CATÓLICA
CEGE · CENTRO DE ESTUDOS EM GESTÃO E ECONOMIA
PORTO

Forecast combination in the frequency domain

Gonçalo Faria
Fabio Verona

CEGE WORKING PAPER SERIES WP24.01
January 2024

fct Fundação para a Ciência e a Tecnologia
This work was supported by Fundação para a Ciência e Tecnologia (through project UIDB/00731/2020).

Forecast combination in the frequency domain*

Gonçalo Faria[†] Fabio Verona[‡]

8 January 2024

Abstract

We propose a new forecasting method – forecast combination in the frequency domain – that takes into account the fact that predictability is time *and* frequency dependent. We use this method to forecast the equity premium and real GDP growth rate. Combining forecasts in the frequency domain produces markedly more accurate predictions relative to the standard forecast combination in the time domain, both in terms of statistical and economic measures of out-of-sample predictability. In a real-time forecasting exercise, the flexibility of this method allows to capture remarkably well the sudden and abrupt drops associated with recessions and further improve predictability.

Keywords: forecast combination, frequency domain, equity premium, GDP growth, Haar filter

JEL classification: C58, G11, G17

* We are grateful for the comments of participants at several conferences and seminars. The views expressed here are those of the authors and do not necessarily reflect those of the Bank of Finland. Faria gratefully acknowledges financial support from FCT through project UIDB/00731/2020. Research at *CEF.UP* has been financed through FCT in the framework of the project with reference UIDB/04105/2020.

† Universidade Católica Portuguesa, Católica Porto Business School, Research Centre in Management and Economics, and University of Vigo – *RGEA* (gfaria@ucp.pt).

‡ Bank of Finland – Monetary Policy and Research Department, and University of Porto – *CEF.UP* (fabio.verona@bof.fi). Corresponding author.

1 Introduction

Multiple forecasts of the same variable are often available, and decision makers have to figure out how to exploit best the information of each individual forecast. This is challenging as predictability is time dependent: no individual variable is clearly superior to other variables consistently throughout time (see e.g. Stock and Watson, 2004 and Henkel et al., 2011).

The fact that the best forecasting variable changes over time renders individual variables unreliable predictors. A method proposed to overcome this problem is forecast combination. Since Bates and Granger's (1969) seminal paper, it has been know that combining forecasts across models often produces a forecast that performs better than the best individual model. Forecast combination achieves a compromise between smoothing out the excessive noise in the individual forecasts and the need to retain some of the volatility that allows to capture the time-varying behavior of the variable of interest. Recent contributions include Rapach et al. (2010), Pettenuzzo and Ravazzolo (2016), and Pirschel and Wolters (2018).[1]

More recent empirical literature has shown that predictability is *also* frequency dependent. Some frequencies of a variable might be good predictors for the variable of interest, others might not. For instance, Faria and Verona (2020) show that the low frequency of the term spread has good predictive power (for equity returns), while the remaining frequencies don't. Likewise, some frequencies of the targeted variable need to be forecasted well. Faria and Verona (2021) and Martins and Verona (2021) show that it is crucial to predict well the low frequencies of the equity premium and inflation, respectively, while the other frequencies mainly bring noise to the forecast exercise.

In this paper we propose a method that improves upon existing methods by reducing the forecast noise simultaneously in the time and in the frequency domain. It is a forecast combination method

[1] Time-varying parameter models (e.g. Dangl and Halling, 2012) also allow to overcome the problem of instability over time of the predictor. Other methods that incorporate information from a large set of potential predictors in a predictive regression framework include Bayesian model averaging (Cremers, 2002) and factor models (Stock and Watson, 2002 and Caggiano et al., 2011).

that takes the frequency dependence between target variables and predictors into account. We apply it to forecast a financial variable – the equity premium – and a macroeconomic variable – real GDP growth rate. In a nutshell, the method goes as follows. We decompose our target variables and a given set of n (n=15) predictor variables into f (f=4) time series frequency components, each of them capturing the oscillations of the original variable within a specific frequency band. We then forecast, separately, each of the f frequency component of the target variables using the corresponding frequency component f of one predictor at a time. We obtain n forecasts for each frequency f of the target variables. Subsequently, the forecast of each frequency component f of the target variables is computed as the (mean) combination of the n forecasts of that frequency component from the n predictors. Finally, the overall forecast of each target variable is computed as the sum of the f forecasts of its f frequency components.

We find that combining forecasts in the frequency domain produces markedly more accurate predictions relative to existing alternatives, both in terms of statistical and economic measures of out-of-sample predictability. An advantage of this approach is its flexibility, as it allows to exclude some of the frequencies of the target variables when computing the forecasts. In particular, we show that, in recessions, it is crucial to ignore the forecast of the low-frequency components of the target variables in order to have more accurate forecasts. When used in a real-time exercise (which mimics the real-life situation of a forecaster), the possibility of ignoring these low-frequency components in recessions allows to improve the forecasting results even further, as the forecasts capture noticeably well the sudden and abrupt drops of the equity market and of the real economy associated with recessions. Besides providing more accurate forecasts, this method is easy to implement, has very broad applicability, and can be used in combination with machine learning methods to take advantage of the large datasets currently easily available to forecasters and researchers.[2]

The remaining of this paper is organized as follows. In section 2 we briefly review how filters have

[2] A non-exhaustive list of papers using machine learning in economics and finance applications are Risse (2019), Gu et al. (2020), Bianchi et al. (2021), Medeiros et al. (2021), Kynigakis and Panopoulou (2022), and Leippold et al. (2022).

been used in the macro-finance forecasting literature. In section 3 we present the data and the band-pass filter used to extract the frequency components from the original variables. Section 4 outlines the econometric methodology. The out-of-sample forecasting results are reported in section 5. We analyze the predictability over the business cycle in section 6. Robustness tests are briefly described in section 7, and section 8 concludes.

2 The use of filters in forecasting: a brief tour of the macro-finance literature

Filters have been used in forecasting applications in the macro-finance field in two main ways: either to filter the variable(s) on the right-hand-side (i.e. the predictor(s)), or to filter the variable on the left-hand-side (i.e. the variable to forecast).[3]

In the literature on forecasting with gaps, the right-hand-side variable is typically detrended or filtered to isolate the business-cycle component which is then used in the forecasting exercise. For instance, in the literature on forecasting inflation with Phillips curves, output or unemployment gap are commonly used variables as proxies for slack. These gaps are usually computed by linearly or quadratic detrending output or unemployment, or by filtering them with e.g. the Christiano and Fitzgerald (2003) band-pass filter (see e.g. Banbura and Bobeica, 2023). Similarly, Cooper and Priestley (2009) show that the output gap is a strong predictor of stock and bond market returns. Whether these gap measures predict well or not depends, however, on the specific application and econometric model used (see e.g. Rossi and Sekhposyan, 2010).

Likewise, differencing is a filter that eliminates low-frequency fluctuations. When working with large macroeconomic databases such as the "Stock-Watson dataset" (developed since Stock and

[3] We refer the reader to Petropoulos et al. (2022, sections 3.3.2 and 3.3.13) for exhaustive reviews on forecasting GDP and stock returns, respectively, and Timmermann (2006), Petropoulos et al. (2022, Section 2.6.1), and Wang et al. (2023) for extensive reviews on forecast combination methods.

Watson, 1996) or the more recent FRED-MD database (McCracken and Ng, 2016), it is common to consider several transformations of the original data such as differences and second differences (as well as gaps computed with the Hodrick and Prescott, 1997 filter), and then use those predictors to forecast the variable of interest.

Then, there is a literature where filters are applied to the left-hand-side variable. Clark and Doh (2014), for instance, compare the forecast accuracies of a wide array of models of trend inflation, which captures the low-frequency variations of inflation. Similarly, researchers (e.g. Berge, 2018) are quite often interested in forecasting core inflation, which is a very specific filtered version of inflation, obtained not through a statistical tool but rather by eliminating / filtering the most volatile components of inflation (energy and food prices).

In this paper we filter both the predictors and the variables to forecast. By filtering the right-hand side variables, we can better exploit the information embedded in each individual predictors. By filtering the left-hand side variables, we can better forecast all the individual frequency components of the variables of interest so as to have more accurate forecast for those variables. Our method builds on Faria and Verona (2018) but is conceptually and methodologically very different. In particular, there are three main differences. First, Faria and Verona (2018) consider univariate regression models that take into account the information from one (frequency of one) predictor only, while in this paper we use information from multiple frequencies of multiple predictors in a forecast combination setup. Second, Faria and Verona (2018) sum the forecasts of three components of stock returns, while our proposed method in this paper sums the weighted forecasts of the equity premium and GDP growth. Third, the forecasting method in Faria and Verona (2018) has a more limited applicability (as it can only be used to forecast stock returns), while our proposed method can be used to forecast any variable.

3 Data

We follow Rapach et al. (2010) and use quarterly data. The sample period span from 1947:Q1 until 2022:Q4. The target variables are the U.S. equity premium and the U.S. quarter-over-quarter real GDP growth rate. The equity premium in quarter t is measured by the difference between the log (total) return of the S&P500 index in quarter t and the log return on a three-month Treasury bill at the beginning of quarter t. Data for the S&P500 index and the U.S. real GDP growth rate are obtained from the Goyal and Welch (2008) updated dataset and the U.S. Bureau of Economic Analysis, respectively.

As predictors, we use fifteen variables from Goyal and Welch (2008) updated dataset. Specifically, we use the log dividend-price ratio (DP), the log dividend yield (DY), the log earnings-price ratio (EP), the log dividend-payout ratio (DE), the stock variance (SVAR), the book-to-market ratio (BM), the net equity expansion (NTIS), the Treasury bill rate (TBL), the long-term bond yield (LTY), the long-term bond return (LTR), the term spread (TMS), the default yield spread (DFY), the default return spread (DFR), the lagged inflation rate (INFL), and the lagged investment rate (IK). While these predictors have been extensively used to forecast equity returns, several of them have predictive ability with respect to real GDP growth as well (see e.g. Stock and Watson, 2003). A classical example is the slope of the yield curve (proxied by the term spread), which has widely been used as a predictor for recessions (see e.g. Estrella and Hardouvelis, 1991).

We explain these predictors in appendix 1. The time series of the target variables and of the predictors are plotted in figure 1 and 2, respectively. Table 1 reports summary statistics for all the variables. We note here that both target variables are negatively skewed, suggesting that both the real economy and the equity market have more crashes than what would happen if they were normally distributed.

