Decomposing the yield curve. In *AFA 2010 Atlanta Meetings Paper*.

Consonni, V., Baccolo, G., Gosetti, F., Todeschini, R., & Ballabio, D. (2021). A MATLAB toolbox for multivariate regression coupled with variable selection. *Chenometrics and Intelligent Laboratory Systems, 213*, 104313. https://doi.org/10.1016/j.chemolab.2021.104313

Cooper, I., & Priestly, R. (2009). Time-varying risk premiums and the output gap. *Review of Financial Studies, 22*, 2801–2833. https://doi.org/10.1093/rfs/hhn087

Dai, Z., Dong, X., Kang, J., & Hong, L. (2020). Forecasting stock market returns: New technical indicators and two-step economic constraint method. *The North American Journal of Economics and Finance, 53*, 101216. https://doi.org/10.1016/j.najef.2020.101216

Dai, Z., Zhu, H., & Kang, J. (2021). New technical indicators and stock returns predictability. *International Review of Economics & Finance, 71*, 127–142. https://doi.org/10.1016/j.iref.2020.09.006

Dangl, T., & Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics, 106*(1), 157–181. https://doi.org/10.1016/j.jfineco.2012.04.003

Daskalaki, C., & Skiadopoulos, G. (2011). Should investors include commodities in their portfolios after all? New evidence. *Journal of Banking & Finance, 35*(10), 2606–2626. https://doi.org/10.1016/j.jbankfin.2011.02.022

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies, 22*(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075

Diris, B., Palm, F., & Schotman, P. (2015). Long-term strategic asset allocation: An out-of-sample evaluation. *Management Science, 61*(9), 2185–2202. https://doi.org/10.1287/mnsc.2014.1924

Dichtl, H., Drobetz, W., Neuhierl, A., & Wendt, V. S. (2021). Data snooping in equity premium prediction. *International Journal of Forecasting, 37*(1), 72–94. https://doi.org/10.1016/j.ijforecast.2020.03.002

Driesprong, G., Jacobsen, B., & Maat, B. (2008). Striking oil: Another puzzle? *Journal of Financial Economics, 89*, 307–327. https://doi.org/10.1016/j.jfineco.2007.07.008

Elliott, G., Gargano, A., & Timmermann, A. (2013). Complete subset regressions. *Journal of Econometrics, 177*(2), 357–373. https://doi.org/10.1016/j.jeconom.2013.04.017

Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature, 46*(1), 3–56. https://doi.org/10.1257/jel.46.1.3

Fama, E. F. (1990). Term-structure forecasts of interest rates, inflation and real returns. *Journal of Monetary Economics, 25*(1), 59–76. https://doi.org/10.1016/0304-3932(90)90045-6

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Ferreira, M. A., & Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics, 100*(3), 514–537. https://doi.org/10.1016/j.jfineco.2011.02.003

Feunou, B., Jahan-Parvar, M. R., & Okou, C. (2018). Downside variance risk premium. *Journal of Financial Econometrics, 16*(3), 341–383. https://doi.org/10.1093/jjfinec/nbx020

Fisher, J. D., Pettenuzzo, D., & Carvalho, C. M. (2020). Optimal asset allocation with multivariate Bayesian dynamic linear models. *Annals of Applied Statistics, 14*(1), 299–338. https://doi.org/10.1214/19-AOAS1303

Fugazza, C., Guidolin, M., & Nicodano, G. (2015). Equally weighted vs. long-run optimal portfolios. *European Financial Management, 21*(4), 742–789. https://doi.org/10.1111/eufm.12042

Gao, L., Han, Y., Li, S. Z., & Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics, 129*, 394–414. https://doi.org/10.1016/j.jfineco.2018.05.009

Gao, X., & Nardari, F. (2018). Do commodities add economic value in asset allocation? New evidence from time-varying moments. *Journal of Financial and Quantitative Analysis, 53*(1), 365–393. https://doi.org/10.1017/S002210901700103X

