Conditional Mean Reversion of Financial Ratios and the Predictability of Returns

Boucher, C.[*], Jasinski, A.[†], Tokpavi, S.[‡]

December, 2021

Abstract

While traditional predictive regressions for stock returns using financial ratios are empirically proven to be valuable at long-term horizons, evidence of predictability at few-month horizons is still weak. In this paper, based on the empirical regularity of a typical dynamic of stock returns following the occurrence of a mean reversion in the US Shiller CAPE ratio when the latter is high, we propose a new predictive regression model that exploits this stylized fact. In-sample regressions approximating the occurrence of mean reversion by the smoothed probability from a regime-switching model show superior predictive powers of the new specification at few-month horizons. These results also hold out-of-sample, exploiting the link between the term spread as business cycle variable and the occurrence of mean reversion in the US Shiller CAPE ratio. For instance, the out-of-sample R-squared of the new predictive regression model is ten (four) times bigger than that of the traditional predictive model at a 6 (12) month horizon. Our results are robust with respect to the choice of the valuation ratio (CAPE, excess CAPE or dividend yield), and countries (Canada, France, Germany and the UK). We also conduct a mean-variance asset allocation exercise which confirms the superiority of the new predictive regression in terms of utility gain.

JEL Classification: G12, G17, C53, E32.
Keywords: Return predictability, Valuation ratios, Mean reversion, Business cycle, Term spread.

[*]christophe.boucher@parisnanterre.fr, EconomiX-UPL, CNRS, University Paris Nanterre
[†]ajasinski@parisnanterre.fr, EconomiX-UPL, CNRS, University Paris Nanterre
[‡]Corresponding author, sessi.tokpavi@univ-orleans.fr, LEO, University of Orléans.

1 Introduction

The predictability of stock returns is of great importance not only for practitioners but also for academics with important implications for financial models of risk and return. In this paper, we provide evidence that financial ratios can predict in-sample and out-of-sample returns at a few month horizons, when considering that these financial ratios are persistent during some business cycle phases, while mean-reverting around other phases, and exploiting the informational content of the term spread about these phases.

The Cyclically-Adjusted-Price to Earnings (CAPE) ratio of Campbell and Shiller (1988), is well-known in characterising the strong relationship between an inflation adjusted earnings-price ratio and subsequent long-term returns. It has now become an often cited and followed measure of long-term equity market valuation by both academics and practitioners. More generally, a prolific number of academic papers have focussed on the usefulness of financial ratios for forecasting future stock market returns at multi-year horizons, including price-earnings (P/E), CAPE, dividend yield as well as book-to-market ratios (Rozeff, 1984; Fama and French, 1988; Campbell and Shiller, 1988; Cochrane, 1991; Hodrick, 1992; Goetzmann and Jorion, 1993; Lewellen, 2004, etc.). Moreover, these studies conclude that growth rates of fundamentals, such as dividends or earnings, are much less forecastable than returns, suggesting that most of the variation of financial ratios is due to variations in expected returns through mean reversion.

Most of the existing empirical evidence has shown that this relation holds only for long-term stock market returns, with the consequence that P/E ratios revert to their historical average values over long horizons (Campbell and Shiller, 1998; Weigand and Irons, 2007, etc.). For instance, Campbell and Shiller (1998) in their seminal paper showed that P/E ratios have considerable explanatory power in predicting only long-horizon future returns, with an explanatory power (as measured by the R-squared) of 20% or more in regressing future 4- and 5-year stock returns on initial P/Es. See also Weigand and Irons (2007) who studied a very long dataset (1871-2004) and found that high P/E ratios in the stock market are generally followed by a decade of lower than average real returns.[1]

However, as shown in Lettau and Van Nieuwerburgh (2008), the short-term predictive content in the valuation ratios can be recovered, if one relaxes the assumption of a fixed steady state mean of the economy. In other words, by assuming that the mean of a valuation ratio is regime-specific rather than global, short-term mean reversions can arise with a statistically significant predictive ability of the financial ratio for short-horizon returns. They indeed reported that dealing with regime changes by adjusting valuation ratios to their steady-state values increases their power in predicting returns over the next year.

[1] A wide literature exists for the dividend price (dividend yield) as valuation ratio. See for example Campbell and Shiller (1988), Cochrane (1991), Fama and French (1988), Rozeff (1984), Lewellen (2004). These papers also conclude in predictability for long-horizon returns.

These findings have been analyzed from a more structural point of view by recent papers which try to link regime changes in the dynamic of valuation ratios to variables measuring the state of the business cycle. For instance, Arnott et al. (2017) by assuming that P/Es mean-revert toward levels that are suggested by macroeconomic conditions, rather than toward long-term averages, found that moderate rather than rock-bottom levels of inflation and real interest rates are associated with the highest valuation multiples. By incorporating these features in predictive regressions, they obtained significative improvements in the short-term forecasting power of the Shiller CAPE ratio for the US and other developed markets. See also Boucher (2006).

Against this background, our goal in this paper is to achieve short-term predictability of stocks' returns, but using a different approach. The core of our approach is that if one succeeds in identifying the *occurrence* of mean reversion in valuation ratios, the short-term predictability of returns can be recovered, based on the idea that the dynamic of returns following the occurrence of a mean reversion is usually different from the overall one. Our empirical investigations reveal indeed that average multi-period returns following a mean reversion in the US Shiller CAPE ratio are negative and range from -1.77% at 1 month up to -3.35% at 6 months. Episodes of mean reversion are identified by the levels of the smoothed probability estimated from a regime switching version of the mean reversion model of Jegadeesh (1991). Interestingly, this pattern appears more typical, when mean reversion episodes are associated with high levels of the CAPE ratio, with subsequent average multi-period returns of -16.46% at twelve months approximately. Predictive regressions exploiting the latter stylized fact show clear-cut superior predictive power at short-term horizons compared to the traditional predictive regression. For illustration, while the adjusted R-squared of the traditional predictive regression ranges from 0.03% (1 month) to 3.18% (12 months), the same statistic ranges from 7.65% (one month) to 11.83% (12 months).

One limitation of the above results is that the predictive powers are evaluated in-sample using the level of smoothed probabilities as an indicator of mean reversion regimes. Hence, they are not exploitable out-of-sample, because the smoothed probabilities estimated from the regime switching model are based on the whole available sample. To keep the power of our predictive regression out-of-sample, we use a simple strategy that consists of using a business cycle variable with a strong early-warning property regarding the occurrence of mean reversion in the US Shiller CAPE ratio, i.e. the US term spread.

The rationale of using the term spread springs from two pieces of evidence, i.e. the contemporaneous link between mean reversion in valuation ratios and economic recession, and the predictive power of term spread on the occurrence of economic recession. The first evidence is corroborated by some works that reported the strong predictive power of valuation ratios during recession (Rapach et al., 2010; Henkel et al., 2011; Dangl and Halling, 2012). As for the second evidence, there is an abundant literature that highlights the early-warning nature of term spread on the occurrence of economic recession (Stock and Watson, 1989; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998). The theoretical basis of this typical behaviour goes back to the work of Kessel (1965) who reported the cyclical behavior of the term spread and investigated the common variation of the term structure of interest rates and business cycles. Fama (1986) argued that this relationship could be consistent with the liquidity preference hypothesis and could be explained in an intertemporal CAPM framework. Harvey (1988) provided analytical evidence that the term spread was related to future consumption growth under the Consumption CAPM (CCAPM) framework. More recently, in a dynamic model with rational expectations, Estrella (2005) showed that the term spread contains information about expectations of future activity and is affected by current monetary policy.

Hence, we use lagged values of the term spread as an indicator of mean reversion in the US Shiller CAPE ratio, and observe that the forecast ability of our new predictive regression over short-term horizons continues to hold out-of-sample. Indeed, while the predictive powers of the traditional predictive regression, as measured by the out-of-sample R-squared, are very low, even negative at very short-term horizons, the new predictive regression shows higher predictive powers at the same horizons. Specifically, at the horizon of 1 (12) month, the out-of-sample R-squared is equal to -0.72% (2.28%) for the traditional regression, while it is equal to 0.64% (8.73%) for the new predictive regression. These findings also hold when considering other financial variables including the excess CAPE yield and the dividend yield, and other countries (Canada, France, Germany and the UK).

We also conduct a mean-variance asset allocation exercise which confirms the superiority of the new predictive regression in terms of utility gain. For instance, at the 1 month horizon and with a relative risk aversion parameter equal to 3, the utility gains or the annual portfolio management fees that an investor would be willing to pay to switch from the traditional model to the new proposed model, are equal to 2.06%, 3.51%, 0.90%, 2.04% and 1.54% for the US, the UK, France, Germany and Canada, respectively.

Our contribution can be linked to a branch of the literature which sets the objective of increasing the short-term predictive power of valuation ratios regarding risk premiums, using models with time-varying parameters that fit business cycles, and specifically recession and expansion phases (Rapach et al., 2010; Henkel et al., 2011; Dangl and Halling, 2012; Gomez Cram, 2021). For example, Rapach et al. (2010) and Dangl and Halling (2012) documented that excess stock return predictability by the dividend-price ratio and the earnings-price ratio concentrates mostly in recessions, with valuation ratios having higher predictive power during recessions. Henkel et al. (2011) also provided the same evidence with the short-horizon performance of aggregate return predictors, such as the dividend yield and the short rate, that appear non-existent during business cycle expansions, but sizeable during contractions, with the phenomenon related to countercyclical risk premiums as well as the time-variation in the dynamic of the predictors. Similarly, we relate the predictability of returns based on valuation ratios to the state of the business cycle, with the latter approximated via the early-warning property of the term spread regarding mean reversion in financial ratios. This contrasts with the contributions cited above which impose tight parametric restrictions on how predictive coefficients in their dynamic models evolve over time. In our framework, the dynamic comes from the term spread which helps in identifying a financial ratio's mean reversion in a forward looking manner.

Our results suggest that using the informational content of the term spread regarding the occurrence of mean reversion in valuation ratios, helps to improve the short-term predictability of stock returns. In this line, our paper shares the same objective as that of Moench and Tobias (2021) which confirmed the importance of the term spread for equity premium forecasts. Using recession probability forecasts based on the term spread as an explanatory variable, they achieved improvement in the equity premium predictability at short-term horizons. While the improvement as measured by the out-of-sample R-squared is of the same order as in our predictive regression, our out-of-sample approach is simpler, because it is based on a single step regression with only observed and non-estimated variables, and is therefore robust to estimation risk.

The rest of the paper proceeds as follows. Section 2 develops a dynamic mean reversion model in valuation ratios to identify a mean reversion regime, and analyses stocks' returns dynamic following mean reversion. Based on the empirical findings, our new predictive regression is introduced in this section and its predictive power is evaluated in-sample. Section 3 investigates the out-of-sample forecast ability of the new predictive regression, and Section 4 is devoted to robustness checks regarding the choice of the financial ratio. Section 5 deals with international evidence, and Section 6 analyzes the implications for asset allocation. The last section concludes.

2 Mean reversion in valuation ratios and in-sample short-term predictability of returns

This first section analyses the informational content of mean reversion in valuation ratios for the short-term dynamic of returns. The first part of the section provides a non-structural model for the occurrence of mean reversion in valuation ratios, and the second part evaluates to what extent this occurence has predictive power (in-sample) for the short-term dynamic of stock prices.

2.1 Dynamic model for mean reversion in valuation ratios

For the description of our model of mean reversion in valuation ratios, let x_t be the natural logarithm of a given valuation ratio, here the US Shiller CAPE ratio recorded at month t. Unit root tests are the usual tools to check for mean reversion in a time series. Indeed, if x_t is nonstationary, it will exhibit no tendency to return to a long-run mean. This is the approach followed by Becker et al. (2012). Using unit roots and multiple structural break tests, they show that the P/E ratio is nonstationary globally, but is stationary around multiple breaks, which implies that this ratio will eventually revert to some local long-run means, confirming the regime-specific dynamic of valuation ratios as stressed by Lettau and Van Nieuwerburgh (2008). Although this approach is the most used in the literature, we do not follow it because it is about stationarity, and does not provide a model for the occurrence of the mean reversion phenomenon. We instead follow the methodology in Jegadeesh (1991) which provides a simple way to model mean reversion in a given time series through linear regression. For x_t the regression writes:

$$dx_t = \alpha_k + \beta_k \left(\sum_{s=1}^{k} dx_{t-s} \right) + \epsilon_{k,t}, \tag{1}$$

where $dx_t = x_t - x_{t-1}$ is the first difference of x_t, k is the holding period[2] and $\epsilon_{k,t}$ is an error term. The parameter of interest is β_k indexed by the holding period. Indeed, mean reversion occurs when $\beta_k < 0$, with the current value of x_t which adjusts to the past value x_{t-1} with regards to the level of lagged multi-period variations, i.e. $\sum_{s=1}^{k} dx_{t-s}$.

Table 1: Estimation of the mean reversion parameter for the Shiller CAPE ratio

k	3	6	12	24	36	48	60	120
Estimates	0.0603***	0.0487***	0.0263***	0.0039	0.0037	0.0027	−0.0002	−0.0007
Std. Err.	0.0187	0.0126	0.0096	0.0074	0.0065	0.0048	0.0042	0.0030
R^2 (in %)	1.48	2.10	1.46	0.06	0.08	0.05	0.00	0.01

Notes: For different values of the holding period k, the table displays the parameter estimates of the slope parameter β_k in the mean reversion equation (1), followed by the Newey-West robust standard errors. The table also reports the explanatory power as given by the R-squared. *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively.

Table 1 displays the estimates of β_k using monthly data of the US Shiller CAPE ratio over a very long dataset (February, 1881 to April, 2020). The evolution of this ratio is displayed in figure 1. For $k \in \{3, 6, 12\}$, the parameter β_k is positive and statistically significant at the 1% nominal risk level. For the other values of k, the same parameter is not significant at the usual nominal risk level. These results suggest the absence of mean reversion in the US

[2]In the empirical applications, we will consider different values of the holding periods, $k \in \{3, 6, 12, 24, 36, 48, 60, 120\}$, corresponding to one quarter, one semester, and one, two, three, four, five and ten years, respectively.