To decompose the variables into their time series frequency components, we band-pass the data with the Haar filter. Besides its simplicity and wide use (see e.g. Bandi et al., 2019, Kilponen and

Verona, 2022, Martins and Verona, 2023, and Stein, 2024), the Haar filter makes a clean connection to temporal aggregation as the filter coefficients are simply differences of moving averages. We consider four frequency components: the first one (D_1) captures fluctuations of the original variable with a period between 2 and 4 quarters, while components D_2 and D_3 capture fluctuations with a period of 1-2 and 2-4 years, respectively. Finally, component D_4 captures fluctuations with a period longer than 4 years. We note that the sum of these four time series frequency components gives exactly the time series of the original variable. We refer the reader to appendix 2 for further details on the Haar filter.

As an example, figure 3 shows the time series of investment rate (upper plot) and of its time series frequency components (remaining plots). Component D_1 captures the high-frequency movements of investment rate (the noisy component) and frequency component D_4 its trend, while the remaining frequencies (D_2 - D_3) broadly capture the short end of business cycle frequency fluctuations.

4 Econometric methodology

The one-step ahead out-of-sample (OOS) forecasts are generated using a sequence of expanding windows. We use an initial in-sample period (1947:Q1 to 1964:Q4) to make the first one-step ahead OOS forecast. The in-sample period is then increased by one observation and a new one-step ahead OOS forecast is produced. We proceed in this way until the end of the sample.

4.1 Predictive regression model

Let r be the target variable (the equity premium or real GDP growth rate). For each predictor $x_i, i = 1, \ldots, n$ ($n = 15$), the predictive regression is

$$r_{x_i,t} = \alpha_{x_i} + \beta_{x_i} x_{i,t-1} + \varepsilon_t \ , \tag{1}$$

and the corresponding forecasts are given by

$$\hat{r}_{x_i,t+1} = \hat{\alpha}_{x_i} + \hat{\beta}_{x_i} x_{i,t} \ , \tag{2}$$

where $\hat{\alpha}_{x_i}$ and $\hat{\beta}_{x_i}$ are the Ordinary Least Squares (OLS) estimates of α_{x_i} and β_{x_i} in equation (1), respectively, using data from the beginning of the sample until quarter *t*.[4]

4.2 Forecast combination in the time domain

The forecast combination of *r* in the time domain (FC-TD) made at time *t* for *t+1*, denoted $\hat{r}_{t+1}^{FC-TD}$, is the mean of the n (n=15) individual forecasts based on equation (2):

$$\hat{r}_{t+1}^{FC-TD} = \frac{1}{n} \sum_{i=1}^{n} \hat{r}_{x_i,t+1} \ . \tag{3}$$

We have also considered other combination methods (median, trimmed mean, as well as discounted mean square prediction error). As in previous literature (e.g. Rapach et al., 2010), results were usually not better than the mean average.

4.3 Forecast combination in the frequency domain

The first step of our method consists in decomposing all variables into their time series frequency components (D_1 - D_4). We then estimate, for each predictor x_i, a model like (1) for each frequency *f*. That is, we estimate – separately – each frequency component D_f of *r* using the corresponding frequency component of the predictor x_i:

$$r_t^{D_f,x_i} = \alpha_{t,f}^{x_i} + \beta_{t,f}^{x_i} x_{i,t-1}^{D_f} + \varepsilon_t \ . \tag{4}$$

[4] We do not report the results of a multiple regression forecasting model that includes all potential predictors – the so-called "kitchen sink" model – as it performs much worse than the historical average forecast.

This setup is akin to the band spectrum regression proposed by Engle (1974) and have been used by e.g. Gallegati et al. (2011), Gallegati and Ramsey (2013), Ortu et al. (2013), Faria and Verona (2018, 2021), and Martins and Verona (2023). As we use a two-sided filter in the OOS exercise, we use real-time filtering and recompute the time series frequency components of the variables recursively at each iteration of the OOS forecasting process using data from the start of the sample through the quarter at which the forecasts are made. This step ensures that our method does not have a look-ahead bias, as the forecasts are made with current and past information only. When using a two-sided filter some assumptions regarding how to deal with the observations at the beginning and at the end of the sample have to be made. The literature suggests several types of boundary treatment rules to deal with boundary effects (e.g. periodic rule, reflection rule, zero padding rule, and polynomial extension). Here, we use a reflection rule, whereby the original time series are reflected symmetrically at the boundaries before filtering them.

We use the estimation results in (4) to produce the one-step ahead OOS forecast of the corresponding frequency component of r:

$$\hat{r}_{t+1}^{D_f,x_i} = \hat{\alpha}_{t,f}^{x_i} + \hat{\beta}_{t,f}^{x_i} x_{i,t}^{D_f} \ ,$$

where $\hat{\alpha}_{t,f}^{x_i}$ and $\hat{\beta}_{t,f}^{x_i}$ are the OLS estimates of $\alpha_{t,f}^{x_i}$ and $\beta_{t,f}^{x_i}$, respectively, using data from the beginning of the sample until quarter t.

We then compute the forecasts of each frequency components $\mathrm{D_f}$ of r as the mean forecast combination for that frequency f:

$$\hat{r}_{c,t+1}^{D_f} = \frac{1}{n} \sum_{i=1}^{n} \hat{r}_{t+1}^{D_f,x_i} \ .$$

Finally, the overall forecast of r made at time t for $t+1$ in the frequency domain (FC-FD), denoted $\hat{r}_{t+1}^{FC-FD}$, is obtained by summing the forecasts of the f individual frequencies of r:

$$\hat{r}_{t+1}^{FC-FD} = \sum_{f=1}^{4} \hat{r}_{c,t+1}^{D_f} = \sum_{f=1}^{4} \left(\frac{1}{n} \sum_{i=1}^{n} \hat{r}_{t+1}^{D_f,x_i} \right) \ . \tag{5}$$

4.4 Forecast evaluation

4.4.1 Statistical performance

The forecasting performances of the forecast combination models are evaluated using the Campbell and Thompson (2008) R^2_{OS} statistic. The R^2_{OS} statistic measures the proportional reduction in the mean squared forecast error (MSFE) for the predictive model ($MSFE_{PRED}$) relative to the benchmark model ($MSFE_{BENCHMARK}$) and is given by

$$R^2_{OS} = 100\left(1 - \frac{MSFE_{PRED}}{MSFE_{BENCHMARK}}\right) = 100\left[1 - \frac{\sum_{t=t_0}^{T-1}\left(r_{t+1} - \hat{r}_{t+1}\right)^2}{\sum_{t=t_0}^{T-1}\left(r_{t+1} - \hat{r}_{t+1}^{BENCHMARK}\right)^2}\right] ,$$

where $\hat{r}_{t+1}$ is the forecast for *t+1* from the FC-TD or the FC-FD model (equation (3) and (5), respectively) and r_{t+1} is the realized equity premium / GDP growth from t to *t+1*. A positive (negative) R^2_{OS} indicates that the predictive model outperforms (underperforms) the benchmark model in terms of MSFE. The benchmark model to forecast the equity premium is the average equity premium up to time t, and to forecast the GDP growth is an AR(p) model, where p is chosen recursively according to the Akaike information criterion.

The statistical significance of the R^2_{OS} is evaluated using the Clark and West (2007) MSFE-adjusted statistic. This statistic tests the null hypothesis that the MSFE of the benchmark model is less than or equal to the MSFE of the FC-TD or FC-FD model against the alternative hypothesis that the MSFE of the benchmark model is greater than the MSFE of the FC-TD or FC-FD model ($H_0 : R^2_{OS} \leq 0$ against $H_A : R^2_{OS} > 0$).

Additionally, we compute the R^2_{OS} and its statistical significance of our proposed method (FC-FD) against the forecast combination in the time domain (FC-TD) as the benchmark.

4.4.2 Economic performance

We also analyze stock return forecasts with utility-based metrics, which provide a more direct measure of the value of forecasts to decision makers. In these exercises, stock return forecasts are used as inputs for asset allocation decisions derived from expected utility maximization problems. A leading utility-based metric is the average utility gain for a mean-variance investor, who allocates the portfolio between equities and risk-free bills. At the end of quarter t, the investor optimally allocates a share $w_t = \hat{R}_{t+1}/\left(\gamma\hat{\sigma}^2_{t+1}\right)$ of the portfolio to equity for period *t+1*, where γ is the investor's relative risk aversion coefficient, $\hat{R}_{t+1}$ is the time t (FC-TD or FC-FD) model forecast of the equity premium, and $\hat{\sigma}^2_{t+1}$ is the forecast of the variance of the equity premium. We assume a relative risk aversion coefficient of three, use a five-year moving window of past equity premium to estimate the variance forecast and constrain the weights w_t to lie between -0.5 and 1.5. These constraints limit the possibilities of short selling and leveraging the portfolio to realistic levels.

The realized portfolio return at time *t+1*, RP_{t+1}, is given by $RP_{t+1} = w_t R_{t+1} + RF_{t+1}$, where R_{t+1} is the equity market return from time t to *t+1* and RF_{t+1} denotes the risk-free return from time t to *t+1* (*i.e.* the market rate, which is known at time t). The average utility (or certainty equivalent return, CER) is computed as $CER = \overline{RP} - 0.5\gamma\sigma^2_{RP}$, where $\overline{RP}$ and σ^2_{RP} are the sample mean and variance of the portfolio return, respectively. We report the annualized utility gain, which is computed as the difference between the CER for an investor that uses the FC-TD or FC-FD model to forecast the equity premium and the CER for an investor who uses the historical mean benchmark for forecasting. The difference is multiplied by 4 to annualize quarterly performance, which allows to interpret it as the annual portfolio management fee that an investor would accept to pay to have access to the alternative forecasting model versus the benchmark model forecast. Following Gargano et al. (2019) and Bianchi et al. (2021), we use a Diebold and Mariano (1995) test to assess if the annualized CER gains are statistically greater than zero.

We also compute the CER gains (and its statistical significance) of the FC-FD method against the

FC-TD method as the benchmark.

5 Results

For equity premium predictability, the OOS period spans from 1965:Q1 to 2022:Q4 (that is, 232 quarterly forecasts). We include the COVID-19 recession in the evaluation period as the drop and recover of the equity market during that recession were comparable with (or even less abrupt than) those in previous recessions (see the upper graph in figure 1). Furthermore, as we are using quarterly data, all high frequency stock market fluctuations are, by construction, smoothed.