Gargano, A., Pettenuzzo, D., & Timmermann, A. (2019). Bond return predictability: Economic value and links to the macroeconomy. *Management Science, 65*(2), 508–540. https://doi.org/10.1287/mnsc.2017.2829

Ganie, I. R., Wani, T. A., & Yadav, M. P. (2022). Impact of COVID-19 outbreak on the stock market: An evidence from select economies. *Business Perspectives and Research*. https://doi.org/10.1177/22785337211073635

Goetzmann, W. N., Li, L., & Rouwenhorst, K. G. (2001). *Long-term global market correlations* (No. w8612). National Bureau of Economic Research. https://www.nber.org/papers/w8612

Golez, B., & Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics, 127*(2), 248–263. https://doi.org/10.1016/j.jfineco.2017.12.007

Goyal, A., Welch, I., & Zafirov, A. (2023). A Comprehensive 2022 look at the empirical performance of equity premium prediction (September 28, 2023). *Swiss Finance Institute Research Paper*, No. 21-85. https://doi.org/10.2139/ssrn.3929119.

Guidolin, M., & Hyde, S. (2012). Can VAR models capture regime shifts in asset returns? A long horizon strategic asset allocation perspective. *Journal of Banking & Finance, 36*(3), 695–716. https://doi.org/10.1016/j.jbankfin.2011.10.011

Guo, H. (2006). On the out-of-sample predictability of stock market returns. *Journal of Business, 79*, 645–670. https://doi.org/10.1086/499134

Habbab, F. Z., Kampouridis, M., & Voudouris, A. A. (2022). Optimizing mixed-asset portfolios involving REITs. In *IEEE Symposium on Computational Intelligence for Financial Engineering and Economics*,1–8. https://doi.org/10.1109/CIFEr52523.2022.9776074

Hjalmarsson, E. (2010). Predicting global stock returns. *Journal of Financial and Quantitative Analysis, 45*, 49–80. https://doi.org/10.1017/S0022109009990469

Hammerschmid, R., & Lohre, H. (2018). Regime shifts and stock return predictability. *International Review of Economics & Finance, 56*, 138–160. https://doi.org/10.1016/j.iref.2017.10.021

Hill, R. A., & Rodrigues, P. M. M. (2022). Forgetting approaches to improve forecasting. *Journal of Forecasting, 41*(7), 1356–1371. https://doi.org/10.1002/for.2877

Hirshleifer, D., Hou, K., & Teoh, S. H. (2009). Accruals, cash flows, and aggregate stock returns. *Journal of Financial Economics, 91*, 389–406. https://doi.org/10.1016/j.jfineco.2007.11.009

Hong, H., Torous, W., & Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics, 83*, 367–396. https://doi.org/10.1016/j.jfineco.2005.09.010

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies, 28*(3), 791–837. https://doi.org/10.1093/rfs/hhu080

Huynh, T. D., & Xia, Y. (2020). Climate change news risk and corporate bond returns. *Journal of Financial and Quantitative Analysis, 56*(6), 1985–2009. https://doi.org/10.1017/S0022109020000757

Jagannathan, R., & Liu, B. (2019). Dividend dynamics, learning, and expected stock index returns. *The Journal of Finance, 74*(1), 401–448. https://doi.org/10.1111/jofi.12731

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics, 132*(1), 126–149. https://doi.org/10.1016/j.jfineco.2018.10.001

Johannes, M., Korteweg, A., & Polson, N. (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance, 69*(2), 611–644. https://doi.org/10.1111/jofi.12121

Karathanasopoulos, A., Theofilatos, K. A., Sermpinis, G., Dunis, C., Mitra, S., & Stasinakis, C. (2016). Stock market prediction using evolutionary support vector machines: An application to the ASE20 index. *European Journal of Finance, 22*(12), 1145–1163. https://doi.org/10.1080/1351847x.2015.1040167

Kilic, M., & Shaliastovich, I. (2019). Good and bad variance premia and expected returns. *Management Science, 65*(6), 2522–2544. https://doi.org/10.2139/ssrn.3024086