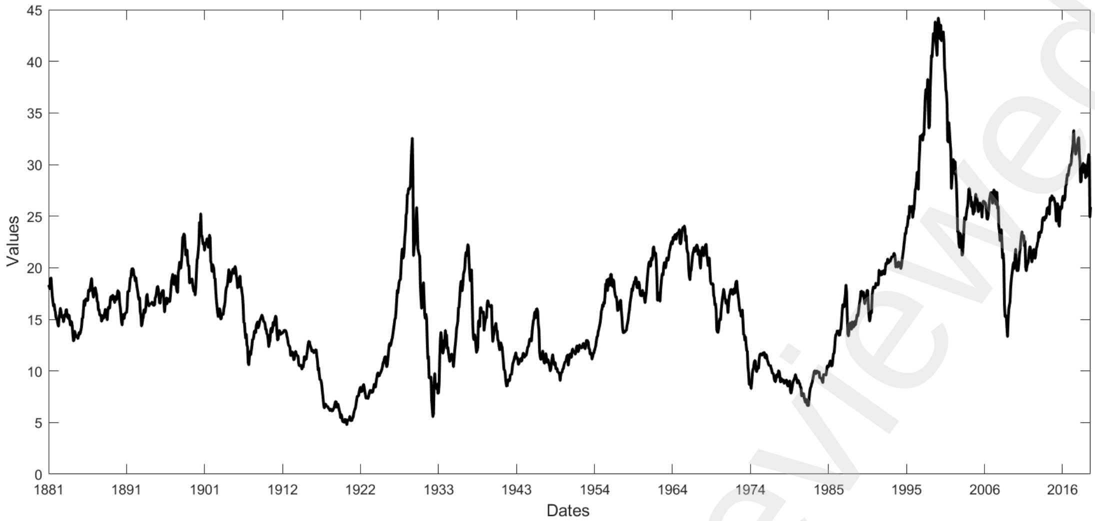

Figure 1: Dynamic of the US Shiller CAPE ratio: 1881/02-2020/04

Notes: The ratio is computed based on a dataset that consists of monthly stock index prices, earnings data and the consumer price index (to allow conversion to real values). Monthly earnings data are computed from the S&P four-quarter totals for the quarter since 1926, with linear interpolation to monthly figures. Earnings data before 1926 are from Cowles and associates, interpolated from annual data. Stock price data are monthly averages of daily closing prices. The CPI-U (Consumer Price Index-All Urban Consumers) published by the US Bureau of Labor Statistics begins in 1913; the years before 1913 come from the CPI Warren and Pearson's price index.

Shiller CAPE ratio over the whole sample. Moreover, the explanatory power of the mean reversion equation as given by the value of the adjusted R-squared is overall very low.

This absence of mean reversion in the valuation ratio can result from the existence of instabilities in the estimated relationship, materialized by regime changes. To capture regime shifts, we consider the following Markov-switching extension (Goldfeld and Quandt, 1973; Hamilton, 1989, 1994; Kim and Nelson, 1999) of the mean reversion equation (1)

$$dx_t = \alpha_{k,S_{k,t}} + \beta_{k,S_{k,t}} \left(\sum_{s=1}^{k} dx_{t-s} \right) + \epsilon_{k,t}, \tag{2}$$

with $S_{k,t} \in \{0, 1\}$ a latent binary state variable which takes value 0 (1) when the first (second) regime is at stake. This state variable follows a first order Markov chain with the following transition matrix:

$$P = \begin{bmatrix} \Pr(S_{k,t} = 0 \,|S_{k,t-1} = 0) & \Pr(S_{k,t} = 1 \,|S_{k,t-1} = 0) \\ \Pr(S_{k,t} = 0 \,|S_{k,t-1} = 1) & \Pr(S_{k,t} = 1 \,|S_{k,t-1} = 1) \end{bmatrix} = \begin{bmatrix} p_{k,00} & p_{k,01} \\ p_{k,10} & p_{k,11} \end{bmatrix} \tag{3}$$

where $p_{k,ij}$, $(i, j = 0, 1)$ denote the transition probabilities of $S_{k,t} = j$ given that $S_{k,t-1} = i$, with the equality $p_{k,i0} + p_{k,i1} = 1$. The transition matrix governs the random behavior of the state variable and is characterized by only two parameters, $p_{k,00}$ and $p_{k,11}$. For the estimation, we make the assumption of a zero-mean Gaussian distribution for the random error term $\epsilon_{k,t}$, with regime-specific variances, i.e. $\epsilon_{k,t} \sim (0, \sigma_{k,S_t})$. The full set of parameters is given by the vector $\theta = (\alpha_{k,0}, \alpha_{k,1}, \beta_{k,0}, \beta_{k,1}, \sigma_{k,0}, \sigma_{k,1}, p_{k,00}, p_{k,11})'$. This vector of parameters can be estimated by the method of quasi-maximum likelihood.

Table 2: Regime-switching estimation of the mean reversion equation

k	$\alpha_{k,0}$	$\alpha_{k,1}$	$\beta_{k,0}$	$\beta_{k,1}$	$\sigma_{k,0}$	$\sigma_{k,1}$	$dur_{k,0}$	$dur_{k,1}$
3	0.0037***	−0.0231***	0.0613***	−0.0044	0.0009***	0.0060***	41.54	7.57
	(0.0008)	(0.0060)	(0.0131)	(0.0356)	(0.0000)	(0.0003)		
6	0.0024***	0.0009	0.0519***	0.0404	0.0009***	0.0069***	44.12	7.12
	(0.0008)	(0.0073)	(0.0074)	(0.0305)	(0.0000)	(0.0005)		
12	0.0037***	−0.0251***	0.0232***	−0.0117	0.0009***	0.0062***	38.14	6.59
	(0.0008)	(0.0068)	(0.0046)	(0.0153)	(0.0000)	(0.0004)		
24	0.0051***	−0.0334***	0.0016	−0.0366***	0.0009***	0.0057***	37.04	6.83
	(0.0008)	(0.0071)	(0.0033)	(0.0112)	(0.0000)	(0.0004)		
36	0.0049***	−0.0355***	0.0073***	−0.0435***	0.0009***	0.0055***	35.66	6.58
	(0.0009)	(0.0065)	(0.0027)	(0.0082)	(0.0000)	(0.0004)		
48	0.0050***	−0.0326***	0.0095***	−0.0431***	0.0008***	0.0053***	31.64	6.31
	(0.0008)	(0.0059)	(0.0023)	(0.0089)	(0.0000)	(0.0003)		
60	0.0041***	−0.0008	0.0057***	−0.0031	0.0008***	0.0063***	36.91	7.40
	(0.0009)	(0.0059)	(0.0021)	(0.0106)	(0.0000)	(0.0005)		
120	0.0051***	−0.0233***	0.0007	0.0124	0.0008***	0.0060***	31.26	6.31
	(0.0009)	(0.0057)	(0.0016)	(0.0088)	(0.0000)	(0.0004)		

Notes: For different values of the holding period k, the table displays the results of the estimation of the regime-switching mean reversion equation as given in (2). The parameter estimates are given followed by the standard error in parentheses. *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively.

The estimation results (except for the transition probabilities to save space) of this model are displayed in table 2 for the different values of the holding period parameter k. Focusing on the mean reversion parameters $\beta_{k,0}$ and $\beta_{k,1}$, we can observe that a regime's separation with the absence (presence) of mean reversion in the first (second) regime only occurs with $k \in \{24, 36, 48\}$. Indeed in these cases, $\beta_{k,1}$ is statistically significant and negative suggesting mean reversion in the second regime given by $S_{k,t} = 1$. At the same time $\beta_{k,0}$ is either statistically insignificant or significant and positive, indicating the absence of mean reversion in the first regime $S_{k,t} = 0$. In particular, for $k = 36, 48$, this parameter is positive and statistically significant. This means that rather than being mean reverting, the process of the US Shiller CAPE ratio is persistent for these values of the holding period k. Indeed, as largely discussed by Marques (2004), mean reversion and persistence are inversely related, as high persistence implies low mean reversion and vice-versa. Note also that the regime-specific variances of the error term are different, with estimated values ten times larger in the mean reversion regime $S_{k,t} = 1$.

Looking at the magnitude of the estimated values of $\beta_{k,1}$, results suggest that the strongest mean reversion phenomenon occurs with $k = 36$ (36 months or 3 years). With this value of the holding period k, the estimated values of the parameters $p_{k,00}$ and $p_{k,11}$ are equal to 97.20% and 84.80%, respectively. This means that the unconditional probability in staying in the first (second) no mean reversion (mean reversion) regime is equal to 97.20% (84.80%). The estimated values of expected durations for the two regimes are thus equal to:

$$dur_{k,0} = \frac{1}{1 - \widehat{p}_{k,00}} = 35.66, \tag{4}$$

$$dur_{k,1} = \frac{1}{1 - \widehat{p}_{k,11}} = 6.58. \tag{5}$$

Thus, compared to the absence of mean reversion, the presence of mean reversion is a short-lived event that lasts approximately half-a year. Figure 2 which displays the estimates of the smoothed probabilities $\widehat{\Pr}(S_{k,t} = i \Big| \Omega_T; \widehat{\theta})$, $i = 0, 1$ of the two regimes, confirms this stylized fact, with the probability of staying in the first regime (absence of mean reversion) taking values close to one for a much longer period.[3]

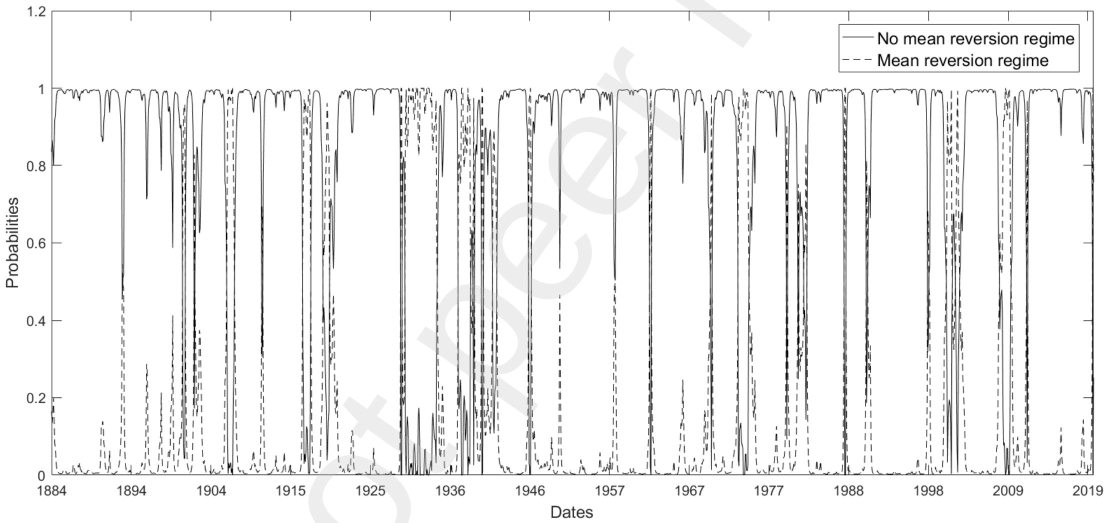

Figure 2: Dynamic of smoothed probabilities: 1884/02-2020/04

Notes: The figure displays the smoothed probabilities of the two regimes (no mean reversion/mean reversion) that result from the estimation of the regime-switching mean reversion equation in (2). The estimation sample ranges from February, 1884 to April, 2020, with a total of 1635 monthly observations.

2.2 Return's dynamic following mean reversion and in-sample predictive regressions

A question of interest, which is at the heart of our methodology, is to evaluate the dynamic of the returns on the stock index in the period following the occurrence of a mean reversion state or regime, and to compare it to the same dynamic in the opposite state or regime (no mean reversion). Formally, let τ be a given horizon in month and $\widehat{\Pr}(S_{k,t} = i \Big| \Omega_T; \widehat{\theta})$

[3]The smoothed probability of a given regime corresponds to the likelihood of this regime at a given time t conditional to the set Ω_T of all available information from $t = 1$ to $t = T$, with T the sample length which here equals $T = 1635$ (monthly data from 1884/02 to 2020/04).

the estimated smoothed probability of regime i at time t, with θ the vector of parameters.[4] Denote $Z_{t,i} = \mathbb{I}\left(\widehat{\Pr}(S_{k,t} = i \middle| \Omega_T; \widehat{\theta}) > \gamma\right)$ the dummy variable taking value one when the smoothed probability of regime i is large and higher than a threshold γ at time t, with $\gamma \in \{0.5, 0.6, 0.7\}$. Thus, this variable indicates whether regime i prevails or nor at time t. For a fixed time t with $Z_{t,i} = 1$, let us compute the multi-period return of the stock index in the subsequent period, i.e.

$$r_{t+1:t+\tau,i} = \sum_{s=t+1}^{t+\tau} r_s, \tag{6}$$

with r_s the monthly log-return of the index. By denoting n_i the number of observations (over the entire sample of length T) with $Z_{t,i} = 1$, we can compute the average value of $r_{t+1:t+\tau,i}$ given by:

$$\overline{r}_{i,\tau} = \frac{1}{n_i} \sum_{s=1}^{n_i} r_{t_s+1:t_s+\tau,i}. \tag{7}$$

Our goal is to compare $\overline{r}_{0,\tau}$ and $\overline{r}_{1,\tau}$ for the different values of τ, where $\overline{r}_{0,\tau}$ ($\overline{r}_{1,\tau}$) provides the average value of multi-period returns following a state without (with) mean reversion in the US Shiller CAPE ratio.

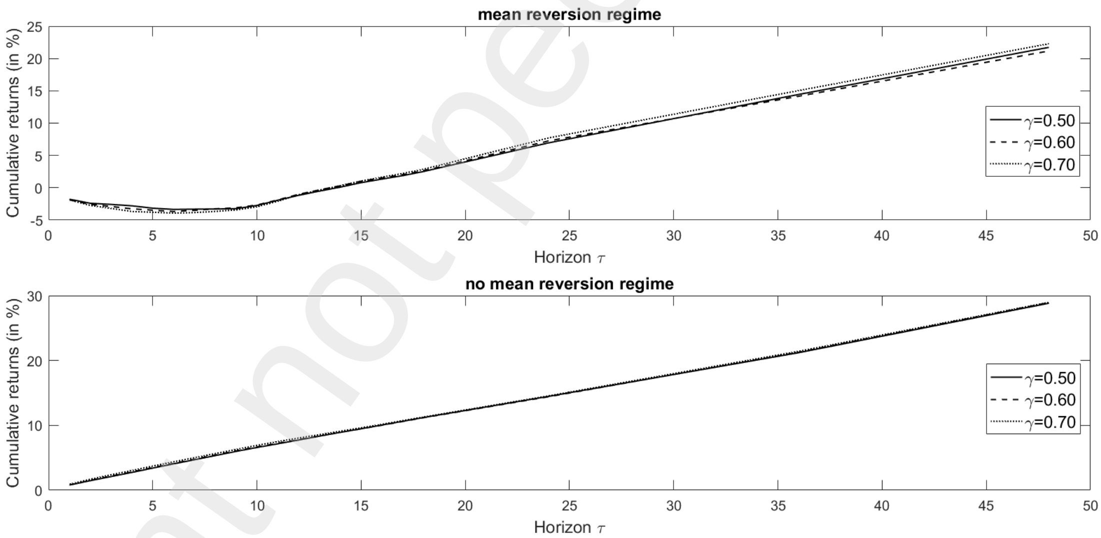

Figure 3: Prevaling regimes and subsequent S&P 500 average returns

Notes: The figure compares subsequent S&P 500 average multi-period returns following the prevalence of a given regime (presence or absence of mean reversion) in the US Shiller CAPE ratio. Multi-period returns are indexed by the horizon τ in months, and the prevalence of a regime is given by the related smoothed probability exceeding a high threshold value $\gamma \in \{0.5, 0.6, 0.7\}$. The top (bottom) panel displays the average multi-period returns for the mean (no mean) reversion regime. Smoothed probabilities displayed in figure 2 are obtained from the estimation of the regime-switching mean reversion model in (2) using monthly data from February, 1884 to April, 2020, with a total of 1635 monthly observations.