On the other hand, as one can see from the lower graph in figure 1, quarter-to-quarter real GDP growth experienced fluctuations ranging from -8.9 % in 2020:Q2 to 7.6 % in 2020:Q3. In comparison, in the global financial crises, it fluctuated between -2.2 % in 2008:Q4 and 1.1 % in 2009:Q4. These huge fluctuations in 2020 render forecast (evaluation) extremely difficult, especially when using the MSFE as evaluation criteria. Hence, for real GDP growth rate, we end the evaluation period in 2019:Q4 (so that we produce 220 quarterly forecasts).

5.1 Equity premium

The second through seventh columns of table 2 reports the results for equity premium predictions for the full OOS forecast evaluation period (second and third columns), and separately for expansions (fourth and fifth columns) and recessions (sixth and seventh columns). Panel A shows the results for individual predictive regression, and panel B and C for different forecast combination models compared to the benchmark model and the FC-TD model, respectively. We do not report the results of the so-called kitchen sink model, which corresponds to a multiple predictive regression model that includes all 15 predictors, as it performs poorly.

Results in panel A are in line with those in previous literature (Rapach et al., 2010). The R^2_{OS}

statistics in the second column clearly show that individual predictive regression forecasts of the equity premium frequently fail to beat the benchmark model in terms of MSFE. Indeed, 12 of the 15 R^2_{OS} statistics are negative, and only one (IK) of the three predictors with a positive R^2_{OS} is statistically significant. That is, only one of these 15 variables displays statistically significant OOS predictive ability at conventional levels. The third column reports the average utility gains. Relative to the R^2_{OS} statistics, the individual predictive regression forecasts appear more valuable from an economic point of view, as 9 of the 15 variables offer positive gains. However, only TMS has statistically significant positive annualized gains (320 basis points).

When analyzing stock return predictability over the business cycle, the R^2_{OS} statistics in the fourth and sixth columns confirm the well-known fact that predictability is higher in recessions than in expansions (see e.g. Cujean and Hasler, 2017). Only one variable (LTR) delivers a positive and statistically significant R^2_{OS} in expansions, while four variables (DP, DY, TMS, and IK) are good predictors of equity returns in recessions. Similar differences in predictability over business cycle phases are visible when looking at the utility gains (in the fifth and seventh columns).

Overall, the results in panel A show that no single variable is clearly and constantly better (from a statistical or economic point of view) in all sub-samples than the others as equity premium predictor.

In panel B (between columns two to seven) are reported the R^2_{OS} statistics and average utility gains for the combining methods for equity premium predictions. The first row in panel B demonstrates the usefulness of the forecast combination in the time domain. The FC-TD model delivers positive and statistically significant R^2_{OS} and positive (but not statistically significant) CER gains for the full forecast evaluation period, as well as in both expansion and recession periods.

The second row in panel B shows the results of the forecast combination in the frequency domain (FC-FD). From a statistical point of view, the FC-FD model performs slightly better than the FC-TD model over the full OOS period and in each subsamples. From an economic point of view, the

FC-FD model delivers larger utility gains regardless of the sample period considered. Over the full OOS period, the CER gains generated by the FC-FD model are sizable, statistically significant (at conventional levels), and twice as large as than those of the FC-TD model (261 versus 126 basis points). However, when looking at the results in the first row in panel C, both the R^2_{OS} statistic and the CER gains of the FC-FD model are not statistically better than those of the FC-TD model.

Figure 4 provides information on the behavior of the portfolios based on the forecasts from the benchmark model and from the forecast combination models. Panel A and B depicts equity weights and the log cumulative wealth, respectively, over the forecast evaluation period.

The equity weight for the portfolio based on the benchmark model (black line) is relatively stable throughout the OOS period, which reflects the fact that the historical mean benchmark forecast is very smooth. The equity weight for the portfolio based on the FC-TD model (red line) exhibits substantial fluctuations around the weight of the benchmark portfolio, especially until 1990. After that, the weight closely tracks the one of the benchmark portfolio.

The equity weight for the portfolio based on the FC-FD model (blue line) exhibits even more fluctuations, especially around recessions. The enhanced portfolio performance of the FC-FD model, quite evident from the log cumulative wealth in panel B, is due to its better market timing, as it allows to quickly reduce the exposure to the equity market around recessions.

5.2 GDP growth

The eighth through tenth columns of table 2 reports the results for real GDP growth rate predictions for the full OOS forecast evaluation period (eighth column), and separately for expansions (ninth column) and recessions (tenth column). Panel A shows the results for individual predictive regression, and panel B and C for different forecast combination models compared to the benchmark model and the FC-TD model, respectively.

Over the full OOS period, results are similar to the equity premium ones: only one of the individual predictor (NTIS) displays statistically significant out-of-sample predictive ability. Differently from the equity premium case, predictability of GDP growth rates is higher in expansions than in recessions: seven variables are good predictors of GDP growth in expansions, while only two in recessions.

Looking at the forecast combination methods in panel B, the FC-TD model performs well over the full OOS period but the FC-FD model performs even better. The R^2_{OS} for the FC-FD model is 7.4 %, which is much larger than the 2.9 % R^2_{OS} for the FC-TD model (both statistically significant at the 1 % level). Both forecast combination methods perform well in expansions (with the FC-FD model being better than the FC-TD model) but poorly in recessions (negative R^2_{OS} statistics). Moreover, when looking at the results in the first row in panel C, both the R^2_{OS} statistic and the CER gains of the FC-FD model are statistically better than those of the FC-TD model over the entire sample period as well as in expansions and recessions.

We report the actual forecasts in figure 5. The enhanced performance of the FC-FD model is due to its ability to better capture both the trend of GDP growth and some of its higher frequency fluctuations. It is also evident that none of the forecast combination methods is nevertheless able to capture the sudden and abrupt drops associated with recessions.

5.3 Placebo test

To demonstrate that our procedure does not mechanically generate predictability, we run the following placebo test. We generate 1000 datasets, each of them containing 15 variables, and each variable having the same persistence and standard deviation as the respective variable in the real dataset. Innovations in the simulated datasets are produced by a random number generator so they are independent from the true data. Then we run the forecast with our FC-FD model to each of these simulated datasets, and record the R^2_{OS} and CER gains.

Figure 6 shows the distribution of the R^2_{OS} and CER gains for equity premium predictions (left and middle graph, respectively) and the R^2_{OS} for GDP growth rate predictions (right graph). For all measures and variables, the medians (red lines within each box) as well as the 90 % confidence interval of the distributions (vertical dashed lines) are below the results with the FC-FD model with the original data (black dots). This placebo analysis thus shows that the predictability power of the FC-FD model is thus unlikely to be driven by a mechanical bias.

5.4 Why is it important to take into account the frequency domain (in forecast combination)?

The benefits of using forecast combination methods are well known in the literature. As stressed in the seminal paper by Goyal and Welch (2008), the inconsistent out-of-sample performance of individual predictive regression models is due to structural instability. The graphs in the top row in figure 7 give a visual impression of the changing nature of the relationships between the target variables (equity premium on the left column and GDP growth on the right column) and three individual predictors (DP, black lines; TMS, blue lines; IK, red lines). The figure depicts the OLS estimates from expanding windows that start with the sample 1947:Q1-1964:Q4 and recursively add one quarter through 2022:Q4. The regression coefficients, which are ultimately used to produce the OOS forecasts with the FC-TD model, fluctuate substantially over the period, and there are even instances where the relationship switches sign. Given this instability over time, averaging across individual forecasts gives more stable and, ultimately, better forecasts.

The remaining rows in figure 7 report the OLS estimates in each frequency bands, which are used to produce the OOS forecasts with the FC-FD model. We emphasize three features about time- and frequency-varying changing relationships. First, as it happens with the original variables, there is a clear time variation of each coefficient within a specific frequency, and switching sign is also quite common. Second, there are cases where the sign of the relationship between aggregate variables differs from the sign between the same variables at different frequencies. For instance, the estimated OLS coefficients are negative (positive) between the equity premium and IK (GDP growth and TMS), but the sign flips at frequency D_2 (D_2 and D_4) during most of the sample period. Third, for a given predictor, the magnitude of the estimated coefficients also significantly varies across frequencies and quite often differs from the magnitude using the original series.[5]

Overall, figure 7 suggests important structural instabilities in the relationships between the target variables and these predictors not only over time, but also across frequencies. These findings support the relevance of taking the frequency domain into account, as the magnitude and sign of the estimated coefficients (that are used to make the forecasts) are time- and frequency-specific.

Another way to understand why forecast combination models perform better than individual predictive regression models is to look at the Theil (1971) MSFE decomposition into the squared forecast bias and a remainder term (as proposed by Rapach et al., 2010). The latter term depends, among other things, on the forecast volatility, and limiting forecast volatility helps to reduce the remainder term. A model's forecasting performance ultimately depends on the trade-off between the reduction in bias and variance. To get a sense of this bias-efficiency trade-offs in the forecasts, figure 8 is a scatterplot depicting the MSFE decomposition into the squared forecast bias and the remainder term for the individual predictive regression models, the benchmark models, and the forecast combination models for the full OOS period.

Looking at the equity premium forecast (left graph), several forecasting methods produce relatively unbiased return predictions, many of them even better than the historical mean benchmark. However, their performance relative to the historical mean is negatively affected by their higher remainder term.

The FC-TD forecast has low forecast variance and a relatively small squared forecast bias (close to the smallest squared biases of the individual predictive regression model, IK). When compared

[5] It is beyond the scope of this paper to analyze why, for some of these variables, the sign and magnitude at some frequencies are different than those with the original series. However, this fact has already been emphasized in other applications. For instance, in the context of the Q theory of investment, Gallegati and Ramsey (2013) and Verona (2020) show that the investment-Q sensitivity is not always positive at all points in time and for all frequencies.

with the historical mean benchmark, both the squared bias and the remainder term are substantially below. Hence, the FC-TD model achieves a higher R^2_{OS} (that is, a smaller MSFE) than the historical mean benchmark and any of the individual predictive regression models (except IK).

The FC-FD model delivers more accurate forecasts (higher R^2_{OS} and smaller MSFE) than the FC-TD model due to its ability to further reduce both the forecast bias and the remainder term.

