UNIVERSITÄT TRIER

Felix Haase

Sum-of-the-Parts Revised: Economic Regimes and Flexible Probabilities

Research Papers in Economics
No. 10/24

Sum-of-the-Parts Revised: Economic Regimes and Flexible Probabilities*

Felix Haase†

Department of Economics, University of Trier

First Draft: September 9, 2024
This Version: September 27, 2024

Abstract

Building on the success of Ferreira and Santa-Clara (2011) in separately forecasting the return components of the stock market, this paper examines the links between economic regimes and these components to predict the aggregate U.S. stock market. We propose a three-step methodology that we call the *flexible regime* approach. First, we estimate the regime dynamics of ten macro-financial variables using Markov-switching regressions. Second, we treat the regime filtering results from the Hamilton filter as views and test the predicted regime classification, the predicted regime probabilities, and the conditional and mixture densities as view generators. We use entropy pooling to re-weight the historical distribution and to derive posterior probabilities. Finally, we link these probabilities to the realized outcomes of earnings growth and changes in the price-earnings multiple to form the sum-of-the-parts forecast. Our results demonstrate significant predictability from a statistical and economic perspective. We emphasize the role of default spreads and interest rates in predicting earnings growth and stock market volatility and inflation in predicting multiple growth. Finally, our results suggest that the predictability of both return components varies over time and is affected by the business cycles. While earnings growth is more predictable during periods of expansion, forecasting multiple growth is more advantageous during recessions.

JEL Codes: C53; G11; G17.

Keywords: Economic Restrictions; Entropy Pooling; Flexible Probabilities; Markov-switching Models; Return Predictability; Stock Market Regimes; Sum-of-the Parts.

*I am grateful to Matthias Neuenkirch and Thomas Osowski for their helpful comments on a previous version of this article.

†Corresponding author: Trier University, Department of Economics, Universitätsring 15, 54296 Trier, Germany, E-mail: mail@felixhaase.de

1 Introduction

"The sum of the parts is more than the whole" – with this insight, Ferreira and Santa-Clara (2011) established a new strand in the literature on return predictability. Decomposing the complex return process into its economic components helps to reduce estimation uncertainty and increase predictability. In their basic model, they assume that earnings growth has only a low-frequency predictable trend, that the changes in the price-earnings ratio are zero, and that the dividend-price ratio follows a random walk. There is ample evidence that this simple forecasting strategy is hard to beat, especially for the U.S. stock market, see for instance McMillan and Wohar (2011) and Dichtl et al. (2021). So far, extensions have either predicted the return components using fundamental variables and/or technical indicators as regressors (Ferreira and Santa-Clara 2011; Baetje and Menkhoff 2016; Yi et al. 2019; Dai and Zhou 2020; Dichtl et al. 2021) or modeled the dynamics of the components endogenously (Faria and Verona 2018; Dai and Zhu 2020). It is evident that earnings growth is linked to the business cycle (Longstaff and Piazzesi 2004; Van Binsbergen and Koijen 2010) and that the price-earnings multiple tends to revert in the medium to long term (Shiller 2015). Accordingly, the dynamics of these components should exhibit some cyclicality, which may be related to different economic conditions. Nevertheless, a comprehensive evaluation of the predictive relationship between economic regimes and the return components is still lacking in the literature.

This paper fills this gap by combining economically motivated model restrictions with regimes. We propose a three-step methodology, which we call the *flexible regime approach*. First, we filter the regime dynamics of ten macro-financial predictors using a two-state Markov-switching autoregressive model. Second, we derive flexible probabilities of economic regimes to re-weight the empirical distribution. In this context, we employ entropy pooling to derive the posterior probabilities, treating the information about the regime dynamics as a view. We test the predicted regime classification, the predicted regime probabilities, the conditional and mixture densities as view generators. We compute expected values of earnings and price-earnings growth as a weighted average of historical realizations and regime-implied probabilities. Finally, we combine the return component forecasts with the current dividend-price ratio to predict excess returns. We examine whether the predictability of changes in earnings or the price-earnings ratio contributes more to the aggregate return predictability. We also analyze whether their impact depends on the state of the economy and which macro-financial predictor is helpful in this regard.

The empirical results demonstrate significant predictability of the *flexible regime* approach in terms of statistical accuracy and economic benefits for the United States over a 75-year history (1947/12 to 2022/12). Depending on the specification, almost all forecasts have both positive R^2_{OS} and certainty-equivalent gains. At their best, these strategies outperform the historical average by 1.96% in terms of R^2_{OS}, and a moderately risk-averse investor would be willing to pay an average management fee of 2.17% p.a. at most. Testing different weighting schemes suggests that binary regime classification and predicted state probabilities are more advantageous in recessions, and density-smoothed approaches, whether as a conditional state density or as a mixture of two states, are more beneficial in expansions. In this context, a Bayesian perspective using entropy pooling significantly improves the trade-off between forecast variance and forecast accuracy. The impact of regime-filtered variables depends on whether the economy is in an expansion or a recession and whether the predictability stems from changes in earnings or the price-earnings ratio. The results emphasize the role of default spreads or interest rates in predicting earnings growth and stock market volatility or inflation in predicting price-earnings multiple growth. We provide evidence that the predictability of both return components varies over time and is affected by the business cycles. While earnings growth is more predictable during expansion, forecasting changes in the price-earnings ratio is more advantageous during recessions. In addition, our regime-based trading strategies are robust to transaction costs and various investment constraints and are particularly helpful for investors with low and moderate risk aversion.

This paper contributes to several strands of the literature. First, we add to the overwhelming body of research on the predictability of stock returns. Finding adequate predictors or approximating the stock market dynamics is difficult for many reasons (Welch and Goyal 2008). Continuous learning by investors and competition among market participants often ensure that predictable patterns quickly disappear once they have been discovered. From an empirical perspective, parameter instability, model and estimation uncertainty complicate this task (Pesaran and Timmermann 1995). Therefore, successful strategies should reduce estimation uncertainty, detect time-varying patterns, rely on robust economic or behavioral relationships, and exploit frictions in the financial market structure. According to the survey of Rapach and Zhou (2013), four promising paths are able to at least partially meet these requirements: (i) economically motivated model restrictions (Campbell and Thompson 2008; Ferreira and Santa-Clara 2011; Pettenuzzo et al. 2014; Zhang et al. 2019), (ii) modeling regime shifts (Guidolin and Timmermann 2007; Dangl and Halling 2012), (iii) dimension reduction (Ludvigson and Ng 2007; Çakmaklı and Dijk 2016), and (iv) forecasting combination (Rapach et al. 2010; Neely et al. 2014). There are also examples that combine some (Zhu and Zhu 2013; Hammerschmid and Lohre 2018; Haase and Neuenkirch 2023). However, studies that employ economic restrictions and regime-switching models are rare. Since restrictions aim to reduce estimation uncertainty, this is usually at the expense of timely consideration of regime changes. We resolve this trade-off through a novel combination scheme that separates the evaluation of observations with economic regimes from a decomposed forecasting model. Instead of using complex models for the return components, we propose a flexible but simple approach in which the expected return components are expressed as a weighted average of past realizations and of a flexible probability vector. The vector is derived using entropy pooling, which modifies a prior distribution to ensure that the moments resulting from the respective regime prediction are satisfied. This procedure captures regime shifts in a timely manner without adding too much variability to the forecasts due to lower estimation uncertainty.

Second, our results provide a more nuanced view of the drivers of stock market predictability. There is a extensive academic debate as to whether the discount rate or cash flow channel drives the predictability of stock returns. Based on the present value decomposition of Campbell and Shiller (1988), proponents of the former argue that changes in dividends are independent and identically distributed (i.i.d.) such that the current dividend yield results only from expected returns (Cochrane 2011). In contrast, proponents of the cash flow channel say that fluctuations in the dividend yields are mainly due to the uncertainty in future cash flows and long-term growth expectations (Bansal and Yaron 2004). Our results suggest that the role of each channel is state-dependent and influenced by the dynamics of different variables. The cash flow channel (approximated by earnings growth) is more predictable during expansions with the regime filtering of monetary policy and default spreads, while the dynamics of the discount rate channel (using price-earnings multiple growth as a proxy) is more predictable during recessions.

Third, we contribute to the empirical strand of the literature that emphasizes the different information content of (past) observations. Based on the assumption that historical observations do not have the same relevance for the future path of a variable, a variety of weighting schemes are possible that can be used for parameter estimation and forecasting. In this context, nonlinear models such as smooth transition, Markov-switching, or kernel estimation are popular choices (Gisbert 2003; Guidolin et al. 2009; Cheng et al. 2019). More recently, Czasonis et al. (2020) addressed another avenue by measuring the relevance of observations as the sum of similarity (to the last data point) and informativeness (expressed as deviation from expected normal conditions). They use this measure as a decision criterion to select the relevant subset for the prediction task. Closely related to this approach, Meucci (2010) introduces a flexible, non-parametric way to model the future distribution of returns. He derives flexible probabilities for historical scenarios conditioning on time, market states, or some prospective market views. While his primary focus lies on risk management and stress testing tasks, this idea can be applied to forecasting and asset allocation exercises, see for example Meucci et al. (2014) and Pedersen (2017). We build on this strand of the literature by using the filtration from Markov-switching models to systematically evaluate the utility of economic regimes for weighting observations. To the best of our knowledge, we are the first to pursue precisely this direction.

The remainder of this paper is organized as follows: Section 2 reviews the sum-of-the-parts literature. Section 3 presents the three steps of our *flexible regime* methodology and the evaluation measures. Section 4 introduces the dataset. Section 5 illustrates how our approach works in-sample, and Section 6 reports the out-of-sample results using a recursive forecasting setup. Finally, Section 7 concludes.

2 The Sum-of-the-Parts Method

Ferreira and Santa-Clara (2011) introduce their sum-of-the-parts (SOP) methodology by decomposing the gross stock market return $(1 + R_{t+1})$ into gross capital gains $(1 + CG_{t+1})$, and the dividend yield (DY_{t+1}):

$$1 + R_{t+1} = 1 + CG_{t+1} + DY_{t+1} = \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t}. \tag{1}$$

The capital gain is defined as the percentage change in the stock price $(CG_{t+1} = P_{t+1}/P_t - 1)$ and the dividend yield is the ratio of future dividends to the lagged stock price $(DY_{t+1} = D_{t+1}/P_t)$. If both the numerator and the denominator are expanded by the gross growth rate of earnings $(1 + GE = E_{t+1}/E_t)$, then the gross capital gain can be expressed as the product of the gross growth rate of the price-earnings multiple $(1 + GM = \frac{P_{t+1}/E_{t+1}}{P_t/E_t})$ and that of earnings:[1]

$$\begin{aligned} 1 + CG_{t+1} = \frac{P_{t+1}}{P_t} &= \frac{P_{t+1}/E_{t+1}}{P_t/E_t}\frac{E_{t+1}}{E_t} \\ &= (1 + GM_{t+1})(1 + GE_{t+1}) \end{aligned} \tag{2}$$

Similarly, the dividend yield can be decomposed as the product of the dividend-price ratio D_{t+1}/P_{t+1} and the gross capital gain:

$$\begin{aligned} DY_{t+1} = \frac{D_{t+1}}{P_t} &= \frac{D_{t+1}}{P_{t+1}}\frac{P_{t+1}}{P_t} \\ &= \frac{D_{t+1}}{P_{t+1}}\left(\frac{P_{t+1}/E_{t+1}}{P_t/E_t}\frac{E_{t+1}}{E_t}\right) \\ &= DP_{t+1}(1 + GM_{t+1})(1 + GE_{t+1}) \end{aligned} \tag{3}$$

1. As Ferreira and Santa-Clara (2011) note, the earnings E_{t+1}/E_t and the multiple P_{t+1}/E_{t+1} can be replaced by the book value, the free cash flow, or another fundamental measure and their respective multiple. Our choice of earnings growth is motivated by its close relationship with macroeconomic variables and the limited data availability over a long history for the other measures.

Combining the two parts from Eq. (1) shows that the total stock return is equal to the product of one plus the dividend-price ratio and the gross growth rate of the price-earnings multiple and that of earnings:

$$\begin{aligned} 1 + R_{t+1} &= \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t} \\ &= (1 + GM_{t+1})(1 + GE_{t+1}) + DP_{t+1}(1 + GM_{t+1})(1 + GE_{t+1}) \\ &= (1 + GM_{t+1})(1 + GE_{t+1})(1 + DP_{t+1}) \end{aligned} \tag{4}$$

Taking the logarithm of both sides makes the term additive:

$$r_{t+1} = gm_{t+1} + ge_{t+1} + dp_{t+1} \tag{5}$$

The SOP method suggests that the components of stock returns should be forecasted separately. Ferreira and Santa-Clara (2011) make assumptions about reasonable estimates of the three components. In their simplest version, they assume that the dividend-price ratio follows a random walk and that there are no expected changes in the price-earnings multiple. Based on Van Binsbergen and Koijen (2010)'s findings that earnings growth has only a low-frequency predictable component, they propose a 20-year moving average as a proxy for future earnings growth. As a result, their excess return forecast ($\hat{\mu}$) over the risk-free rate (rf) is as follows:

$$\hat{\mu}_{t+1} = \hat{r}_{t+1} - rf_{t+1} = \bar{ge}_t^{20Y} + dp_t - rf_{t+1}. \tag{6}$$

According to Ferreira and Santa-Clara (2011), this forecasting model can also be interpreted as a constrained predictive regression with the (excess) earnings growth as intercept and the dividend-price ratio as predictor with a slope of one.[2] They show that this simple model robustly predicts stock returns. The main advantages are the lower estimation uncertainty compared to predictive regression models and a better ability to capture the long-term cyclicality of the economy relative to the historical average. International evidence suggests that the SOP approach works particularly well for the equity markets in the UK, Japan, and the U.S. (Ferreira and Santa-Clara 2011; McMillan and Wohar 2011). Therefore, we use this version as the ultimate benchmark strategy throughout the paper and refer to it as FSC.

Ferreira and Santa-Clara (2011) test two extensions themselves. First, they forecast multiple growth or its reversion using predictive regression with fundamental predictors. And second, they use analysts' earnings forecasts instead of realized earnings to better capture the expected part. By applying shrinkage to the regression estimates, they achieved improvements in predictability in both cases. Particularly helpful were

2. Since such a constraint could be purely coincidental, Ferreira and Santa-Clara (2011) verified the robustness and could allayed these concerns.

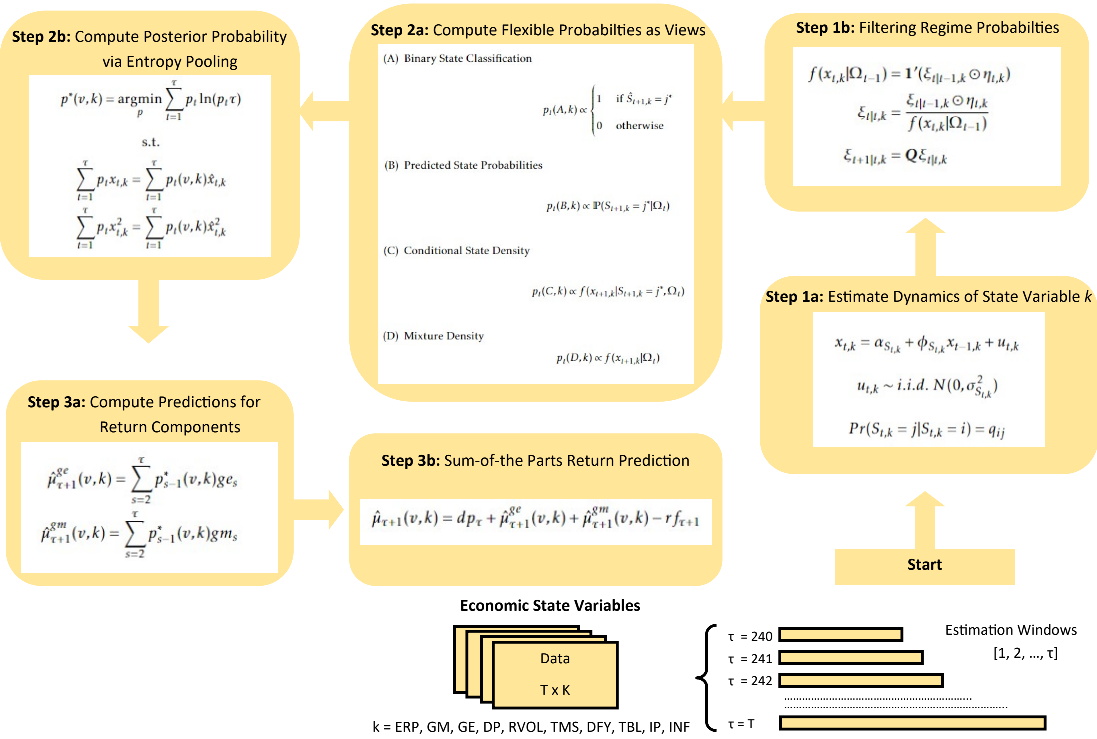
term spreads, bond yields, net equity expansions, and default yield spreads. Alternative estimates of the three components are possible. First, some authors focus on the components' dynamics and decompose the frequency spectrum using wavelet analysis (Faria and Verona 2018) or empirical mode decomposition (Dai and Zhu 2020). Second, other researchers model the expected return components with external predictors. Baetje and Menkhoff (2016) extend the set of fundamental predictors with technical indicators, apply predictive regression to all components, and pool the univariate forecasts to account for model uncertainty. Yi et al. (2019) use predictive regression with cyclically decomposed predictors. Dai and Zhou (2020) predict multiple growth with economic restrictions based on momentum rules of past predictability, trimming extreme forecasts or restricting the forecast sign. All of these extensions are able to find certain strategies that can significantly outperform the basic FSC in terms of statistical accuracy and economic value. Of course, data snooping and multiple testing can be a problem, but Dichtl et al. (2021) show that the SOP method (including some extensions) still provides robust results even after controlling for these effects.

3 Methodology

A large body of research shows that regimes matter for stock markets (Guidolin and Timmermann 2007; Henkel et al. 2011; Hammerschmid and Lohre 2018). Therefore, it would be a logical consequence that their components are also, to some extent, regime-dependent. The relevant question is how to empirically account for regimes in a forecasting setting.

We address this question through a three-step procedure that combines regime models with economically motivated restrictions to forecast stock returns. Figure 1 summarizes the methodology of our semi-parametric approach, which we call the *flexible regime* approach. Instead of relating the predictors directly to the return components, we employ Markov-switching models on ten economic variables k to filter their regime dynamics (Step 1a). The Hamilton filter provides several alternatives for the weighting vector to reflect our regime view v (Step 1b). We test the predicted state classification (A), the predicted regime probabilities (B), the conditional (C), and mixture densities (D) as view generators (Step 2a). Then, we apply entropy pooling to derive flexible probabilities $p^*(v,k)$ that minimize the distance relative to the historical distribution, using our views as constraints (Step 2b). This step helps to reduce the variance of the forecasts by smoothing the observation weights. We link the resulting posterior probabilities to the realized earnings outcomes and multiple growth to form conditional expectations on both components (Step 3a). Finally, we aggregate their predictions with the current dividend-price ratio to obtain the sum-of-the-parts forecast (Step 3b). We test different specifications to evaluate the marginal benefits of the economically motivated restriction and assess whether the predictability of earnings growth or multiple growth contributes more to the aggregate return predictability. Additionally, we combine the information from individual regime forecasts as part of our extensions.