Figure 3 compares these values (in %) for $\tau \in \{1 : 18, 24, 36, 48\}$ corresponding to one

[4] We set here k to value 36 which corresponds to the best separation of the two regimes as displayed in table 2.

to eighteen months, two, three and four years. The figure reveals interesting stylized facts. First, the top panel shows that average multi-period returns following a mean reversion regime are negative from values of τ ranging from 1 month to 13 months. For instance, with $\tau = 1$ (1 month) and $\gamma = 0.5$, the recorded average multi-period returns is equal to -1.77%, and this value decreases up to -3.35% for $\tau = 6$ (6 months), followed by decreases in loss for higher values of τ. Precisely, for $\tau \geq 14$ months, the realized average multi-period returns become positive. Second, the bottom panel shows that in the absence of mean reversion in the US Shiller CAPE ratio, subsequent multi-period returns are always positive, with reported values increasing monotonously with τ.

These results provide strong evidence about short-term predictability of S&P 500 returns based on the occurrence of a mean reversion state in the US Shiller CAPE ratio. In other words, if we know that a mean reversion regime prevails at a given time, we will be able to predict market downturns in the subsequent months (1 month to approximately 1 year), with the most severe cumulative loss in the sixth month. To confirm and evaluate the strength of this predictive power, we estimate the following stock return predictive regression for different values of the prediction horizon τ

$$r_{t+1:t+\tau} = a_0 + a_1 \widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta}) + u_{t+1:t+\tau}, \tag{8}$$

where again x_t is the natural logarithm of the US Shiller CAPE ratio, $\widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta})$ is the estimated probability of mean reversion regime at time t, a_0, a_1 some parameters, and $u_{t+1:t+\tau}$ the error term. It is worth noting that our predictive regression equation in (8) differs from the traditional equation which links multi-period returns to current values of valuation ratios as follows:

$$r_{t+1:t+\tau} = b_0 + b_1 x_t + \xi_{t+1:t+\tau}, \tag{9}$$

with b_0, b_1 some parameters, and $\xi_{t+1:t+\tau}$ the error term. The difference arises from using as predictor the prevalence of a mean reversion regime at time t as given by the estimated smoothed probability $\widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta})$ instead of the level of the valuation ratio. By doing so, our goal is to exploit the stylized facts observed in figure 3 which indicate that high and low values of this probability lead to different subsequent dynamics for stock prices.

Table 3 displays the estimation results of our predictive regression. We report the estimates of parameters a_0 and a_1 in the first two columns and the last two columns display the adjusted R-squared of our predictive regression (Adj. R^2) and that of the traditional predictive regression (Adj. R^2 Tradi.).

Two important trends emerge from the results. First, the parameter a_1 is statistically significant and negative for forecast horizons lower than 15 months, and the absolute values of the estimates appear higher at the prediction horizons $\tau \in \{9, 10, 11\}$, roughly one year. The negative value means that an increase in the probability of mean reversion leads to

Table 3: Estimation results of stock return predictive regressions

tau	a_0	a_1	Adj. R^2 (%)	Adj. R^2 Tradi (%)
1	0.0098***	−0.0333***	5.31	0.03
2	0.0177***	−0.0542***	5.53	0.25
3	0.0245***	−0.0675***	5.38	0.46
4	0.0312***	−0.0792***	5.49	0.64
5	0.0380***	−0.0912***	5.79	0.84
6	0.0445***	−0.1021***	5.92	1.09
7	0.0506***	−0.1102***	5.79	1.38
8	0.0564***	−0.1162***	5.51	1.70
9	0.0620***	−0.1206**	5.16	2.05
10	0.0672***	−0.1227**	4.69	2.41
11	0.0719***	−0.1215**	4.09	2.79
12	0.0763***	−0.1185**	3.49	3.18
13	0.0806***	−0.1145*	2.96	3.57
14	0.0848***	−0.1095*	2.49	3.93
15	0.0889***	−0.1045	2.11	4.27
16	0.0933***	−0.1016	1.88	4.59
17	0.0978***	−0.0992	1.69	4.91
18	0.1021***	−0.0957	1.48	5.25
24	0.1265***	−0.0659	0.51	6.78
36	0.1794***	−0.0363	0.07	8.81
48	0.2391***	−0.0531	0.16	11.22

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (8). The last two columns display the adjusted R-squared of this predictive regression (Adj. R^2) and that of the traditional predictive regression (Adj. R^2 Tradi.). *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

a decrease in short-term returns of the S&P 500 index. Second, the adjusted R-squared of our predictive regression is much higher than the one from the traditional predictive regression, notably at very short horizons. For instance, with $\tau = 1$ month, the adjusted R-squared is 177 times higher (5.31% against 0.03%). The highest predictive power is reached at the horizon $\tau = 6$ months. However, the predictive power of our regression model vanishes at longer time horizons. These results are new and interesting and suggest that in predictive regressions, returns predictability at short-term horizons can be recovered using the occurrence of mean reversion as a predictor rather than the level of valuation ratios. As underlined in the introduction, this result shares some similarities with that of Moench and Tobias (2021), which shows that using the probability of recession in forecasting equity risk premiums increases the predictive power at short-term horizons. In our framework, we rather use the probability of mean-reversion in the US Shiller CAPE ratio, based on the stylized facts reported in figure 3.

Another interesting point that deserves to be investigated is to distinguish the two possible states underlying the prevalence of a mean reversion event, namely an increase (decrease) from low (high) values of the US Shiller CAPE ratio to get back to its average values. To do so, we reproduce the top panel of figure 3 (mean reversion regime) by separating these two states given by the Shiller CAPE ratio being lower or higher than the quantile of order 0.4.[5]

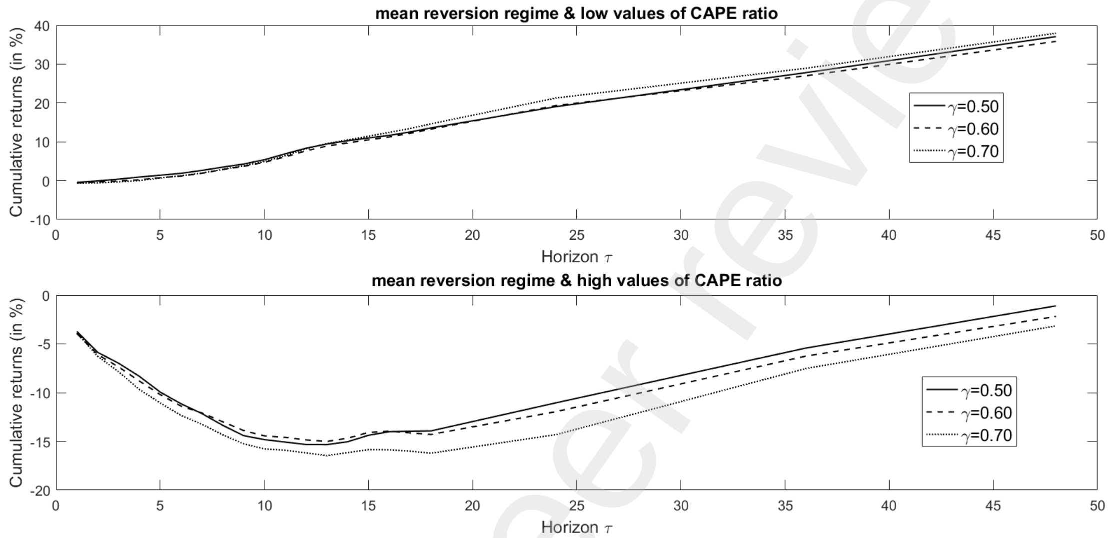

Figure 4: Mean reversion regime, levels of CAPE ratio and subsequent S&P 500 average returns

Notes: The figure compares subsequent S&P 500 average multi-period returns following the prevalence of a mean reversion regime in two states regarding the levels of the US Shiller CAPE ratio. Multi-period returns are indexed by the horizon τ from 1 month to 18 months, and the prevalence of a mean reversion regime is given by the related smoothed probability exceeding a high threshold value $\gamma \in \{0.5, 0.6, 0.7\}$. Smoothed probabilities displayed in figure 2 are obtained from the estimation of the regime-switching mean reversion model in (2) using monthly data from February, 1884 to April, 2020, with a total of 1635 monthly observations. The first (second) panel displays the average multi-period returns for the low (high) state of the CAPE ratio identified by the values of the latter being lower (higher) than the historical quantile of order 0.4.

Figure 4 which displays the results provides evidence that the patterns (negative average multi-period returns) observed in the top panel of figure 3 are attenuated by not taking into account which state (high or low levels of CAPE) of mean reversion is at stake. In other words, disentangling the patterns helps to discover significantly large decreases in the S&P 500 index prices in the months following a mean reversion in the CAPE ratio, when current CAPE ratio is very high (bottom panel). For instance, in this case, the cumulative multi-period returns decrease up to 16.46% for $\tau = 13$ (approximately one year). In the other case with low values of the CAPE ratio (top panel), subsequent multi-period returns are positive or negative, but close to zero.

[5]This threshold value (0.4) is calibrated based on the data to obtain clear-cut differences.

Table 4: Additional estimation results of stock return predictive regressions

τ	a_0	a_1	a_2	Adj. R^2 (%)
1	0.0101***	0.0845*	−0.0455***	7.65
2	0.0182***	0.1748**	−0.0887***	9.03
3	0.0252***	0.2481**	−0.1222***	9.54
4	0.0321***	0.3077**	−0.1498***	10.14
5	0.0389***	0.3617**	−0.1753***	10.85
6	0.0456***	0.4135**	−0.1996***	11.28
7	0.0519***	0.4734**	−0.2259***	11.55
8	0.0578***	0.5377**	−0.2531***	11.72
9	0.0635***	0.6022**	−0.2798***	11.76
10	0.0689***	0.6792**	−0.3104***	11.85
11	0.0738***	0.7660**	−0.3436***	11.91
12	0.0784***	0.8475***	−0.3739***	11.83
13	0.0829***	0.9112***	−0.3970***	11.53
14	0.0870***	0.9536***	−0.4115***	10.99
15	0.0912***	0.9838***	−0.4213***	10.43
16	0.0957***	1.0111***	−0.4307***	10.08
17	0.1002***	1.0464***	−0.4434***	9.91
18	0.1046***	1.0896***	−0.4588***	9.82
24	0.1292***	1.2358***	−0.5038***	8.45
36	0.1819***	1.1613***	−0.4635***	5.18
48	0.2420***	1.3498***	−0.5430***	5.65

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (10). The last column displays the adjusted R-squared of this predictive regression (Adj. R^2). *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

These new results call for a modification of our predictive regression in (8), in order to increase its predictive power reported in table 3. We thus consider the following regression:

$$r_{t+1:t+\tau} = a_0 + a_1 \widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta}) + a_2 x_t \widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta}) + u_{t+1:t+\tau}, \tag{10}$$

where again x_t is the natural logarithm of the US Shiller CAPE ratio, $\widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta})$ is the estimated smoothed probability of mean reversion regime at time t, a_0, a_1 and a_2 the parameters, and $u_{t+1:t+\tau}$ the error term. In this new specification, we have:

$$\frac{\partial r_{t+1:t+\tau}}{\partial \widehat{\Pr}\left(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta}\right)} = a_1 + a_2 x_t, \tag{11}$$

which depends on the level of the US Shiller CAPE ratio, and takes value zero for $x^\star = -a_1/a_2$. With $a_2 < 0$ and $a_1 > 0$, $x^\star$ is positive, and with high (low) values of the Shiller

CAPE ratio, i.e. $x_t > x^\star$ ($x_t < x^\star$), an increase in the occurence of mean reversion leads to negative (positive) subsequent short-term returns, a pattern compatible with the trends in figure 4.

Results in table 4 confirm the expected figures, in the sense that the parameter a_1 (a_2) is positive (negative) and statistically significant for all forecast horizons. This suggests that conditioning the mean reversion regime to the level of the US Shiller CAPE ratio is valuable for predicting short-term horizon stock returns. Figure 5 compares the explanatory power as given by the adjusted R-squared of this new predictive regression (last column in table 4) and the predictive regression in (8) as displayed in the penultimate column of table 3. We observe an increase in the adjusted R-squared, notably for the horizons close to $\tau = 12$, i.e. approximately one year.

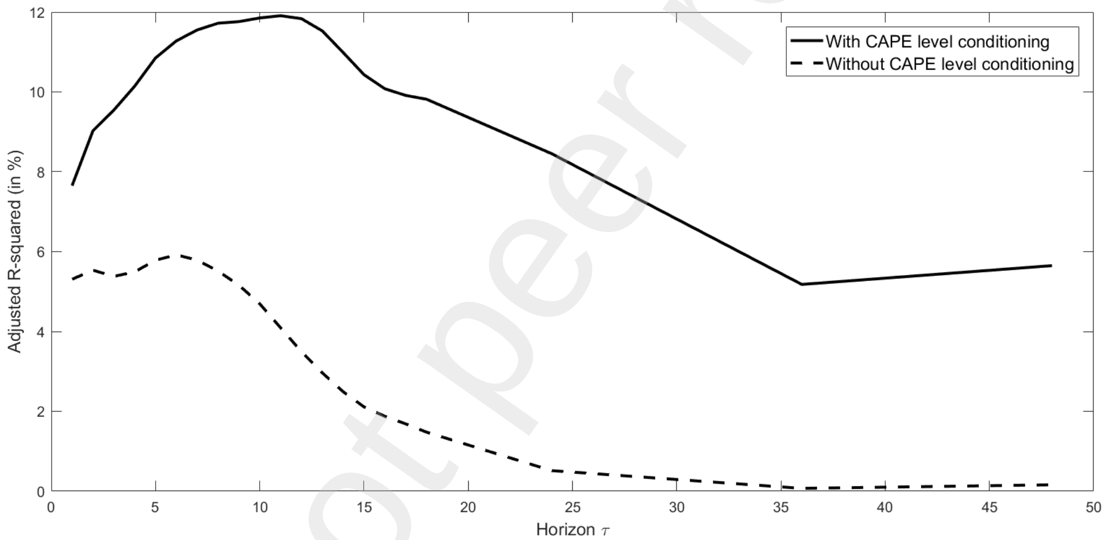

Figure 5: Explanatory powers of stock return predictive regressions

Notes: The figure compares the explanatory powers (adjusted R-squared) of two competing stock return predictive regressions. The first specified in (8) uses the probability of mean reversion in the valuation ratio as the explanatory variable, while the second conditions this latter variable to the level of the ratio (10).

3 Does the predictability hold out-of-sample?

Although interesting, these results are difficult to exploit empirically on an out-of-sample basis, because the identification of the mean reversion regime is based on the smoothed probabilities (see figure 2) which are estimated using Ω_T the information set available over the entire period (t from 1 to T).