Similar conclusions can be drawn from the analysis of the scatterplot for real GDP growth forecast in the right graph of figure 8. Thus, forecasts based on the FC-FD model are generally both less biased and more efficient than all the other forecasts analyzed here, including the forecast combination in the time domain.

6 Predictability over the business cycle

So far we sum the forecasts of all (four) frequencies of the target variables when making the forecast in the frequency domain. However, as shown by Faria and Verona (2021) and Martins and Verona (2021), ignoring some frequencies of the target variable usually leads to better forecasts. For instance, forecasting with a model like $FC-FD_{t+1} = \sum_{f=2}^{3} \hat{r}^{D_f}_{c,t+1}$, which ignores both the highest and the lowest frequency forecasts of the target variable, might produce more accurate forecast than an identical model that sums all (four) frequencies.

We now check if and when it is valuable to ignore some frequencies of the target variables. We start from an ex-post exercise to gain some intuition about the predictability over business-cycle phases. We then move to a real-time exercise where the status of the business cycle is assessed in real time and the forecaster switches between two forecasts according to the state of the economy.

6.1 Ex-post exercise

To investigate how our method performs in recessions and in expansions, we analyse which frequencies of the target variables are important to include (or exclude) to have good forecasts, and whether there are differences between expansions and recessions.

The third row in panel B in table 2 reports the forecasts of the FC-FD model when the low-frequency component (D_4) of the target variable is ignored when making the forecast. This method is denoted as FC-FD (no LF) and its forecast is given by $\sum_{f=1}^{3} \hat{r}_{c,t+1}^{D_f}$. The gains over the full OOS period are not very impressive (for the equity premium) or even really bad (for GDP growth). However, for both target variables, there are huge forecasting gains in ignoring their low-frequency forecasts in recessions. The intuition is that ignoring the forecast of the trend allows to better track the quick and sudden drop associated with recessions (recall that both variables are negatively skewed). However, it is crucial to forecast well the trend (as well as some high-frequency fluctuations) in expansions.[6]

This finding of enhanced return predictability during recessions is ex-post, since the dates of NBER business-cycle peaks and troughs are known retrospectively. The question is then whether we can use this insight in real time. In particular, how large are the statistical / economic gains if we were able to switch between the forecasts of two different frequency combination methods – $\sum_{f=1}^{4} \hat{r}_{c,t+1}^{D_f}$ for expansions and $\sum_{f=1}^{3} \hat{r}_{c,t+1}^{D_f}$ for recessions – in real time according to the perception of the state of the business cycle? We address this question in the next subsection.

[6] Regardless of the forecasts of the other frequency components D_1 - D_3, we find similar results for almost all possible frequency combinations in the spirit of Faria and Verona (2021): whenever we exclude (include) the low-frequency forecasts (D_4) of the target variable, the forecasts in recessions are much better (worse) and in expansions are much worse (better).

6.2 Real-time exercise

To guide switching between forecasts over the business cycle, we rely on a well-known leading indicator of the business cycle – the stock market. In particular, we use the information from some stock market technical indicators (TIs), which are widely employed by practitioners, to compute a real-time indicator of the state of the business cycle. TIs rely on past stock market price and volume patterns to identify trends believed to persist into the future, so they provide useful forward looking information about the business cycle.[7]

Following Neely et al. (2014), we use two moving average indicators and three momentum indicators, which are described and plotted in appendix 3. A value of 0 (1) for each of these indicators implies a sell (buy) signal at the end of quarter *t*, hence quarter *t+1* is considered to be a recession (expansion) according to this specific technical indicator. Relying on a single TI might however generate too many false recession signals. Hence, we introduce a novel business cycle leading indicator, that we name as coincident index, that summarizes the information from the five TIs. In particular, for quarter *t+1* to be considered a recession, all five TIs have to be 0 at the end of quarter *t*. In this case, we use the forecasts for *t+1* (made at the end of quarter *t*) that exclude the low-frequency forecast of the target variable (i.e. we use $\sum_{f=1}^{3} \hat{r}_{c,t+1}^{D_f}$) . The coincident index, plotted in figure 9, captures most of the actual NBER-dated recession quarters, albeit triggering some false recession signals.

The results for the equity premium and GDP growth forecasts in real time, denoted FC-FD real time, are reported in the last row of panel B and C in table 2. Being able to switch forecasts in real time according to the state of the business cycle allows to improve forecast even further when

[7] This method is similar in spirit but much simpler than the one proposed by Aruoba et al. (2009), who use high-frequency data to compute a real-time indicator of economic activity. Other variables commonly used as real-time / leading indicator of the business cycle are the Chicago Fed National Activity Index, the Business Conditions Index, the term spread, and indicators based on survey data (Survey of Professional Forecasters, Livingston Survey, and Purchasing Managers' index). Markov-switching models (see e.g. Guidolin and Timmermann, 2007) provide a different framework for switching between forecasting models according to estimated probabilities of the state of the economy.

compared to both the benchmark models and the FC-FD model. When compared to the benchmark models, for the equity premium (second and third column), the R^2_{OS} statistic and CER gain both are statistically significant at 4.41 % and 385 basis points, respectively. The blue dashed line in figure 4 reports the log cumulative wealth for an investor who trades using the FC-FD real time model. The possibility of being able to switch between forecasts significantly increases wealth throughout the entire OOS period.

Looking at the real GDP growth rate forecasts (eighth column), the R^2_{OS} statistic of the FC-FD real time model is statistically significant and markedly higher than those of the other combination models (14.8 % against 7.4 % and 2.9 % of the FC-FD and FC-TD model, respectively). The blue dashed line in figure 5 shows the forecast of the FC-FD real time model. This method produces more accurate forecast as it allows to capture remarkably well the drops associated with recessions.

Likewise, from the bias-variance scatterplot (figure 8), the improved performance of the FC-FD real time model is due to its ability to decrease the forecast bias almost to zero and to reduce even further the remainder term.

Remarkably, for both equity premium and real GDP growth, the FC-FD real time model produces forecasts that are significantly better (from a statistical and economic point of view) than those of the FC-TD model, as reported in the last row of panel C of table 2.

7 Robustness

We run the following robustness tests.

The choice of the band-pass filter affect both the equity premium and real GDP growth forecast. We run the analysis with the Daubechies filter of length two and four, which is commonly used with quarterly data (e.g. Crowley and Hudgins, 2021, 2023). The choice of the parameters related with the asset allocation exercise only affects the CER gain results. We consider a risk aversion coefficient of 5 (instead of 3), different set of portfolio constraints for the equity weights w_t (no leverage and / or no short selling instead of 50 % leverage and short selling), a ten-year (instead of five-year) moving window of past equity premium to estimate the variance forecast, and CER gains net of transactions costs of 50 basis points. For the real-time exercise, we use different technical indicators to compute the coincident index. Results, not reported here but available upon request, turn out to be robust to all these changes.

We also run the forecasts using the Christiano and Fitzgerald (2003) asymmetric band-pass filter. As in other forecasting applications (e.g. Faria and Verona, 2020), this band-pass filter performs worse than the Haar filter, especially when forecasting the equity premium.

8 Conclusions

In this paper we propose a new forecasting method – forecast combination in the frequency domain – that takes into account the fact that predictability is time *and* frequency dependent. We apply this method to forecast the equity premium and real GDP growth rate. Combining forecasts in the frequency domain produces markedly more accurate predictions relative to the traditional forecast combination in the time domain, both in terms of statistical and economic measures of out-of-sample predictability. This result supports the theoretical findings of Kelly et al. (2024), who show that more complex models (for instance, forecast combination in the frequency domain) tend to deliver larger economic gains than simpler models (for instance, forecast combination in the time domain). This method is flexible enough that it allows to exclude some of the frequencies of the target variables when making the forecasts. In particular, we show that, in recessions, it is of major relevance to ignore the forecast of the low-frequency components of the target variables. This flexibility turns out to be crucial in a real-time forecasting exercise, as the method allows to capture remarkably well the sudden and abrupt drops associated with recessions and further improve predictability of the equity premium and real GDP growth rate. Besides providing more accurate forecasts, this method has very broad applicability, is easy to implement, and can be used in combination with machine learning methods to take advantage of the large dataset nowadays easily available to researchers and forecasters.

References

Aruoba, S. Boragan, Francis X. Diebold, and Chiara Scotti, "Real-Time Measurement of Business Conditions," *Journal of Business & Economic Statistics*, 2009, *27* (4), 417–427.

Banbura, Marta and Elena Bobeica, "Does the Phillips curve help to forecast euro area inflation?," *International Journal of Forecasting*, 2023, *39* (1), 364–390.

Bandi, Federico, Bernard Perron, Andrea Tamoni, and Claudio Tebaldi, "The Scale of Predictability," *Journal of Econometrics*, 2019, *208* (1), 120–140.

Bates, J. M. and C. W. J. Granger, "The Combination of Forecasts," *Operational Research Quarterly*, 1969, *20* (4), 451–468.

Berge, Travis J., "Understanding survey-based inflation expectations," *International Journal of Forecasting*, 2018, *34* (4), 788–801.

Beveridge, Stephen and Charles Nelson, "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'," *Journal of Monetary Economics*, 1981, *7* (2), 151–174.

Bianchi, Daniele, Matthias Buchner, and Andrea Tamoni, "Bond Risk Premiums with Machine Learning," *Review of Financial Studies*, 2021, *34* (2), 1046–1089.

Caggiano, Giovanni, George Kapetanios, and Vincent Labhard, "Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK," *Journal of Forecasting*, 2011, *30* (8), 736–752.

Campbell, John Y. and Samuel B. Thompson, "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, 2008, *21* (4), 1509–1531.

Christiano, Lawrence and Terry Fitzgerald, "The Band Pass Filter," *International Economic Review*, 2003, *44* (2), 435–465.

Clark, Todd E. and K.D. West, "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 2007, *138* (1), 291 – 311.

__ **and Taeyoung Doh**, "Evaluating alternative models of trend inflation," *International Journal of Forecasting*, 2014, *30* (3), 426–448.

Cooper, Ilan and Richard Priestley, "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, July 2009, *22* (7), 2601–2633.

Cremers, K. J. Martijn, "Stock Return Predictability: A Bayesian Model Selection Perspective," *Review of Financial Studies*, 2002, *15* (4), 1223–1249.