Koop, G., & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics, 177*(2), 185–198. https://doi.org/10.1016/j.jeconom.2013.04.007

Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics, 44*(2), 169–203. https://doi.org/10.1016/S0304-405X(97)00002-0

Le Bris, D., Goetzmann, W. N., & Pouget, S. (2019). The present value relation over six centuries: The case of the Bazacle company. *Journal of Financial Economics, 132*(1), 248–265. https://doi.org/10.1016/j.jfineco.2017.03.011

Leardi, R., Boggia, R., & Terrile, M. (1992). Genetic algorithms as a strategy for feature selection. *Journal of Chenometrics, 6*, 267–281. https://doi.org/10.1002/cem.1180060506

Leardi, R., & Gonzáles, A. L. (1998). Genetic algorithms applied to feature selection in PLS regression: How and when to use them. *Chenometrics and Intelligent Laboratory Systems, 41*, 195–207. https://doi.org/10.1016/S0169-7439(98)00051-3

Leigh, W., Purvis, R., & Ragusa, J. M. (2002). Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support. *Decision Support Systems, 32*(4), 361–377. https://doi.org/10.1016/s0167-9236(01)00121-x

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance, 56*(3), 815–849. https://doi.org/10.1111/0022-1082.00347

Ling, D. C., Wang, C., & Zhou, T. (2020). A first look at the impact of COVID-19 on commercial real estate prices: Asset-level evidence. *The Review of Asset Pricing Studies, 10*(4), 669–704. https://doi.org/10.1093/rapstu/raaa014

Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies, 22*(12), 5027–5067. https://doi.org/10.1093/rfs/hhp081

Ma, F., Lu, X., Liu, J., & Huang, D. (2022a). Macroeconomic attention and stock market return predictability. *Journal of International Financial Markets, Institutions and Money, 79*, 101603. https://doi.org/10.1016/j.intfin.2022.101603

Ma, F., Lu, F., & Tao, Y. (2022b). Geopolitical risk and excess stock return predictability: New evidence from a century of data. *Finance Research Letters, 50*, 103211. https://doi.org/10.1016/j.frl.2022.103211

Maeso, J., & Lionel, M. (2020). Measuring portfolio rebalancing benefits in equity markets. *The Journal of Portfolio Management., 46*(4), 94–109. https://doi.org/10.3905/jpm.2020.1.130

Maio, P., & Santa-Clara, P. (2015). Dividend yields, dividend growth, and return predictability in the cross section of stocks. *Journal of Financial and Quantitative Analysis, 50*(1–2), 33–60. https://doi.org/10.1017/S0022109015000058

McMillan, D. (2021). Forecasting U.S. stock returns. *The European Journal of Finance, 27*(1–2), 86–109. https://doi.org/10.1080/1351847X.2020.1719175

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science, 60*(7), 1772–1791. https://doi.org/10.1287/mnsc.2013.1838

Nonejad, N. (2018). Déjà vol oil? Predicting S&P 500 equity premium using crude oil price volatility: Evidence from old and recent time-series data. *International Review of Financial Analysis, 58*, 260–270. https://doi.org/10.1016/j.irfa.2018.03.012

Ozcalici, M., & Bumin, M. (2022). Optimizing filter rule parameters with genetic algorithm and stock selection with artificial neural networks for an improved trading: The case of Borsa Istanbul. *Expert Systems with Applications, 208*, 118120. https://doi.org/10.1016/j.eswa.2022.118120

Parslow, J., Cressie, N., Campbell, E. P., Jones, E., & Murray, L. (2013). Bayesian learning and predictability in a stochastic nonlinear dynamical model. *Ecological Applications, 23*(4), 679–698. https://doi.org/10.1890/12-0312.1

Piatti, I., & Trojani, F. (2019). Dividend growth predictability and the price-dividend ratio. *Management Science*. https://doi.org/10.1287/mnsc.2018.3155

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association, 89*(428), 1303–1313. https://doi.org/10.1080/01621459.1994.10476870