One solution to this problem is to check for the stylized facts reported in figure 3 using the filtered probabilities rather than the smoothed counterparts. Note that the filtered probabilities are defined conditionally to the information set Ω_t available at time t, and hence appear useful (to some extent) for a real-time forecasting exercise. Figure 6 which displays the patterns, indicates that the prevailing regime (based on filtered probabilities higher than a given threshold) has no predictive power on average multi-period returns. Indeed, the latters are positive for both regimes, and this result is robust to the horizon τ and the probability threshold parameter γ.

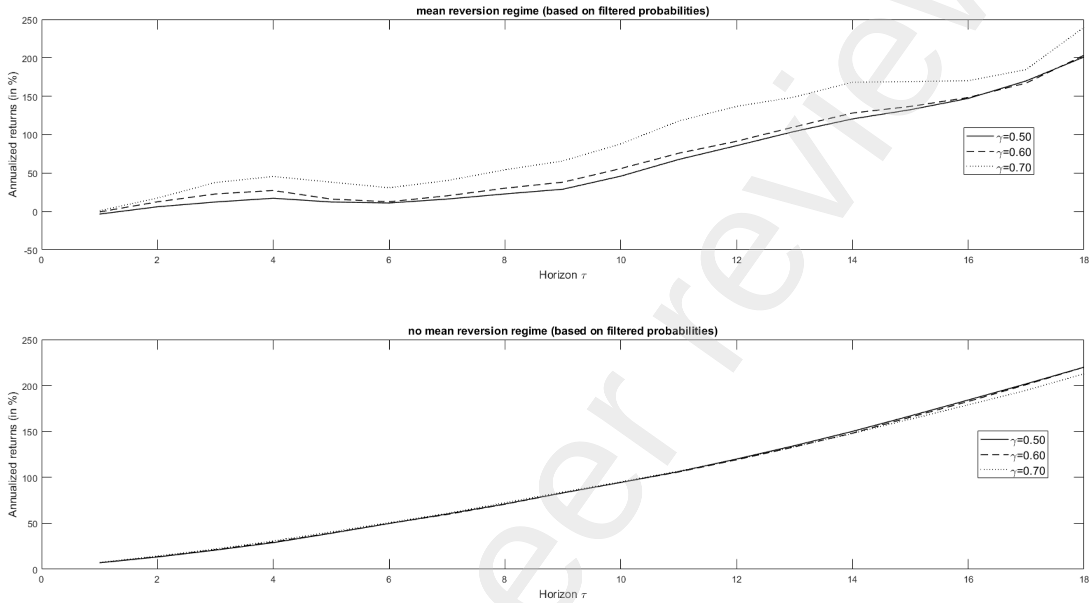

Figure 6: Prevaling regimes and subsequent S&P 500 average returns

Notes: The figure compares subsequent S&P 500 average multi-period returns following the prevalence of a given regime (presence or absence of mean reversion) in the US Shiller CAPE ratio. Multi-period returns are indexed by the horizon τ from 1 month to 18 months, and the prevalence of a regime is given by the related filtered probability exceeding a high threshold value $\gamma \in \{0.5, 0.6, 0.7\}$. The first (second) panel displays the average multi-period returns for the mean (no mean) reversion regime. Filtered probabilities are obtained from the estimation of the regime-switching mean reversion model in (2) using monthly data from February, 1884 to April, 2020, with a total of 1635 monthly observations.

Another solution we retain in this paper and which is at the heart of our contribution, is to identify an early-warning business cycle variable which has high informational content on the smoothed probabilities. This variable can thus be used based on the information available at time t to infer the regime that prevails at that time in order to anticipate future evolutions in the stock market prices. To be more precise, if we denote w_{t-m} the value of such a variable observed at the date $t-m$, with m the lag-order, our predictive regression writes:

$$r_{t+1:t+\tau} = a_0 + a_1 w_{t-m} + a_2 x_t w_{t-m} + v_{t+1:t+\tau}, \tag{12}$$

with $r_{t+1:t+\tau}$ the multi-period returns, a_0, a_1 and a_2 some parameters, and $v_{t+1:t+\tau}$ the error term. Compared to the predictive regression in (10), the above specification considers lagged values of the business cycle variable w_t as a leading indicator for the occurrence of a mean reversion in the US Shiller CAPE ratio as evaluated by the level of the smoothed probability. This new specification thus circumvents the fact that the smoothed probability is based on the whole sample, and allows the deployment of an out-of-sample forecasting exercise based on w_{t-m}.

The choice of the business cycle variable w_t is here critical. We must choose a variable with strong predictive power on the occurrence of mean reversion in the US Shiller CAPE ratio. One approach is to consider a large panel of business cycle variables as regressors in a linear regression model for the smoothed probability. This model estimated along with a selection method like the least absolute shrinkage and selection operator (Lasso) of Tibshirani (1996), would help identify an index function (combination of selected business cycle variables) which can be used as a proxy for the smoothed probability. We do not follow such an approach here, as the selected variables are likely to change over time as well as their combination weights in the index function. We rather consider choosing a single business cycle variable, i.e. the term spread.

Our choice of the term spread is based on two pieces of evidence. On one side, there is a contemporaneous relationship between mean reversion in the US Shiller CAPE ratio and economic recession. Indeed, using the NBER recession indicator and our estimates of the smoothed probability of mean reversion $\widehat{\Pr}(S_{k,t} = 1 \,\Big|\, \Omega_T; \widehat{\theta})$, we observe that in periods of economic recession (expansion) the average value of $\widehat{\Pr}(S_{k,t} = 1 \,\Big|\, \Omega_T; \widehat{\theta})$ is equal to 31.9% (9.8%). This suggests a concomitance between mean reversion in the valuation ratio and economic recession. On the other side, there is an abundant literature that stresses the predictive power of term spread on the occurrence of economic recession. Indeed, it is well known that the behaviour of the yield curve changes across the business cycle. During recessions, upward sloping yield curves not only indicate bad times today, but better times tomorrow. Guided from this intuition, many papers predict GDP growth in OLS regressions with the term spread. Furthermore, the term spread is successful at predicting recessions with dichotomous models (probit and logit models) in a univariate framework (Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Moench and Tobias, 2021). The term spread is also an important variable in the construction of a leading business cycle indicator index (Stock and Watson, 1989). Inversion of the yield curve has come to be viewed as an early leading recession indicator. For example, every recession after the mid-1960s was predicted by an inverted yield curve within 6 quarters of the impending recession. Moreover, there has been only one false positive (an instance of an inverted yield curve that was not followed by a recession) during this time period.

Taken together, these two stylized facts establish the link between the lagged values of the term spread and mean reversion in the US Shiller CAPE ratio. The horizon does indeed appear to be equal to 6 quarters or 18 months approximately, as we can see in figure 7. This figure displays the correlations between lagged values of the term spread and the smoothed probability of mean reversion in the US Shiller CAPE ratio.[6] The correlations are negative with the highest absolute value recorded at the lag-order 17.

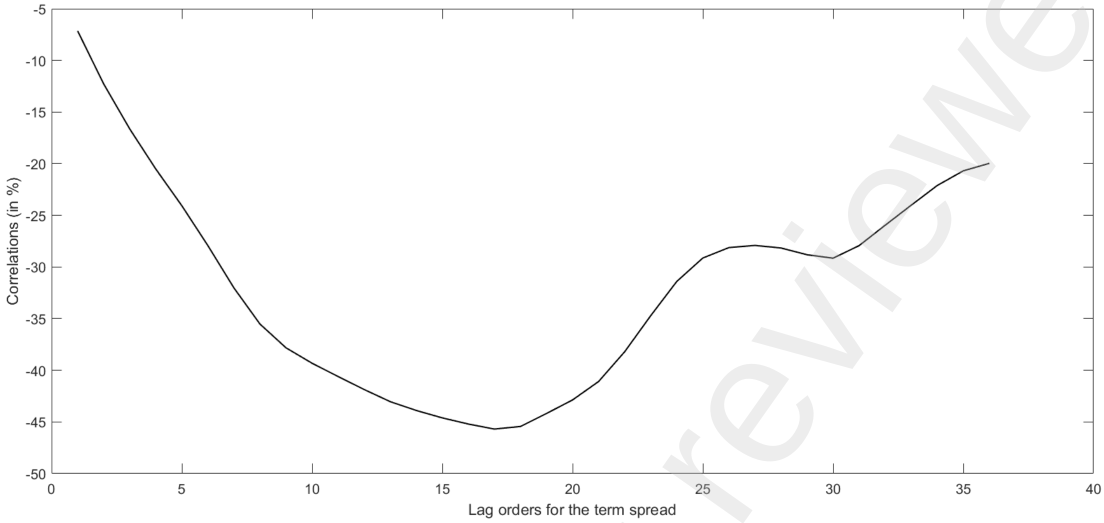

Figure 7: Correlations between the lagged values of the term spread and the smoothed probabilities of mean reversion in the US Shiller CAPE ratio

Notes: The figure displays the correlations between lagged values of the term spread and the smoothed probabilities of mean reversion in the US Shiller CAPE ratio, for different values of lag-order from 1 month to 36 months, i.e. 3 years. The smoothed probabilities are obtained from the estimation of the regime-switching mean reversion model in (2) using monthly data from January, 1971 to April, 2020, with a total of 592 monthly observations.

Based on all the above results, we thus conduct a forecasting exercise to evaluate the out-of-sample power of our predictive model (12) with the business cycle variable w_t corresponding to the term spread, and $m = 17$. Precisely, for a fixed forecasting horizon τ from 1 month to 60 months, i.e. 5 years, and for each month t, we use the available 360 monthly observations to estimate the predictive regression model (12). The estimated parameters are used to forecast the out-of-sample multi-period return $r_{t+1:t+\tau}$. The forecast values together with the realized values are used to compute the out-of-sample R-squared given by:

$$R^2_{OOS}(\tau) = 1 - \frac{\sum_{s=1}^{n_{oos}} (r_s(\tau) - \widehat{r}_s(\tau))^2}{(r_s(\tau) - \bar{r}(\tau))^2} \tag{13}$$

with n_{oos} the number of out-of-sample observations, $r_s(\tau) \equiv r_{t+1:t+\tau}$ the realized multi-period returns, $\widehat{r}_s(\tau)$ the forecast multi-period returns, and $\bar{r}(\tau)$ the average value of realized multi-period returns. Figure 8 displays the out-of-sample R-squared of our predictive regression in (12) with respect to the forecast horizon τ. For comparison, we also display the same statistic for the traditional predictive regression in (9).

Results in figure 8 are interesting as they confirm the trends observed with in-sample estimations. The predictive power of the traditional predictive regression is very low, even

[6]Figure A.1 in Appendix A displays the dynamic of the US term spread.

negative at very short-term horizons, and monotonically increases to reach high levels at long-term horizons. On the contrary, the trend observed for the new predictive regression shows a higher predictive power at short-term horizons, and a monotonic growth until horizon $\tau = 27$ (2 years and 1 quarter), followed by a decrease for higher horizons. For instance, at 12 months the out-of-sample R-squared is equal to 2.28% for the traditional regression, while it is equal to 8.73% for the new predictive regression. Thus, our new predictive regression model, that makes use of the informational content of the term spread regarding the occurrence of mean reversion in the US Shiller CAPE ratio, helps the recovery of short-term predictability of stock returns. In the next section, we conduct additional checks, evaluating the robustness of our results to the choice of the financial ratio.

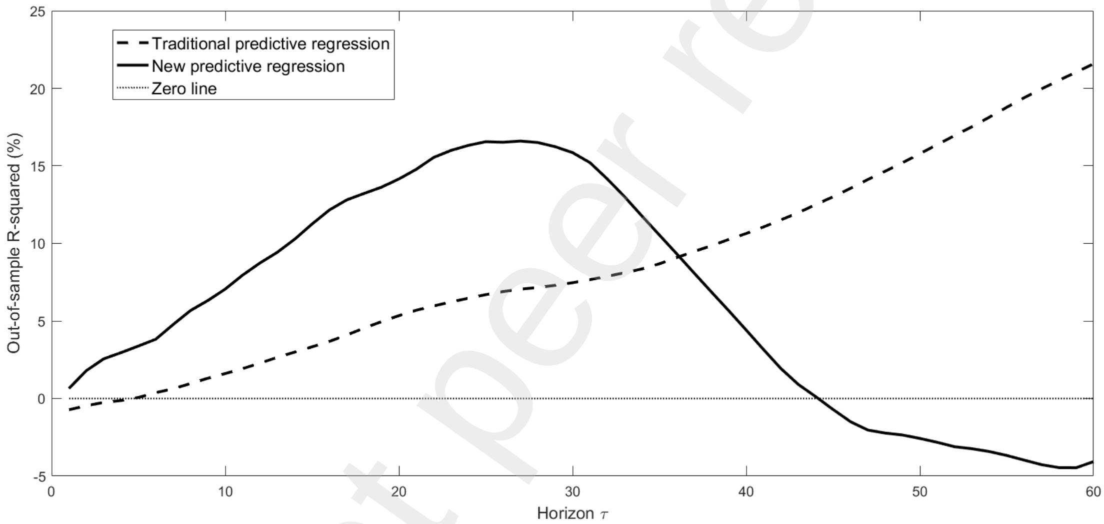

Figure 8: Out-of-sample predictive powers of competing predictive regressions

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the US Shiller CAPE ratio as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the US Shiller CAPE ratio to the occurrence of mean reversion in this valuation ratio as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1971 to April, 2020, with a total of 592 monthly observations.

4 Robustness checks: choice of the valuation ratio

This section evaluates the robustness of our results to the choice of the valuation ratio. First, we consider using the excess CAPE yield instead of the original CAPE ratio. The rationale of using this valuation ratio arises from the historically low interest rates over recent decades, which were playing a potential role in raising CAPE ratios. Theoretically, interest rates are a key component in determining prices. All things being equal, when interest rates fall, the discount rate goes down and stock prices are expected to rise. Thus, the never-ending decrease in interest rates since the beginning of the eighties could be a reason for the increases in stock prices and the CAPE ratio. The level of interest rates is therefore a particularly important element to take into account when evaluating stock prices. In this line, the excess CAPE yield recently introduced by Shiller and which takes into account both stock prices and interest rates, is a good indicator to capture this effect.

The excess CAPE yield is calculated by inverting the CAPE ratio to get a measure of the profits the S&P 500 index is delivering for each dollar investors are paying, minus the ten-year real interest rate (the US 10-year treasury bond). This thus represents the margin that stocks are paying over bonds and characterizes somewhat the risk premium associated with the stock market, taking into account the interaction between long-term changes in stock prices and interest rates. The higher this indicator is, the more attractive stocks are. Figure A.2 in Appendix A displays the dynamic of the US excess CAPE yield.