Figure 1: Graphical Summary of the *Flexible Regime* Approach



Notes: Figure summarizes our three-step methodology. Step 1 estimates the Markov-switching autoregressive models to filter the regime dynamics from the respective variables using the Hamilton filter. $\boldsymbol{Q}$ represents the transition probability matrix with its elements $q_{ij} = Pr(S_{t,k} = j | S_{t-1,k} = i)$, $\xi_{t+1|t,k} = [\mathbb{P}(S_{t+1,k} = 0|\Omega_t), \mathbb{P}(S_{t+1,k} = 1|\Omega_t)]'$ is the vector of predicted probabilities and $\xi_{t|t,k} = [\mathbb{P}(S_{t,k} = 0|\Omega_t), \mathbb{P}(S_{t,k} = 1|\Omega_t)]'$ denotes the filtered probabilities. The conditional state density is defined as $f(x_{t,k}|S_{t,k} = j, \Omega_{t-1}) = \eta_{t,k}$ and the mixture density is given as $f(x_{t,k}|\Omega_{t-1}) = \mathbf{1}'(\xi_{t|t-1,k} \odot \eta_{t,k}) = \sum_{j=0}^{1} \mathbb{P}(S_{t,k} = j|\Omega_{t-1}) f(x_{t,k}|S_{t,k} = j, \Omega_{t-1})$. Step 2 computes the flexible probabilities that re-weight the historical distribution function to account for our regime prediction. In this context, we test four different specifications $v \in \{A, B, C, D\}$ as view generators and use entropy pooling to reduce estimation uncertainty. In the final step, we first map the flexible probabilities to the distribution of subsequent realizations of earnings and multiple growth to form their expected value. Finally, we combine the forecasts of both return components with the most recent dividend-price ratio to obtain the final sum-of-the-parts return prediction. We re-iterate all steps using an expanding window to compute the one-step ahead *flexible regime forecast* for each regime variable. In Step 3, alternative forecasting models are tested, and the state variable for earnings and multiple growth prediction need not be identical. For more details, we refer to Subsection 3.3. The first estimation window ranges from 1927/12 to 1947/11, and the out-of-sample period runs from 1947/12 to 2022/12.

The notation for the next subsections is as follows. The economic variables are given by a $T \times K$ matrix X with the corresponding elements $x_{t,k}$. The number of variables is given by K $(k = 1, 2, ..., K)$, and the number of observations by T $(t = 1, 2, ..., T)$. The total sample length consists of an initial estimation window of length T_0 and an out-of-sample period of length T_1 $(T = T_0 + T_1)$. The hidden state of an observable variable k at time t is denoted by $S_{t,k}$. We assume the existence of two discrete states. In each t, we make one-step ahead predictions for $t+1$ using the available information set Ω_t. We use a recursive approach with τ as the monthly time index representing the expanding estimation window, starting with an initial window of 20 years $(\tau = 240, 241, ..., T - \tau)$.

3.1 Step 1: Filtering the Regime Dynamics

Markov-switching models are useful tools for extracting real-time information about unobservable regimes that can often be explained ex-post (Ang and Timmermann 2012). The use of parametric mixture distributions allows researchers to capture many stylized facts about macroeconomic and financial data (e.g., asymmetric and leptokurtic patterns) in a still analytically tractable framework. Exploiting these capabilities, we apply these models to various stock market predictors to gain insights into the (expected) state of the economy or a particular market.

We specify the data-generating process for each variable independently and try to control for possible short-run persistence in their dynamics. Thus, we propose a first-order Markov-switching autoregressive model, MSAR(1):

$$\begin{aligned} x_{t,k} &= \alpha_{S_{t,k}} + \phi_{S_{t,k}} x_{t-1,k} + u_{t,k} \\ u_{t,k} &\sim i.i.d.\ N(0, \sigma^2_{S_{t,k}}) \\ Pr(S_{t,k} = j | S_{t-1,k} = i) &= q_{ij} \end{aligned} \tag{7}$$

We assume that the hidden regime variable $S_{t,k}$ follows a homogeneous first-order Markov chain, i.e. the current regime j depends only on the realization of the previous regime i. Thus, the transition probability matrix $\boldsymbol{Q}$ drives the entire dynamics $(q_{ij} \in \boldsymbol{Q})$. All parameters in the mean equation $(\alpha_{S_{t,k}}, \phi_{S_{t,k}})$ as well as the variance of the errors $\sigma^2_{S_{t,k}}$ are allowed to vary between regimes. To demonstrate the regime dependence and to avoid unnecessarily complex modeling (due to the increasing number of parameters), we consider only two regimes $(i, j \in \{0, 1\})$. Certainly, a higher number of regimes helps to better understand the regime dynamics and to increase the explanatory power of the process, but the higher estimation uncertainty is usually a hindrance for prediction purposes. Often, two regimes are sufficient (Ang and Timmermann 2012; Haase and Neuenkirch 2023). Therefore, we leave the investigation of the number of regimes to future research.

Since the process $S_{t,k}$ is hidden, it is never known with certainty in which regime you are in t. Hamilton (1989) developed a recursive filter to derive inferences about $S_{t,k}$. Let denote $\xi_{t|t,k} = [\mathbb{P}(S_{t,k} = 0|\Omega_t), \mathbb{P}(S_{t,k} = 1|\Omega_t)]'$ be the vector of filtered probabilities and $\xi_{t+1|t,k} = [\mathbb{P}(S_{t+1,k} = 0|\Omega_t), \mathbb{P}(S_{t+1,k} = 1|\Omega_t)]'$ be the vector of predicted probabilities. We denote $\eta_{t,k}$ as the vector of normally distributed conditional state densities.

$$\eta_{t,k} = \begin{bmatrix} f(x_t|S_{t,k} = 0, \Omega_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_{0,k}} \exp\left(-\frac{x_{t,k}-\alpha_{0,k}-\phi_{0,k}x_{t-1,k}}{2\sigma_{0,k}^2}\right) \\ f(x_t|S_{t,k} = 1, \Omega_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_{1,k}} \exp\left(-\frac{x_{t,k}-\alpha_{1,k}-\phi_{1,k}x_{t-1,k}}{2\sigma_{1,k}^2}\right) \end{bmatrix} \tag{8}$$

According to Hamilton (1994, p.692), we can run the update scheme for each new t based on an initial probability vector $\xi_{1|0,k}$:

$$\begin{aligned} f(x_{t,k}|\Omega_{t-1}) &= \mathbf{1}'(\xi_{t|t-1,k} \odot \eta_{t,k}) \\ \xi_{t|t,k} &= \frac{\xi_{t|t-1,k} \odot \eta_{t,k}}{f(x_{t,k}|\Omega_{t-1})} \\ \xi_{t+1|t,k} &= \boldsymbol{Q}\xi_{t|t,k} \end{aligned} \tag{9}$$

where $\odot$ is the element-by-element product and $\mathbf{1}$ is a vector of ones. This recursive procedure, also known as Hamilton filter, allows us to assign the observation $x_{t,k}$ to a particular regime and provides valuable insight into the regime's dynamics. A nice feature is that the likelihood function $f(x_{t,k}|\Omega_{t-1})$ results as a by-product of the filter. We estimate all model parameters using the maximum likelihood method with numerical optimization.

In order to forecast the regime dynamics, we rely on the predicted probabilities $\xi_{t+1|t,k}$ to classify the next state $\hat{S}_{t+1,k}$ and to compute the predicted densities $\eta_{t+1,k}$ and $f(x_{t+1,k}|\Omega_t)$.

3.2 Step 2: Flexible Regime Probabilities

First introduced by Meucci (2010), the flexible probability approach allows researchers to re-weight historical observations to match specific scenarios of a target variable. This scenario could be determined by a time weighting, the current conditions, or a subjective market view. The flexible probability approach can be considered as a non-parametric way of modeling the distribution of future returns, offering great flexibility at low computational cost. This method is also closely related to weighted kernel estimation techniques (Gisbert 2003) and approaches that evaluate the relevance of observations based on a reference point (Czasonis et al. 2020).[3] According to Meucci

3. They measure the relevance of observations as the sum of similarity (Mahalanobis distance to the most recent data point) and informativeness (Mahalanobis distance to expected normal conditions). They use this measure as a decision criterion to select the relevant subset for prediction. Their so-called

(2010) the empirical distribution of i.i.d. invariants z_t can be generalized with the help of a flexible probability measure p_t:

$$h_z \equiv \sum_{t=1}^{\tau} p_t \delta^{z_t}(z) \tag{10}$$

where $\delta^{z_t}(z)$ is the Dirac delta centered at the generic point z_t.[4]

This general case nests the historical probability density function when we weight the observations equally ($p_t = 1/\tau$). We are flexible in our choice of p_t as long as we re-normalize the probabilities so that they sum to one ($\sum_{t=1}^{\tau} p_t = 1$).[5]

With this framework, we can incorporate *views on economic regimes* to account for the non-uniform weighting of historical observations. We let the flexible probabilities $p_t(v,k)$ be a function of a given economic variable k and its filtered outcomes v. For the latter, we explicitly rely on the recursive Hamilton filter, as given in Eq. (9). We always assume a proportional relationship (denoted by $\propto$) between the flexible probabilities and the predictive regime dynamics obtained from four different specifications ($v \in \{A,B,C,D\}$):[6]

(A) Binary State Classification

$$p_t(A,k) \propto \begin{cases} 1 & \text{if } \hat{S}_{t+1,k} = j^* \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

(B) Predicted State Probabilities

$$p_t(B,k) \propto \mathbb{P}(S_{t+1,k} = j^* | \Omega_t) \tag{12}$$

(C) Conditional State Density

$$p_t(C,k) \propto f(x_{t+1,k} | S_{t+1,k} = j^*, \Omega_t) \tag{13}$$

(D) Mixture Density

$$p_t(D,k) \propto f(x_{t+1,k} | \Omega_t) \tag{14}$$

partial sample regression converges to ordinary linear regression as the subset size approaches to the total sample size.

4. The Dirac delta is a generalized function whose probability mass is concentrated at the generic point and zero elsewhere. We refer to Meucci (2005) for more details on this topic.

5. We use the terms 'flexible probabilities' and 'observation weights' as synonyms in the rest of the paper.

6. We also test the filtered and the smoothed probabilities as weights. Such a choice would change the perspective from the predictive to the current regime dynamics. The results are not reported, but are qualitatively similar and available upon request.

where j^* is the future state with the highest predicted probability ($j^* = \text{argmax}\{\mathbb{P}(S_{t+1,k} = j|\Omega_{t-1}) : j = 0,1\}$) and $\hat{S}_{t+1,k}$ the predicted state classifier which is either 0 or 1. The four approaches differ significantly in their selectivity and, thus, in how much they utilize information from the estimation window. Approach (A) considers only observations belonging to the future regime and weights them equally, while (B) increases the information content by relaxing the binary restriction. The approaches (C) and (D) can be interpreted as kernel smoothing, using almost all observations. Still, they differ in their distribution assumption (conditional normal or mixture of two normal distributions) and thus in their observation-dependent impact.

Since the generalized distribution function for a given historical path of x_k is entirely determined by the set $\{x_k, p(v,k)\}$, all four specifications require a high degree of confidence in the regime measurement of the economic variable. However, both the prediction and the real-time identification of regimes are usually associated with a high amount of uncertainty. Therefore, it is often advisable to adopt a Bayesian perspective (Connor 1997; Pettenuzzo et al. 2014). Meucci (2008) recommends a method to incorporate fully flexible views into a predictive distribution. He suggests the use of *entropy pooling (EP)* to minimize the distance between two distributions using the investor's views as constraints. More formally, he measures the distance between distributions using the relative entropy (aka Kullback-Leibler divergence) $d(p,p^0) = \sum_{t=1}^{\tau} p_t \ln(p_t/p_t^0)$, where p_t represents the posterior probability to be found and p_t^0 a prior distribution.

We treat the weights from (A)–(D) as a view and blend them with the historical distribution function ($p_t^0 = 1/\tau$) as prior. Following Meucci (2008), we apply entropy pooling to shift the probability mass of the prior distribution to a new probability vector p^* that satisfies the first two moments of our regime-based view while being as close as possible to the historical distribution.[7] We rely on numerical optimization to find the posterior probabilities p^*:

$$\begin{aligned} p^*(v,k) &= \underset{p}{\text{argmin}} \sum_{t=1}^{\tau} p_t \ln(p_t \tau) \\ &\text{s.t.} \\ \sum_{t=1}^{\tau} p_t x_{t,k} &= \sum_{t=1}^{\tau} p_t(v,k)\hat{x}_{t,k} \\ \sum_{t=1}^{\tau} p_t x_{t,k}^2 &= \sum_{t=1}^{\tau} p_t(v,k)\hat{x}_{t,k}^2 \end{aligned} \tag{15}$$

7. Since our analysis refers to the full marginal distribution of our regime variables, constraints on higher moments (such as skewness and kurtosis) are also possible. However, we assume that the first two moments are a reasonable approximation. We leave the question of the empirical advantages of including higher moments to future research.

with linear constraints on the first two moments, which are based on the regime-implied distribution $p_t(v,k)$. $\hat{x}_{t,k}$ represent simulated values of the regime variable corresponding to our view. The main advantage of entropy pooling is that it smoothly incorporates the views by changing the prior distribution as "minimally" as possible, significantly reducing forecast uncertainty and increasing robustness. For each new observation, we re-estimate the Markov-switching model to account for changing regime dynamics and update the weighting procedures.

3.3 Step 3: Forecasting with Earnings and Multiple Growth

We start by using the posterior *flexible regime probabilities* as weighting parameters for the return components in each t. Using the historical realizations of earnings and multiple growth, we compute the component forecasts as weighted averages of the flexible regime probabilities $p^*(v,k)$ and the realizations of earnings and multiple growth:

$$\hat{\mu}^{ge}_{\tau+1}(v,k) = \sum_{s=2}^{\tau} p^*_{s-1}(v,k) ge_s \tag{16}$$

$$\hat{\mu}^{gm}_{\tau+1}(v,k) = \sum_{s=2}^{\tau} p^*_{s-1}(v,k) gm_s \tag{17}$$

with the conditional expectation operators $\hat{\mu}^{ge}_{\tau+1}(v,k) = E_\tau[ge_{\tau+1}|\hat{S}_{\tau+1,k} = j^*, v, \Omega_\tau]$ and $\hat{\mu}^{gm}_{\tau+1}(v,k) = E_\tau[gm_{\tau+1}|\hat{S}_{\tau+1,k} = j^*, v, \Omega_\tau]$. It is important to note that we are focusing on predictive relationships since we are always using the subsequent realization of earnings and multiple growth.

In the final step, the forecasts of the return components are combined with the current dividend-price ratio (dp_t), and the logarithmic risk-free rate ($rf_{t+1} = log(1 + Rf_{t+1})$) is subtracted. The latter is known in advance, as in Rapach and Zhou (2013).

Specification (1): In the baseline case, we predict both return components with the same state variable k:

$$\hat{\mu}^{(1)}_{\tau+1}(v,k) = dp_\tau + \hat{\mu}^{ge}_{\tau+1}(v,k) + \hat{\mu}^{gm}_{\tau+1}(v,k) - rf_{\tau+1} \tag{18}$$

However, it is also likely that the regime dynamics of the variable k only has a predictive content for either earnings or multiple growth. Therefore, based on the benchmark model of Ferreira and Santa-Clara (2011), we evaluate the return predictability by introducing two constraints:

Specification (2): We set the multiple growth to zero ($\hat{\mu}^{gm}_{\tau+1} = 0$) and model only the earnings growth with our *flexible regime approach*.

$$\hat{\mu}^{(2)}_{\tau+1}(v,k) = dp_\tau + \hat{\mu}^{ge}_{\tau+1}(v,k) - rf_{\tau+1} \tag{19}$$

Specification (3): As the second constraint, we adapt the assumption of Ferreira and Santa-Clara (2011) that earnings growth has only a low-frequency predictable component ($\hat{\mu}^{ge}_{\tau+1} = \bar{ge}^{20Y}_{\tau}$), but now we allow that multiple growth is regime-dependent.

$$\hat{\mu}^{(3)}_{\tau+1}(v,k) = dp_\tau + \bar{ge}^{20Y}_{\tau} + \hat{\mu}^{gm}_{\tau+1}(v,k) - rf_{\tau+1} \tag{20}$$

To evaluate the benefits of the SOP methodology, we also apply our three-step procedure directly to the aggregated returns rather than to their components.

Specification (4): We reject any component-specific predictability and compute a regime-weighted average of historical returns:

$$\hat{\mu}^{(4)}_{\tau+1}(v,k) = \left(\sum_{s=2}^{\tau} p^*_{s-1}(v,k) r_s\right) - rf_{\tau+1} \tag{21}$$

As benchmark model for our *flexible regime approach*, we use the *FSC* model of Ferreira and Santa-Clara (2011) as in Eq. (6) and the historical average *EW*. In its unrestricted version, the *EW* is equal to the prior distribution of the return components ($p_t = 1/\tau$). We also apply the same restrictions as in specifications (2) and (3).

To further investigate the variable-specific effect on earnings and multiple growth, we combine regimes with different variables. For this purpose, we provide two extensions.

Extension 1: For each choice of v, we test all possible pairs for the ten state variables, where k (l) represents the regime-filtered variable for predicting earnings growth (multiple growth) so that the baseline equation is extended as follows:

$$\hat{\mu}_{\tau+1}(v,k,l) = dp_\tau + \hat{\mu}^{ge}_{\tau+1}(v,k) + \hat{\mu}^{gm}_{\tau+1}(v,l) - rf_{\tau+1} \tag{22}$$

This specification nests the baseline case of (1) for $k = l$ and creates 90 new combinations for $k \neq l$.

Extension 2: Instead of using just one variable for each component, we consider the component forecasts of all variables. The pooled forecast is the weighted average of all variable-specific forecasts.

$$\hat{\mu}^{ge}_{\tau+1,pool}(v) = \sum_{k=1}^{10} \omega_{\tau+1,k} \hat{\mu}^{ge}_{\tau+1}(v,k) \tag{23}$$

$$\hat{\mu}^{gm}_{\tau+1,pool}(v) = \sum_{k=1}^{10} \omega_{\tau+1,k} \hat{\mu}^{gm}_{\tau+1}(v,k) \tag{24}$$

To determine the forecast weights $\omega_{\tau+1,k}$, we apply to both components the average forecast (AVE) with equal weights ($\omega_k = 1/10$) as well as an endogenous weighting scheme, as proposed by Rapach et al. (2010), which adjusts the equal weights based on their past forecast performance. The weights of the Discount Mean Squared Forecast Error (DMSFE) combination are calculated as follows:

$$\omega_{\tau+1,k} = \Phi^{-1}_{\tau+1,k} / \sum_{k=1}^{10} \Phi^{-1}_{\tau+1,k}$$
$$\text{with} \tag{25}$$
$$\Phi_{\tau+1,k} = \sum_{s=m}^{\tau-1} \theta^{\tau-1-s} (y_{s+1} - \hat{y}_{s+1,k})^2$$

where y is a placeholder for GE and GM, $m+1$ is the start of the out-of-sample period, and θ is the discount factor. To account for the momentum of predictability (Dai and Zhou 2020), we give more weight to recent forecasting performance by setting $\theta = 0.9$. The combined forecasts for GE and GM are then used as input for the SOP prediction as in Eq. (22).

3.4 Evaluation Measures

We evaluate the **statistical accuracy** of the one-month ahead excess return forecast $\hat{\mu}_{k,t+1}$ corresponding to the regime variable k with the mean squared forecast error (MSFE).

$$MSFE(\hat{\mu}_{t+1,k}) = \frac{1}{T_1} \sum_{t=T_0}^{T_1} (\mu_{t+1} - \hat{\mu}_{t+1,k})^2 \tag{26}$$

Let T_0 be the in-sample period and T_1 the out-of-sample period. We test the predictability of excess stock returns with the out-of-sample R^2 from Campbell and Thompson (2008), which compares the relative value added in MSFE of our forecasts to the historical average ($MSFE(\bar{\mu}_{t+1})$).

$$R^2_{OS}(\hat{\mu}_{k,t+1}) = 1 - \frac{MSFE(\hat{\mu}_{t+1,k})}{MSFE(\bar{\mu}_{t+1})} \tag{27}$$

A positive R^2_{OS} indicates a lower MSFE of the forecast k and, thus, an improvement of the predictability relative to the historical benchmark. Most empirical studies show that it is very challenging to beat the historical average (Welch and Goyal 2008). Since the predictability can vary significantly over time, we calculate the cumulative differences in the squared forecast error (CDSFE), defined as:

$$CDSFE(\hat{\mu}_{t+1,k}) = \sum_{t=T_0}^{T_1} (\mu_{t+1} - \bar{\mu}_{t+1})^2 - \sum_{t=T_0}^{T_1} (\mu_{t+1} - \hat{\mu}_{t+1,k})^2 \tag{28}$$

This measure is very informative for detecting time-varying predictability. An increase (decrease) in the CDSFE signals a lower (higher) prediction of the regime-based forecasts relative to the historical average for that period.