Crowley, Patrick M. and David Hudgins, "Okun's law revisited in the time-frequency domain: introducing unemployment into a wavelet-based control model," *Empirical Economics*, 2021, *61* (5), 2635–2662.

__ **and** __ , "Resilient Control for Macroeconomic Models," *Computational Economics*, 2023, *61* (4), 1403–1431.

Cujean, Julien and Michael Hasler, "Why Does Return Predictability Concentrate in Bad Times?," *Journal of Finance*, 2017, *72* (6), 2717–2758.

Dangl, Thomas and Michael Halling, "Predictive regressions with time-varying coefficients," *Journal of Financial Economics*, 2012, *106* (1), 157–181.

Diebold, F.X. and R.S. Mariano, “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, 1995, *13* (3), 253–263.

Engle, Robert F., “Band Spectrum Regression,” *International Economic Review*, 1974, *15* (1), 1–11.

Estrella, Arturo and Gikas A Hardouvelis, “The Term Structure as a Predictor of Real Economic Activity,” *Journal of Finance*, 1991, *46* (2), 555–576.

Faria, Gonçalo and Fabio Verona, “Forecasting stock market returns by summing the frequency-decomposed parts,” *Journal of Empirical Finance*, 2018, *45*, 228 – 242.

_ and _ , “The yield curve and the stock market: Mind the long run,” *Journal of Financial Markets*, 2020, *50* (C).

_ and _ , “Out-of-sample time-frequency predictability of the equity risk premium,” *Quantitative Finance*, 2021, *21* (12), 2119–2135.

Gallegati, Marco and James B. Ramsey, “Bond vs stock market’s Q: Testing for stability across frequencies and over time,” *Journal of Empirical Finance*, 2013, *24* (C), 138–150.

_ , Mauro Gallegati, James B. Ramsey, and Willi Semmler, “The US Wage Phillips Curve across Frequencies and over Time,” *Oxford Bulletin of Economics and Statistics*, 2011, *73* (4), 489–508.

Gargano, A., D. Pettenuzzo, and A. Timmermann, “Bond Return Predictability: Economic Value and Links to the Macroeconomy,” *Management Science*, 2019, *65* (2), 508–540.

Goyal, A. and I. Welch, “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies*, 2008, *21* (4), 1455–1508.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, “Empirical Asset Pricing via Machine Learning,” *Review of Financial Studies*, 2020, *33* (5), 2223–2273.

Guidolin, Massimo and Allan Timmermann, "Asset allocation under multivariate regime switching," *Journal of Economic Dynamics and Control*, 2007, *31* (11), 3503–3544.

Henkel, Sam James, J. Spencer Martin, and Federico Nardari, "Time-varying short-horizon predictability," *Journal of Financial Economics*, 2011, *99* (3), 560–580.

Hodrick, Robert J. and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, February 1997, *29* (1), 1–16.

Kelly, Bryan, Semyon Malamud, and Kangying Zhou, "The Virtue of Complexity in Return Prediction," *The Journal of Finance*, 2024, *forthcoming.*

Kilponen, Juha and Fabio Verona, "Investment dynamics and forecast: Mind the frequency," *Finance Research Letters*, 2022, *49* (C).

Kynigakis, Iason and Ekaterini Panopoulou, "Does model complexity add value to asset allocation? Evidence from machine learning forecasting models," *Journal of Applied Econometrics*, 2022, *37* (3), 603–639.

Leippold, Markus, Qian Wang, and Wenyu Zhou, "Machine learning in the Chinese stock market," *Journal of Financial Economics*, 2022, *145* (2, Part A), 64–82.

Martins, Manuel M. F. and Fabio Verona, "Inflation dynamics and forecast : frequency matters," Research Discussion Papers 8/2021, Bank of Finland June 2021.

_ and _ , "Inflation dynamics in the frequency domain," *Economics Letters*, 2023, *231*, 111304.

McCracken, Michael W. and Serena Ng, "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business & Economic Statistics*, 2016, *34* (4), 574–589.

Medeiros, Marcelo C., Gabriel F. R. Vasconcelos, Alvaro Veiga, and Eduardo Zilberman, "Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods," *Journal of Business & Economic Statistics*, 2021, *39* (1), 98–119.

Neely, C., D. Rapach, J. Tu, and G. Zhou, "Forecasting the Equity Risk Premium: The Role of Technical Indicators," *Management Science*, 2014, *60* (7), 1772–1791.

Ortu, Fulvio, Andrea Tamoni, and Claudio Tebaldi, "Long-Run Risk and the Persistence of Consumption Shocks," *Review of Financial Studies*, 2013, *26* (11), 2876.

Petropoulos, Fotios, Daniele Apiletti, Vassilios Assimakopoulos, Mohamed Zied Babai, Devon K. Barrow, Souhaib Ben Taieb, Christoph Bergmeir, Ricardo J. Bessa, Jakub Bijak, John E. Boylan, Jethro Browell, Claudio Carnevale, Jennifer L. Castle, Pasquale Cirillo, Michael P. Clements, Clara Cordeiro, Fernando Luiz Cyrino Oliveira, Shari De Baets, Alexander Dokumentov, Joanne Ellison, Piotr Fiszeder, Philip Hans Franses, David T. Frazier, Michael Gilliland, M. Sinan Gonul, Paul Goodwin, Luigi Grossi, Yael Grushka-Cockayne, Mariangela Guidolin, Massimo Guidolin, Ulrich Gunter, Xiaojia Guo, Renato Guseo, Nigel Harvey, David F. Hendry, Ross Hollyman, Tim Januschowski, Jooyoung Jeon, Victor Richmond R. Jose, Yanfei Kang, Anne B. Koehler, Stephan Kolassa, Nikolaos Kourentzes, Sonia Leva, Feng Li, Konstantia Litsiou, Spyros Makridakis, Gael M. Martin, Andrew B. Martinez, Sheik Meeran, Theodore Modis, Konstantinos Nikolopoulos, Dilek Onkal, Alessia Paccagnini, Anastasios Panagiotelis, Ioannis Panapakidis, Jose M. Pavia, Manuela Pedio, Diego J. Pedregal, Pierre Pinson, Patricia Ramos, David E. Rapach, J. James Reade, Bahman Rostami-Tabar, Michal Rubaszek, Georgios Sermpinis, Han Lin Shang, Evangelos Spiliotis, Aris A. Syntetos, Priyanga Dilini Talagala, Thiyanga S. Talagala, Len Tashman, Dimitrios Thomakos, Thordis Thorarinsdottir, Ezio Todini, Juan Ramon Trapero Arenas, Xiaoqian Wang, Robert L. Winkler, Alisa Yusupova, and Florian Ziel, "Forecasting: theory and practice," *International Journal of Forecasting*, 2022, *38* (3), 705–871.

Pettenuzzo, Davide and Francesco Ravazzolo, "Optimal Portfolio Choice Under Decision-Based Model Combinations," *Journal of Applied Econometrics*, 2016, *31* (7), 1312–1332.

Pirschel, Inske and Maik H. Wolters, "Forecasting with large datasets: compressing information before, during or after the estimation?," *Empirical Economics*, September 2018, *55* (2), 573–596.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *Review of Financial Studies*, 2010, *23* (2), 821–862.

Risse, Marian, "Combining wavelet decomposition with machine learning to forecast gold returns," *International Journal of Forecasting*, 2019, *35* (2), 601–615.

Rossi, Barbara and Tatevik Sekhposyan, "Have economic models' forecasting performance for US output growth and inflation changed over time, and when?," *International Journal of Forecasting*, 2010, *26* (4), 808–835.

Stein, Tobias, "Forecasting the equity premium with frequency-decomposed technical indicators," *International Journal of Forecasting*, 2024, *40* (1), 6–28.

Stock, James H. and Mark W. Watson, "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business & Economic Statistics*, 1996, *14* (1), 11–30.

_ and _ , "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 2002, *20* (2), 147–162.

_ and _ , "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature*, 2003, *41* (3), 788–829.

_ and _ , "Combination forecasts of output growth in a seven-country data set," *Journal of Forecasting*, 2004, *23* (6), 405–430.

Theil, Henri, *Applied Economic Forecasting*, North-Holland, Amsterdam, 1971.

Timmermann, Allan, "Forecast Combinations," in G. Elliott, C. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1 of *Handbook of Economic Forecasting*, Elsevier, 2006, chapter 4, pp. 135–196.

Verona, Fabio, "Investment, Tobin's Q, and Cash Flow Across Time and Frequencies," *Oxford Bulletin of Economics and Statistics*, 2020, *82* (2), 331–346.

Wang, Xiaoqian, Rob J. Hyndman, Feng Li, and Yanfei Kang, "Forecast combinations: An over 50-year review," *International Journal of Forecasting*, 2023, *forthcoming.*

	mean	1st perc.	99th perc.	std. dev.	skew.	kurt.	AR(1)
Equity premium	0.02	-0.24	0.18	0.08	-0.97	4.81	0.07
Real GDP growth	0.76	-2.14	3.79	1.15	-1.29	22.5	0.12
DP	-3.53	-4.47	-2.64	0.45	-0.06	2.20	0.98
DY	-3.51	-4.48	-2.60	0.45	-0.06	2.26	0.98
EP	-2.79	-4.26	-1.88	0.45	-0.50	5.09	0.95
DE	-0.74	-1.23	0.65	0.29	2.77	20.3	0.90
SVAR	0.01	0.00	0.06	0.01	6.91	62.0	0.38
BM	0.52	0.14	1.13	0.25	0.55	2.49	0.98
NTIS	0.01	-0.04	0.04	0.02	-0.82	3.28	0.94
TBL	0.04	0.00	0.15	0.03	0.99	4.25	0.96
LTY	0.06	0.01	0.14	0.03	0.82	3.23	0.98
LTR	0.01	-0.11	0.20	0.05	0.89	5.86	-0.01
TMS	0.02	-0.02	0.04	0.01	-0.06	3.17	0.84
DFY	0.01	0.00	0.02	0.00	1.94	8.78	0.88
DFR	0.00	-0.08	0.06	0.02	-0.52	15.1	-0.12
INFL	0.01	-0.01	0.04	0.01	0.33	5.67	0.41
IK	0.04	0.03	0.04	0.00	0.49	2.81	0.96

Table 1: Summary statistics, U.S. data, 1947:Q1-2022:Q4

This table reports summary statistics for the (log) equity premium, real GDP growth, and for the 15 predictive variables. See appendix A for a description of the predictors.