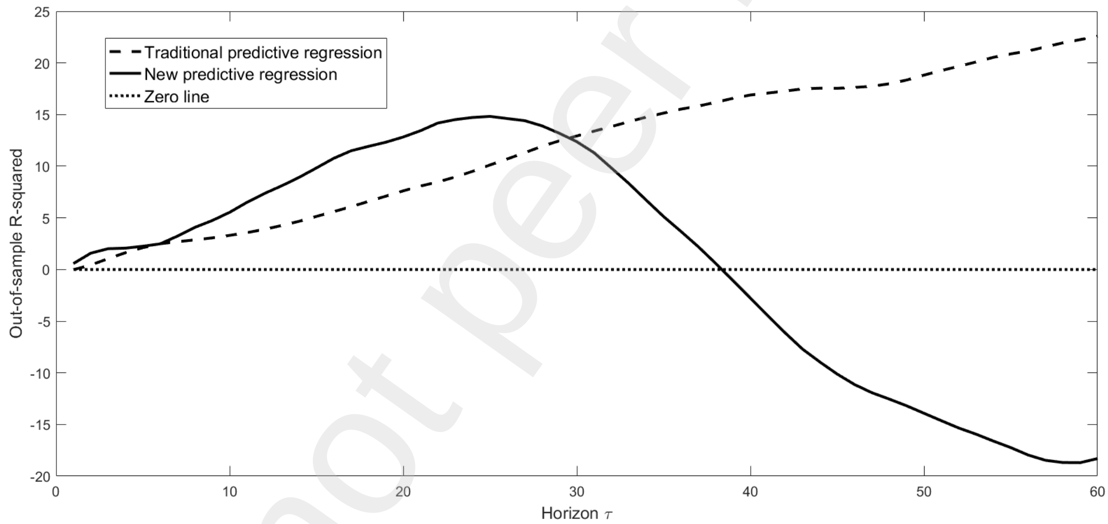

Figure 9: Out-of-sample predictive powers of competing predictive regressions: the US Excess CAPE yield

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the US excess CAPE yield as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the US excess CAPE yield to the occurrence of mean reversion in this valuation ratio as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1971 to April, 2020, with a total of 592 monthly observations.

Results displayed in figure 9 are similar to those obtained when using the original Shiller CAPE ratio. The new regression has very superior predictive powers at short-term horizons. For the one month horizon, the out-of-sample R-squared are equal to -0.02% and 0.57% for the traditional and the new predictive regression. Moreover, at the horizon of 12 (24) months, the out-of-sample R-squared of the new predictive regression is equal to 7.35% (14.71%), while the recorded value for the traditional regression is equal to 3.88% (9.50%).

For more robustness checks, we also consider the US dividend yield as the predictive financial ratio. The choice of this financial ratio arises from its predictive content for long-term multi-period returns on the stock index (Rozeff, 1984; Campbell and Shiller, 1988; Fama and French, 1988; Cochrane, 1991; Lewellen, 2004). For instance, Rozeff (1984), based on the constant dividend growth model and the subsidiary Golden Rule of Accumulation view that real long-term growth equals the real rate of interest, presented evidence that dividend yields are directly related to and predict future stock returns. In particular, the author's predictive tests show that dividend yields provide superior predictions of equity risk premiums in terms of lower bias, a lower mean square error and a lower mean absolute error as compared with the method of using historical realized returns. Similar results are provided by Lewellen (2004) who corrected predictive regressions for small-sample biases, and showed that dividend yield predicted market returns during the period 1946–2000, as well as in various subsamples.[7]

Figure 10 which compares our new specification to the traditional one, confirms once again the robustness of our results to the choice of the financial ratio.

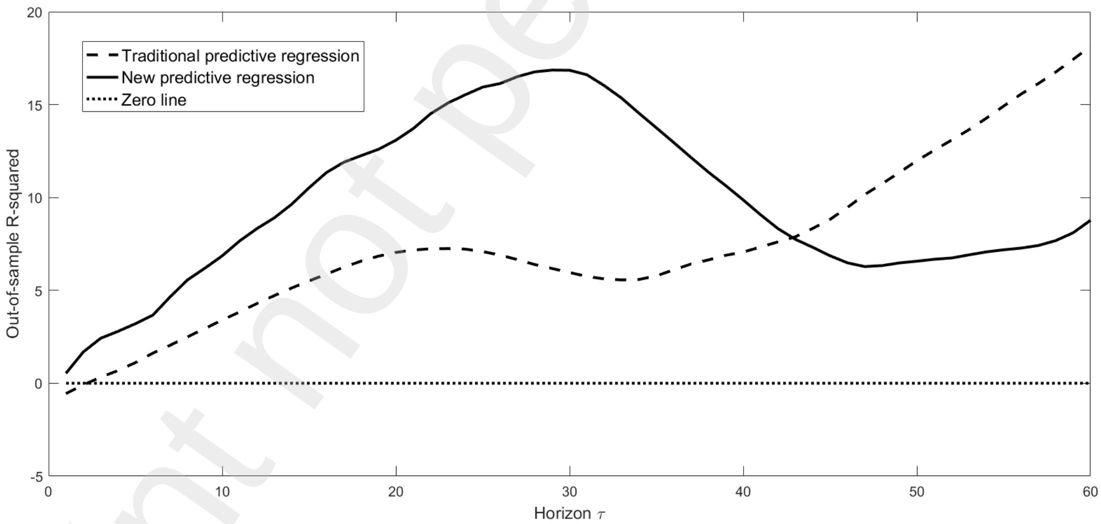

Figure 10: Out-of-sample predictive powers of competing predictive regressions: the US dividend yield

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the US dividend yield as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the US dividend yield to the occurrence of mean reversion in this valuation ratio as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1971 to April, 2020, with a total of 592 monthly observations.

[7]Figure A.3 in Appendix A displays the dynamic of the US dividend yield.

5 International evidence

In this section, we investigate whether our results hold over other countries. We consider four additional countries including Canada, France, Germany and the UK. The first part of the section evaluates the robustness of our results in-sample, and the second part is devoted to out-of-sample analysis.

5.1 In-sample evidence

For in-sample evidence, we estimate for each country the predictive regression (10). Recall that the estimation of this regression requires as regressors the valuation ratio x_t and $\widehat{\Pr}(S_{k,t} = 1 \Big| \Omega_T; \widehat{\theta})$, the estimated smoothed probability of mean reversion in the valuation ratio. The latter predictor is obtained estimating the regime switching mean-reversion model in (2), with k the lag-order which leads to the best separation between the two regimes (mean-reversion versus no-mean reversion). Figure A.4 in Appendix displays the dynamic of the monthly values of the valuation ratio (CAPE) for each country over the sample that ranges from January 1983 to April 2020. Figures A.5-A.8 display the dynamic of the smoothed probabilities for each country. The optimal value of k are equal to 12 for all countries.

Table 5 displays the results of the predictive regression (10) for Canada. The presentation is similar to that of table 4 for the US, with an additional column that reports the adjusted R-squared of the traditional predictive regression for the purpose of comparison. As in table 4 the parameter a_1 (a_2) is positive (negative) and statistically significant for all forecast horizons, except the horizon of 1 month for a_1. These results are proof for this country, that our predictive regression, which relies on the mean reversion probability coupled with the level of the CAPE ratio, is valuable. The comparison of the predictive powers as measured by the adjusted R-squared confirms this statement. At short-term horizons, the values displayed for our approach are higher than those of the traditional model, the inversion of dominance being observed only at long horizons (36 and 48 months). Precisely, while the traditional model shows an adjusted R-squared of -0.09% (2.93%) at 1 (12) months, our model reaches a value of 5.30% (7.13%).

Tables A.1-A.3 report the same results for the remaining countries, i.e. the UK, Germany and France, respectively. For the UK (table A.1), the evidence seems to be stronger, because the explanatory power of our predictive regression (R-squared) is higher than that of the traditional model at all horizons, with the highest dominance at the horizons of 10, 11 and 12 months. Results are also robust for Germany as displayed in table A.2. There is a clear-cut superiority of our approach compared to the traditional one, in particular for the short horizons (1 to 4 months), with a very marked difference at 1 month. Indeed, at this horizon, the adjusted R-squared is equal to 6.11% in our approach against 0.43% for the

Table 5: Estimation results of stock return predictive regressions: Canada

τ	a_0	a_1	a_2	Adj. R^2 (%)	Adj. R^2 Tradi (%)
1	0.0088***	0.1488	−0.0608*	5.30	−0.09
2	0.0156***	0.3525**	−0.1300**	5.95	0.18
3	0.0222***	0.4881**	−0.1757***	6.05	0.47
4	0.0286***	0.6642***	−0.2338***	6.94	0.78
5	0.0349***	0.8344***	−0.2892***	7.85	1.13
6	0.0411***	0.9660***	−0.3312***	8.26	1.47
7	0.0474***	1.0589***	−0.3609***	8.37	1.80
8	0.0532***	1.0937***	−0.3706***	7.66	2.00
9	0.0592***	1.1252***	−0.3799***	7.18	2.19
10	0.0655***	1.1749***	−0.3957***	7.06	2.41
11	0.0716***	1.2397***	−0.4164***	7.10	2.66
12	0.0777***	1.3081***	−0.4376***	7.13	2.93
13	0.0835***	1.3107***	−0.4367***	6.53	3.08
14	0.0887***	1.2939***	−0.4275***	5.79	3.17
15	0.0938***	1.3050***	−0.4275***	5.48	3.24
16	0.0990***	1.3313***	−0.4330***	5.36	3.33
17	0.1045***	1.3904***	−0.4506***	5.58	3.38
18	0.1102***	1.4800***	−0.4787***	6.08	3.48
24	0.1473***	2.0763***	−0.6724***	10.30	3.75
36	0.2315***	1.0223**	−0.3585**	2.47	4.08
48	0.3073***	0.7481**	−0.2796**	1.80	3.51

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (10). The penultimate column displays the adjusted R-squared of this predictive regression (Adj. R^2), and the last column the adjusted R-squared of the traditional regression (Adj. R^2 Tradi). *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

traditional one. Even though the parameters a_1 and a_2 are not significant, similar results are obtained for France (see table A.3), with higher adjusted R-squared up to 3 months.

5.2 Out-of-sample evidence

Out-of-sample robustness across countries (Canada, France, Germany, the UK) are investigated here, using the approach adopted for the US in Section 3. Formally, for a fixed forecasting horizon τ from 1 month to 60 month, i.e. 5 years, and for each month t, we use the available monthly observations to estimate the predictive regression model (12). The estimated parameters are used to forecast the out-of-sample multi-period return $r_{t+1:t+\tau}$, and compute the out-of-sample R-squared as given in (13).

Recall that regression (12) requires the lagged value of the term spread w_{t-m} with m the lag-order.[8] As done in section 3 for the US, we search for the optimal value of m, computing

[8]Figure A.9 displays the dynamic of the term spread for each of the 4 countries.

the correlation between lagged values of the term spread and the smoothed probabilities of mean reversion in the CAPE ratio (see figure 7 for the US). The optimal lag-orders are equal to 21, 25, 6 and 9 for Canada, the UK, Germany and France, with the consequence that the early-warning property of the term spread holds over very short horizons for Germany (two quarters) and France (three quarters).

Figure 11 displays for Canada, the out-of-sample R-squared of our predictive regression with respect to the forecasting horizon. For comparison, the same statistic is reported for the traditional predictive regression. The patterns observed in the figure are qualitatively similar to those reported for the US in figure 8. The new predictive regression has superior (inferior) predictive ability at short-term (long-term) horizons. For illustration, at the horizon of 12 (24) months, the adjusted R-squared is equal to 4.08% (9.09%) for the new approach, whereas it is only equal to −2.84% (0.49%) in the traditional approach.

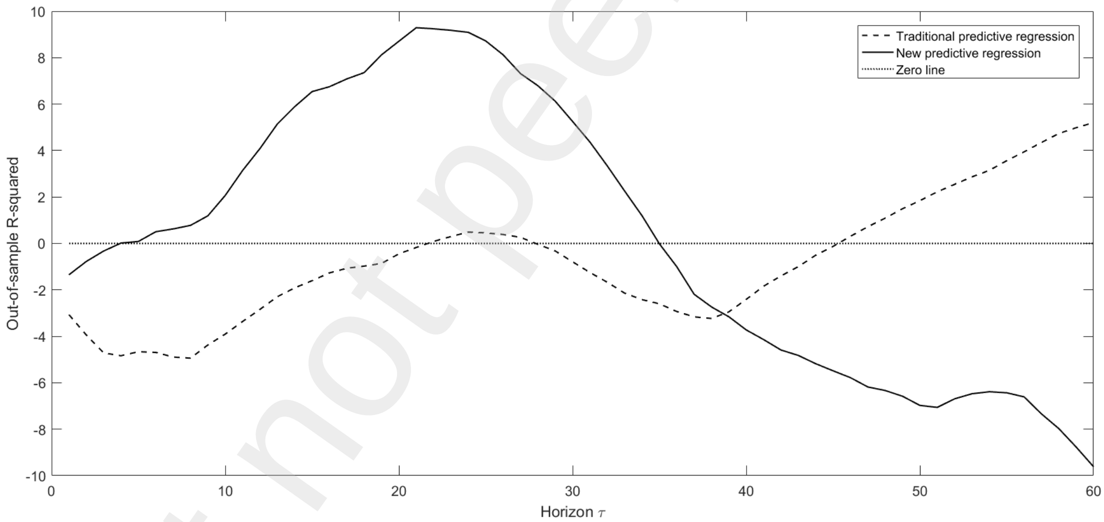

Figure 11: Out-of-sample predictive powers of competing predictive regressions: Canada

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the CAPE ratio as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the CAPE to the occurrence of mean reversion in this valuation ratio, as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1983 to April, 2020, with a total of 448 monthly observations.

Figures A.10-A.12 in Appendix A display the same information for the UK, Germany and France, respectively. Results are similar, although the reported levels of adjusted R-squared are low and negative at very short-term horizons for France and the UK, even for the new predictive regression.

6 Assessment of economic value

This section evaluates the economic value of the new predictive model for equity premium, based on an asset allocation exercise. We can expect that the robust out-of-sample forecast ability obtained for the US with positive out-of-sample adjusted R-squared at all horizons (see figure 8 for the US CAPE, figure 9 for the US Excess CAPE, and figure 10 for the US dividend yield) should lead to economic gains for an investor that allocates his wealth between the S&P 500 and a risk-free instrument. For the other countries, especially the UK and France, for which the new predictive model also dominates the traditional one, but with negative adjusted R-squared at very short horizons (see figures A.10 and A.12), we can still achieve significant economic gains. Indeed, as underlined by Cenesizoglu and Timmermann (2012), the link between statistical and economic measures of forecast performance is positive but of low magnitude, with the consequence that negative out-of-sample adjusted R-squared can be associated with positive economic gains for investors.