To determine the significance of R^2_{OS}, we test the hypothesis $H_0 : R^2_{OS} \leq 0$ against $H_1 : R^2_{OS} > 0$. For nested forecasts, it is common to rely on the adjusted MSFE from Clark and West (2007), which is defined as:

$$\hat{e}_{t+1} = (\mu_{t+1} - \bar{\mu}_{t+1})^2 - [(\mu_{t+1} - \hat{\mu}_{t+1,k})^2 - (\bar{\mu}_{t+1} - \hat{\mu}_{t+1,k})^2] \tag{29}$$

Using the sample mean $\bar{e} = 1/T_1 \sum_{t=T_0}^{T_1} \hat{e}_{t+1}$ and the sample variance $\hat{V} = 1/(T_1 - 1)(\hat{e}_{t+1} - \bar{e})$, the CW test statistic ($\sqrt{T_1}\bar{e}/\hat{V}$) is approximately standard normal distributed. Therefore, we can directly apply the standard critical values for a one-sided hypothesis test.

In order to verify the **economic benefits** of an investor, we translate the predictions of the equity risk premium into a trading strategy that can switch between stocks and cash. We assume a risk-averse agent with mean-variance preferences. By solving the standard expected utility maximization problem, we obtain the optimal equity market weight according to Merton (1969):

$$w_{t,k} = \frac{\hat{\mu}_{t+1,k}}{\gamma \hat{\sigma}^2_{t+1}} \tag{30}$$

We assume a moderate (relative) risk aversion ($\gamma = 3$), and we use the historical five-year variance as an estimate of the expected variance $\hat{\sigma}^2_{t+1}$. The portfolio return $RP_{t+1,k}$ of the trading strategy k is then:

$$RP_{t+1,k} = w_{t,k} R_{t+1} + Rf_{t+1} \tag{31}$$

To design a walk-forward backtest of the timing strategies that is consistent with the constraints of practitioners, we do not allow short selling and leverage, so that the equity weight varies only between 0% and 100%. In addition, we account for proportional transaction costs by assuming a roundtrip fee of 50 basis points.[8]

To evaluate the trading strategies, we use the certainty equivalent return (CER) as the average utility gain for a mean-variance investor:

$$CER = \mu_{RP(k)} - \frac{1}{2}\gamma \sigma^2_{RP(k)} \tag{32}$$

$\mu_{RP(k)}$ ($\sigma^2_{RP(k)}$) is the average return (variance) of the strategy for the out-of-sample period. To assess the value added of the timing strategies, we calculate the CER gain

8. As a part of our robustness testing, we check the sensitivity of our results to variations in risk aversion ($\gamma = 1, 3, 5$), to the possibility of partial short selling ($w_{min} = -50\%$), and to partial leveraging ($w_{max} = 150\%$).

as the difference between the strategy's CER and the CER of the historical average (henceforth Δ^k_{CER}). The Δ^k_{CER} has the practical interpretation as annual management fee an investor is willing to pay to participate in the forecast-based strategy (Rapach and Zhou 2013). To evaluate the investor's utility over time, we compute the rolling five-year certainty equivalent gain ($\Delta^{k,5Y}_{t,CER}$), which corresponds to a typical investment horizon of a medium-term investor.

4 Data

We use the monthly dataset of Welch and Goyal (2008) to examine the impact of regimes on the predictability of return components.[9] The sample begins in December 1927 and ends in December 2022. We use continuously compounded stock market returns including dividends of the Standard & Poor's (S&P) 500 Index in excess of the risk-free rate. Dividends and earnings are calculated as trailing 12-month sum. The corresponding return components are the dividend-price ratio (DP), earnings growth (GE), and the price-earnings multiple growth (GE), all expressed in logs.[10]

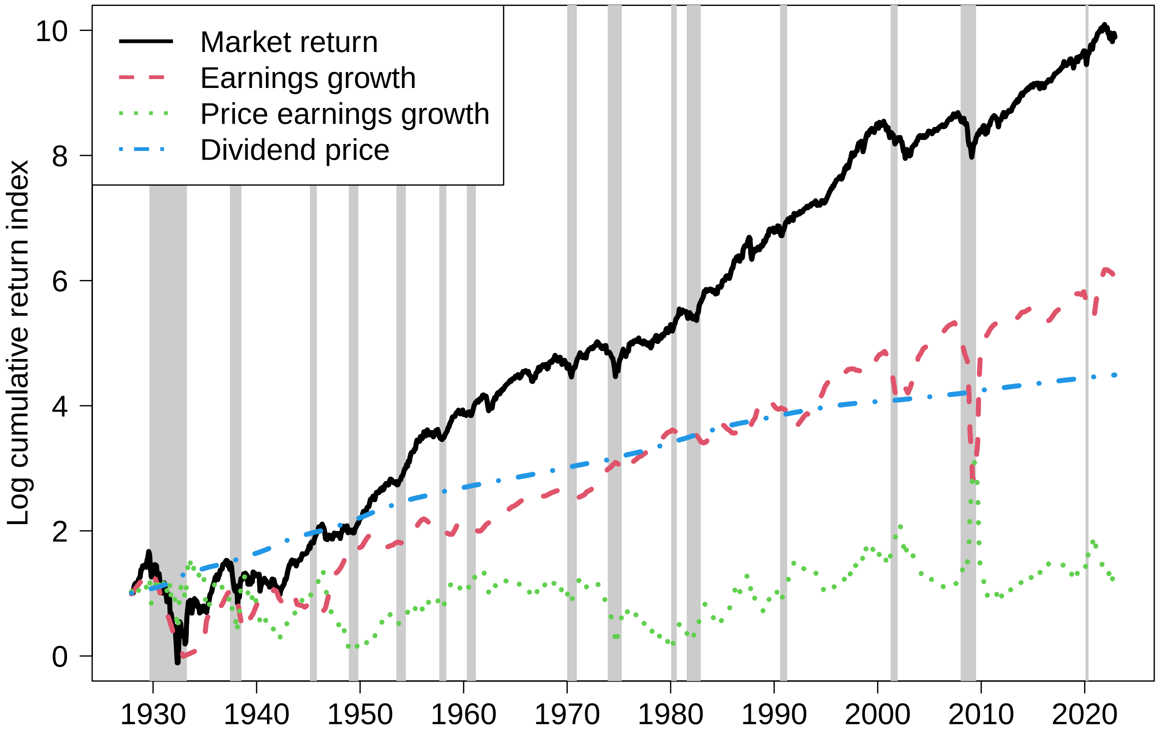
Table A1 presents the summary statistics. Over the 95-year period, the annualized equity risk premium is 6.1%, with a volatility of 18.8%. The largest variability in returns results from the multiple and earnings growth, with annual standard deviations of 23.1% and 14.9%, respectively. Multiple growth exhibits only slight persistence, while the strong autocorrelation of earnings growth can be attributed to overlapping observations due to the measurement of earnings. The dividend-price ratio and the risk-free rate demonstrate similar descriptive properties and are highly persistent, with an AR(1) coefficient of 0.98.

Figure 2 illustrates the trajectory of the stock market and its components. The cumulative series of the dividend-price ratio is characterized by a smooth and steady increase, while earnings growth exhibits an upward trend with cyclical dips associated with NBER recessions. The data indicate that multiple growth does not demonstrate a distinct trend over the entire period. Rather, it frequently exhibits peaks at the end of recessions.

9. We thank Amil Goyal for providing and updating this valuable data source (https://sites.google.com/view/agoyal145, accessed on 10/26/2023). If no separate source is indicated, the time series were taken from this data set.

10. For reasons of clarity, we deviate from the notation in Section 2 and always use the variable abbreviations in capital letters.

Figure 2: Stock Market and its Component



Notes: Cumulative stock market return and the corresponding components of the S&P 500 according to the decomposition of Eq. (5). The period is from 1927/12 to 2022/12. Gray shaded areas indicate NBER recessions.

To infer the impact of economic regimes on stock market predictability, we first examine the excess return, ERP, and the return components GE, GM, DP. This allows to analyze the endogenous dynamics.

1. **Equity risk premium** (ERP) is defined as the monthly excess return of the aggregate stock market over the one-month risk-free rate.

2. **Earnings growth** (GE) is denoted as the monthly log change of the trailing 12-month sum of the S&P 500's earnings per share.[11]

3. **Multiple growth** (GM) is expressed as a monthly log change in the price-earnings ratio of the S&P 500.

4. **Dividend-price ratio** (DP) is defined as the logarithm of the ratio between the trailing 12-month sum of the S&P 500 dividends and the current price plus one.

11. Since monthly changes in the 12-month sum of earnings produces very short-lasting regimes, we slightly deviate from this variable definition and use annual log changes as input for the Markov-switching model.

To capture exogenous data sources, we also consider six additional macro-financial variables ($RVOL$, TMS, DFY, TBL, IP, INF), which are well-documented state variables that appear to be promising for predicting returns or their components, see for instance Rapach and Zhou (2013) and references therein.

5. **Realized stock market volatility** ($RVOL$) is estimated using the method of Mele (2007): $RVOL_t = \sqrt{\frac{\pi}{2}}\sqrt{12}\frac{1}{12}\sum_{i=1}^{12}|r_{t+1-i}|$.
6. **Term spread** (TMS) is defined as the yield differential between a 10-year Treasury bond and a 3-month treasury bill rate.
7. **Default spread** (DFY) is denoted as the yield differential between Moody´s BAA and AAA-rated corporate bonds.
8. **Short-term interest rate** (TBL) is represented by the 3-month Treasury bill rate.
9. **Industrial production** (IP) is measured as the month-over-month growth rate in Industrial Production.[12]
10. **Inflation** (INF) is expressed as the month-over-month growth rate in the Consumer Price Index (All Items in U.S. City Average)

Given that the level of many of these variables is already reflected in the current price, our focus is on the changes in the state variables. Accordingly, the variables are either already measured as percentage changes (GE, GM, ERP) or have been transformed to first differences (DP, $RVOL$, TMS, DFY, TBL).[13] The descriptive statistics clearly show non-normal patterns in the data, highlighted by skewed and often leptokurtic unconditional distributions. In addition, the serial correlation inherent in the level of many exogenous series can be reduced by the data transformation.[14] From the correlation matrix, it can be seen that the contemporaneous relationship between many state variables and earnings or multiple growth is stronger than for the excess return ($RVOL$, TBL, IP, INF). This provides some first descriptive evidence that the SOP approach can improve the predictability.

12. We obtain this data from FRED database (https://fred.stlouisfed.org/series/INDPRO, Accessed on 10/26/2023)

13. We leave IP and INF unchanged from the definition above, but we lag both macroeconomic variables by one month to account for the publication lag.

14. The serial correlation inherent in the level of many exogenous series can be reduced by the data transformation. The remaining autocorrelation is controlled for by using an AR(1) coefficient in the Markov-switching model.

5 In-Sample Results

We start with the in-sample (IS) estimation results of the regime models (Table 1). It becomes evident that the dispersion parameter is the most effective discriminator for the regime-dependent distributions. The turbulent (calm) regime is associated with negative (positive) changes for ERP, GM, TBL and an increase (decrease) for TMS. For some processes, the state-dependent mean is comparable between states ($RVOL$, IP, INF). All economic variables show the presence of (highly) persistent regimes with an expected duration between 9 (DFY) and 103 months (INF). The less volatile regime is typically more persistent than the more volatile one. Despite the transformation of the economic variables, some time dependence in the mean process persists, which is addressed by incorporating the state-dependent autoregressive term ϕ (insignificant only for ERP and changes in $RVOL$).

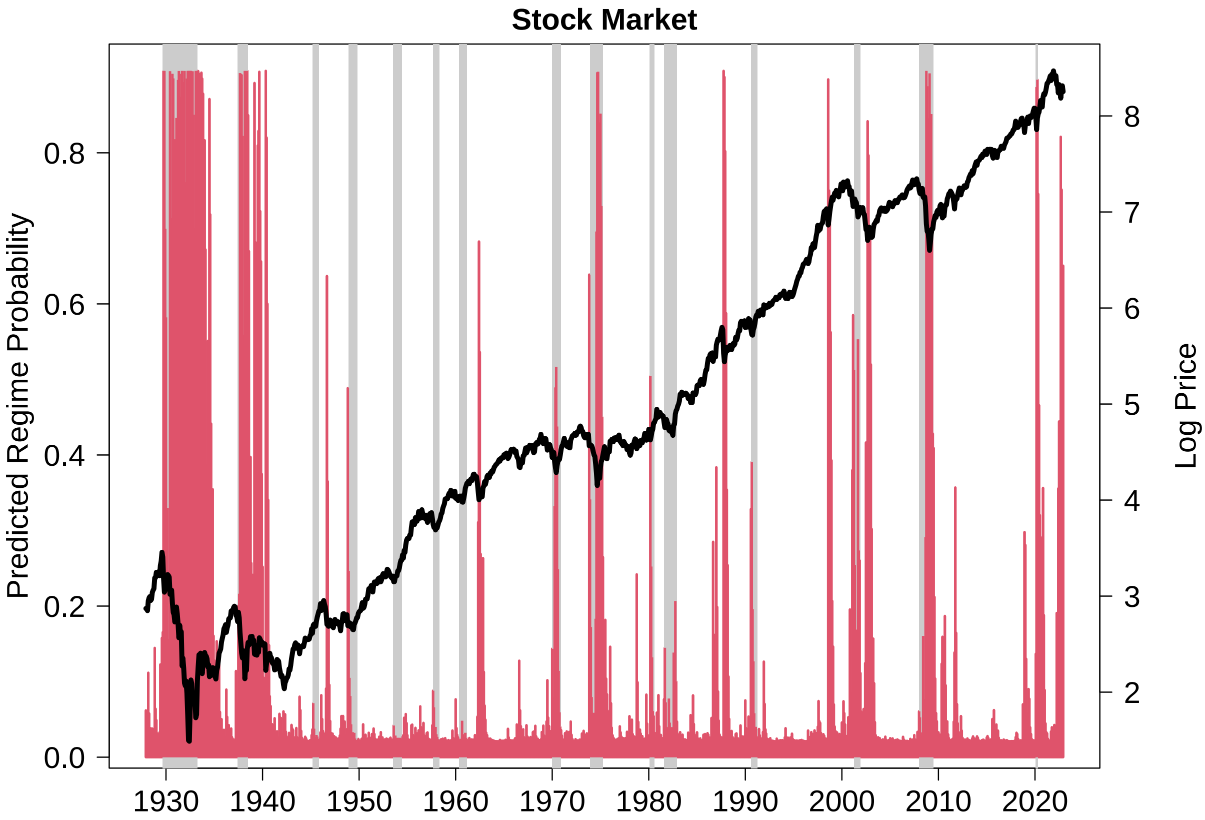
Table 1: Estimation of Regime Models

Parameters of univariate MSAR(1) models								
Economic Variable	$\hat{q}_{00}$	$\hat{q}_{11}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\sigma}_0$	$\hat{\sigma}_1$
ERP	0.985	0.908	0.892***	−1.71*	−0.011	0.088	3.792	10.881
GM	0.989	0.937	0.182	−0.641	0.116***	0.341***	4.096	13.279
GE	0.984	0.905	0.001	0.003	0.978***	0.979***	0.020	0.170
DP	0.986	0.908	−0.001	0.002	0.057	0.211***	0.011	0.066
RVOL	0.990	0.937	0.009	0.008	−0.039	−0.023	1.433	4.409
TMS	0.928	0.951	−0.013***	0.008	−0.005	0.106***	0.091	0.458
DFY	0.973	0.886	−0.002	0.011	0.263***	0.207***	0.049	0.326
TBL	0.958	0.938	0.011***	−0.012	0.231***	0.328***	0.074	0.527
IP	0.976	0.920	0.174***	0.104	0.346***	0.531***	0.634	3.003
INF	0.990	0.979	0.139***	0.126***	0.439***	0.507***	0.264	0.735

Notes: Table shows the estimation results of the two-state Markov-switching autoregressive model, MSAR(1), using the full-sample (1927/12 to 2022/12). The estimation is done by numerical optimization. State 1 (0) always indicates the volatile (calm) regime. ERP: excess stock return, GM: price-earnings ratio growth, GE: earnings growth, DP: change in log dividend-price ratio, $RVOL$: change in realized 12M volatility, TMS: change in 10Y–3M term spread, DFY: change in BAA–AAA default spread, TBL: change in 3M Treasury bill rate, IP: industrial production growth, INF: CPI inflation. The Data are transformed before estimation, as described in Section 4. ***/**/* denotes significance at the 1%/5%/10% level.

In **Step 1** of the analysis, the regime dynamics of the excess stock return are filtered by an MSAR(1) model, with the estimated parameters from Table 1. Figure 3 illustrates the evolution of the log price of the S&P 500, together with the corresponding predicted probabilities and the NBER recession periods. It has been observed that there is a high correlation between economic crises and financial market turmoil. A surge in regime probabilities is frequently observed in the period preceding or during economic recessions. Typically, the stock market begins to recover around the time of an economic turning point. However, it should be noted that not all economic recessions are associated with a decline in stock prices and vice versa.[15]

Figure 3: Economic Regimes of Excess Return (Step 1)



Notes: Figure displays the predicted regime probability (B) of the monthly excess stock returns. The predicted probabilities of the turbulent regime is depicted in red (left axis). For reasons of illustration, we plot the log stock prices over time. The in-sample period runs from 1927/12 to 2022/12. Gray shaded areas indicate NBER recessions.

The subsequent step is to compute the flexible probabilities in accordance with our regime view (**Step** 2). The predicted state probabilities indicate a continuation of the turbulent regime, which suggests a weighting according to the upper left panel of Figure 4. Given the uncertainty associated with our view, our objective is to identify a weighting vector that is consistent with the fundamental characteristics of our view (represented by the first two moments) while remaining as close as possible to the historical distribution. Consequently, entropy pooling is employed to derive a posterior probability vector p^* (illustrated in the upper right panel). By combining the flexible probabilities with historical scenarios (resampled by bootstrapping), we can compare the historical distribution (depicted in red) with both the view-based density (lower

15. Figure A1 in the Appendix A displays the ex post regime identification for all ten economic regime variables.

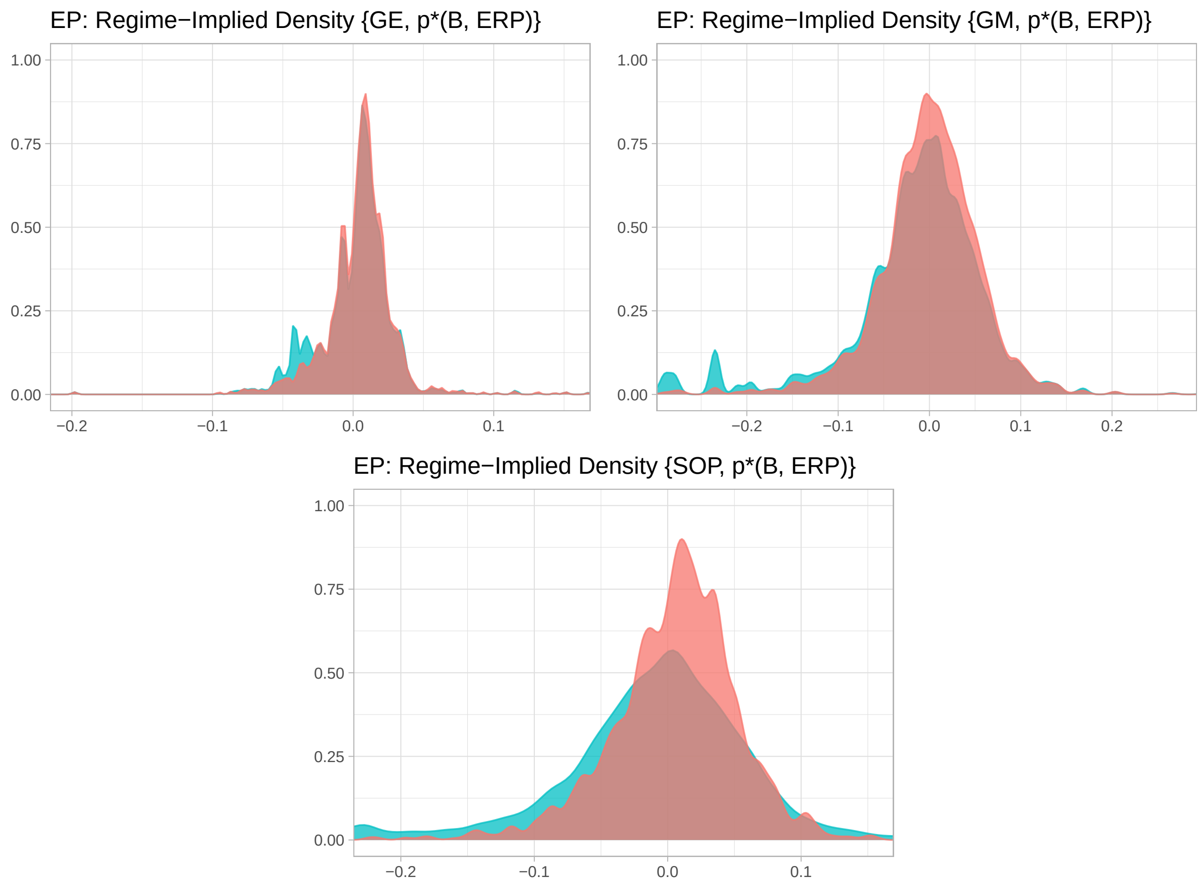
left panel of Figure 4 in green) and with the posterior density (lower right panel in green). It can, therefore, be concluded that entropy pooling compresses the historical distribution and assigns greater probability mass to negative outcomes. Relative to the original view-based density, these observations occur to an even greater extent.