	Equity premium						Real GDP growth rate		
column	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
predictor / method	1965:Q1-2022:Q4		Expansions		Recessions		1965:Q1-2019:Q4	Expansions	Recessions
	R^2_{OS}	CER gain	R^2_{OS}	CER gain	R^2_{OS}	CER gain	R^2_{OS}	R^2_{OS}	R^2_{OS}
Panel A: Individual predictive regression (vs benchmark models)									
DP	-0.15	-1.46	-4.88	-3.78	6.24***	9.74**	-5.2	13.1***	-34.0
DY	-0.17	-0.72	-5.84	-3.83	7.47***	14.6**	-4.6	14.9***	-35.3
EP	-1.23	0.43	-2.40	-0.56	0.34	5.06	-6.2	10.3***	-32.1
DE	-1.36	0.30	-2.77	-1.13	0.54	7.39	-10.7	-2.1	-24.2
SVAR	-8.70	-0.82	-12.8	-0.21	-3.12	-4.18	-42.6	-77.1	11.7***
BM	-1.87	-0.44	-2.26	-0.64	-1.34	0.24	-6.9	-1.5	-15.3
NTIS	-1.97	-0.79	-0.08	0.26**	-4.53	-5.54	2.1***	12.9***	-14.8
TBL	-1.87	2.07	-2.92	-0.05	-0.46	12.6	-7.9	-10.7	-3.5
LTY	-1.72	1.37	-2.19	-0.59	-1.09	11.1	-7.5	-6.0	-9.8
LTR	0.03	0.17	1.47*	0.27	-1.92	-0.39	-26.3	-11.6	-49.3
TMS	-2.75	3.20**	-9.57	0.25	6.47**	17.8**	-10.3	4.5***	-33.8
DFY	-2.10	-0.21	-2.03	0.05*	-2.19	-1.35	-3.2	-13.7	13.2*
DFR	-2.22	0.06	-4.63	-1.64	1.02	8.57	-6.8	0.6***	-18.3
INFL	0.50	0.70	-0.24	-0.14	1.49	5.12	-0.6	7.8***	-13.8
IK	2.43***	2.24	-1.58	-0.59	7.86***	16.3*	-21.0	-29.6	-7.4
Panel B: Forecast combination regression (vs benchmark models)									
FC-TD	2.35***	1.26	1.52**	0.42	3.47**	5.56	2.9***	11.2***	-10.0
FC-FD	2.78***	2.61*	1.75**	0.91	4.18**	11.0	7.4***	15.6***	-5.5
FC-FD (no LF)	1.51***	-0.23	-5.05	-5.91	10.4***	29.0**	-76.0	-158	54.1***
FC-FD real time	4.41***	3.85*					14.8***		
Panel C: Forecast combination regression (vs FC-TD model)									
FC-FD	0.44	1.35	0.23	0.49	0.73	5.46	4.6***	5.0***	4.1**
FC-FD (no LF)	-0.86	-1.49	-6.67	-6.33	7.15***	23.4**	-81.4	-191	58.3***
FC-FD real time	2.11**	2.60*					12.3***		

Table 2: Equity premium and real GDP growth rate out-of-sample forecasting results

R^2_{OS} is the Campbell and Thompson (2008) out-of-sample R^2 statistic. CER gain is the portfolio management fee (in annualized percentage return) that an investor with mean-variance preferences and risk aversion coefficient of three would be willing to pay to have access to the forecasting model given in column (2), (4), or (6) relative to the benchmark forecasting model. Statistical significance for the R^2_{OS} statistic is based on the p-value for the Clark and West (2007) out-of-sample MSFE-adjusted statistic; the statistic corresponds to a one-sided test of the null hypothesis that the competing forecasting model has equal expected square prediction error relative to the benchmark forecasting model against the alternative hypothesis that the competing forecasting model has a lower expected square prediction error than the benchmark forecasting model. Statistical significance for the CER gains is based on a one-sided Diebold and Mariano (1995) test. $*$, $**$, and $***$ indicate significance at the 10 %, 5 %, and 1 % levels, respectively.

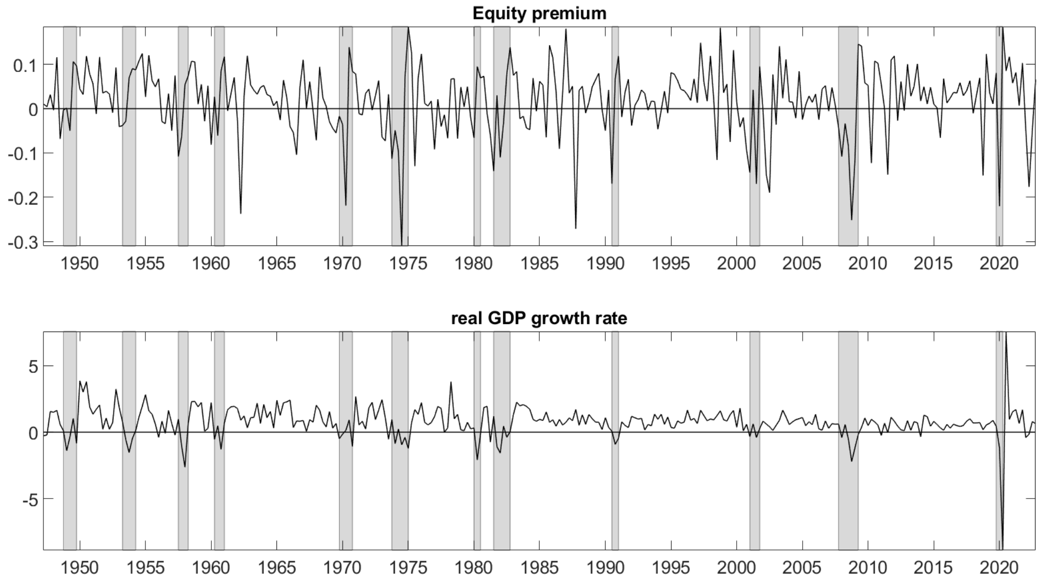


Figure 1: Equity premium and real GDP growth rate, U.S. data, 1947:Q1-2022:Q4

Time series of the quarterly log equity premium (upper graph) and quarterly real GDP growth rate (lower graph). Grey bars depict NBER-dated recessions.

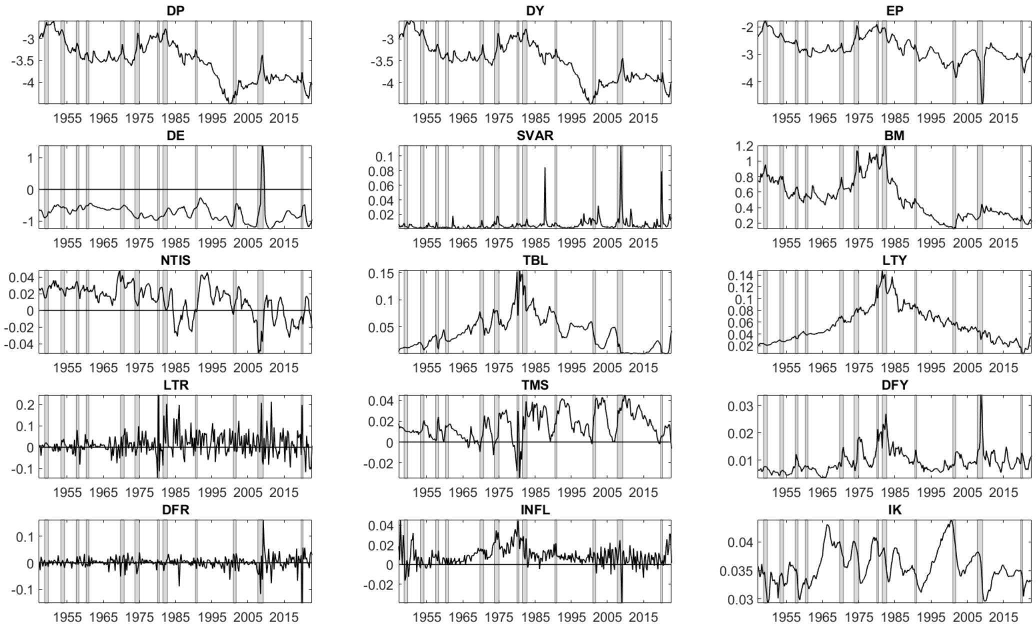


Figure 2: Time series of the predictors, U.S. data, 1947:Q1-2022:Q4
Grey bars depict NBER-dated recessions.

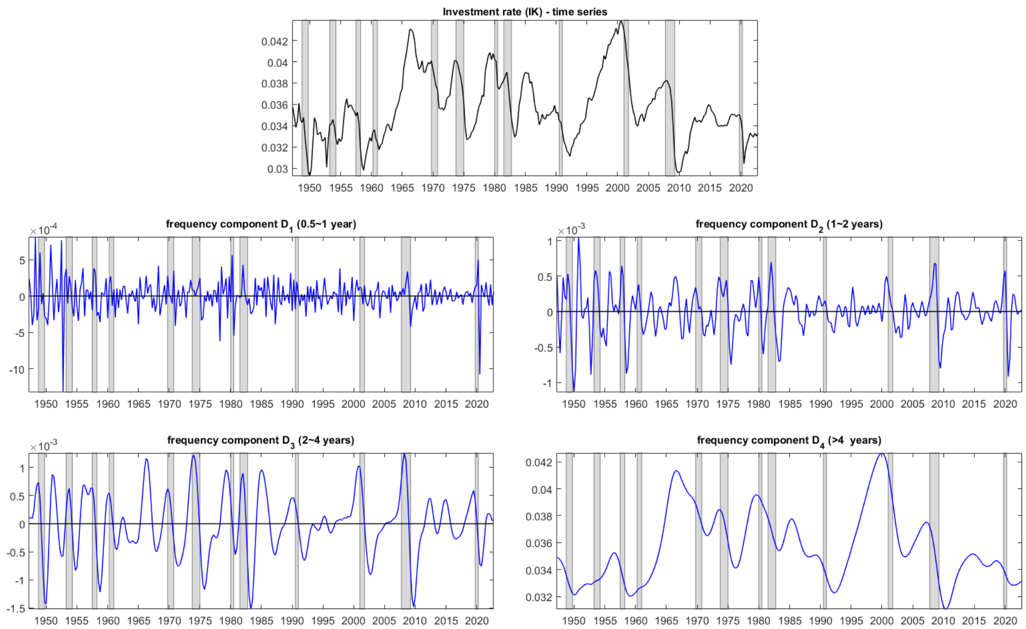
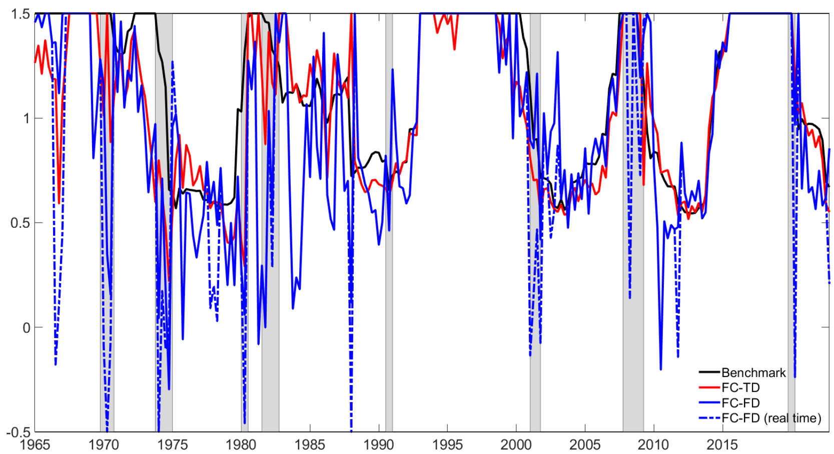