Based on this, and for each country, we consider the standard mean-variance portfolio choice of an investor who chooses a portfolio in the universe of two instruments, i.e. the country stock market index and the risk-free asset (cash). Denote by r_t the returns on the stock index at month t. The investor has a rebalancing horizon τ for his portfolio that coincides with the forecast horizon for the risk premium. At time t, if we denote the optimal share of the wealth allocated to the stock index as w_t, we have:

$$\widehat{w}_t = \frac{1}{\gamma}\frac{\widehat{R}_{t+\tau}}{\widehat{\sigma}^2_{t+\tau}}, \tag{14}$$

with γ the relative risk aversion parameter, $\widehat{R}_{t+\tau}$ the forecast risk premium using a given predictive model (the traditional or the improved one) for the returns r_t, and $\widehat{\sigma}^2_{t+\tau}$ the variance of the portfolio returns computed here as the sample variance over a 10-year rolling window of past data, following Rapach et al. (2010) and Moench and Tobias (2021). Hence, $\widehat{w}_t$ differs only by the predictive model retained to forecast the risk premium $\widehat{R}_{t+\tau}$, and this allows a fair comparison between alternative models. The realized monthly portfolio return at time j between t and $t+\tau$ is given by:

$$r_{p,t+j} = w_t r_{t+j}. \tag{15}$$

If we consider a proportional transaction cost c, the portfolio's net return is given by:

$$\tilde{r}_{p,t+j} = r_{p,t+j} - c\left|\widehat{w}_t - \widehat{w}_t^+\right|, \tag{16}$$

where $\widehat{w}_t^+$ is the weight in the risky stock index at time t before rebalancing. The economic value of a given predictive model for risk premium can be evaluated based on the realized certainty equivalent return (CER) given by:

$$CER_p = \widehat{\mu}_p - \frac{1}{\gamma}\widehat{\sigma}^2_p, \tag{17}$$

Table 6: CER and differences in CER: the US.

	New model		Traditional model		Utility gain	
	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$
$\tau = 1$	2.75	1.40	0.69	−2.15	2.06	3.55
$\tau = 2$	2.53	1.22	0.65	−2.26	1.88	3.48
$\tau = 3$	2.60	1.36	2.12	−0.44	0.48	1.80
$\tau = 4$	1.81	0.74	1.99	−0.35	−0.18	1.09
$\tau = 5$	0.67	0.21	0.69	−2.08	−0.02	2.30
$\tau = 6$	1.76	0.93	2.07	−0.34	−0.31	1.27
$\tau = 7$	0.85	0.48	1.65	−0.87	−0.80	1.34
$\tau = 8$	0.82	−0.15	1.50	−0.73	−0.68	0.58
$\tau = 9$	0.31	−0.14	1.54	−0.64	−1.23	0.50
$\tau = 10$	−0.58	−0.81	1.03	−0.86	−1.61	0.05
$\tau = 11$	−0.08	−0.69	1.46	−0.76	−1.55	0.07
$\tau = 12$	0.91	−0.46	1.47	−0.51	−0.57	0.04

Notes: For different values of the relative risk aversion parameter γ and the forecast horizon τ, the table displays the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The last two columns display the differences in CER (utility gain).

with $\widehat{\mu}_p$ and $\widehat{\sigma}_p^2$ the mean and variance of the net portfolio's returns $\tilde{r}_{p,t+j}$. As in Moench and Tobias (2021), this value of CER is multiplied by 12 to interpret it as the annual risk-free rate that an investor would be willing to accept to not hold the risky portfolio. For each horizon τ we report the difference in CERs between our predictive model for risk premium and the baseline traditional model. This difference corresponds to the utility gain, i.e. the annual portfolio management fee that an investor would be willing to pay to switch from the traditional model to the new proposed model.

Table 6 displays for the US, the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The utility gain is also reported. We consider different values for the forecast horizon between 1 and 12, and two values for the relative risk aversion parameter γ. The portfolio weight w_t is restricted to lie between 0 and 1, thus excluding short-selling and leveraging.

For $\gamma = 3$, our new predictive model has positive CERs that globally decrease with the forecast horizon, with positive and high values at very short horizons (1 month to 3 months) and negative values at the highest horizons (10, 11 and 12 months). For the traditional model, the CERs are positive at all forecast horizons, but lower for the values reported for the new model at the very short horizons. Hence, the utility gains are positive at these horizons (1 month to 3 months). For example, at the horizon of 1 month (3 months), the annual portfolio management fee that an investor would be willing to pay to switch from the traditional model to the new proposed model is equal to 2.06% (0.48%). For $\gamma = 5$ the utility gains are positive at all horizons, but globally decrease with the forecast horizons. All these results confirm the statistical evidence, i.e. the superior predictive power of the new model that decreases with the forecast horizon.

Table 7 displays the same results for the UK. Overall, we observe the same patterns, with high utility gains at the very short horizons that decrease with the horizons. Tables A.4-A.6 reports the same information for France, Germany and Canada, respectively. For France (table A.4), we no longer observe the regularity of the evolution of CERs and utility gains as a function of the forecast or rebalancing horizon τ. However, the robustness of the results still holds with positive utility gains across all horizons. Finally, results in tables A.5 and A.6 show positive CERs at very short-term horizons, i.e. 1 to 3 months for Germany for both values of the relative risk aversion parameter γ, and 1 to 2 months for Canada with $\gamma = 5$. For $\gamma = 3$, the results are more erratic, even if we observe overall that all the CERs up to $\tau = 8$ are positive, except for $\tau = 3$.

Table 7: CER and differences in CER: the UK

	New model		Traditional model		Utility gain	
	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$
$\tau = 1$	2.47	1.77	-1.05	-1.32	3.51	3.08
$\tau = 2$	2.94	2.16	-0.83	-1.41	3.77	3.57
$\tau = 3$	3.56	2.50	-0.54	-0.85	4.10	3.34
$\tau = 4$	3.81	2.82	0.43	-0.22	3.39	3.04
$\tau = 5$	2.90	2.23	1.53	0.09	1.37	2.14
$\tau = 6$	3.89	2.60	2.06	0.91	1.83	1.68
$\tau = 7$	2.66	1.91	0.96	0.10	1.70	1.81
$\tau = 8$	3.52	2.46	1.82	0.98	1.70	1.48
$\tau = 9$	2.70	1.55	2.14	1.60	0.56	-0.06
$\tau = 10$	3.78	2.32	3.01	2.26	0.77	0.06
$\tau = 11$	3.10	1.91	2.01	1.41	1.10	0.50
$\tau = 12$	2.54	1.45	2.87	1.94	-0.33	-0.49

Notes: For different values of the relative risk aversion parameter γ and the forecast horizon τ, the table displays the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The last two columns display the differences in CER (utility gain).

7 Conclusion

Valuation is an important determinant of future returns, and the literature reported evidence of forecast ability at long-term horizons. Evidence at short-term horizons is still weak, and the available contributions reached short-term predictability by relaxing the assumption of a fixed steady regime of the economy. Recent papers achieved this task through models with time-varying parameters that fit business cycles, and specifically recession and expansion phases. However, these specifications usually impose tight parametric restrictions on how predictive coefficients in their dynamic models evolve over time.

In this article, we contribute to this literature proposing a new predictive regression model based on the observed dynamics of stock returns following the occurrence of a mean reversion in the US Shiller CAPE ratio, when the latter is high. First, the occurrence of mean reversion is approximated by the smoothed probability from a regime-switching version of the mean reversion model of Jegadeesh (1991). Second, to avoid model misspecification and allow our predictive regression to be operational for an out-of-sample exercise, we exploit the link between the term spread and mean reversion in valuation ratios. Both in-sample and out-of-sample predictions show large and significant improvement relative to the traditional predictive regression. We show that our results are robust with respect to the choice of the valuation ratio (CAPE, excess CAPE and dividend yield), and report robustness across countries (Canada, France, Germany and the UK). We also conduct a mean-variance asset allocation exercise which confirms the superiority of the new predictive regression in terms of utility gain.

These results have important implications regarding the understanding of asset price dynamics and mean reversion in relation with the business cycle (bad and good times) and then, practically, on dynamic asset allocations.

A Appendix A: Additional Figures and Tables

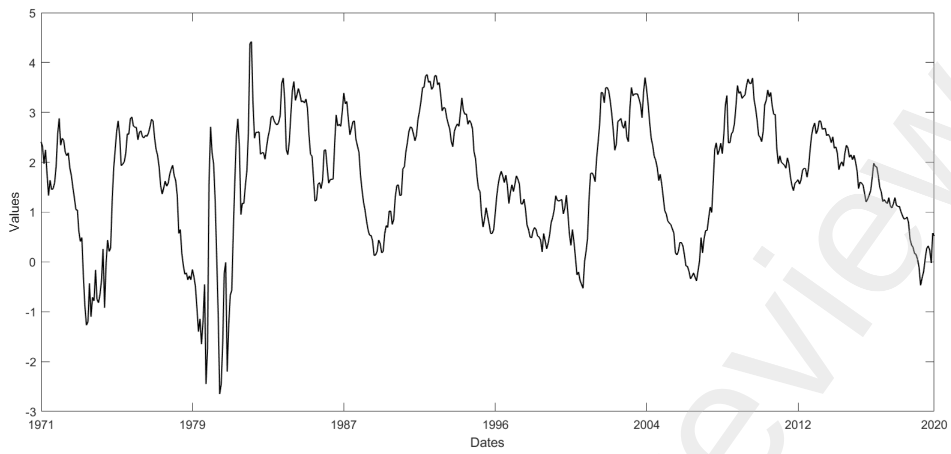

Figure A.1: Dynamic of the US term spread: 1971/01-2020/04

Notes: The term spread is calculated as the difference between 10-Year Treasury Constant Maturity and 3-Month Treasury Constant Maturity.

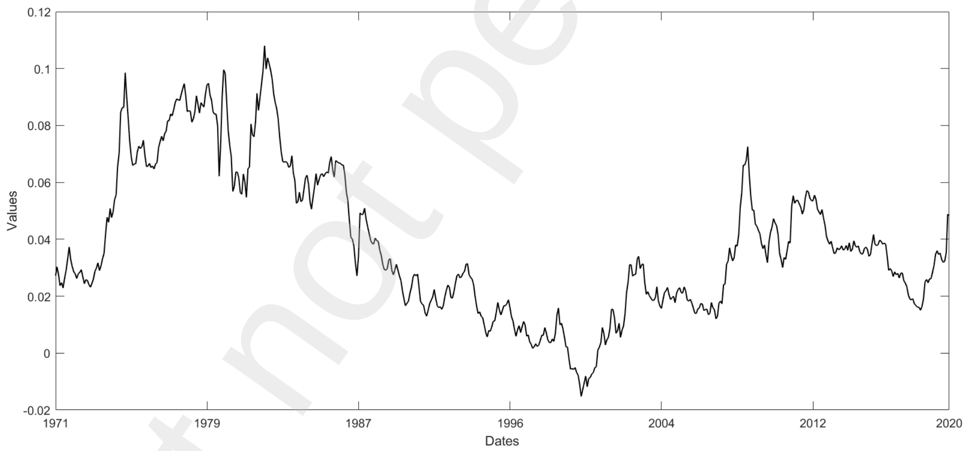

Figure A.2: Dynamic of the US Excess CAPE yield: 1971/01-2020/04

Notes: The ratio is computed by inverting the CAPE ratio to get a measure of the profits the S&P 500 index is delivering for each dollar investors are paying, minus the ten-year real interest rate (the US 10-year treasury bond).

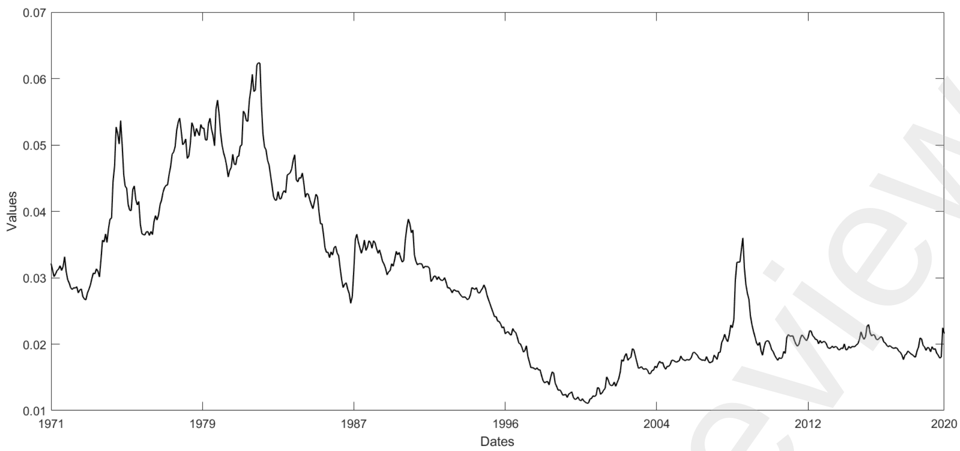

Figure A.3: Dynamic of the US dividend yield: 1971/01-2020/04

Notes: Monthly dividend data are computed from the S&P four-quarter totals, with linear interpolation to monthly figures.

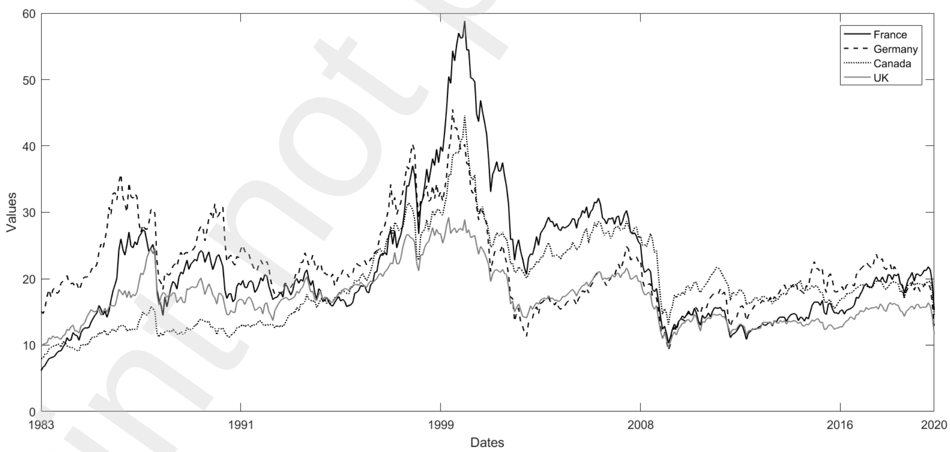

Figure A.4: Dynamic of the CAPE ratio: 1983/01-2020/04

Notes: The ratio for each country is computed based on a dataset that consists of monthly stock index prices, earnings data and the consumer price index (to allow conversion to real values). Monthly earnings are computed from the latest annualised earnings of the last financial year or derived from an aggregation of quarterly earnings. Stock price data are monthly data of the first day of the month closing prices. The CPI-U (Consumer Price Index-All Urban Consumers) are published by the National Institute of Statistics and Economic Studies (INSEE) for France, the Federal Statistical Office (Statistisches Bundesamt) for Germany, the Office for National Statistics for the UK and by Statistics Canada, for Canada.