Figure 4: Flexible Probabilities and Entropy Pooling (Step 2)

Notes: Figure shows the flexible probabilities (upper panel) and regime-implied densities (lower panel in green) for the excess stock return, both relative to the historical prior with equal weights (red line and red density, respectively). The flexible probabilities are determined by the predicted turbulent state probabilities (B) as of 2022/12. In the left plots, the flexible probabilities are fully captured by the state probabilities, while in the right plots the entropy pooled weights are used. The density simulations are generated by bootstrapping with 10,000 draws. The interval ranges from 0.5% to 99.5% quantile.

In the final **Step 3**, we retain the posterior probability vector and bind it with the realizations of the return components *GE* and *GM*. The result is the regime-implied probability distribution (in green) depicted in Figure 5. In comparison to the historical distribution, it is evident that there is a greater concentration towards the left side, with a corresponding reduction in probability mass in the middle. This accounts for the characteristics of the turbulent market view, which has a higher likelihood in 2022M12. Finally, we combine the distributions of *GE*, and *GM* with the current dividend-price ratio dp and subtract the risk-free rate to get the flexible regime distribution of the SOP forecasts. The expected value of this distribution is then the prediction for the next excess return.

Figure 5: Regime-implied Densities for the SOP Predictions (Step 3)



Notes: Figure shows the regime-implied densities (green) and its historical density (red) for earnings growth (*GE*), multiple growth (*GM*), and the combined sum-of-the-parts (SOP) predictive distribution of the unrestricted specification (1). The flexible probabilities are determined by the predicted turbulent state probabilities (B) as of 2022/12. The density simulations are generated by bootstrapping with 10,000 draws. The interval ranges from 0.5% to 99.5% quantile.

6 Out-of-Sample Results

Starting with an initial estimation window of 20 years, we employ a recursive approach that updates the regime filtering and probability weighting every month to forecast the excess stock return with a one-month horizon. The total out-of-sample (OS) period spans approximately 75 years, from December 1947 to December 2022.

6.1 Individual Economic Regimes

The **statistical accuracy** of our return forecasts is presented in Table 2. We can demonstrate that our *flexible regime* approach is an effective method for predicting stock returns. Irrespective of the economic variable, weighting method, or restriction under examination, the R^2_{OS} is found to be significantly positive in most cases (only three exceptions exist). Compared to the historical average, the benefit in R^2_{OS} is frequently around 1% and reaches its maximum at 1.69%. The predictability is generally higher in NBER recessions (average R^2_{OS} of 1.9%) than in expansions (average R^2_{OS} of 0.6%), which is consistent with the extensive evidence in the literature (Henkel et al. 2011; Rapach and Zhou 2013).

Applying our methodology directly to returns, rather than to return components, leads to worse results, as specification (4) illustrates. In more than 90% of all cases, the SOP method yields higher R^2_{OS} than the direct return forecasts (for the unrestricted as well as all restricted cases). Even if the forecast quality itself is not inadequate, no economic regime exists where the direct approach dominates all three SOP specifications. In light of these results, we identify the SOP approach as an important element in the forecasting process. Consequently, subsequent attention will be devoted to specifications (1) to (3).

The benchmark approaches also offer high predictability of returns with R^2_{OS} between 0.81% (unrestricted EW) and 1.27% (FSC). Compared to these references, however, a substantial number of models provide better forecasts (highlighted in bold entries). More precisely, the number of superior forecasting models ranges from six using specifications (1) and (3) to 19 in the specification (2). In this context, the value-added of the density-weighting approaches (C) and (D) is particularly evident. Therefore, replacing the moving average earnings estimate with a regime-weighted one and setting the expected multiple growth to zero is an advantageous strategy. This finding is not surprising given the volatile nature of multiple changes and the well-established relationship between corporate earnings and the business cycle (Longstaff and Piazzesi 2004).

Table 2: Statistical Accuracy (R^2_{OS} in %)

Variable	Overall (1)	Overall (2)	Overall (3)	Overall (4)	Expansion (1)	Expansion (2)	Expansion (3)	Expansion (4)	Recession (1)	Recession (2)	Recession (3)	Recession (4)
Panel 1: Benchmark												
FSC	-	1.27***		-	-	0.83***		-	-	2.52***		-
EW	0.81***	1.08***	1.04***	-	0.3*	0.57**	0.6**	-	2.25**	$\mathbf{2.56}^{***}$	2.3***	-
Panel 2: Binary State Classification (A)												
ERP	$\mathbf{1.55}^{***}$	0.81***	$\mathbf{1.69}^{***}$	$\mathbf{1.27}^{***}$	$\mathbf{1.15}^{***}$	0.63**	$\mathbf{1.03}^{**}$	$\mathbf{1.04}^{***}$	$\mathbf{2.7}^{**}$	1.31*	$\mathbf{3.58}^{**}$	1.93*
GM	$\mathbf{1.33}^{***}$	1.13***	$\mathbf{1.28}^{***}$	1.08***	0.75**	$\mathbf{1.01}^{***}$	0.38*	0.66**	$\mathbf{3}^{***}$	1.49**	$\mathbf{3.84}^{***}$	2.26**
GE	0.92***	$\mathbf{1.32}^{***}$	0.86***	0.69***	0.43**	0.81***	0.43**	0.41**	2.32***	$\mathbf{2.79}^{***}$	2.09***	1.52*
DP	0.87***	1.06***	1.1***	0.68**	0.34*	0.52**	0.65**	0.37**	2.38**	$\mathbf{2.61}^{***}$	2.39***	1.56*
RVOL	$\mathbf{1.31}^{***}$	1.12***	$\mathbf{1.39}^{***}$	0.97***	$\mathbf{1.07}^{***}$	0.68**	$\mathbf{1.11}^{***}$	$\mathbf{0.9}^{***}$	1.98**	2.39***	2.17***	1.16
TMS	0.62***	0.92***	0.35**	0.28**	0.32*	0.38**	0.08*	0.16*	1.48**	2.47***	1.14**	0.61
DFY	−0.01	$\mathbf{1.39}^{***}$	−0.04	−0.22	0.27**	$\mathbf{1.31}^{***}$	−0.22	0.22**	−0.82	1.62**	0.47	−1.49
TBL	1.06***	$\mathbf{1.4}^{***}$	0.78***	0.79***	$\mathbf{1.04}^{***}$	$\mathbf{1.23}^{***}$	0.44*	$\mathbf{0.95}^{***}$	1.1*	1.9**	1.74**	0.32
IP	0.71***	1.19***	0.75***	0.51**	0.22*	$\mathbf{0.94}^{***}$	0.04	0.2*	2.11**	1.89**	$\mathbf{2.77}^{***}$	1.39*
INF	0.96***	1.11***	1.08***	0.72***	0.46**	0.68**	0.54**	0.39**	2.41**	2.36***	$\mathbf{2.63}^{***}$	1.66*
Panel 3: Predicted Probabilities (B)												
ERP	$\mathbf{1.47}^{***}$	0.84***	$\mathbf{1.69}^{***}$	1.2***	$\mathbf{0.97}^{***}$	0.65**	$\mathbf{0.95}^{***}$	$\mathbf{0.88}^{***}$	$\mathbf{2.88}^{**}$	1.37*	$\mathbf{3.81}^{***}$	2.12**
GM	1.21***	1.01***	$\mathbf{1.32}^{***}$	0.96***	0.72**	0.8**	0.6**	0.64**	2.6***	1.62**	$\mathbf{3.36}^{***}$	1.86**
GE	0.88***	1.23***	0.94***	0.67**	0.35*	0.64**	0.54**	0.34*	2.39***	$\mathbf{2.92}^{***}$	2.06***	1.58**
DP	0.85***	1.09***	1.06***	0.66**	0.32*	0.56**	0.6**	0.34*	2.37**	2.6***	2.39***	1.55*
RVOL	$\mathbf{1.49}^{***}$	1.14***	$\mathbf{1.57}^{***}$	1.18***	$\mathbf{1.15}^{***}$	0.72**	$\mathbf{1.18}^{***}$	$\mathbf{1.02}^{***}$	2.44***	2.34***	$\mathbf{2.68}^{***}$	1.63*
TMS	0.55**	0.97***	0.53**	0.28*	0.21*	0.45**	0.22*	0.15	1.54**	$\mathbf{2.46}^{***}$	1.39**	0.67
DFY	0.94***	$\mathbf{1.51}^{***}$	0.68**	0.68***	0.66**	$\mathbf{1.33}^{***}$	0.1	0.58**	1.75**	2.03***	$\mathbf{2.33}^{**}$	0.97
TBL	0.98***	$\mathbf{1.31}^{***}$	0.88***	0.76***	0.59**	$\mathbf{0.89}^{***}$	0.45**	0.58**	2.08**	2.5***	2.09***	1.28*
IP	0.69***	1.19***	0.82***	0.51**	0.18*	$\mathbf{0.85}^{***}$	0.21	0.19*	2.14**	2.16***	$\mathbf{2.57}^{***}$	1.4*
INF	0.95***	1.23***	1.01***	0.7**	0.45**	0.8***	0.49**	0.37**	2.39**	2.48***	$\mathbf{2.52}^{***}$	1.62*
Panel 4: Conditional State Density (C)												
ERP	1.07***	$\mathbf{1.44}^{***}$	0.72***	0.75***	0.76***	$\mathbf{1.28}^{***}$	0.1	0.6**	1.95**	1.87***	2.48***	1.17*
GM	1.02***	$\mathbf{1.3}^{***}$	0.95***	0.66***	0.77***	$\mathbf{1.04}^{***}$	0.5**	0.56**	1.75**	2.05***	2.21***	0.96
GE	0.79***	1.16***	0.91***	0.58**	0.32*	0.69**	0.46**	0.32*	2.14**	2.51***	2.19***	1.34*
DP	0.78***	1.11***	0.92***	0.57**	0.32*	0.69**	0.44**	0.31*	2.1**	2.29***	2.28***	1.32*
RVOL	0.97***	$\mathbf{1.37}^{***}$	0.83***	0.43**	0.79***	$\mathbf{1.07}^{***}$	0.49**	0.35**	1.47**	2.22***	1.8***	0.65
TMS	0.87***	1.19***	0.91***	0.63**	0.34**	$\mathbf{0.84}^{***}$	0.25*	0.29*	2.39***	2.16***	$\mathbf{2.8}^{***}$	1.59**
DFY	0.76***	$\mathbf{1.52}^{***}$	0.4**	0.45**	0.71***	$\mathbf{1.37}^{***}$	0.01	0.56***	0.92*	1.94***	1.51**	0.15
TBL	1.16***	$\mathbf{1.55}^{***}$	0.81***	0.86***	$\mathbf{0.9}^{***}$	$\mathbf{1.38}^{***}$	0.27*	0.79***	1.9**	2.03**	2.37***	1.07
IP	0.71***	$\mathbf{1.38}^{***}$	0.53**	0.48**	0.26*	$\mathbf{1.14}^{***}$	−0.15	0.2*	2.01**	2.07***	2.44**	1.28*
INF	0.96***	1.26***	1***	0.7***	0.55**	$\mathbf{0.89}^{***}$	0.52**	0.46**	2.11**	2.3***	2.37***	1.37*
Panel 5: Mixture Density (D)												
ERP	1.02***	$\mathbf{1.49}^{***}$	0.68**	0.69***	0.67**	$\mathbf{1.28}^{***}$	0.09	0.5**	2.03**	2.07***	2.38***	1.24*
GM	0.99***	$\mathbf{1.38}^{***}$	0.86***	0.67***	0.65**	$\mathbf{1.05}^{***}$	0.4*	0.49**	1.97**	2.32***	2.17***	1.16*
GE	0.67***	0.92***	1.02***	0.47**	0.2*	0.49**	0.52**	0.21*	1.99**	2.16***	$\mathbf{2.44}^{***}$	1.19*
DP	0.78***	1.04***	0.97***	0.58**	0.32*	0.61**	0.51**	0.31*	2.1**	2.29***	2.28***	1.32*
RVOL	1.05***	$\mathbf{1.38}^{***}$	0.9***	0.57**	$\mathbf{0.86}^{***}$	$\mathbf{1.04}^{***}$	0.59**	0.5**	1.61**	2.35***	1.81***	0.77
TMS	1.26***	$\mathbf{1.52}^{***}$	0.9***	1.01***	0.81***	$\mathbf{1.26}^{***}$	0.27*	0.77***	$\mathbf{2.54}^{***}$	2.27***	$\mathbf{2.71}^{***}$	1.7**
DFY	1.14***	$\mathbf{1.59}^{***}$	0.62**	0.82***	$\mathbf{0.83}^{***}$	$\mathbf{1.42}^{***}$	0.03	0.67***	2.04***	2.08***	2.28***	1.25*
TBL	$\mathbf{1.32}^{***}$	$\mathbf{1.51}^{***}$	1.01***	0.97***	$\mathbf{0.9}^{***}$	$\mathbf{1.18}^{***}$	0.48**	0.74***	2.51***	2.45***	$\mathbf{2.53}^{***}$	1.64**
IP	0.66**	$\mathbf{1.39}^{***}$	0.54**	0.46**	0.17*	$\mathbf{1.12}^{***}$	−0.13	0.15	2.08**	2.15***	2.46**	1.34*
INF	0.94***	$\mathbf{1.35}^{***}$	0.83***	0.65**	0.47**	$\mathbf{1.04}^{***}$	0.22*	0.34*	2.29***	2.23***	$\mathbf{2.55}^{***}$	1.54*

Notes: Table shows the out-of-sample R^2. In the unrestricted version (1) the SOP forecast consists of variable-specific regime forecasts for earnings and multiple growth. (2) assumes no multiple growth ($\hat{\mu}^{gm} = 0$) and uses only the *flexible regime* forecasts for earnings growth, (3) assumes that earnings growth follows a 20-year moving average ($\bar{ge}^{20Y}$) and the expected multiple growth are is predicted by the *flexible regime* forecasts. Finally (4) computes the return forecasts directly. Panel 1 displays the benchmark models FSC (Ferreira and Santa-Clara 2011), and the SOP forecasts assuming equal weights (EW). The Panels 2–5 show the accuracy with a weighting according to the different views. Superior results compared to FSC are highlighted in bold. ***/**/* denotes significance at the 1%/5%/10% level according to the CW statistics.

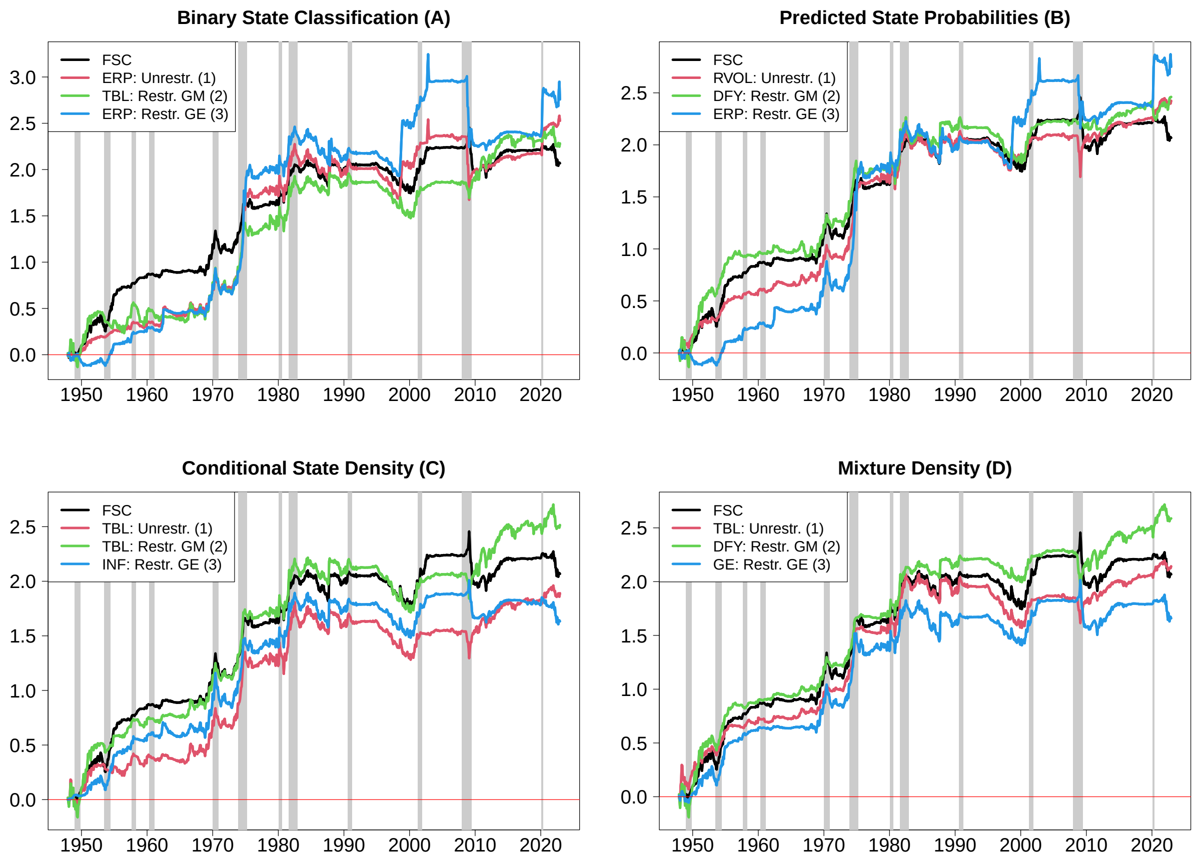
However, it is challenging to identify a superior weighting method, as the results are dependent on the chosen specification, the state of the business cycle, and the economic variable. While the binary state classification (A) and the predicted probabilities (B) prove their relative usefulness mainly during NBER recessions with R^2_{OS} of up to 3.84%, the forecasting power of the density-smoothed approaches (C) and (D) is stronger during economic expansions (R^2_{OS} of more than 1%). The regime dynamics of ERP and $RVOL$, derived from the binary state classification or the predicted probabilities, yield promising results for predicting multiple growth. Changes in the price-to-earnings multiple may be regarded as an approximation of the discount rate channel, which is more affected by higher uncertainty or risk aversion. Typically, these two factors experience a surge during periods of economic contraction and can be approximated using the filtered regimes of ERP and $RVOL$. Concerning the cash flow channel, earnings growth can be predicted with the dynamics of DFY and TBL, as the results of specification (2) indicate. Over all weighting approaches they serve as adequate state variables as they capture aggregated credit risk and monetary policy, which are typically correlated with cash flow news.

A review of the forecast performance over time reveals that the average statistical accuracy does not guarantee predictability over the entire test period. The occurrence of structural breaks and the presence of continuous learning investors are the primary factors that contribute to the emergence and decline of predictability. Figure 6 provides empirical evidence for this argumentation and shows which time periods are responsible for the observed predictability. The graphs show the best forecasting strategies according to the CDSFE for each weighting method and highlight the best strategies for the unconstrained and the two constrained cases. The benchmark approach of Ferreira and Santa-Clara (2011) is represented by the black line and indicates that much of its average predictability is concentrated from the beginning of the test period until the mid-80s. Over the past four decades, the forecast accuracy has moved more or less sideways. There are just two brief periods when FSC has outperformed the historical average: from 2001 to 2003 and in 2008.