Figure 3: Investment rate (IK), time series and frequency decomposition, U.S. data, 1947:Q1-2022:Q4

The top panel shows the time series of quarterly U.S. investment rate, while the remaining panels show the four time series frequency components into which the investment rate series is decomposed. Grey bars depict NBER-dated recessions.

A. Equity weights



B. Log cumulative wealth

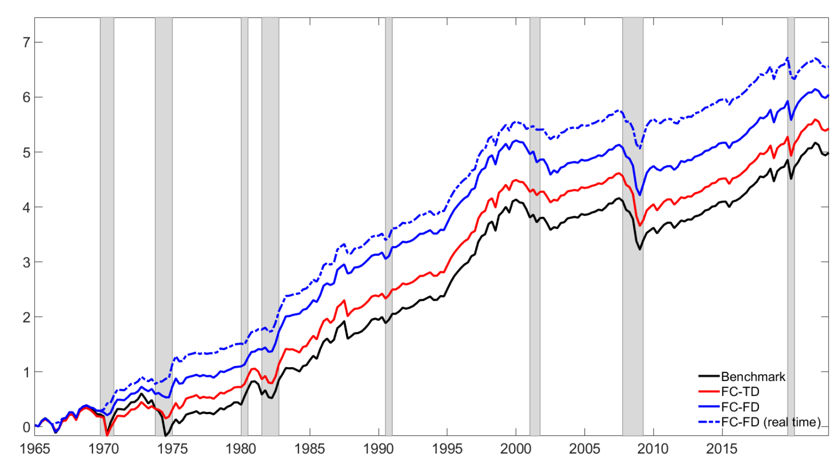


Figure 4: Equity weights and log cumulative wealth, 1965:Q1-2022:Q4

Panel A plots the dynamics of the equity weights for a mean-variance investor with relative risk aversion coefficient of three who allocates quarterly between equities and risk-free bills using a predictive regression excess return forecast based on the benchmark forecast model (black solid line), the FC-TD model (red solid line), the FC-FD model (blue solid line), or the FC-FD real time model (blue dashed line). The equity weights are constrained to lie between -0.5 and 1.5. Panel B delineates the corresponding log cumulative wealth for the investor, assuming that the investor begins with 1€ and reinvests all proceeds. Grey bars denote NBER-dated recessions.

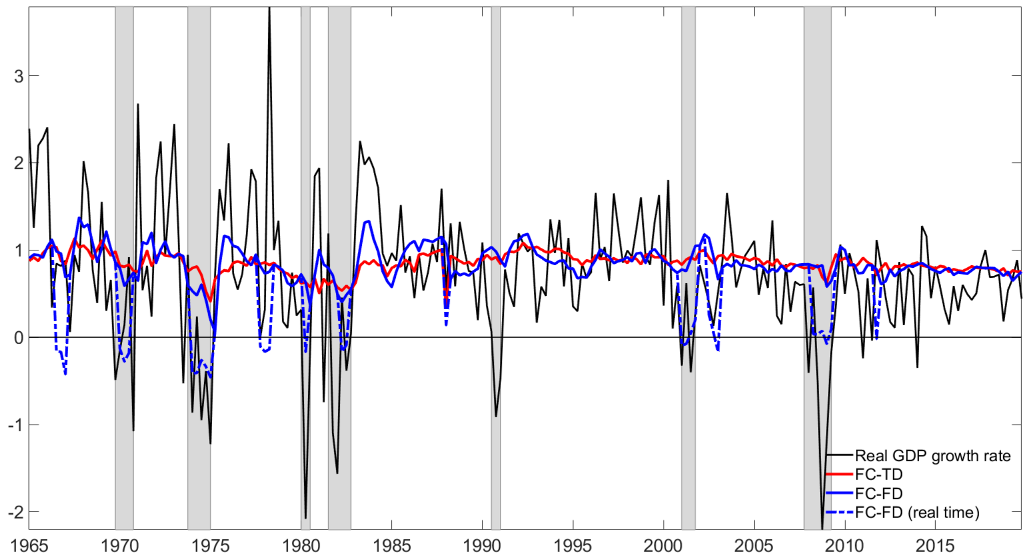


Figure 5: Real GDP growth rate, realized and out-of-sample forecasts, 1965:Q1-2019:Q4

Quarterly U.S. real GDP growth rate (black solid line) and its out-of-sample forecasts based on the FC-TD model (red solid line), the FC-FD model (blue solid line), or the FC-FD real time model (blue dashed line). Grey bars depict NBER-dated recessions.

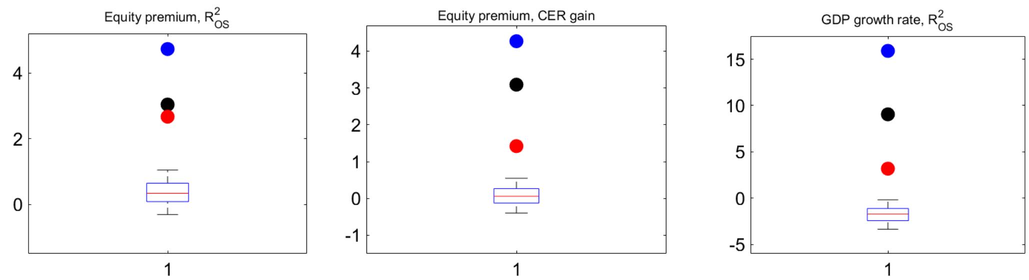


Figure 6: Placebo test R^2_{OS} and CER gains

In each box, the red line displays the median across models, the boundaries of the box depict the 25 % and 75 % percentiles, and the whiskers outside of the box mark the 90 % confidence interval of the distribution. Red, black, and blue dots denote the results (as reported in table 2) with the FC-TD model, the FC-FD model, and the FC-FD real time model, respectively.

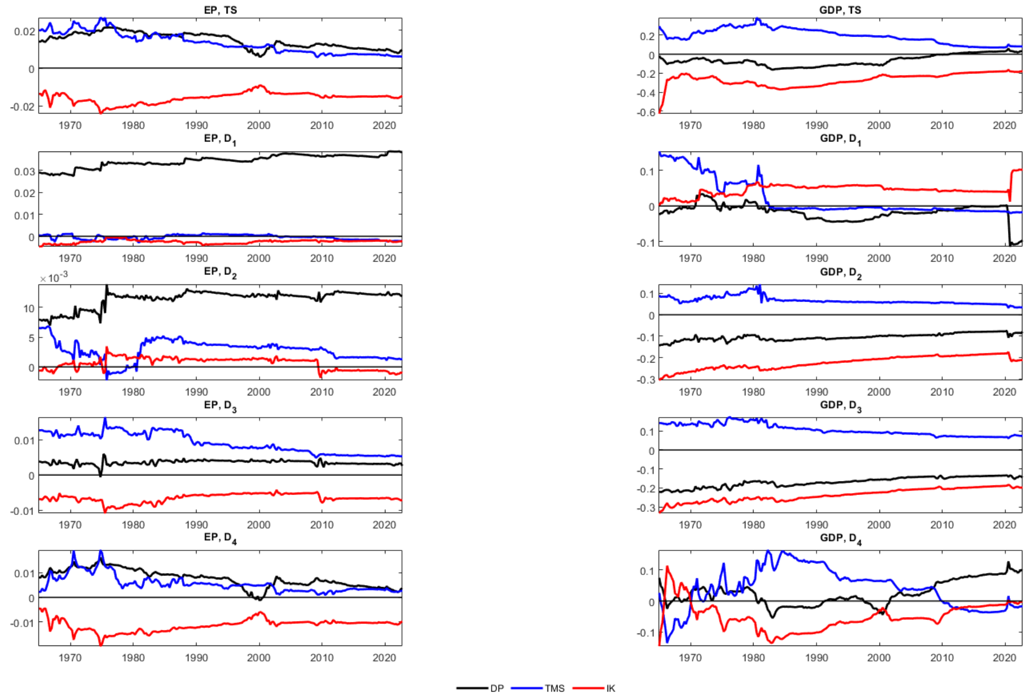


Figure 7: OLS regression coefficients between the target variables and three individual predictors based on expanding window estimates

Left (right) column: OLS regression coefficients (based on expanding window estimates starting in 1947:Q2-1964:Q4Q4, recursively including one additional quarter through 2022:Q4) between the equity premium (real GDP growth) and three individual predictors (DP, black lines; TMS: blue lines; IK, red lines). Top row: time series. Remaining rows: coefficients in each frequency band. Each predictor variable is standardized to have a standard deviation of one before running the estimation.