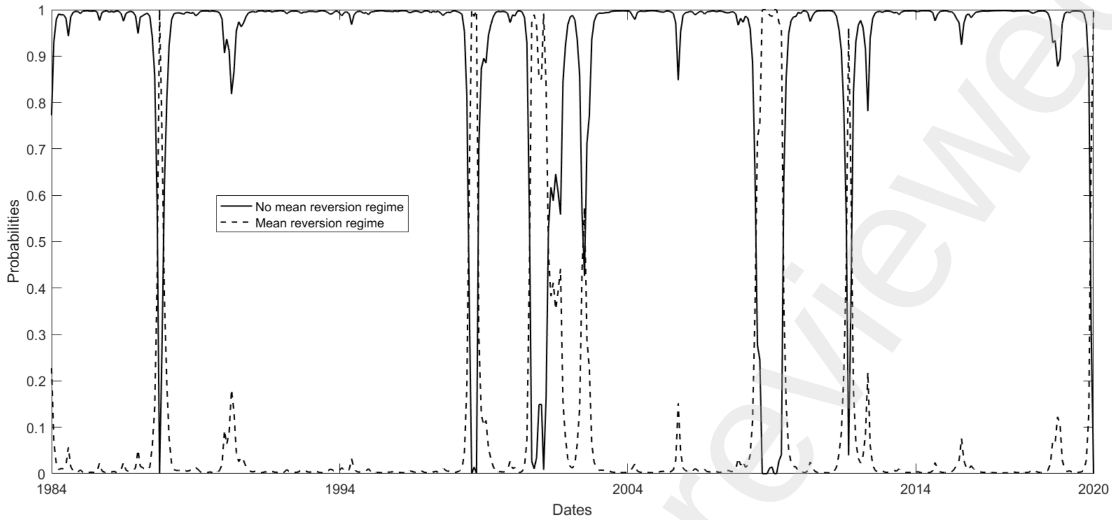

Figure A.5: Dynamic of smoothed probabilities (Canada): 1984/02-2020/04

Notes: The figure displays the smoothed probabilities of the two regimes (no mean reversion/mean reversion) that result from the estimation of the regime-switching mean reversion equation in (2). The estimation sample ranges from February, 1984 to April, 2020, with a total of 435 monthly observations.

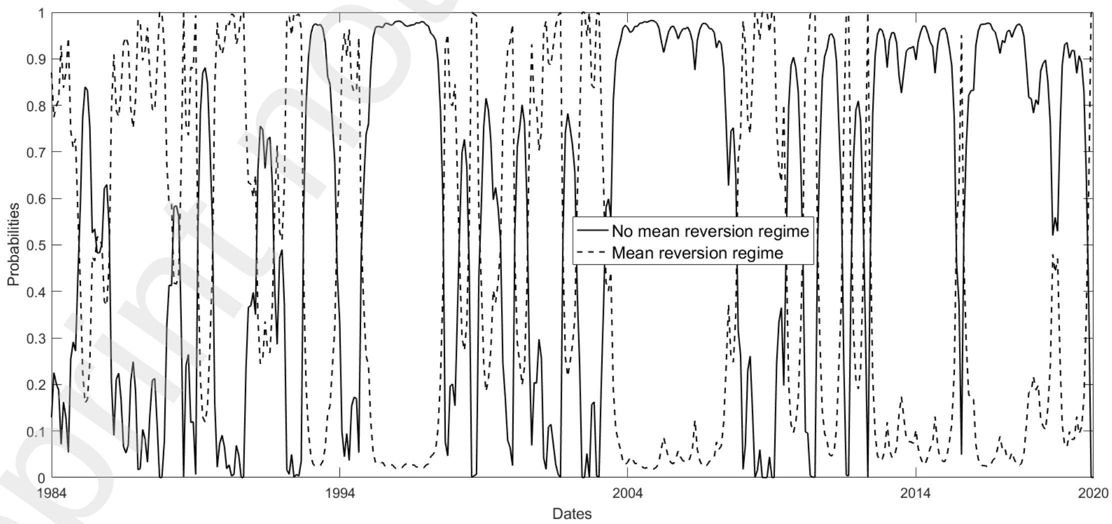

Figure A.6: Dynamic of smoothed probabilities (the UK): 1984/02-2020/04

Notes: The figure displays the smoothed probabilities of the two regimes (no mean reversion/mean reversion) that result from the estimation of the regime-switching mean reversion equation in (2). The estimation sample ranges from February, 1984 to April, 2020, with a total of 435 monthly observations.

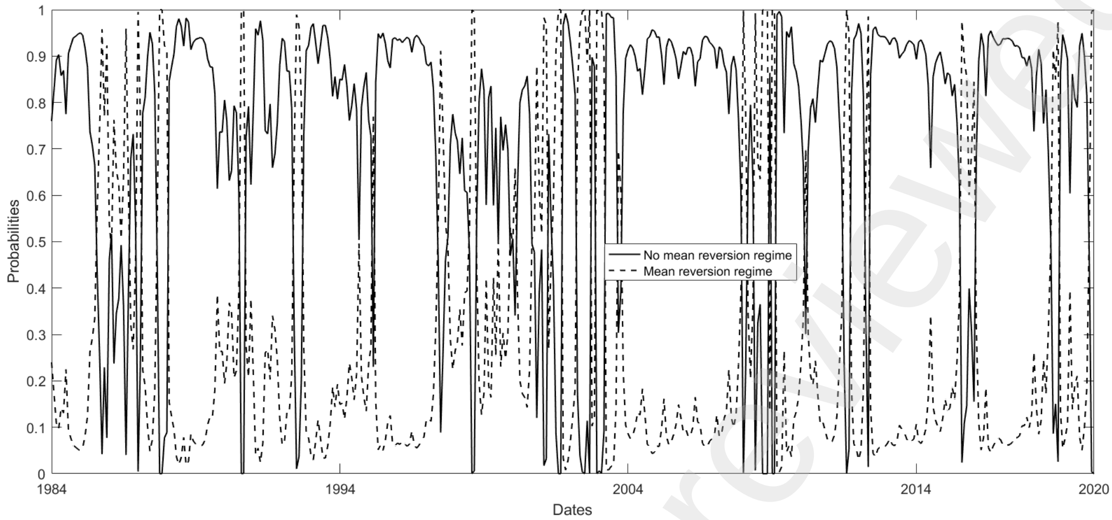

Figure A.7: Dynamic of smoothed probabilities (Germany): 1984/02-2020/04

Notes: The figure displays the smoothed probabilities of the two regimes (no mean reversion/mean reversion) that result from the estimation of the regime-switching mean reversion equation in (2). The estimation sample ranges from February, 1984 to April, 2020, with a total of 435 monthly observations.

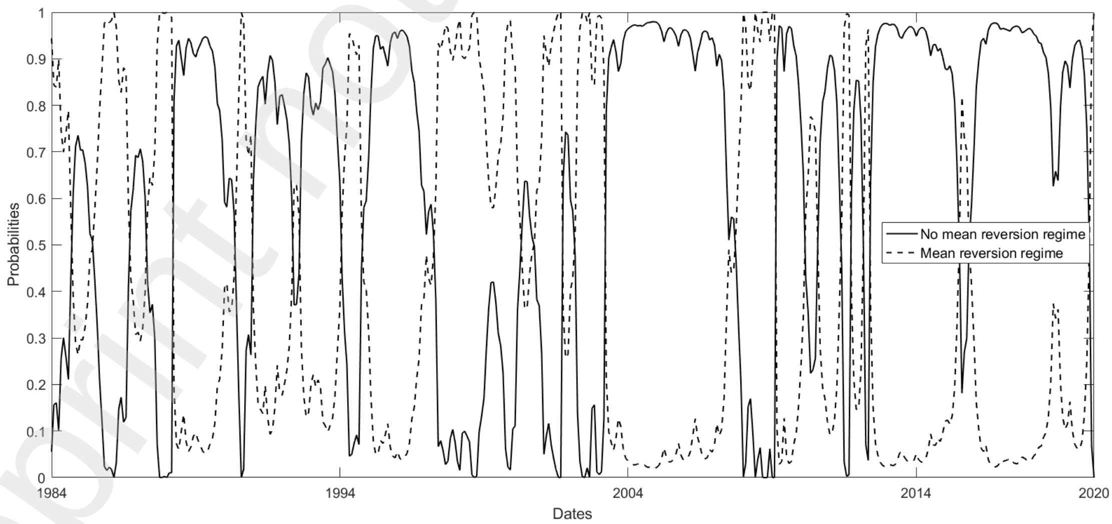

Figure A.8: Dynamic of smoothed probabilities (France): 1984/02-2020/04

Notes: The figure displays the smoothed probabilities of the two regimes (no mean reversion/mean reversion) that result from the estimation of the regime-switching mean reversion equation in (2). The estimation sample ranges from February, 1984 to April, 2020, with a total of 435 monthly observations.

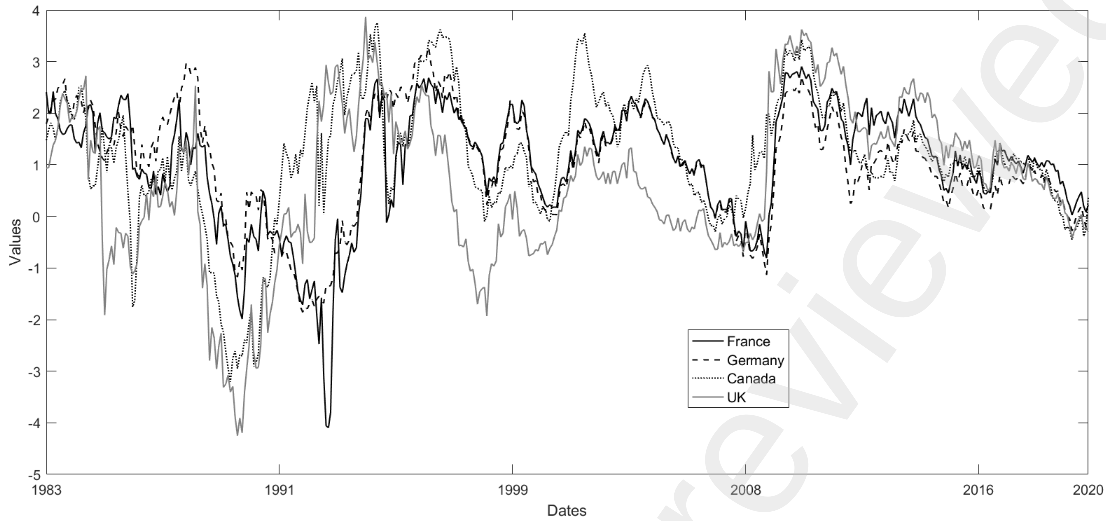

Figure A.9: Dynamic of term spreads: 1983/01-2020/04

Notes: The term spreads are calculated as the spread between the rates of 10-Year Government Bonds Constant Maturity and 3-Month Government Bonds Constant Maturity. Data are from the OECD.

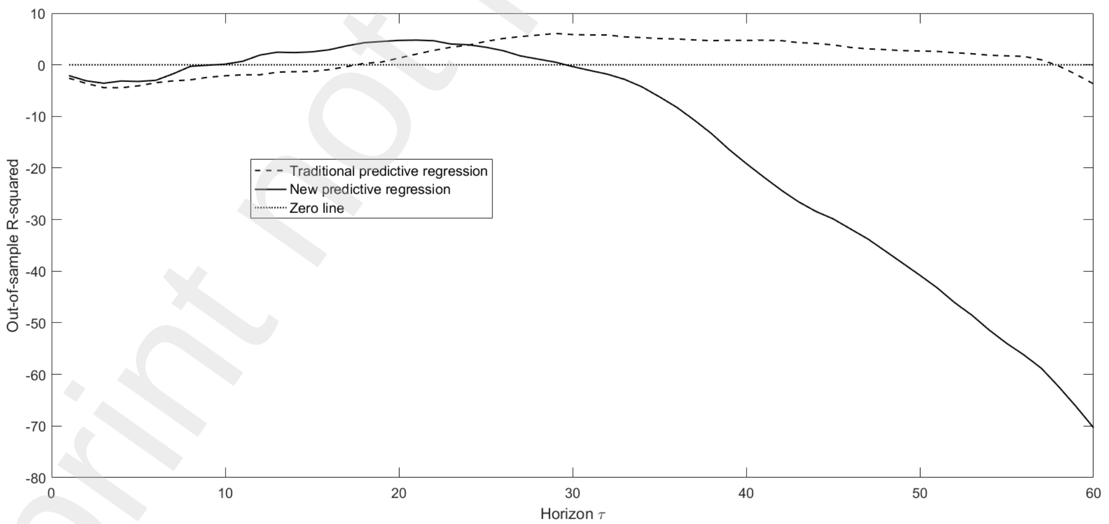

Figure A.10: Out-of-sample predictive powers of competing predictive regressions: the UK

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the CAPE ratio as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the CAPE to the occurrence of mean reversion in this valuation ratio, as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1983 to April, 2020, with a total of 448 monthly observations.

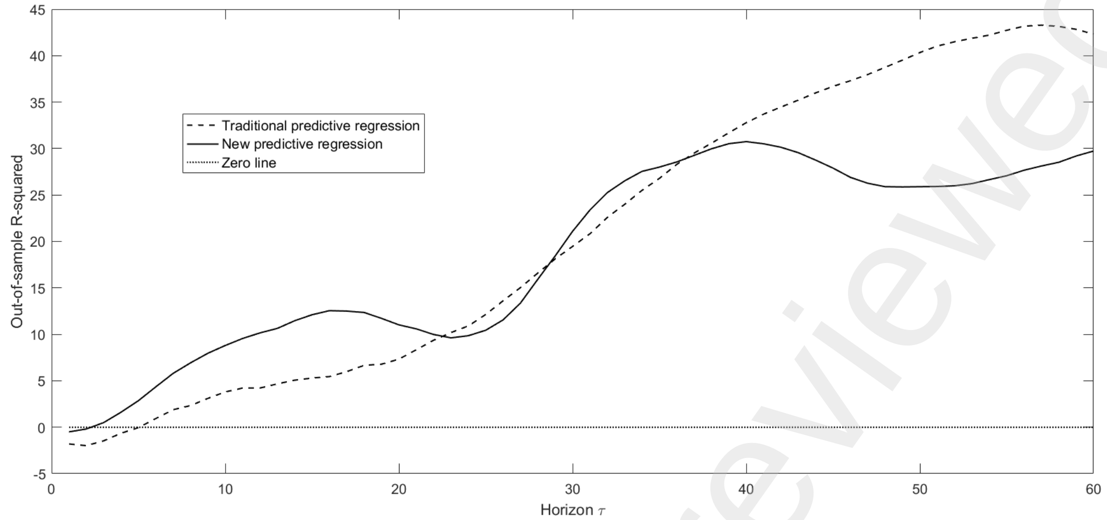

Figure A.11: Out-of-sample predictive powers of competing predictive regressions: Germany

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the CAPE ratio as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the CAPE to the occurrence of mean reversion in this valuation ratio, as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1983 to April, 2020, with a total of 448 monthly observations.