In comparison, our regime-based forecasts have a similar trajectory up to the financial crisis of 2008, with a rising course mostly around NBER recessions. The unrestricted return forecasts in red slightly underperform the FSC benchmark at the beginning before they manage to catch up and beat the benchmark in 3 out of 4 cases by the end of 2022. Return forecasts that focus on modeling earnings growth (green lines) and using the weighting approaches (C) and (D) produce smooth and steadily increasing CDSFE curves. In particular, they have provided superior forecasts since the end of the financial crisis in 2009. The forecasting power of approaches that predict only multiple growth (blue lines) was rather low at the beginning of the test period but increased sharply in the wake of the oil crisis in the mid-1970s and remained so until

1982. Since the late 1990s, the flexible probability weighting scheme has played a more prominent role. From this time until the end of 2022, approaches using regime probabilities or binary classification as weighting have sudden spikes of outperformance in 1999, 2002, and 2020. However, all charts illustrate the difficulty of predicting returns in the decade from the late 1980s to the late 1990s, as documented by numerous studies (Welch and Goyal 2008; Dangl and Halling 2012; Hammerschmid and Lohre 2018). Our findings illustrate the benefits of economic regimes, especially in the recent past, and partly explain the success of research focusing on regime predictability over the last 25 years (Haase and Neuenkirch 2023, e.g.,).

Figure 6: Statistical Accuracy over time (CDSFE)



Notes: Figure shows the cumulative differences in squared forecast errors (CDSFE) of the best unrestricted (in red), *GM*-restricted (in green), and *GE*-restricted (in blue) regime-based forecasts. As a benchmark, the *FSC* forecast of Ferreira and Santa-Clara (2011) is added in black. A rising (falling) line indicates a lower (higher) prediction error than the historical average for a given point in time. Graph (A) indicates the weighting scheme using the binary state classification, (B) uses the predicted state probabilities, (C) the conditional state density and (D) the mixture state density. The test period runs from 1947/12 to 2022/12. Gray shaded areas indicate NBER recessions. All values are multiplied by 100.

We evaluate the **economic benefit** from a mean-variance investor's perspective using Merton's rule according to Eq. (30) (with no short-selling and no leverage) as an asset allocation model derived from the return forecasts. The corresponding trading strategies always consider proportional transaction costs of 50 basis points and a relative risk aversion of three to enable realistic backtesting. Table 3 shows the Δ_{CER} of our regime-based strategies. All models exceed the historical average, which has a certainty equivalent return of 7.2%. Across the test sample, an investor is willing to pay an annual management fee between 0.5 and 1.97% to participate in the individual regime-based strategies. As previously demonstrated in the statistical accuracy assessment, the investor benefit is more pronounced during periods of economic contraction (average Δ_{CER} of 6.7%) than during periods of economic growth (average Δ_{CER} of 0.4%). The best strategy uses the regime dynamics of the interest rate (TBL) with the binary state classification as the view generator and utilizes specification (2), which focuses on forecasting earnings growth.

Among all variables and all weighting approaches, the dynamics of GM and ERP provide the most robust results for the unrestricted specification (1), TBL and DFY for specification (2), and INF and $RVOL$ for specification (3). This demonstrates that the economic benefit of stock market forecasts is contingent upon the influence of the economic variable on the various return components. Some economic regimes are more conducive to approximating the cash flow channel, while others are more suited to capture changes in the discount rate. Most models provide higher gains than the equally weighted benchmark (EW) with a Δ_{CER} of 1.22 to 1.41%. Considering the FSC benchmark ($\Delta_{CER} = 1.72\%$) and highlighted by bold entries, we have ten superior regime-based strategies. A review of the economic cycle reveals that the number of models demonstrating a higher Δ_{CER} increases to 26 in expansions and 16 in recessions, with each panel producing at least one strategy that dominates all benchmarks.

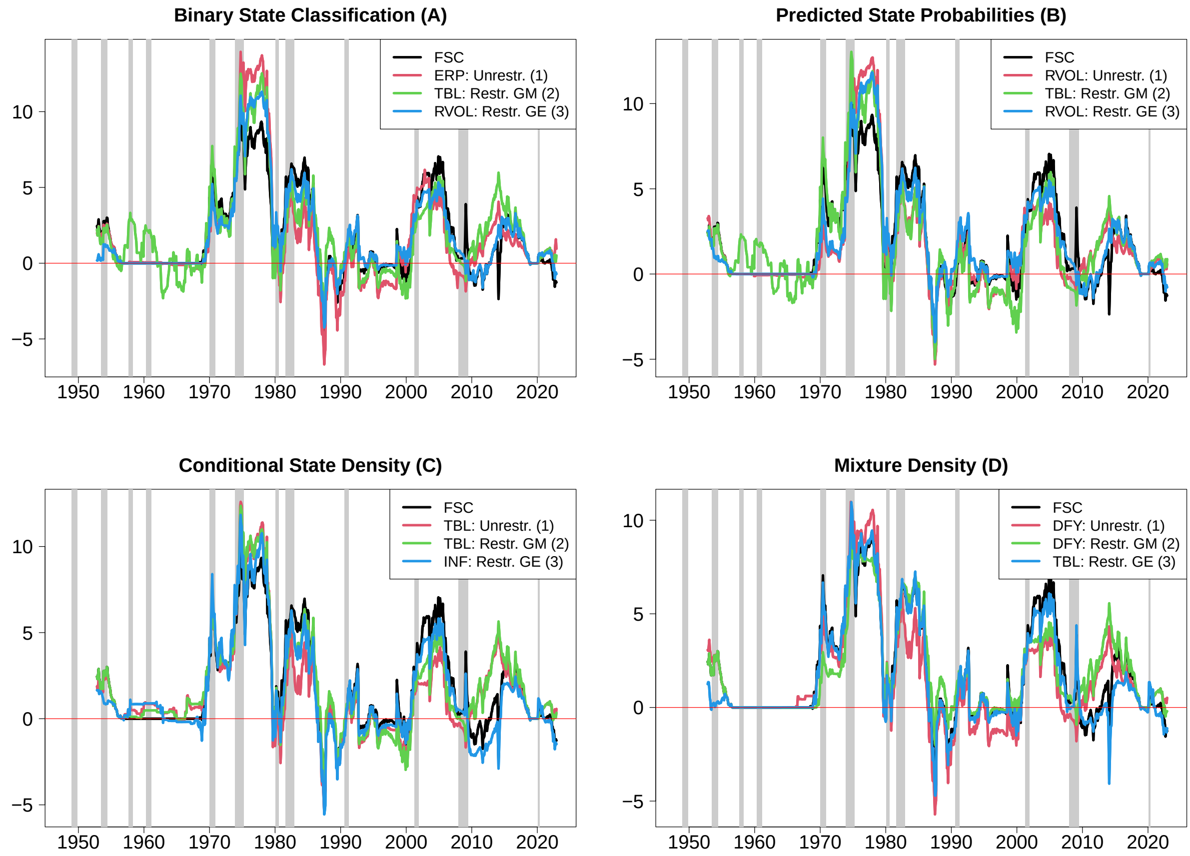
The robustness test in the Appendix C (Tables C1–C3) shows how sensitive the Δ_{CER} behaves if we vary the risk aversion or if we relax the investment restrictions (such as partial leveraging or short selling). In the baseline case, we always assume $\gamma = 3$ and $w = [0, 1]$. Reducing risk aversion does not result in a notable change in the absolute certainty equivalent gains. However, it increases the probability of outperforming the FSC benchmark. Conversely, an increase in risk aversion to a value of five results in a discernible reduction in the Δ_{CER}. This finding can be explained by the fact that our timing strategies cannot achieve an appropriate risk reduction. Furthermore, a relaxation of investment restrictions leads to higher certainty equivalent gains. The marginal effect is largest when the leverage restriction is removed. These results confirm the findings of Baltas and Karyampas (2018), who analyzed the performance of return forecasting models in different regimes and concluded that the economic benefits for high risk-averse and leverage-constrained investors are diminishing.

Table 3: Economic Value (Δ_{CER} in % p.a.)

	Overall				**Expansion**				**Recession**			
Variable	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Panel 1: Benchmark												
FSC	-	1.72		-	-	0.67		-	-	8.20		-
EW	1.22	1.34	1.41	-	0.25	0.30	0.48	-	7.16	7.75	7.13	-
Panel 2: Binary State Classification (A)												
ERP	1.68	1.02	1.51	1.17	0.62	0.22	0.30	0.27	8.19	6.01	**9.04**	6.72
GM	1.67	1.15	1.31	1.15	0.59	0.41	0.09	0.19	**8.39**	5.75	**8.84**	7.07
GE	1.31	1.55	1.33	0.91	0.37	0.47	0.40	0.16	7.12	**8.20**	7.07	5.53
DP	1.23	1.36	1.56	0.93	0.23	0.31	0.62	0.17	7.40	7.80	7.37	5.67
RVOL	1.62	1.24	1.66	1.13	**0.71**	0.24	**0.74**	0.42	7.19	7.42	7.35	5.55
TMS	0.67	0.97	0.88	0.56	−0.06	−0.14	0.32	0.00	5.16	7.83	4.32	4.01
DFY	1.27	1.58	0.69	0.84	0.47	**0.96**	−0.32	0.30	6.25	5.44	6.95	4.22
TBL	1.33	**1.97**	1.00	1.01	**0.81**	**1.09**	0.10	0.66	4.55	7.45	6.56	3.20
IP	1.16	1.57	0.97	0.70	0.19	**0.71**	−0.20	−0.08	7.11	6.84	8.19	5.55
INF	1.36	1.55	1.61	0.95	0.33	0.58	0.44	0.15	7.68	7.54	**8.84**	5.88
Panel 3: Predicted Probabilities (B)												
ERP	1.64	0.93	1.51	1.13	0.59	0.14	0.31	0.24	8.15	5.83	**8.95**	6.64
GM	1.68	1.07	1.48	1.12	0.62	0.24	0.35	0.19	**8.22**	6.14	**8.48**	6.86
GE	1.22	1.50	1.27	0.87	0.23	0.33	0.42	0.11	7.35	**8.72**	6.47	5.58
DP	1.19	1.35	1.47	0.90	0.19	0.31	0.52	0.13	7.38	7.80	7.36	5.64
RVOL	1.70	1.27	**1.83**	1.24	**0.76**	0.34	**0.94**	0.50	7.46	6.99	7.33	5.78
TMS	0.79	1.10	0.88	0.50	0.16	0.03	0.42	0.07	4.69	7.74	3.71	3.17
DFY	1.28	1.65	0.97	0.94	0.54	**1.00**	−0.02	0.42	5.89	5.67	7.07	4.21
TBL	1.20	**1.76**	0.92	0.71	0.34	**0.75**	0.07	0.04	6.49	8.03	6.14	4.85
IP	1.04	1.57	1.06	0.67	0.09	0.61	−0.01	−0.09	6.94	7.56	7.73	5.36
INF	1.36	1.71	1.40	0.97	0.42	0.62	0.42	0.21	7.22	**8.42**	7.45	5.65
Panel 4: Conditional State Density (C)												
ERP	1.37	1.64	1.11	0.95	0.48	**0.86**	−0.05	0.27	6.87	6.48	**8.31**	5.14
GM	1.30	1.65	1.16	0.85	0.43	**0.77**	0.23	0.21	6.71	7.04	6.96	4.83
GE	1.41	**1.75**	1.18	0.87	0.43	**0.74**	0.27	0.11	7.43	7.97	6.79	5.57
DP	1.23	1.22	1.32	0.78	0.32	0.27	0.42	0.07	6.89	7.11	6.92	5.20
RVOL	1.14	1.58	1.02	0.75	0.44	**0.69**	0.32	0.14	5.45	7.12	5.33	4.55
TMS	1.03	1.39	1.41	0.76	0.04	0.43	0.09	−0.04	7.17	7.31	**9.54**	5.70
DFY	1.30	1.70	0.83	0.92	0.54	**0.95**	−0.20	0.35	6.04	6.33	7.19	4.48
TBL	1.62	**1.91**	1.18	1.16	**0.78**	**1.01**	0.06	0.48	6.79	7.48	8.10	5.40
IP	1.12	**1.83**	0.65	0.65	0.20	**0.96**	−0.46	−0.08	6.76	7.25	7.58	5.16
INF	1.45	1.55	1.43	0.98	0.49	0.59	0.41	0.20	7.39	7.46	7.72	5.76
Panel 5: Mixture Density (D)												
ERP	1.53	**1.77**	1.11	0.96	0.58	**0.90**	−0.05	0.22	7.38	7.16	**8.29**	5.55
GM	1.43	1.71	1.15	0.83	0.51	**0.75**	0.22	0.10	7.13	7.65	6.91	5.32
GE	1.11	1.48	1.41	0.65	0.15	0.50	0.39	−0.10	7.02	7.53	7.70	5.30
DP	1.24	1.18	1.39	0.80	0.32	0.22	0.50	0.09	6.89	7.11	6.92	5.20
RVOL	1.43	1.71	1.18	0.93	0.63	**0.75**	0.46	0.22	6.36	7.62	5.64	5.30
TMS	1.34	1.68	1.32	1.07	0.32	**0.74**	0.01	0.29	7.63	7.46	**9.48**	5.86
DFY	1.56	**1.83**	0.99	1.08	**0.68**	**0.96**	−0.16	0.39	7.00	7.19	8.08	5.36
TBL	1.40	1.54	1.49	1.01	0.52	0.64	0.37	0.33	6.85	7.10	**8.46**	5.20
IP	0.95	**1.76**	0.81	0.57	0.03	**0.83**	−0.36	−0.15	6.65	7.55	8.04	5.05
INF	1.39	**1.73**	1.26	0.83	0.31	**0.71**	0.10	−0.03	8.00	8.02	**8.42**	6.16

Notes: Table shows the certainty equivalent gains. In the unrestricted version (1) the SOP forecast consists of variable-specific regime forecasts for earnings and multiple growth. (2) assumes no multiple growth ($\hat{\mu}^{gm} = 0$) and uses only the *flexible regime* forecasts for earnings growth, (3) assumes that earnings growth follows a 20-year moving average ($\bar{ge}^{20Y}$) and the expected multiple growth are is predicted by the *flexible regime* forecasts. Finally (4) computes the return forecasts directly. Panel 1 displays the benchmark models FSC (Ferreira and Santa-Clara 2011), and the SOP forecasts assuming equal weights (EW). The Panels 2–5 show the economic value with an observation weighting according to the different views. Superior results compared to FSC are highlighted in bold.

Figure 7: Economic Value over Time (Δ_{CER}^{5Y})



Notes: Figure shows the rolling five-year certainty equivalent gains (Δ_{CER}^{5Y}) of the best unrestricted (in red), *GM*-restricted (in green), and *GE*-restricted (in blue) regime-based trading strategies. As a benchmark, the *FSC* forecast of Ferreira and Santa-Clara (2011) is added in black. Graph (A) indicates the weighting scheme using the binary state classification, (B) uses the predicted state probabilities, (C) the conditional state density and (D) the mixture state density. The Δ_{CER} measures the maximum management fee an investor is willing to pay to participate on the forecast-based strategy (in % p.a.). Gray shaded areas indicate NBER recessions.

So far, we have evaluated the economic value only on average. The empirical evidence suggests, however, that the benefits are time-dependent (Dangl and Halling 2012; Baltas and Karyampas 2018). Therefore, Figure 7 presents the rolling 5-year average Δ_{CER}^{5Y} of the best strategies for the constrained and unconstrained cases, separated by the weighting method (A)–(D). There is no trading strategy that strongly dominates all others, nor has a positive Δ_{CER}^{5Y} over the evaluation period. The relatively flat performance observed at the beginning and end of the 1960s is attributable to the leverage restriction and illustrates comparable certainty equivalent gains between the regime strategies and the naïve forecast strategy. From this point onwards, the Δ_{CER}^{5Y} increased substantially, indicating that it is advantageous to follow a timing strategy based on our *flexible regime* approach until the 1980s. The rolling average certainty equivalent gain demonstrates a benefit of up to 12% per year. Similar to the results of the CDSFE curves, it is challenging for the models to achieve economic gains between the 1980s and the end of the 1990s, even if a Δ_{CER}^{5Y} of approximately 3% can be attained during the recessionary phase at the beginning of the 1990s. From the 2000s onwards, the majority of timing strategies are once again capable of generating continuous utility gains for a risk-averse investor. While the amount is limited to a maximum of 5%, it is notable that at the most recent measurement date in December 2022, specifications (1) and (2) still managed to generate slightly positive certainty equivalent gains over the past five years. In contrast, the FSC benchmark was unable to achieve this.

6.2 Combining Economic Regimes

We provide two extensions by combining the results from the individual variable analysis. We start with a test of all 90 combinations using different predictors for earnings growth and multiple growth. Table 4 presents the out-of-sample results with the choice of the multiple (earnings) predictor in the rows (columns). With bold entries, we highlight the best GM predictor $\mu^{gm}(v,l)$ for a given GE predictor $\mu^{ge}(v,k)$.

The results show that it is rarely the best choice to use the same variable to predict earnings growth and multiple changes. Typically, a combination of different variables is needed to achieve the lowest prediction errors or the highest economic benefit. Using the binary state classification as a view generator, the R_{OS}^2 of the bold entries ranges between 1.26% (ERP+TMS) and 1.89% (ERP+DFY), whereas the predicted probability weighting produces R_{OS}^2 of 1.30% ($RVOL$+TMS) and 1.96% (ERP+DFY). The results of the density-weighted approaches are qualitatively similar but somewhat weaker with R_{OS}^2 of 1.32% (in (C) with INF+DFY) and 1.41% (in (D) with TBL+DFY). The best pairs, according to the Δ_{CER}, range from 1.69% (in (D) with TBL+DFY) to 2.17% (in (A) with $RVOL$+TBL). These benefits provide substantial improvements relative to the previous subsection. Analyzing the patterns according to each regime variable, we emphasize the robust evidence of predicting multiple growth with ERP, $RVOL$, INF, and GE and the usefulness of DFY, TBL, and TMS as earnings growth predictors.