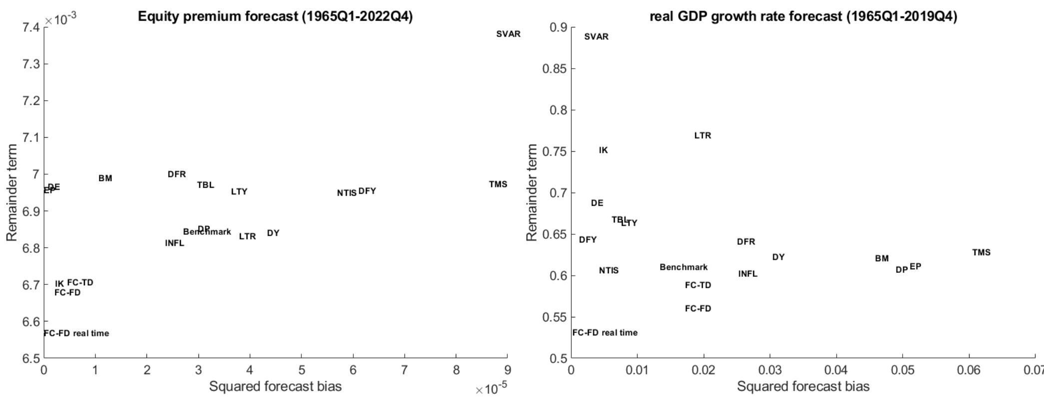


Figure 8: Scatterplot of the Theil (1971) MSFE decomposition into the squared forecast bias and a remainder term

Left (right) graph: equity premium (real GDP growth rate) forecast. Benchmark corresponds to the benchmark forecast model (historical mean for equity premium, and AR(p) for GDP growth), and FC-TD, FC-FD, and FC-FD real time denote the combination forecast in the time domain, in the frequency domain, and in the frequency domain in the real time exercise in section 6.2, respectively. The other points correspond to the individual predictive regression model forecasts.

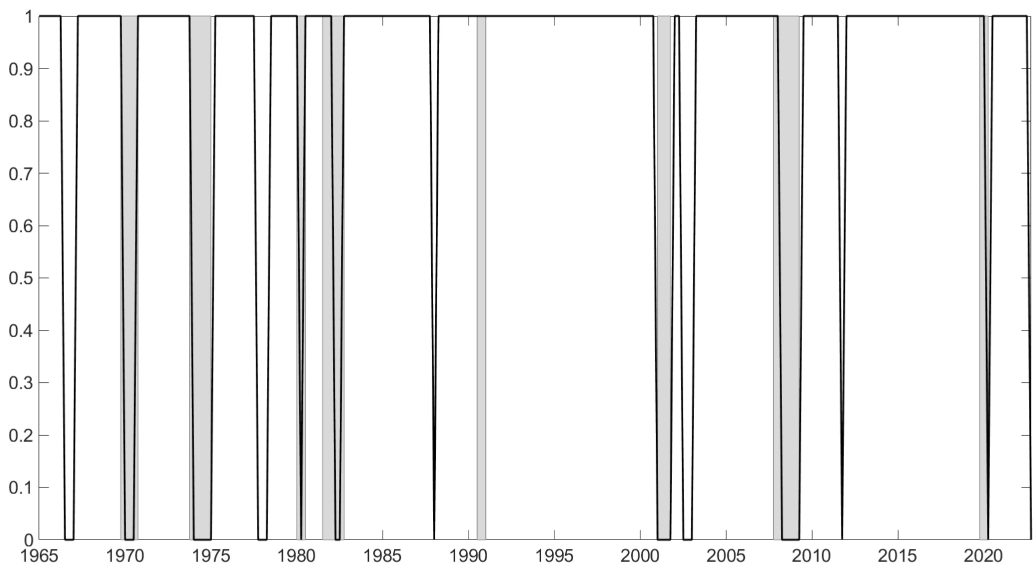


Figure 9: Coincident index and NBER-dated recessions (grey bars)

The coincident index is computed from five stock market technical indicators. A value of 0 indicates a recessions. See appendix 2 for a description of the technical indicators.

Appendix 1. List of predictors

- Log dividend-price ratio (DP): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of prices (S&P 500 index).
- Log dividend yield (DY): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of lagged prices (S&P 500 index).
- Log earnings-price ratio (EP): difference between the log of earnings (12-month moving sums of earnings on S&P 500) and the log of prices (S&P 500 index price).
- Log dividend-payout ratio (DE): difference between the log of dividends (12-month moving sums of dividends paid on S&P 500) and the log of earnings (12-month moving sums of earnings on S&P 500).
- Stock variance (SVAR): sum of squared daily returns on the S&P 500.
- Book-to-market ratio (BM): ratio of book value to market value for the Dow Jones Industrial Average.
- Net equity expansion (NTIS): ratio of 12-month moving sums of net equity issues by NYSE-listed stocks to the total end-of-year NYSE market capitalization.
- Treasury bill rate (TBL): three-month Treasury bill rate.
- Long-term yield (LTY): long-term government bond yield.
- Long-term return (LTR): long-term government bond return.
- Term spread (TMS): difference between the long-term government bond yield and the T-bill.
- Default yield spread (DFY): difference between Moody's BAA- and AAA-rated corporate bond yields.

- Default return spread (DFR): difference between long-term corporate bond and long-term government bond returns.
- Inflation rate (INFL): calculated from the Consumer Price Index (CPI) for all urban consumers.
- Investment to capital ratio (IK): ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy.

Appendix 2. Haar filter

To extract the different frequency components from the data, we use the Maximal Overlap Discrete Wavelet Transform (MODWT). This approach permits decomposition of any variable, regardless of its time series properties, into a trend and several cycles in a manner similar to the traditional Beveridge and Nelson (1981) time series trend-cycle decomposition approach.

By using the MODWT with the Haar filter, any variable X_t can be decomposed as:

$$X_t = \sum_{j=1}^{J} D_{j,t} + S_{J,t} \ , \tag{6}$$

where $D_{j,t}$ are the wavelet coefficients at scale j, and $S_{J,t}$ is the scaling coefficient. These coefficients are given by

$$D_{j,t} = \frac{1}{2^j}\left[\sum_{i=0}^{2^{(j-1)}-1} X_{t-i} - \sum_{i=2^{(j-1)}}^{2^j-1} X_{t-i}\right] \tag{7}$$

and

$$S_{J,t} = \frac{1}{2^J}\sum_{i=0}^{2^J-1} X_{t-i} \ . \tag{8}$$

In particular, in our analysis we compute a J=3 level decomposition. Hence the corresponding time series components are given by

$$\begin{aligned}
D_{1,t} &= \frac{X_t - X_{t-1}}{2} \\
D_{2,t} &= \frac{X_t + X_{t-1} - (X_{t-2} + X_{t-3})}{4} \\
D_{3,t} &= \frac{X_t + X_{t-1} + X_{t-2} + X_{t-3} - (X_{t-4} + X_{t-5} + X_{t-6} + X_{t-7})}{8} \\
S_{3,t} &= \frac{X_t + X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4} + X_{t-5} + X_{t-6} + X_{t-7}}{8} \ .
\end{aligned}$$

In the paper we denote component S_3 as D_4.

Appendix 3. List of technical indicators

Let P_t be the stock price index in quarter t.

Moving average indicator. The MA rule generates a buy ($S_{i,t} = 1$) or sell ($S_{i,t} = 0$) signal at the end of quarter t by comparing two moving averages:

$$S_{i,t} = \begin{cases} 1 & if\, MA_{short,t} \geq MA_{long,t} \\ 0 & if\, MA_{short,t} < MA_{long,t} \end{cases}$$

where

$$MA_{j,t} = \frac{1}{j} \sum_{i=0}^{j-1} P_{t-i} \; for \; j = short, long$$

and *short* (*long*) is the length of the short (long) MA (*short* < *long*). The MA indicator with MA lengths *short* and *long* is denoted as MA(*short*,*long*). Intuitively, the MA rule detects changes in stock price trends because the short MA will be more sensitive to recent price movement than will the long MA. In the paper we use MA indicators with *short*=1 and *long*=3,4.

Momentum indicator. The momentum rule generates the following buy ($S_{i,t} = 1$) or sell ($S_{i,t} = 0$) signal at the end of quarter t:

$$S_{i,t} = \begin{cases} 1 & if\, P_t \geq P_{t-m} \;, \\ 0 & if\, P_t < P_{t-m} \;. \end{cases}$$

Intuitively, a current stock price that is higher than its level m periods ago indicates positive momentum and relatively high expected excess returns, thereby generating a buy signal. The momentum indicator that compares P_t to P_{t-m} is denoted by MOM(m) and we compute momentum indicators for m=3, 4, 6.

These technical indicators are plotted in figure 10.

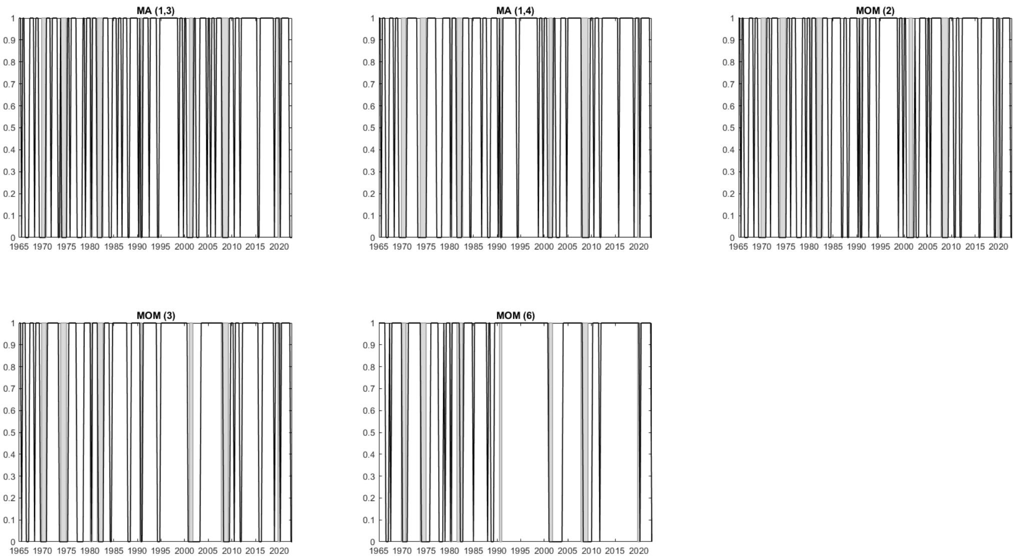


Figure 10: Time series of the technical indicators, U.S. data, 1965:Q1-2022:Q4
Grey bars depict NBER-dated recessions.