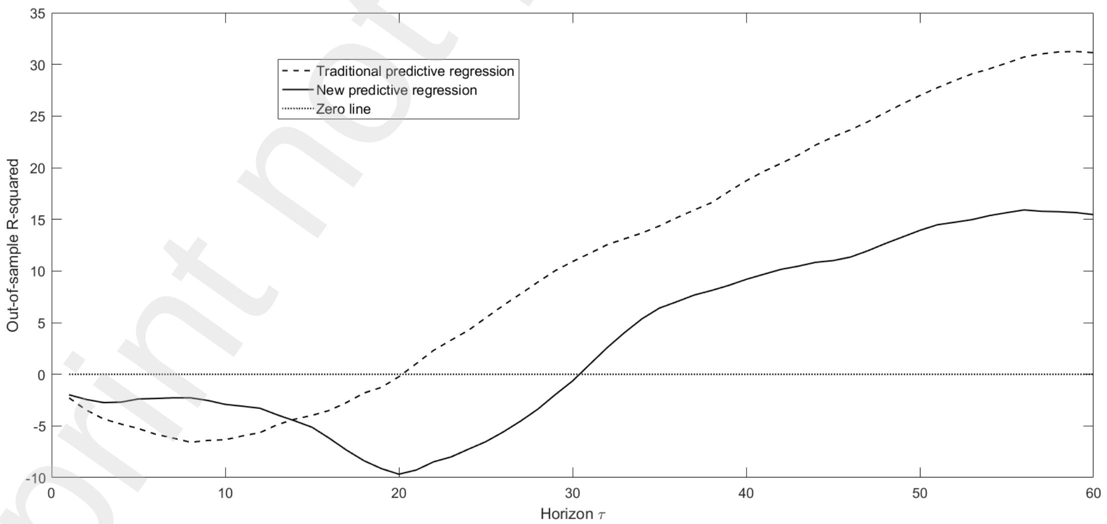

Figure A.12: Out-of-sample predictive powers of competing predictive regressions: France

Notes: The figure displays the out-of-sample predictive powers of alternative predictive regressions: the traditional predictive regression in (9) with the CAPE ratio as the explanatory variable, and the new predictive regression in (12) that conditions the influence of the CAPE to the occurrence of mean reversion in this valuation ratio, as approximated by lagged values of the term spread. Forecasts are obtained using monthly data from January, 1983 to April, 2020, with a total of 448 monthly observations.

Table A.1: Estimation results of stock return predictive regressions: the UK

τ	a_0	a_1	a_2	Adj. R^2 (%)	Adj. R^2 Tradi (%)
1	0.0117***	0.0915	−0.0374*	2.20	0.71
2	0.0185***	0.2268**	−0.0857**	3.59	2.00
3	0.0244***	0.3384**	−0.1248**	4.87	3.03
4	0.0297***	0.4491***	−0.1631***	6.23	3.98
5	0.0342***	0.6022***	−0.2154***	8.49	5.19
6	0.0386***	0.7626***	−0.2703***	11.01	6.44
7	0.0441***	0.8983***	−0.3173***	13.22	7.59
8	0.0502***	1.0201***	−0.3601***	15.15	8.55
9	0.0574***	1.1390***	−0.4027***	17.21	9.43
10	0.0637***	1.2576***	−0.4445***	19.20	10.26
11	0.0690***	1.3596***	−0.4798***	20.49	11.06
12	0.0743***	1.4410***	−0.5079***	21.20	11.81
13	0.0786***	1.5017***	−0.5280***	21.27	12.51
14	0.0824***	1.5561***	−0.5454***	21.24	13.19
15	0.0857***	1.6075***	−0.5614***	21.19	13.85
16	0.0884***	1.6847***	−0.5860***	21.73	14.52
17	0.0924***	1.7856***	−0.6199***	22.86	15.27
18	0.0970***	1.8818***	−0.6525***	24.03	15.97
24	0.1339***	2.2908***	−0.7969***	27.90	19.22
36	0.1791***	2.9484***	−1.0090***	32.53	24.58
48	0.2337***	2.9306***	−0.9980***	29.35	24.83

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (10). The penultimate column displays the adjusted R-squared of this predictive regression (Adj. R^2) and the last column the adjusted R-squared of the traditional regression (Adj. R^2 Tradi). *, **, and *** denote traditional significance at 10%, 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

Table A.2: Estimation results of stock return predictive regressions: Germany

τ	a_0	a_1	a_2	Adj. R^2 (%)	Adj. R^2 Tradi (%)
1	0.0212***	0.0286	−0.0254*	6.11	0.43
2	0.0326***	0.1576	−0.0712**	5.04	1.51
3	0.0433***	0.2422	−0.1016**	4.86	2.61
4	0.0530***	0.3204	−0.1287***	4.63	3.75
5	0.0632***	0.4233	−0.1642***	4.96	4.98
6	0.0722***	0.5529**	−0.2071***	5.49	6.17
7	0.0813***	0.6462**	−0.2383***	5.76	7.33
8	0.0932***	0.7185**	−0.2660***	6.50	8.41
9	0.1076***	0.8693***	−0.3221***	8.56	9.66
10	0.1189***	1.0604***	−0.3877***	10.44	10.94
11	0.1277***	1.2021***	−0.4344***	11.21	12.25
12	0.1388***	1.3061***	−0.4715***	11.98	13.58
13	0.1489***	1.4801***	−0.5302***	13.43	14.98
14	0.1591***	1.6093***	−0.5747***	14.52	16.21
15	0.1686***	1.6702***	−0.5962***	14.66	17.34
16	0.1773***	1.7545***	−0.6247***	15.01	18.47
17	0.1866***	1.8441***	−0.6556***	15.59	19.60
18	0.1969***	1.9433***	−0.6910***	16.60	20.62
24	0.2507***	2.6067***	−0.9155***	22.94	26.33
36	0.3305***	2.9883***	−1.0301***	23.63	34.43
48	0.4207***	3.3066***	−1.1317***	26.40	38.90

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (10). The penultimate column displays the adjusted R-squared of this predictive regression (Adj. R^2) and the last column the adjusted R-squared of the traditional regression (Adj. R^2 Tradi). *. **. and *** denote traditional significance at 10%. 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

Table A.3: Estimation results of stock return predictive regressions: France

τ	a_0	a_1	a_2	Adj. R^2 (%)	Adj. R^2 Tradi (%)
1	0.0164***	−0.0097	−0.0050	2.12	0.27
2	0.0257***	0.0477	−0.0254	1.88	0.97
3	0.0329***	0.1166	−0.0475	1.91	1.65
4	0.0400***	0.1772	−0.0670	2.10	2.29
5	0.0453***	0.2599	−0.0922	2.47	3.05
6	0.0497***	0.3553	−0.1205*	3.09	3.87
7	0.0536***	0.4414*	−0.1454*	3.73	4.70
8	0.0583***	0.5030*	−0.1633*	4.08	5.45
9	0.0637***	0.5810**	−0.1870**	4.74	6.25
10	0.0696***	0.6645**	−0.2130**	5.58	7.02
11	0.0754***	0.7522**	−0.2403**	6.41	7.93
12	0.0816***	0.8383***	−0.2673***	7.22	8.87
13	0.0874***	0.9232***	−0.2937***	8.00	9.79
14	0.0937***	1.0035***	−0.3192***	8.75	10.66
15	0.1002***	1.0735***	−0.3415***	9.35	11.50
16	0.1054***	1.1526***	−0.3658***	10.09	12.31
17	0.1106***	1.2495***	−0.3959***	11.15	13.12
18	0.1153***	1.3423***	−0.4241***	12.13	13.96
24	0.1454***	1.7898***	−0.5591***	16.84	18.80
36	0.2016***	2.3111***	−0.7158***	22.13	24.68
48	0.3002***	2.4208***	−0.7747***	21.49	31.15

Notes: For different values of the prediction horizon τ, the table displays the estimation results of the stock return predictive regression as specified in (10). The penultimate column displays the adjusted R-squared of this predictive regression (Adj. R^2) and the last column the adjusted R-squared of the traditional regression (Adj. R^2 Tradi). *. **. and *** denote traditional significance at 10%. 5% and 1% levels, respectively. Inference is conducted with the robust Newey-West standard error.

Table A.4: CER and differences in CER: France

	New model		Traditional model		Utility gain	
	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$
$\tau = 1$	0.46	0.08	-0.44	-0.55	0.90	0.62
$\tau = 2$	1.45	0.68	-0.50	-0.85	1.95	1.53
$\tau = 3$	0.39	-0.17	-2.12	-2.06	2.51	1.89
$\tau = 4$	0.40	-0.14	-0.87	-0.71	1.27	0.57
$\tau = 5$	0.67	0.17	-2.10	-1.60	2.77	1.77
$\tau = 6$	1.28	0.52	-1.18	-1.22	2.46	1.74
$\tau = 7$	0.92	0.21	-2.19	-2.27	3.11	2.48
$\tau = 8$	1.30	0.45	-0.64	-0.79	1.94	1.24
$\tau = 9$	0.99	0.35	-0.44	-0.76	1.43	1.10
$\tau = 10$	-0.02	0.12	-0.26	-0.82	0.24	0.94
$\tau = 11$	1.14	0.25	0.75	0.06	0.39	0.20
$\tau = 12$	1.51	0.40	-0.54	-1.20	2.04	1.59

Notes: For different values of the relative risk aversion parameter γ and the forecast horizon τ, the table displays the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The last two columns display the differences in CER (utility gain).

Table A.5: CER and differences in CER: Germany

	New model		Traditional model		Utility gain	
	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$
$\tau = 1$	2.56	1.11	0.51	-0.65	2.04	1.76
$\tau = 2$	2.10	0.85	0.54	-0.54	1.56	1.39
$\tau = 3$	1.07	0.27	-0.48	-0.95	1.55	1.22
$\tau = 4$	1.03	-0.04	1.27	0.41	-0.24	-0.45
$\tau = 5$	1.01	0.13	1.69	0.11	-0.68	0.01
$\tau = 6$	1.41	0.50	2.25	1.04	-0.84	-0.54
$\tau = 7$	-0.09	-0.64	0.79	-0.83	-0.87	0.19
$\tau = 8$	0.03	-0.77	2.37	0.78	-2.34	-1.56
$\tau = 9$	0.25	-0.96	3.48	2.46	-3.23	-3.42
$\tau = 10$	0.71	-0.83	3.48	2.30	-2.77	-3.13
$\tau = 11$	0.85	-1.12	3.29	2.32	-2.44	-3.44
$\tau = 12$	-0.38	-2.24	3.25	2.15	-3.64	-4.39

Notes: For different values of the relative risk aversion parameter γ and the forecast horizon τ, the table displays the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The last two columns display the differences in CER (utility gain).

Table A.6: CER and differences in CER: Canada

	New model		Traditional model		Utility gain	
	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 3$	$\gamma = 5$
$\tau = 1$	3.59	2.57	2.04	0.72	1.54	1.85
$\tau = 2$	3.46	1.99	2.65	1.65	0.82	0.34
$\tau = 3$	2.96	1.73	2.98	2.15	-0.01	-0.41
$\tau = 4$	3.09	1.69	2.31	1.92	0.78	-0.23
$\tau = 5$	4.13	2.24	3.16	2.56	0.97	-0.32
$\tau = 6$	4.23	2.22	3.05	2.51	1.18	-0.29
$\tau = 7$	3.61	1.74	2.70	1.40	0.90	0.34
$\tau = 8$	3.08	1.46	3.00	2.06	0.07	-0.61
$\tau = 9$	2.09	0.60	2.83	2.00	-0.74	-1.40
$\tau = 10$	2.63	1.13	2.93	1.91	-0.29	-0.79
$\tau = 11$	2.17	0.81	3.17	1.94	-1.01	-1.14
$\tau = 12$	1.68	0.87	2.70	1.62	-1.02	-0.75

Notes: For different values of the relative risk aversion parameter γ and the forecast horizon τ, the table displays the annualized value of CER (in %) of the new predictive regression, followed by the same statistic for the traditional predictive model. The last two columns display the differences in CER (utility gain).

References

R. D. Arnott, D. B. Chaves, and T.-m. Chow. King of the Mountain: The Shiller P/E and Macroeconomic Conditions. *The Journal of Portfolio Management*, 44(1):55–68, 2017.

R. Becker, J. Lee, and B. E. Gup. An empirical analysis of mean reversion of the S&P 500 P/E ratios. *Journal of Economics and Finance*, 36:675–690, 2012.

C. Boucher. Stock prices-inflation puzzle and the predictability of stock market returns. *Economics Letters*, (90):205–212, 2006.

J. Y. Campbell and R. J. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1:195–227, 1988.

J. Y. Campbell and R. J. Shiller. Valuation ratios and the long-run stock market outlook. *The Journal of Portfolio Management*, 24(2):11–26, 1998.

T. Cenesizoglu and A. Timmermann. Do return prediction models add economic value? *Journal of Banking & Finance*, (36):2974–2987, 2012.

J. H. Cochrane. Explaining the variance of price-dividend ratios. *Review of Financial Studies*, 5:243–280, 1991.

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

A. Estrella. Why does the yield curve predict output and inflation? *The Economic Journal*, 115(505):722–744, 2005.

A. Estrella and G. A. Hardouvelis. The term structure as a predictor of real economic activity. *The Journal of Finance*, 46(2):555–576, 1991.

A. Estrella and F. S. Mishkin. Predicting US recessions: financial variables as leading indicators. *Review of Economics and Statistics*, 1:45–61, 1998.

E. Fama and K. French. Permanent and temporary components of stock prices. *Journal of Political Economy*, 96:246–273, 1988.

E. F. Fama. Term premiums and default premiums in money markets. *Journal of Financial Economics*, 17(1):175–196, 1986.

W. Goetzmann and P. Jorion. Testing the predictive power of dividend yields. *Journal of Finance*, 48(2):663–679, 1993.

S. M. Goldfeld and R. E. Quandt. A Markov model for switching regressions. *Journal of Econometrics*, 1:3–16, 1973.

R. Gomez Cram. Late to recessions: Stocks and business cycle. *Journal of Finance*, (Forthcoming), 2021.

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

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

C. R. Harvey. The real term structure and consumption growth. *Journal of Financial Economics*, 22(2):305–333, 1988.

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

R. J. Hodrick. Dividend yields and expected stock returns: alternative procedures for inference and measurement. *Review of Financial Studies*, 5(3):357–386, 1992.

N. Jegadeesh. Seasonality in Stock Price Mean Reversion: Evidence from the US and the UK. *Journal of Finance*, 46(4):1427–1444, 1991.

R. A. Kessel. The Cyclical Behavior of the Term Structure of Interest Rates. 1965.

C. J. Kim and C. R. Nelson. *State Space Models with Regime Switching, Classical and Gibbs Sampling Approaches with Applications*. Cambridge, MA: MIT Press, 1999.

M. Lettau and S. Van Nieuwerburgh. Reconciling the Return Predictability Evidence. *The Review of Financial Studies*, 21(4):1607–1652, 2008.

J. W. Lewellen. Predicting returns with financial ratios. *Journal of Financial Economics*, 74:209–235, 2004.

C. Marques. Inflation persistence: Facts or artefacts? *European Central Bank Working Paper*, 07 2004.

E. Moench and S. Tobias. Equity premium predictability over the business cycle. *Discussion Paper, Deutsche Bundesbank*, (No 25/2021), 2021.

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

M. S. Rozeff. Dividend yields are equity risk premiums. *The Journal of Portfolio Management*, 11(1):68–75, 1984.

J. H. Stock and M. W. Watson. *New indexes of coincident and leading indicators.* In: Blanchard, O., Fischer, S. (Eds.), 1989 NBER Macroeconomics Annual. MIT Press, Cambridge, MA, 1989.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58(1):267–288, 1996.

R. A. Weigand and R. Irons. The market p/e ratio, earnings trends, and stock return forecasts. *Journal of Portfolio Management*, 33(4):87–101, 2007.