Table 4: Performance for Pairs of Economic Regimes (continued on next page)

Weighting with Binary State Classification (A) | *Weighting with Predicted Probabilities (B)*

Panel 1: Out-of-Sample R-Squared (R^2_{OS} in %)

GM Predictor	A: GE Predictor ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF	B: GE Predictor ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF
ERP	**1.55**	**1.61**	**1.61**	**1.32**	**1.50**	**1.26**	**1.89**	**1.72**	**1.52**	**1.45**	**1.47**	**1.45**	1.51	1.32	**1.50**	1.26	**1.96**	1.57	**1.49**	**1.53**
GM	0.94	1.33	1.27	0.97	1.10	0.87	1.43	1.33	1.11	1.06	0.96	1.21	1.21	1.02	1.15	0.93	1.55	1.24	1.13	1.18
GE	0.41	0.73	0.92	0.60	0.70	0.45	1.01	0.92	0.72	0.65	0.51	0.68	0.88	0.68	0.78	0.56	1.19	0.89	0.78	0.82
DP	0.65	0.96	1.12	0.87	0.94	0.71	1.25	1.19	0.97	0.89	0.64	0.81	1.00	0.85	0.92	0.72	1.31	1.05	0.93	0.96
RVOL	0.96	1.31	1.48	1.22	1.31	1.06	1.48	1.52	1.27	1.21	1.20	1.39	**1.55**	**1.41**	1.49	**1.30**	1.81	**1.62**	1.47	1.51
TMS	-0.02	0.34	0.59	0.37	0.34	0.62	0.54	0.96	0.46	0.35	0.18	0.36	0.57	0.45	0.49	0.55	0.85	0.79	0.56	0.57
DFY	-0.48	-0.24	-0.17	-0.48	-0.28	-0.50	-0.01	-0.01	-0.15	-0.32	0.25	0.36	0.51	0.31	0.47	0.25	0.94	0.57	0.51	0.53
TBL	0.34	0.65	0.81	0.55	0.62	0.50	0.93	1.06	0.75	0.61	0.47	0.63	0.82	0.67	0.75	0.60	1.15	0.98	0.80	0.82
IP	0.32	0.55	0.62	0.32	0.51	0.32	0.95	0.82	0.71	0.49	0.40	0.50	0.64	0.44	0.61	0.39	1.10	0.71	0.69	0.68
INF	0.65	0.93	1.05	0.77	0.89	0.68	1.25	1.18	0.96	0.96	0.60	0.74	0.91	0.73	0.85	0.65	1.27	0.98	0.89	0.95

Panel 2: Out-of-Sample Certainty Equivalent Gains (Δ_{CER} in %)

GM Predictor	A: GE Predictor ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF	B: GE Predictor ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF
ERP	**1.68**	1.53	1.27	1.06	1.26	0.48	**1.72**	1.45	1.57	1.28	**1.64**	1.41	1.15	1.06	1.39	0.80	1.76	1.28	1.39	1.32
GM	1.32	**1.67**	1.24	0.86	1.11	0.64	1.30	1.44	1.34	1.25	1.18	**1.68**	1.42	1.16	1.41	1.05	1.59	1.45	1.38	1.44
GE	0.70	0.97	1.31	1.01	1.04	0.53	1.14	1.50	1.31	1.17	0.61	0.79	1.22	1.06	1.05	0.82	1.30	1.32	1.22	1.23
DP	0.86	0.95	1.23	1.23	1.05	0.56	1.50	1.47	1.44	1.26	0.78	0.76	1.21	1.19	1.08	0.70	1.53	1.31	1.23	1.40
RVOL	0.95	1.39	**1.64**	**1.54**	**1.62**	**1.31**	1.60	**2.17**	**1.65**	**1.60**	1.24	1.55	**1.70**	**1.71**	**1.70**	**1.58**	**1.82**	**2.05**	**1.79**	**1.78**
TMS	0.44	0.60	0.47	0.46	0.50	0.67	0.97	1.45	0.89	0.84	0.35	0.50	0.60	0.47	0.56	0.79	0.91	1.05	0.73	0.90
DFY	0.50	0.41	0.42	-0.13	0.31	-0.28	1.27	0.53	0.83	0.36	0.58	0.43	0.57	0.37	0.69	0.30	1.28	0.69	0.85	0.80
TBL	0.45	0.51	0.63	0.53	0.52	0.29	0.96	1.33	0.81	0.79	0.43	0.45	0.76	0.62	0.65	0.60	1.11	1.20	0.75	0.92
IP	0.53	0.50	0.45	0.13	0.41	-0.26	1.25	0.62	1.16	0.50	0.49	0.56	0.58	0.46	0.69	0.29	1.13	0.76	1.04	0.75
INF	1.01	1.12	1.18	1.03	1.10	0.47	1.47	1.46	1.49	1.36	0.82	0.94	1.03	0.92	1.11	0.80	1.30	1.38	1.28	1.36

Table 4: OS Performance for Pairs of Economic Regimes (continued from previous page)

Weighting with Conditional State Density (C) | ***Weighting with Mixture Density (D)***

Panel 1: Out-of-Sample R-Squared (R^2_{OS} in %)

GM Predictor	GE Predictor (C): ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF	GE Predictor (D): ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF
ERP	1.07	0.76	0.47	0.51	0.87	0.73	1.12	1.09	0.90	0.67	1.02	0.78	0.03	0.43	0.79	1.03	1.18	1.05	0.81	0.73
GM	1.16	**1.02**	0.87	0.83	**1.09**	0.89	1.24	1.26	1.07	0.91	1.12	0.99	0.34	0.66	0.98	1.14	1.25	1.14	0.97	0.91
GE	1.12	0.95	0.79	0.73	1.02	0.81	1.20	1.17	1.01	0.85	1.23	**1.12**	0.67	0.78	1.11	1.25	1.34	1.25	1.13	1.07
DP	1.15	0.92	0.64	0.78	1.01	0.81	1.23	1.16	0.98	0.80	1.21	1.04	0.29	0.78	1.03	1.24	1.38	1.24	1.02	0.96
RVOL	0.83	0.90	**1.05**	**0.84**	0.97	0.76	0.90	1.14	0.94	0.89	1.06	1.06	**0.67**	**0.79**	1.05	1.13	1.12	1.10	1.04	1.04
TMS	1.18	0.92	0.71	0.68	1.01	0.87	1.22	1.23	1.09	0.90	1.25	1.01	0.37	0.64	1.02	1.26	1.39	1.26	1.06	0.98
DFY	0.73	0.44	0.13	0.17	0.53	0.37	0.76	0.73	0.56	0.32	0.96	0.71	-0.10	0.34	0.71	0.97	1.14	0.98	0.73	0.63
TBL	1.09	0.83	0.59	0.62	0.93	0.78	1.13	1.16	0.96	0.78	**1.29**	1.11	0.52	0.78	**1.12**	**1.32**	**1.41**	**1.32**	**1.14**	**1.08**
IP	0.90	0.55	0.20	0.26	0.64	0.49	0.95	0.86	0.71	0.44	0.89	0.63	-0.18	0.25	0.63	0.89	1.07	0.90	0.66	0.54
INF	**1.25**	1.00	0.76	0.77	1.09	**0.93**	**1.32**	**1.28**	**1.11**	**0.96**	1.11	0.94	0.29	0.61	0.94	1.14	1.25	1.14	0.96	0.94

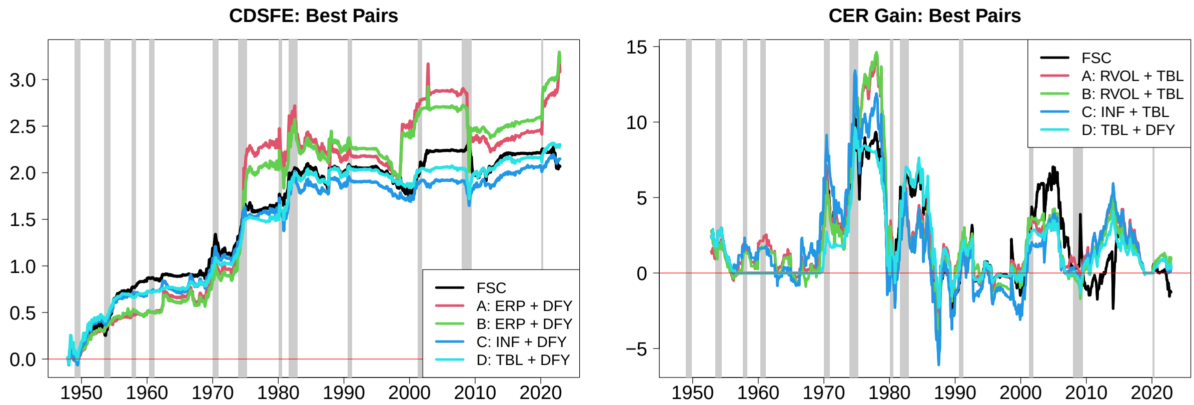
Panel 2: Out-of-Sample Certainty Equivalent Gains (Δ_{CER} in %)

GM Predictor	GE Predictor (C): ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF	GE Predictor (D): ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP	INF
ERP	1.37	1.22	0.84	0.69	1.13	0.90	1.42	1.60	1.38	0.92	1.53	1.33	0.36	0.62	1.29	1.35	1.60	1.29	1.34	1.17
GM	1.34	1.30	1.30	1.09	1.27	0.95	1.39	1.38	1.47	1.07	1.49	1.43	0.78	1.04	1.40	1.35	1.53	1.27	1.46	1.36
GE	1.36	1.29	**1.41**	0.86	1.23	1.00	1.39	1.44	1.44	1.09	1.54	1.41	**1.11**	1.02	1.39	1.35	1.60	**1.33**	1.50	1.38
DP	1.40	1.28	0.77	**1.23**	1.30	0.93	1.50	1.40	1.34	0.88	1.45	1.42	0.41	**1.24**	1.33	1.40	1.62	1.33	1.33	1.37
RVOL	1.03	1.10	1.29	1.05	1.14	0.96	1.12	1.38	1.13	1.22	1.41	1.41	1.05	1.24	1.43	1.44	1.44	1.39	1.46	1.45
TMS	1.42	1.31	1.06	0.70	1.23	1.03	1.49	1.61	1.57	1.21	1.51	1.33	0.57	0.52	1.27	1.34	1.60	1.19	1.40	1.19
DFY	1.17	0.92	0.45	0.38	0.84	0.77	1.30	1.20	1.15	0.64	1.46	1.27	-0.13	0.38	1.15	1.28	1.56	1.09	1.21	0.99
TBL	1.29	1.23	1.13	0.82	1.21	0.95	1.33	1.62	1.39	1.08	**1.64**	**1.55**	0.97	1.07	**1.60**	**1.61**	**1.69**	**1.40**	**1.68**	**1.63**
IP	1.06	0.74	0.12	0.16	0.72	0.56	1.19	1.10	1.12	0.42	1.25	1.00	-0.17	0.15	0.84	1.00	1.36	0.90	0.95	0.73
INF	**1.61**	**1.51**	1.33	0.91	**1.51**	**1.22**	**1.70**	**1.79**	**1.66**	**1.45**	1.39	1.27	0.80	0.89	1.28	1.29	1.53	1.23	1.36	1.39

Notes: Table shows the statistical (Panel 1) and economic performance (Panel 2) using all pairs of economic regimes for the return components *GE* (columns) and *GM* (rows). The results are sorted according to the different regime weightings. The statistical accuracy (economic value) of the return forecasts is measured by the out-of-sample R^2_{OS} (Δ_{CER}) relative to the historical average.

The course of the most dominant strategies according to the statistical accuracy and the economic value is depicted in Figure 8. The CDSFE curves indicate that a combination of different pairs of variables has a higher forecasting accuracy than the historical average over long periods of time. Of particular note is the outstanding forecast accuracy observed during the 1970s, from the late 1990s until the 2008 financial crisis, as well as from 2010 to 2022. In light of the rolling five-year CER gain, it is evident that the considerable benefit gains observed in the 1970s, with average values over 10%, cannot be replicated. This is largely attributable to the intensifying competition to exploit market inefficiencies, which reduce economic profits. Nevertheless, even modestly positive Δ_{CER}, ranging from 0.5 to 1%, at the end of 2022 can be regarded as a noteworthy achievement.

Figure 8: Forecasting Performance over Time for the Best Pairs



Notes: Figure shows the cumulative differences in the squared forecast errors (CDSFE, left graph) and the rolling five-year certainty equivalent gain (Δ^{5Y}_{CER}, right graph) of the best forecasting pairs according to each weighting scheme. The label reads as follows: (A) indicates the weighting scheme using the binary state classification, (B) the predicted state probabilities, (C) the conditional state density and (D) the mixture state density. The variable in the first place is the best forecast for multiple growth (GM), and the variable in the second place the best forecast for earnings growth (GE). As benchmark, the FSC forecasts of Ferreira and Santa-Clara (2011) is added in black. Gray shaded areas indicate NBER recessions. All values are expressed in percentage points.

As a second extension, we do not consider only one economic variable; we rather pool all forecasts together to predict both return components separately. Such an approach has the benefit of hedging against model uncertainty and is popular in the return predictability literature (Rapach et al. 2010). Table 5 presents the out-of-sample performance for the pooled forecasts. We see that the forecast quality has deteriorated compared to the best individual forecasts. Whether we rely on the equally weighted combination or 'tilt' the weights according to their past performance (DMSFE), we obtain an R^2_{OS} of around 1%. During expansions, this value ranges between 0.62 and 0.75%, while during recessions, values of around 2% are achieved. On average, these results still indicate a significant predictability of returns, but they are all worse than the FSC or EW benchmarks. The situation is similar if we measure the benefit of an investor. The pooled timing strategies deliver a total Δ_{CER} of 1.37−1.53% (0.44–0.6% in expansions and 7.05–7.44% in recessions). In addition, we cannot document major differences in the forecasting performance across the equally weighted average and the DMSFE weighting. This result suggests a high similarity and low dispersion of the component forecasts. For comparison purposes, the pooling of forecasts was also employed for *flexible regime forecasts* for aggregate returns directly. Neither a significant improvement in forecast errors nor a higher economic value is achieved. This demonstrates the need for decomposing returns.

Table 5: Performance of Pooled Forecasts

	Forecast Combination					
	R^2_{OS} (in %)			Δ_{CER} (in % p.a.)		
	Overall	Expansion	Recession	Overall	Expansion	Recession
Panel 1: Binary State Classification (A)						
SOP-DMSFE	1.09***	0.75***	2.07**	1.51	0.57	7.33
R-DMSFE	−0.21	0.25*	−1.5	0.05	0.07	−0.12
SOP-AVE	1.09***	0.75***	2.07**	1.50	0.57	7.27
R-AVE	−0.21	0.25*	−1.53	0.05	0.07	−0.15
Panel 2: Predicted Probabilities (B)						
SOP-DMSFE	1.07***	0.62**	2.33***	1.53	0.58	7.36
R-DMSFE	−0.19	0.19*	−1.28	0.09	0.09	0.07
SOP-AVE	1.06***	0.62**	2.33***	1.53	0.60	7.33
R-AVE	−0.19	0.19*	−1.29	0.09	0.09	0.07
Panel 3: Conditional State Density (C)						
SOP-DMSFE	1.01***	0.68**	1.96**	1.37	0.44	7.06
R-DMSFE	−0.46	−0.02	−1.74	0.06	0.07	0.02
SOP-AVE	1.02***	0.69**	1.97**	1.37	0.45	7.05
R-AVE	−0.46	−0.01	−1.75	0.07	0.07	0.02
Panel 4: Mixture Density (D)						
SOP-DMSFE	1.04***	0.65**	2.17***	1.47	0.50	7.44
R-DMSFE	−0.43	−0.02	−1.62	0.08	0.07	0.09
SOP-AVE	1.04***	0.65**	2.15***	1.48	0.53	7.34
R-AVE	−0.43	−0.01	−1.63	0.08	0.07	0.09

Notes: Table shows the statistical and economic performance of out-of-sample pooled forecasts. Two approaches are used. SOP-DMSFE computes the weights for both return components (GE and GM) based on the realized forecast performance with a discount factor of 0.9, and SOP-AVE comines all component forecasts with a equal weight. R-DMSFE and R-AVE apply forecast combination to regime-based return forecasts instead of the return components. Panel 1−4 separate the combination approaches according to the different views. The statistical accuracy (economic value) of return forecasts is measured with the out-of-sample R^2_{OS} (Δ_{CER}) relative to the historical average.

6.3 Benefits of Entropy Pooling

Assuming that regime filtering has some predictive power, the question arises whether entropy pooling is necessary for the prediction at all. Instead, the information on regime dynamics from the four approaches (A)–(D) could be used directly to compute the observation weights. The answer to this question is given in Table B1. In statistical (R^2_{OS}) and economic (Δ_{CER}) terms, entropy pooling provides an average improvement of 39 basis points (bps) over the plain view approach. Across the different weighting approaches, the average difference in R^2_{OS} is 77 bps, respectively 67 bps for Δ_{CER} in (A), 50 and 48 bps in (B), 10 and 18 bps in (C), and 22 and 27 bps in (D). Since the density-weighted approaches already generate relatively smooth flexible probabilities, the added value of entropy pooling is significantly lower there. The most significant improvements are documented in specification (2) using (A) and (B) as view generators. ERP, GM, $RVOL$, and TMS achieve performance increases of at least 100 and at most 330 bps.

Another argument in favor of entropy pooling is the relatively low dispersion in forecasting power across all economic variables and across all weighting schemes, which can be seen in Tables 2 and 3. Certainly, the assumption of a common prior distribution affects the results. However, minimizing the relative entropy can mitigate significant discrepancies while accentuating particular aspects identified within the regime views.

Figure 9: Forecast Variance vs. Squared Forecast Bias

Notes: Figure shows the trade-off between forecast bias and forecast variance for the earnings growth and the price-earnings multiple growth. (A) indicates the weighting scheme using the binary state classification, (B) uses the predicted state probabilities, (C) the conditional state density and (D) the mixture state density. We add the assumption of the basic model of Ferreira and Santa-Clara (2011) to both return components. *20Y*: 20-year rolling average of earnings growth and *0*: no multiple growth (constant). For comparison, we also print the results of the historical average for both components (*HA*).

A reason for the success of our *flexible regime* approach is the low forecast variability. According to Theil et al. (1966) and Rapach et al. (2010), a forecaster can typically improve the $MSFE$ either by reducing the squared forecast bias or by decreasing the forecast variance.[16] Figure 9 demonstrates that the primary advantage is the reduction of forecast variance for earnings growth and, to a lesser extent, the reduction of forecast bias for multiple changes (using the unrestricted case as an example). Concerning the different weighting methods, (C) and (D) exhibit particularly low forecast variance, whereas forecasts from (A) and (B) are more volatile.

In addition to predictive performance, the information content is often a key issue when considering flexible probabilities (Meucci 2012). If too few observations are used, the analysis might depend heavily on single observations and is therefore very sensitive to small data changes. Since the four specifications emphasize different fractions of the historical sample, we want to evaluate the informativeness of a given flexible probability choice. For this purpose, we rely on the *effective number of scenarios* (ENS), as Meucci (2012) suggests.[17]

Table B2 in the Appendix B compares the ENS of the view-based flexible probabilities with the entropy pooled counterparts, whereby the average out-of-sample ENS is expressed as the fraction of the upper limit. Along with all methods, we see an improvement in the information content of the entropy pooling approaches relative to the plain regime view. The ENS is always around 90% of the equally weighted reference and the entropy pooling increases the information content by more than 10% relative to the view-based approaches. Additionally, we find that the predictable probability weighting (B) has the highest proportional ENS on average in both approaches.

We conclude that entropy pooling improves the predictability and increases the robustness of our results. Similar to other Bayesian approaches (Connor 1997), introducing some degree of shrinkage helps to reduce estimation uncertainty. This is valuable, especially in applications with high uncertainty, such as financial market forecasting.

7 Conclusion

Predicting stock returns is of great interest to investors and academics. Despite their relevance, it is difficult to find successful strategies that are robust to market frictions, such as transaction costs and data snooping (Pesaran and Timmermann 1995; Welch

16. Theil et al. (1966) approximate the $MSFE$ of a variable y using $MSFE \approx (\bar{y} - \bar{\hat{y}})^2 + \sigma_{\hat{y}} + \sigma_y$ assuming only a weak correlation between the predicted and the realized return. To deal with the bias-variance trade-off, reducing the squared bias $(\bar{y} - \bar{\hat{y}})^2$ typically comes at the cost of increasing the forecast variance $\sigma_{\hat{y}}$.

17. The ENS is defined as exponent of the entropy $ENS_t = exp(-\sum_{t=1}^{\tau} p_t \ln(p_t))$, and measures the "concentration of probability mass" (Meucci 2012) in the probability vector p. The effective number of scenarios is maximal for an equally-weighted historical average and minimal when using only a single data point.

and Goyal 2008). As Dichtl et al. (2021) show, the SOP method of Ferreira and Santa-Clara (2011) succeeds in overcoming this challenge.

This paper revises the SOP method by combining economic regimes with model restrictions. We propose a three-step methodology that deals with parameter instability and estimation uncertainty to separately predict the return components before aggregation. We use Markov-switching models (Hamilton 1989) and flexible probabilities (Meucci 2008, 2010) to semi-parametrically forecast earnings growth and changes in the price-earnings multiple using filtered regime dynamics with a Bayesian approach. We provided a comprehensive evaluation and achieved economically and statistically significant results that outperformed the SOP benchmark of Ferreira and Santa-Clara (2011) with R^2_{OS} and Δ_{CER} of more than 1.5%. We highlighted the role of DFY and TBL in predicting earnings growth and $RVOL$, ERP, and INF in predicting multiple growth. Finally, our results suggest that the predictability of both return components varies over time and is affected by the business cycles. While earnings growth was more predictable during periods of expansion, forecasting multiple changes was more advantageous during recessions.

Based on our results, there are many avenues for future research. In this paper, we only use the expected value of the distribution implied by our *flexible regime* approach. Thus, incorporating predictive information about higher (regime-dependent) moments (variance, skewness, kurtosis) or specific tail measures may improve our results and be helpful for risk management and portfolio insurance purposes. In addition, the number of regimes can be augmented, and the regime dynamics can be jointly filtered using multivariate models. Another promising avenue would be to follow Faria and Verona (2018) and conduct a wavelet analysis of the return components to relate the economic regimes only to the frequency-decomposed parts. Finally, it would be interesting to test the robustness of our methodology for other equity markets or asset classes.

References

Ang, A., and A. Timmermann. 2012. "Regime Changes and Financial Markets." *Annual Review of Financial Economics* 4 (1): 313–337.

Baetje, F., and L. Menkhoff. 2016. "Predicting the equity premium via its components." *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Empirical Finance II, No. E12-V2.*

Baltas, N., and D. Karyampas. 2018. "Forecasting the equity risk premium: The importance of regime-dependent evaluation." *Journal of Financial Markets* 38:83–102.

Bansal, R., and A. Yaron. 2004. "Risks for the long run: A potential resolution of asset pricing puzzles." *The Journal of Finance* 59 (4): 1481–1509.

Çakmaklı, C., and D. van Dijk. 2016. "Getting the most out of macroeconomic information for predicting excess stock returns." *International Journal of Forecasting* 32 (3): 650–668.

Campbell, J. Y., and R. J. Shiller. 1988. "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies* 1 (3): 195–228.