Politis, D. N., & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews, 23*(1), 53–70. https://doi.org/10.1081/etc-120028836

Pontiff, J., & Schall, L. D. (1998). Book-to-market ratios as predictors of market returns. *Journal of Financial Economics, 49*(2), 141–160. https://doi.org/10.1016/S0304-405X(98)00020-8

Rangvid, J. (2006). Output and expected returns. *Journal of Financial Economics, 81*, 595–624. https://doi.org/10.1016/j.jfineco.2005.07.010

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *The Review of Financial Studies, 23*(2), 821–862. https://doi.org/10.1093/rfs/hhp063

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: What is the role of the United States? *The Journal of Finance, 68*(4), 1633–1662. https://doi.org/10.1111/jofi.12041

Rapach, D. E., & Zhou, G. (2013). Forecasting stock returns. *In Handbook of economic forecasting* (Vol. 2, pp. 328–383). Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00006-2

Rapach, D. E., Ringgenberg, M. C., & Zhou, G. (2016). Short interest and aggregate stock returns. *Journal of Financial Economics, 121*, 46–65. https://doi.org/10.1016/j.jfineco.2016.03.004

JP Morgan/Reuters (1996). *RiskMetrics^{TM} - Technical Document*, 4th Ed., New York. Available at: https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Santos, T., & Veronesi, P. (2006). Labor income and predictable stock returns. *Review of Financial Studies, 19*, 1–44. https://doi.org/10.1093/rfs/hhj006

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society.* https://doi.org/10.2307/1912017

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*(3), 375–421. https://doi.org/10.1016/S0304-405X(99)00041-0

Szczygielski, J. J., Charteris, A., Bwanya, P. R., & Brzeszczyński, J. (2023). Which COVID-19 information really impacts stock markets? *Journal of International Financial Markets, Institutions and Money, 84*, 101592. https://doi.org/10.1016/j.intfin.2022.101592

Thornton, D. L., & Valente, G. (2012). Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *Review of Financial Studies, 25*(10), 3141–3168. https://doi.org/10.1093/rfs/hhs069

Tian, Z., Li, H., & Li, F. (2021). A combination forecasting model of wind speed based on decomposition. *Energy Reports, 7*, 1217–1233. https://doi.org/10.1016/j.egyr.2021.02.002

Timmermann, A. (2006). Forecast combinations. *Handbook of Economic Forecasting, 1*, 135–196. https://doi.org/10.1016/S1574-0706(05)01004-9

Van Binsbergen, J. H., & Koijen, R. S. (2010). Predictive regressions: A present-value approach. *The Journal of Finance, 65*(4), 1439–1471. https://doi.org/10.1111/j.1540-6261.2010.01575.x

Wang, Y., Liu, L., Ma, F., & Diao, X. (2018). Momentum of return predictability. *Journal of Empirical Finance, 45*, 141–156. https://doi.org/10.1016/j.jempfin.2017.11.003

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies, 21*(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering, 26*(1), 97–107. https://doi.org/10.1109/TKDE.2013.109

Xu, Y., Liang, C., & Wang, J. (2023). Financial stress and returns predictability: Fresh evidence from China. *Pacific-Basin Finance Journal, 78*, 101980. https://doi.org/10.1016/j.pacfin.2023.101980

Zhang, Y., Ma, F., Shi, B., & Huang, D. (2018). Forecasting the prices of crude oil: An iterated combination approach. *Energy Economics, 70*, 472–483. https://doi.org/10.1016/j.eneco.2018.01.027

Zhang, Y., Zeng, Q., Ma, F., & Shi, B. (2019). Forecasting stock returns: Do less powerful predictors help? *Economic Modelling, 78*, 32–39. https://doi.org/10.1016/j.econmod.2018.09.014

Zhu, B., & Lizieri, C. (2022). Local beta: Has local real estate market risk been priced in REIT returns? *The Journal of Real Estate Finance and Economics.* https://doi.org/10.1007/s11146-022-09890-4