Campbell, J. Y., and S. B. Thompson. 2008. "Predicting excess stock returns out of sample: Can anything beat the historical average?" *The Review of Financial Studies* 21 (4): 1509–1531.

Cheng, T., J. Gao, and O. Linton. 2019. "Nonparametric predictive regressions for stock return prediction." *Working Paper.*

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

Cochrane, J. H. 2011. "Presidential address: Discount rates." *The Journal of Finance* 66 (4): 1047–1108.

Connor, G. 1997. "Sensible return forecasting for portfolio management." *Financial Analysts Journal* 53 (5): 44–51.

Czasonis, M., M. Kritzman, and D. Turkington. 2020. "Addition by Subtraction: A Better Way to Forecast Factor Returns (and Everything Else)." *The Journal of Portfolio Management* 46 (8): 98–107.

Dai, Z., and H. Zhou. 2020. "Prediction of stock returns: sum-of-the-parts method and economic constraint method." *Sustainability* 12 (2): 541.

Dai, Z., and H. Zhu. 2020. "Forecasting stock market returns by combining sum-of-the-parts and ensemble empirical mode decomposition." *Applied Economics* 52 (21): 2309–2323.

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

Dichtl, H., W. Drobetz, A. Neuhierl, and V.-S. Wendt. 2021. "Data snooping in equity premium prediction." *International Journal of Forecasting* 37 (1): 72–94.

Faria, G., and F. Verona. 2018. "Forecasting stock market returns by summing the frequency-decomposed parts." *Journal of Empirical Finance* 45:228–242.

Ferreira, M. A., and P. Santa-Clara. 2011. "Forecasting stock market returns: The sum of the parts is more than the whole." *Journal of Financial Economics* 100 (3): 514–537.

Gisbert, F. J. G. 2003. "Weighted samples, kernel density estimators and convergence." *Empirical Economics* 28:335–351.

Guidolin, M., S. Hyde, D. McMillan, and S. Ono. 2009. "Non-linear predictability in stock and bond returns: When and where is it exploitable?" *International Journal of Forecasting* 25 (2): 373–399.

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

Haase, F., and M. Neuenkirch. 2023. "Predictability of bull and bear markets: A new look at forecasting stock market regimes (and returns) in the US." *International Journal of Forecasting* 39 (2): 587–605.

Hamilton, J. D. 1989. "A new approach to the economic analysis of nonstationary time series and the business cycle." *Econometrica* 57 (2): 357–384.

Hamilton, J. D. 1994. *Time Series Analysis.* Vol. 2. Princeton University Press Princeton, NJ.

Hammerschmid, R., and H. Lohre. 2018. "Regime shifts and stock return predictability." *International Review of Economics and Finance* 56 (C): 138–160.

Henkel, S. J., J. S. Martin, and F. Nardari. 2011. "Time-varying short-horizon predictability." *Journal of Financial Economics* 99 (3): 560–580.

Kim, C.-J. 1994. "Dynamic linear models with Markov-switching." *Journal of Econometrics* 60 (1-2): 1–22.

Longstaff, F. A., and M. Piazzesi. 2004. "Corporate earnings and the equity premium." *Journal of Financial Economics* 74 (3): 401–421.

Ludvigson, S. C., and S. Ng. 2007. "The empirical risk–return relation: A factor analysis approach." *Journal of Financial Economics* 83 (1): 171–222.

McMillan, D. G., and M. E. Wohar. 2011. "Sum of the parts stock return forecasting: international evidence." *Applied Financial Economics* 21 (12): 837–845.

Mele, A. 2007. "Asymmetric stock market volatility and the cyclical behavior of expected returns." *Journal of Financial Economics* 86 (2): 446–478.

Merton, R. C. 1969. "Lifetime portfolio selection under uncertainty: The continuous-time case." *The Review of Economics and Statistics*, 247–257.

Meucci, A. 2005. *Risk and Asset Allocation.* Vol. 1. Springer.

———. 2008. "Fully flexible views: Theory and practice." *Risk* 21:97–102.

———. 2010. "Historical scenarios with fully flexible probabilities." *GARP Risk Professional*, 47–51.

———. 2012. "Effective number of scenarios in fully flexible probabilities." *GARP risk professional*, 32–35.

Meucci, A., D. Ardia, and M. Colasante. 2014. "Portfolio construction and systematic trading with factor entropy pooling." *Risk Magazine* 27 (5): 56–61.

Neely, C. J., D. E. Rapach, J. Tu, and G. Zhou. 2014. "Forecasting the equity risk premium: the role of technical indicators." *Management Science* 60 (7): 1772–1791.

Pedersen, J. B. 2017. *Essays on Financial Risk Management and Asset Allocation.* Institut for Økonomi, Aarhus Universitet. Phd Thesis.

Pesaran, M. H., and A. Timmermann. 1995. "Predictability of stock returns: Robustness and economic significance." *The Journal of Finance* 50 (4): 1201–1228.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2014. "Forecasting stock returns under economic constraints." *Journal of Financial Economics* 114 (3): 517–553.

Rapach, D., and G. Zhou. 2013. "Forecasting stock returns." In *Handbook of Economic Forecasting*, 2:328–383. Elsevier.

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

Shiller, R. J. 2015. "Irrational exuberance: Revised and expanded third edition." In *Irrational exuberance.* Princeton university press.

Theil, H., G. Beerens, C. Tilanus, and C. B. De Leeuw. 1966. *Applied economic forecasting.* Vol. 4. North-Holland Publishing Company Amsterdam.

Van Binsbergen, J. H., and R. S. Koijen. 2010. "Predictive regressions: A present-value approach." *The Journal of Finance* 65 (4): 1439–1471.

Welch, I., and A. Goyal. 2008. "A comprehensive look at the empirical performance of equity premium prediction." *The Review of Financial Studies* 21 (4): 1455–1508.

Yi, Y., F. Ma, Y. Zhang, and D. Huang. 2019. "Forecasting stock returns with cycle-decomposed predictors." *International Review of Financial Analysis* 64:250–261.

Zhang, Y., Y. Wei, F. Ma, and Y. Yi. 2019. "Economic constraints and stock return predictability: A new approach." *International Review of Financial Analysis* 63:1–9.

Zhu, X., and J. Zhu. 2013. "Predicting stock returns: A regime-switching combination approach and economic links." *Journal of Banking & Finance* 37 (11): 4120–4133.

Appendix

Appendix A: Full-Sample Results

Table A1: Summary Statistics of Economic Variables

Panel A: Univariate statistics

Variable	Mean	Median	Std	Minimum	Maximum	Skewness	Kurtosis	AR(1)
ERP	0.51	0.97	5.42	−34.88	35.60	−0.43	7.53	0.07
GM	0.03	0.13	6.67	−72.38	53.61	−0.57	20.53	0.26
GE	0.44	0.64	4.30	−52.83	70.39	1.92	98.53	0.77
DP	0.31	0.28	0.15	0.09	1.27	1.23	3.45	0.98
RVOL	16.59	14.44	9.20	5.44	68.17	2.61	8.92	0.97
TMS	1.69	1.72	1.29	−3.65	4.55	−0.23	0.15	0.96
DFY	1.12	0.90	0.68	0.32	5.64	2.53	9.28	0.97
TBL	3.30	2.74	3.08	0.01	16.30	1.13	1.37	0.99
IP	0.24	0.28	1.81	−14.37	15.32	−0.09	17.08	0.50
INF	0.25	0.24	0.52	−2.05	5.88	1.14	14.22	0.49

Panel B: Correlations

Variable	ERP	GM	GE	DP	RVOL	TMS	DFY	TBL	IP
ERP									
GM	0.77								
GE	0.07	−0.58							
DP	−0.86	−0.69	−0.01						
RVOL	−0.06	0.03	−0.13	0.05					
TMS	−0.01	0.01	−0.02	0.03	0.06				
DFY	−0.26	−0.12	−0.13	0.33	0.14	0.05			
TBL	−0.05	−0.10	0.09	0.04	−0.05	−0.77	−0.14		
IP	0.03	−0.09	0.18	−0.02	−0.17	−0.09	−0.13	0.08	
INF	−0.04	−0.12	0.16	0.04	−0.01	0.01	−0.02	0.03	0.18

Notes: Panel A shows the univariate statistics of the economic variables before transformation (in %). ERP: excess stock return, GM: price-earnings ratio growth, GE: earnings growth, DP: log dividend-price ratio, $RVOL$: realized 12M volatility, TMS: 10Y–3M term spread, DFY: BAA–AAA default spread, TBL: 3M Treasury bill rate, IP: industrial production growth, INF: CPI inflation. Panel B presents the contemporaneous correlation of the transformed variables. For more information about the variable definitions and the transformation, we refer to Section 4. The sample period runs from 1927/12 to 2022/12.

Figure A1: Economic Regimes over Time (continued on next page)

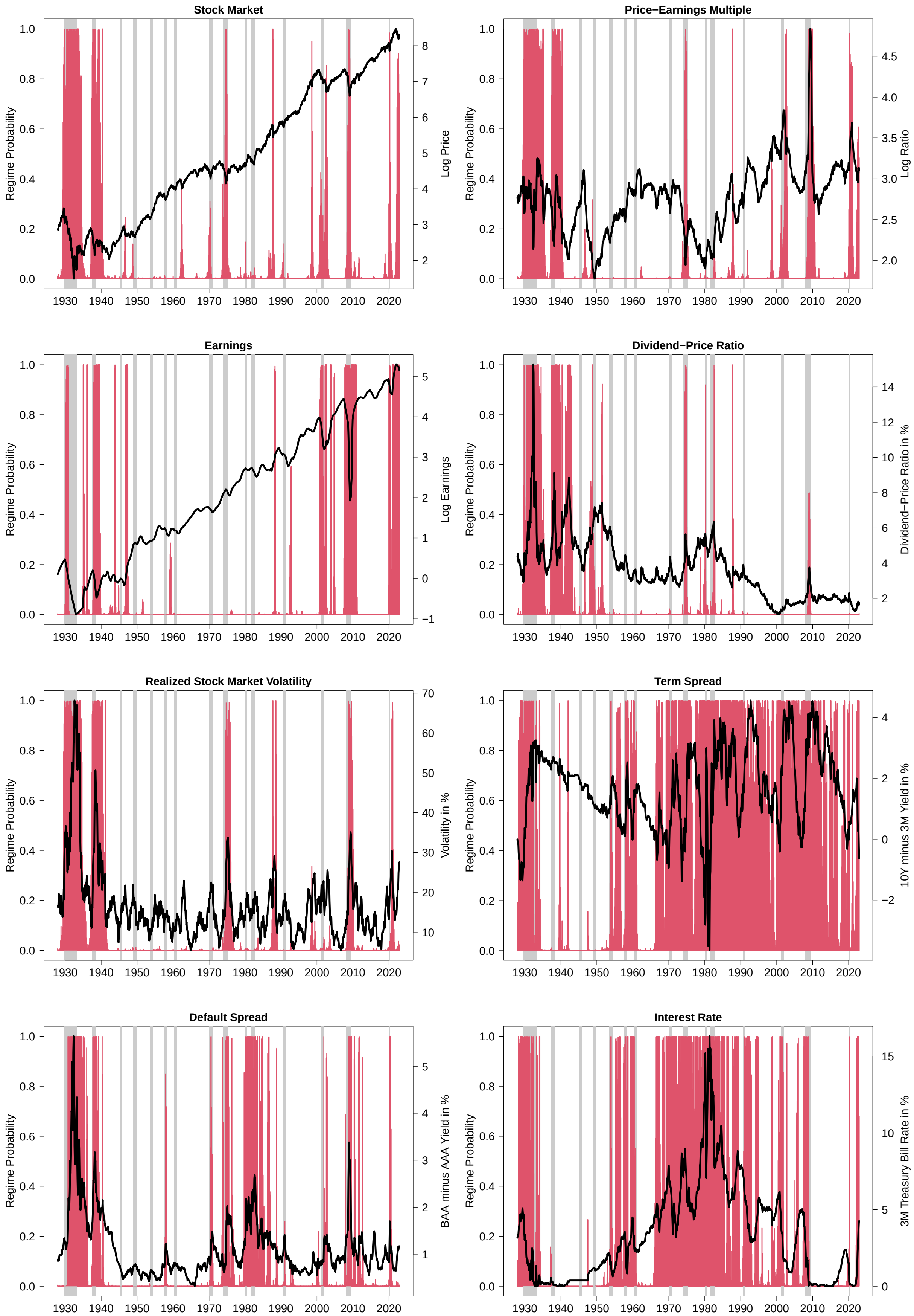
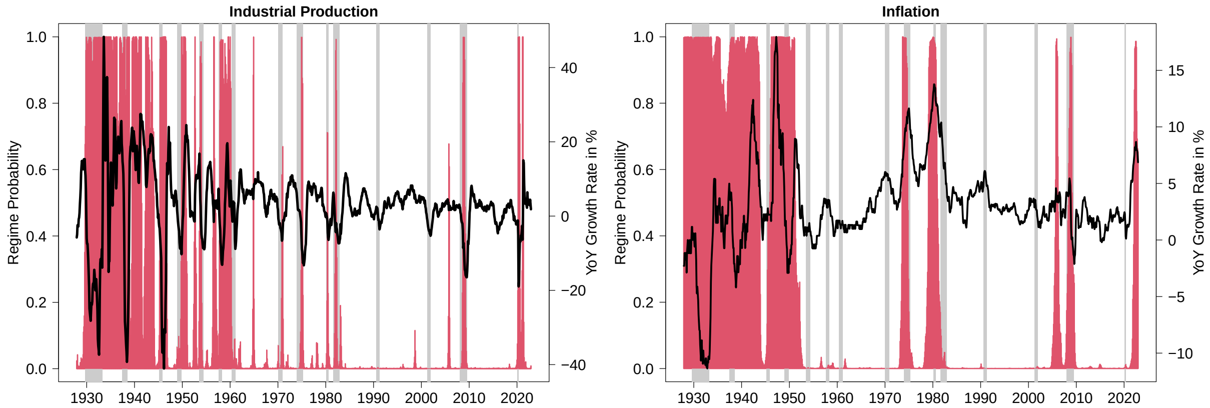
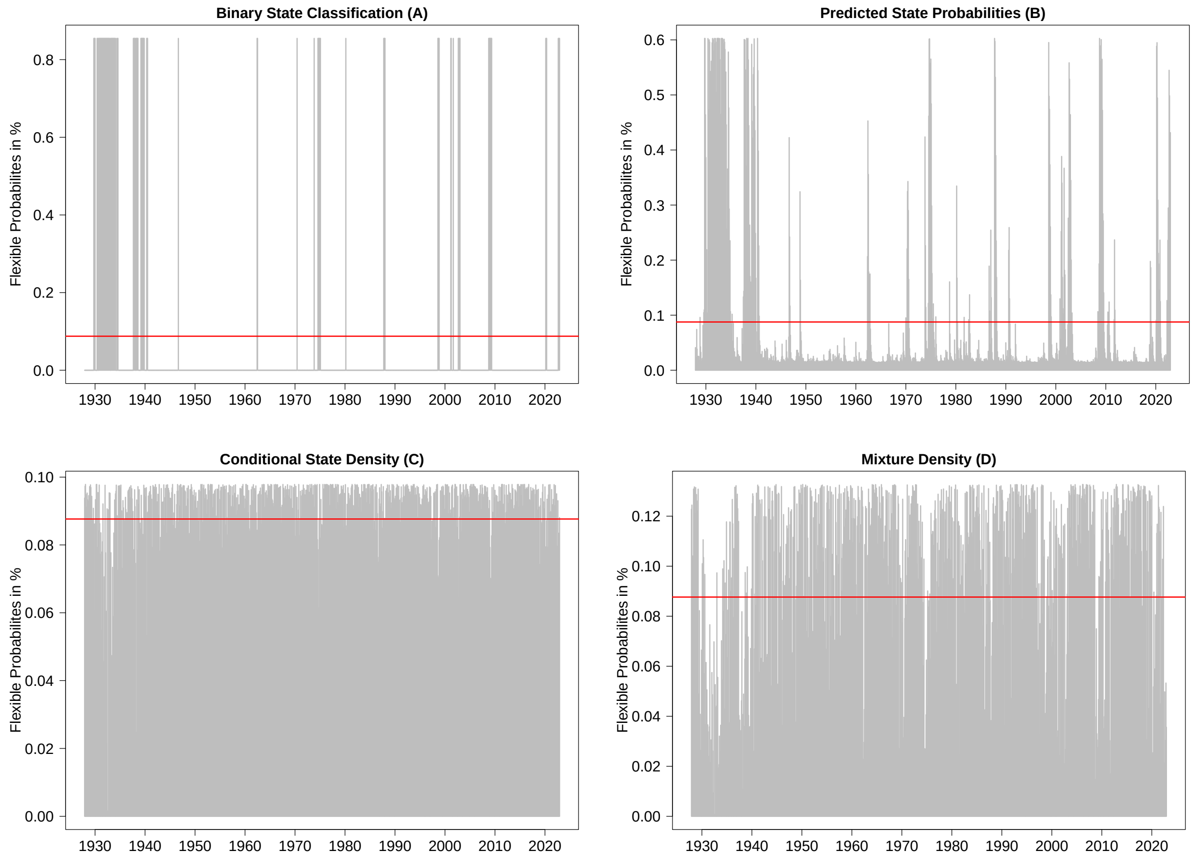


Figure A1: Economic Regimes over Time (continued from previous page)



Notes: Figure displays the regime identification of the ten economic variables. The smoothed probabilities of the turbulent regime is depicted in red (left axis) and is calculated according to the Kim smoother (Kim 1994). For reasons of illustration, we plot the variables in levels or annual growth rate in black (right axis). The actual regime filtering is applied on the transformed variables. We refer to the Section 4 for further details. The in-sample period runs from 1927/12 to 2022/12. Gray shaded areas indicate NBER recessions.

Figure A2: Flexible Probabilities for Excess Returns (ERP)



Notes: Figure shows the flexible probabilities (before entropy pooling) for the turbulent regime of excess stock returns over the entire sample. Four different weighting schemes are used. Panel A uses the state classifier, Panel B the predicted probability, Panel C the conditional state density, and Panel D the mixture density. For comparison, the red line shows an equal-weighting. The in-sample period runs from 1927/12 to 2022/12.

Appendix B: Additional Out-of-Sample Results

Table B1: EP vs. View: Differences in Forecasting Performance

Variable	$R^2_{OS}(EP) - R^2_{OS}(View)$ (1)	(2)	(3)	(4)	$\Delta_{CER}(EP) - \Delta_{CER}(View)$ (1)	(2)	(3)	(4)
Panel 1: Binary State Classification (A)								
ERP	1.12	3.11	1.04	0.99	0.90	0.30	0.97	0.69
GM	0.78	1.93	−0.12	0.67	0.95	0.78	−0.02	0.67
GE	1.72	0.34	1.06	1.59	2.35	0.46	1.51	1.93
DP	0.09	−0.05	0.2	0.09	0.05	0.04	0.25	0.17
RVOL	1.00	1.72	0.53	0.84	0.81	0.47	0.58	0.70
TMS	0.82	3.32	−0.49	0.58	0.92	2.56	0.26	0.56
DFY	−0.34	1.69	−0.17	−0.35	0.09	−0.10	0.34	−0.02
TBL	1.16	1.00	1.68	0.99	1.56	1.41	1.61	1.08
IP	0.48	0.47	0.11	0.40	0.99	−0.06	0.24	0.45
INF	0.10	0.4	0.07	0.16	−0.03	0.36	−0.04	−0.01
Panel 2: Predicted Probabilities (B)								
ERP	0.98	1.54	0.68	0.90	0.87	0.07	0.64	0.67
GM	0.78	1.01	0.11	0.69	1.01	0.46	0.11	0.74
GE	0.88	0.17	0.53	0.77	1.80	0.24	0.67	1.20
DP	0.07	−0.02	0.12	0.07	0.02	0.03	0.18	0.11
RVOL	1.26	1.01	0.58	1.17	0.93	0.34	0.52	0.85
TMS	0.37	2.07	−0.82	0.25	0.91	2.59	−0.42	0.33
DFY	0.18	0.82	0.00	0.14	0.02	0.15	0.34	−0.04
TBL	0.57	0.65	0.85	0.50	0.94	1.26	0.86	0.20
IP	0.08	0.34	−0.09	0.07	0.02	0.05	−0.03	0.04
INF	0.16	0.38	−0.10	0.20	0.04	0.45	−0.20	0.12
Panel 3: Conditional State Density (C)								
ERP	0.02	0.03	0.00	0.03	0.02	0.01	0.03	0.02
GM	0.03	0.00	0.03	0.02	0.05	−0.01	0.04	0.04
GE	1.49	0.04	0.86	1.28	2.67	0.48	1.24	1.99
DP	0.01	−0.38	0.51	0.05	−0.02	−0.48	0.61	0.08
RVOL	−0.01	−0.01	−0.01	0.01	0.10	−0.01	0.09	0.08
TMS	−0.10	0.06	−0.09	−0.04	0.02	0.01	0.02	0.01
DFY	−0.02	0.09	−0.10	−0.10	0.11	0.07	−0.02	0.01
TBL	0.05	0.00	0.26	0.04	−0.02	−0.17	0.40	0.00
IP	0.20	0.05	0.14	0.16	0.32	0.03	0.15	0.20
INF	−0.21	0.09	−0.24	−0.10	−0.48	0.08	−0.28	−0.34
Panel 4: Mixture Density (D)								
ERP	0.01	−0.05	0.22	−0.02	0.16	0.01	0.32	0.08
GM	−0.13	−0.07	−0.03	−0.10	0.00	−0.04	−0.02	−0.10
GE	3.47	0.50	1.42	2.97	3.54	1.33	1.76	2.41
DP	0.01	−0.40	0.50	0.06	−0.02	−0.52	0.68	0.09
RVOL	−0.30	−0.08	−0.19	−0.30	0.00	0.01	−0.15	−0.03
TMS	−0.10	0.00	0.00	−0.05	−0.20	−0.07	0.13	−0.08
DFY	0.17	0.10	0.00	0.17	0.31	0.06	0.17	0.11
TBL	0.10	0.00	0.63	0.02	−0.16	0.04	0.89	−0.06
IP	0.22	−0.03	0.30	0.21	0.20	−0.13	0.44	0.22
INF	−0.20	0.10	−0.30	−0.10	−0.27	0.08	−0.24	−0.25

Notes: Table shows the differences in forecasting performance between entropy pooling (EP) and the plain view-based weighting. In the unrestricted version (1) the SOP forecast consists of variable-specific regime forecasts for earnings and multiple growth. (2) assumes no multiple growth ($\hat{\mu}^{gm} = 0$) and uses only the *flexible regime* forecasts for earnings growth, (3) assumes that earnings growth follows a 20-year moving average ($\bar{ge}^{20Y}$) and the expected multiple growth are is predicted by the *flexible regime* forecasts. Finally (4) computes the return forecasts directly.

Table B2: Average Effective Number of Scenarios (ENS)

	ENS: Flexible Probability Weighting							
		View $\mathbf{p}(\mathbf{v}, \mathbf{k})$				**Entropy Pooling $\mathbf{p}^*(\mathbf{v}, \mathbf{k})$**		
Variable	(A)	(B)	(C)	(D)	(A)	(B)	(C)	(D)
ERP	81.4	87.7	85.7	82.5	94.2	94.9	83.9	89.8
GM	75.5	83.4	81.4	77.2	93.6	94.2	84.9	89.9
GE	56.3	76.4	65.4	62.8	93.6	94.2	84.9	89.9
DP	99.8	99.9	99.8	99.8	99.8	99.8	97.7	97.7
RVOL	77.8	85.5	83.5	78.9	94.0	94.6	84.3	89.3
TMS	50.8	73.6	63.6	58.2	94.0	94.6	84.3	89.3
DFY	70.2	83.0	76.0	73.7	88.2	93.6	84.6	87.5
TBL	60.1	77.3	67.5	64.7	88.1	95.2	87.6	83.7
IP	78.9	89.7	84.3	81.4	90.4	94.7	89.1	92.2
INF	66.6	78.1	75.3	70.6	89.1	91.4	86.6	90.7
Mean	71.7	83.5	78.3	75.0	92.5	94.7	86.8	90.0

Notes: Table shows the average out-of-sample effective number of scenarios for the different flexible probability choices in % of the maximal information content (with $p_t = 1/\tau$). (A) indicates the weighting scheme using the binary state classification, (B) uses the predicted state probabilities, (C) the conditional state density and (D) the mixture state density. In addition, *Mean* shows the average over all variables for each weighting scheme.

Appendix C: Out-of-Sample Robustness

Table C1: Robustness in Δ_{CER} (in %) for unrestricted model (1)

	$w = [0,1]$			$w = [0,1.5]$			$w = [-0.5,1.5]$		
Variable	$\gamma = 1$	$\gamma = 3$	$\gamma = 5$	$\gamma = 1$	$\gamma = 3$	$\gamma = 5$	$\gamma = 1$	$\gamma = 3$	$\gamma = 5$
Panel 1: Benchmarks									
FSC	0.78	1.72	1.46	2.21	2.36	1.78	2.12	2.28	1.74
EW	**1.00**	1.22	0.61	1.92	1.24	0.73	**2.20**	1.44	0.89
Panel 2: Binary State Classification (A)									
ERP	**0.99**	1.68	1.24	1.95	2.05	1.27	**2.22**	2.23	1.42
GM	**0.96**	1.67	1.13	1.89	1.94	1.30	**2.17**	2.12	1.45
GE	**0.92**	1.31	0.86	1.94	1.56	1.02	**2.24**	1.73	1.17
DP	**0.93**	1.23	0.65	1.79	1.29	0.79	2.11	1.57	1.00
RVOL	**0.93**	1.62	1.12	1.87	1.78	1.16	2.04	1.91	1.27
TMS	0.50	0.67	0.14	1.09	0.29	0.14	1.22	0.42	0.24
DFY	0.57	1.27	1.08	1.36	1.83	1.36	1.13	1.59	1.15
TBL	**1.03**	1.33	0.68	1.79	1.26	0.63	2.02	1.39	0.73
IP	**0.86**	1.16	0.46	1.61	1.01	0.52	1.87	1.21	0.69
INF	0.77	1.36	0.86	1.61	1.57	0.89	2.11	1.94	1.14
Panel 3: Predicted Probabilities (B)									
ERP	**1.08**	1.64	1.14	1.91	1.91	1.16	**2.21**	2.09	1.30
GM	**1.00**	1.68	1.08	1.91	1.86	1.21	**2.20**	2.03	1.36
GE	**0.99**	1.22	0.77	1.92	1.42	0.93	**2.20**	1.59	1.09
DP	**0.92**	1.19	0.63	1.77	1.26	0.77	2.10	1.54	0.98
RVOL	**1.06**	1.70	1.32	1.94	2.17	1.50	**2.15**	**2.30**	1.63
TMS	0.60	0.79	0.28	1.10	0.54	0.42	1.26	0.68	0.52
DFY	**0.78**	1.28	1.06	1.55	1.83	1.27	1.60	1.81	1.28
TBL	0.76	1.20	0.71	1.47	1.31	0.86	1.80	1.56	1.05
IP	**0.85**	1.04	0.39	1.70	0.92	0.50	1.95	1.13	0.67
INF	0.70	1.36	0.83	1.56	1.51	0.87	2.05	1.88	1.12
Panel 4: Conditional State Density (C)									
ERP	**0.91**	1.37	0.99	1.60	1.59	1.18	1.71	1.71	1.28
GM	**0.94**	1.30	0.88	1.56	1.42	1.05	1.67	1.55	1.16
GE	**0.97**	1.41	0.61	1.91	1.26	0.68	**2.25**	1.47	0.85
DP	**0.85**	1.23	0.61	1.61	1.20	0.68	1.88	1.37	0.83
RVOL	0.22	1.14	0.88	0.69	1.26	0.65	0.84	1.37	0.72
TMS	0.52	1.03	0.67	1.24	1.21	0.80	1.54	1.38	0.95
DFY	**0.83**	1.30	0.91	1.57	1.50	1.14	1.73	1.62	1.25
TBL	**1.08**	1.62	1.07	2.17	1.86	1.06	**2.55**	2.08	1.24
IP	0.75	1.12	0.50	1.56	1.03	0.53	1.85	1.20	0.68
INF	**1.07**	1.45	0.87	2.00	1.53	0.92	**2.30**	1.71	1.07
Panel 5: Mixture Density (D)									
ERP	**0.99**	1.53	0.93	1.83	1.60	1.02	1.97	1.73	1.13
GM	**0.82**	1.43	0.91	1.57	1.56	0.97	1.73	1.70	1.09
GE	**1.07**	1.11	0.46	1.91	1.00	0.51	**2.21**	1.17	0.66
DP	**0.85**	1.24	0.62	1.61	1.21	0.69	1.88	1.38	0.83
RVOL	0.53	1.43	0.96	1.38	1.61	0.85	1.52	1.75	0.95
TMS	**1.10**	1.34	0.96	1.99	1.58	1.38	**2.17**	1.73	1.51
DFY	**0.86**	1.56	1.06	1.83	1.82	1.15	2.02	1.95	1.28
TBL	0.74	1.40	0.90	1.57	1.48	1.20	1.72	1.62	1.32
IP	**0.80**	0.95	0.35	1.61	0.83	0.46	1.86	1.02	0.61
INF	**0.97**	1.39	0.83	1.80	1.45	0.89	2.08	1.61	1.03

Notes: Table shows the certainty equivalent gains for varying coefficients of risk aversion (γ) and investment constraints (w). Panel 1 displays the two benchmark models FSC (Ferreira and Santa-Clara 2011) and EW ($p_t = \frac{1}{\tau}$). The Panels 2–5 present the results for different weightings using the regime-dynamics of the ten state variables. The unrestricted SOP forecasting model (1) consists of variable-specific regime forecasts for earnings and multiple growth. Superior results compared to FSC are highlighted in bold.

Table C2: Robustness in Δ_{CER} (in %) for restricted version (2) with $\hat{\mu}^{gm} = 0$

	$w=[0,1]$			$w=[0,1.5]$			$w=[-0.5,1.5]$		
Variable	$\gamma=1$	$\gamma=3$	$\gamma=5$	$\gamma=1$	$\gamma=3$	$\gamma=5$	$\gamma=1$	$\gamma=3$	$\gamma=5$
Panel 1: Benchmarks									
FSC	0.78	1.72	1.46	2.21	2.36	1.78	2.12	2.28	1.74
EW	**1.57**	1.34	0.98	**2.72**	1.76	1.20	**2.95**	1.92	1.34
Panel 2: Binary State Classification (A)									
ERP	−0.13	1.02	0.85	0.61	1.33	1.22	0.06	1.07	1.05
GM	0.25	1.15	1.06	1.07	1.68	1.41	1.04	1.73	1.46
GE	**1.32**	1.55	1.36	**2.39**	2.26	1.69	**2.64**	**2.45**	**1.84**
DP	**1.59**	1.36	0.97	**2.76**	1.75	1.18	**3.02**	1.91	1.32
RVOL	**0.90**	1.24	0.98	1.89	1.60	1.32	2.00	1.73	1.43
TMS	**0.85**	0.97	0.63	1.54	1.14	0.82	2.03	1.47	1.07
DFY	0.73	1.58	1.39	1.85	2.18	**1.83**	1.91	2.23	**1.85**
TBL	**2.00**	**1.97**	1.30	**3.06**	2.25	1.18	**3.55**	**2.55**	1.42
IP	**0.80**	1.57	1.20	1.88	1.88	1.24	1.99	2.02	1.33
INF	**1.24**	1.55	1.15	2.06	1.94	1.31	**2.27**	2.08	1.43
Panel 3: Predicted Probabilities (B)									
ERP	−0.06	0.93	0.80	0.65	1.24	1.15	0.37	1.12	1.08
GM	0.48	1.07	1.04	1.18	1.66	1.38	1.08	1.63	1.36
GE	**1.41**	1.50	1.27	**2.60**	2.15	1.58	**2.78**	**2.31**	1.70
DP	**1.59**	1.35	0.99	**2.76**	1.77	1.21	**3.02**	1.93	1.34
RVOL	0.76	1.27	1.04	1.88	1.67	1.38	1.95	1.79	1.49
TMS	**1.04**	1.10	0.73	1.91	1.28	0.96	**2.32**	1.52	1.15
DFY	**1.17**	1.65	1.36	**2.38**	2.13	**1.85**	**2.52**	2.21	**1.90**
TBL	**1.88**	**1.76**	1.23	**2.95**	2.17	1.24	**3.32**	**2.39**	1.41
IP	**1.21**	1.57	1.19	**2.43**	1.97	1.30	**2.62**	2.14	1.43
INF	**1.36**	1.71	1.26	**2.44**	2.11	1.40	**2.62**	2.25	1.51
Panel 4: Conditional State Density (C)									
ERP	0.62	1.64	1.20	1.63	1.86	1.43	1.79	1.99	1.52
GM	**0.86**	1.65	1.18	1.93	1.87	1.42	2.09	2.01	1.53
GE	**1.34**	**1.75**	1.13	**2.67**	2.01	1.27	**2.92**	2.20	1.41
DP	0.74	1.22	0.79	1.73	1.32	1.21	1.89	1.44	1.31
RVOL	**0.99**	1.58	1.16	1.93	1.85	1.45	2.08	2.00	1.56
TMS	0.76	1.39	1.08	1.69	1.72	1.24	1.81	1.84	1.35
DFY	0.65	1.70	1.38	1.67	2.11	1.70	1.82	2.20	**1.78**
TBL	**1.39**	**1.91**	**1.53**	**2.51**	**2.50**	1.59	**2.68**	**2.64**	1.71
IP	0.69	**1.83**	1.34	1.78	2.03	1.42	1.96	2.18	1.53
INF	**1.16**	1.55	1.14	2.19	1.82	1.34	**2.29**	1.94	1.44
Panel 5: Mixture Density (D)									
ERP	0.72	**1.77**	1.25	1.82	1.93	1.45	1.98	2.07	1.54
GM	**0.93**	1.71	1.29	2.09	2.03	1.55	**2.22**	2.18	1.66
GE	**1.33**	1.48	0.77	**2.59**	1.47	0.89	**2.75**	1.64	1.03
DP	0.74	1.18	0.73	1.73	1.23	1.12	1.89	1.35	1.22
RVOL	**1.09**	1.71	1.21	**2.24**	1.93	1.49	**2.36**	2.07	1.60
TMS	**1.04**	1.68	1.17	1.97	1.78	1.46	2.10	1.92	1.56
DFY	0.58	**1.83**	1.27	1.69	2.01	1.51	1.86	2.15	1.60
TBL	**0.82**	1.54	1.03	1.67	1.60	1.29	1.78	1.72	1.39
IP	**0.94**	**1.76**	1.23	2.20	1.90	1.40	**2.31**	2.04	1.51
INF	**1.25**	**1.73**	1.23	**2.56**	2.04	1.48	**2.68**	2.17	1.59

Notes: Table shows the certainty equivalent gains for varying coefficients of risk aversion (γ) and investment constraints (w). Panel 1 displays the two benchmark models FSC (Ferreira and Santa-Clara 2011) and EW ($p_t = \frac{1}{\tau}$). The Panels 2–5 present the results for different weightings using the regime-dynamics of the ten state variables. The forecasting model is restricted on the multiple growth $\hat{\mu}^{gm} = 0$ such that the SOP forecast consists of variable-specific regime forecasts only for earnings growth. Superior results compared to FSC are highlighted in bold.

Table C3: Robustness in Δ_{CER} (in %) restricted version (3) with $\hat{\mu}^{ge} = \bar{ge}^{20Y}$

	$w=[0,1]$			$w=[0,1.5]$			$w=[-0.5,1.5]$		
Variable	$\gamma=1$	$\gamma=3$	$\gamma=5$	$\gamma=1$	$\gamma=3$	$\gamma=5$	$\gamma=1$	$\gamma=3$	$\gamma=5$
Panel 1: Benchmarks									
FSC	0.78	1.72	1.46	2.21	2.36	1.78	2.12	2.28	1.74
EW	0.60	1.41	1.12	1.69	1.84	1.40	1.75	1.94	1.48
Panel 2: Binary State Classification (A)									
ERP	**0.93**	1.51	1.21	1.97	1.99	1.27	2.09	2.11	1.37
GM	**0.82**	1.31	1.06	1.80	1.70	1.44	1.88	1.80	1.52
GE	0.78	1.33	0.98	1.69	1.64	1.12	1.69	1.70	1.17
DP	0.76	1.56	1.22	1.74	2.02	1.49	1.86	2.15	1.57
RVOL	0.63	1.66	1.25	1.82	1.97	1.55	1.95	2.09	1.63
TMS	0.59	0.88	0.24	1.49	0.47	−0.03	1.49	0.50	−0.01
DFY	0.16	0.69	0.53	0.89	1.02	0.77	0.72	0.88	0.63
TBL	0.66	1.00	0.86	1.45	1.53	0.93	1.39	1.54	0.97
IP	0.51	0.97	0.75	1.45	1.39	0.99	1.62	1.51	1.11
INF	**1.14**	1.61	1.26	**2.33**	2.08	1.48	**2.47**	2.21	1.56
Panel 3: Predicted Probabilities (B)									
ERP	**0.96**	1.51	1.24	1.97	2.00	1.40	2.10	2.12	1.50
GM	0.75	1.48	1.24	1.76	2.02	1.53	1.85	2.12	1.61
GE	0.73	1.27	0.97	1.61	1.59	1.22	1.64	1.69	1.30
DP	0.68	1.47	1.17	1.59	1.93	1.44	1.71	2.06	1.53
RVOL	0.72	**1.83**	**1.51**	1.96	**2.38**	**1.82**	2.07	**2.49**	**1.90**
TMS	0.59	0.88	0.54	1.33	0.88	0.62	1.35	0.92	0.65
DFY	0.36	0.97	0.73	1.24	1.31	0.96	1.25	1.39	1.02
TBL	0.54	0.92	0.79	1.46	1.28	1.13	1.39	1.30	1.15
IP	0.72	1.06	0.86	1.55	1.49	1.11	1.69	1.60	1.21
INF	**0.92**	1.40	1.12	1.92	1.80	1.38	**2.15**	1.96	1.48
Panel 4: Conditional State Density (C)									
ERP	0.75	1.11	0.85	1.59	1.52	1.09	1.58	1.53	1.12
GM	0.66	1.16	1.02	1.62	1.62	1.30	1.55	1.64	1.32
GE	0.58	1.18	0.97	1.47	1.61	1.21	1.44	1.64	1.25
DP	0.62	1.32	1.04	1.46	1.77	1.27	1.72	1.98	1.43
RVOL	−0.13	1.02	0.97	0.15	1.39	0.99	0.15	1.40	1.00
TMS	**1.02**	1.41	0.97	2.21	1.78	1.20	**2.16**	1.80	1.23
DFY	0.43	0.83	0.56	1.29	1.09	0.78	1.28	1.19	0.86
TBL	**0.89**	1.18	0.97	1.97	1.65	1.21	1.89	1.66	1.23
IP	0.38	0.65	0.54	1.20	1.05	0.77	1.25	1.10	0.85
INF	**1.01**	1.43	1.18	2.18	1.98	1.37	**2.27**	2.12	1.47
Panel 5: Mixture Density (D)									
ERP	0.76	1.11	0.87	1.70	1.54	1.08	1.64	1.53	1.10
GM	0.43	1.15	0.95	1.41	1.56	1.21	1.33	1.57	1.24
GE	**0.89**	1.41	1.14	1.83	1.93	1.41	1.82	1.96	1.44
DP	0.66	1.39	1.11	1.56	1.89	1.34	1.80	2.09	1.50
RVOL	0.12	1.18	1.05	0.72	1.61	1.04	0.84	1.69	1.09
TMS	**1.16**	1.32	1.02	2.20	1.80	1.33	**2.20**	1.76	1.32
DFY	**0.78**	0.99	0.79	1.74	1.44	0.99	1.75	1.43	1.01
TBL	0.75	1.49	1.16	2.06	1.97	1.43	2.03	1.96	1.43
IP	0.41	0.81	0.64	1.31	1.23	0.86	1.36	1.26	0.92
INF	0.39	1.26	1.00	1.36	1.67	1.25	1.23	1.66	1.23

Notes: Table shows the certainty equivalent gains for varying coefficients of risk aversion (γ) and investment constraints (w). Panel 1 displays the both benchmark models FSC (Ferreira and Santa-Clara 2011) and EW ($p_t = \frac{1}{\tau}$). The Panels 2–5 present the results for different weightings using the regime-dynamics of the ten state variables. The forecasting model is restricted on the earnings growth ($\bar{ge}^{20Y}$) such that the SOP forecast consists of variable-specific regime forecasts only for multiple growth. Superior results compared to FSC are highlighted in bold.