Bear Market Predictability: S&P 500

Abstract:

This study investigates the predictive power of specific series related to the S&P 500 and the broader US macroeconomy in forecasting bear markets. Using a Markov switching autoregressive model to estimate bear market probabilities, the research employs Ordinary Least Squares (OLS) and Beta Regression techniques to conduct both in-sample and out-of-sample tests. These tests demonstrate the predictive capabilities of various series during bearish periods. The empirical analysis identifies significant predictors within the US stock market framework, highlighting the robust predictive power of financial series such as dividends and macroeconomic variables such as inflation rates. In-sample analyses reveal that dividends and total return-scaled earnings consistently predict future periods across horizons from one to 24 months. Out-of-sample tests further underscore the effectiveness of inflation rates, dividends, long-term interest rates, bond prices, total returns, and total return-scaled earnings in forecasting bear markets across horizons one to 24 months ahead. These findings underscore the potential of these variables to inform predictions within financial markets.

Section 1: Introduction

In the intricate domain of financial markets, the desire to predict stock returns has stood as a fundamental endeavor, driving researchers to create and innovate strategies and methodologies to increase prediction accuracy. When seeking a benchmark reflective of the broader US stock market, researchers often turn to Standard and Poor’s S&P 500 index as a reliable gauge. Comprising the 500 largest US corporations, the S&P 500 serves as a pivotal indicator of the US stock market, holding nearly 80% of all US market capitalization (S&P Global).

However, the interest of using publicly available information to make reliable predictions of stock returns is a violation of the hypothesis of semi-strong market efficiency. Despite this, previous frameworks have identified deviations to this hypothesis by identifying bear markets predictors as a trading strategy. Shen (2003) shows that by forecasting bear markets, investors can exploit profitable opportunities by timing their portfolios, rather than using a buy-and-hold strategy. Along with assisting portfolio management, identifying bear markets may also aid in policy altering. Estrella and Mishkin (1998) have established predicting swings in the market provides useful information regarding business cycles, which can deliver useful information to policymakers. Rigobon and Sack (2003) and Bohl et al. (2007) have documented evidence that monetary authorities may respond to the stock market. Therefore, this paper is focused on predicting bear markets.

To predict what causes a bear market, a bear market must first be identified. Academic literature lacks a unanimous definition of what truly identifies a bear market; however, previous frameworks have found parametric Bayesian approach Markov switching autoregressive models to best elucidate bear and bull markets. (Maheu McCurdy, 2000; Chen, 2008; Chen, Chen, and Chou, 2017; and Hammerschmidt and Lohre, 2018) As justified by Ang and Timmerman (2012), Markov switching autoregressive models are adept to a diverse array of market subtleties in asset returns across regimes, including differing autocorrelation, means, and volatilities. This versatility allows the capturing of heteroskedasticity, fat fails, skewness, and time-varying correlations inherent in financial markets.

Utilizing a regime-switching model offers a robust framework for effectively assessing bull and bear markets, given their ability to encapsulate the intricate dynamics and behaviors of financial markets. In this study, probabilities of bear and bull markets are estimated by using a two-state Markov switching autoregressive model and Bayesian filtering methods. Subsequently, predictive regression is employed to explore the potential enhancement of bear market forecasting using various financial and macroeconomic series. Through both in-sample and out-of-sample testing, the predictive performance is evaluated at horizons of 1, 2, 3, 6, 12, 18, and 24 periods (months) using OLS and Beta Regression, proposed by Ferrari and Cribari-Neto (2004). Additionally, finding are compared with similar frameworks from Chen (2008), and Chen, Chen, and Chou (2017) by performing an additional out-of-sample restricted model to compare to the unrestricted model. This identifies if the series improves the accuracy of the forecast.

Empirical results from Rapach et al. (2005) and Chen (2008) of monthly S&P 500 data suggest inflation rates are reliable and useful predictors of bear markets, shown by in-sample and out-of-sample tests. The series tested in this study are inflation, long-range interest rate, bond price, dividend, total return, total return scaled earnings, and total return scaled cyclically adjusted price-earnings ratio. In-sample results showed dividend, inflation, and total return scaled earnings as the best predictors for all horizons, while long-range interest rate and bond price showed predictive accuracy for horizons of one to three months. Out-of-sample forecasting results showed dividend, inflation, and total return scaled earnings as the best predictors, while bond price and total return showed enhancement of the forecast at horizons of one and two months. Back-casting results showed inconsistencies with forecasting results, showing long-range interest rate and bond price as the best back-casting series at horizons one to three, with

TRCAPE and total return scaled earnings as the best back-casting series for horizons spanning 12 to 24 months.

Section 2 of this paper offers a literature review of previous frameworks. Section 3 of this paper demonstrates the economic framework under investigation, specifically the assumptions and violations of the semi-strong Efficient Market Hypothesis. Section 4 introduces the data and their statistical properties, along with modifications to ensure data integrity and the assumptions of time series analysis. Section 5 explores the intricacies of a Markov switching model and its capacity to identify bear markets. Section 6 represents the in-sample predictive methods using OLS and Beta Regression, and their corresponding results. Section 7 shows the out-of-sample predictive estimation using the forecasting and back-casting results of an unrestricted model, in comparison to forecasting and back-casting using a restricted model. Finally, section 8 offers concluding remarks and suggestions for future studies.

Section 2: Literature Review

Chen (2008) examines the impact of series specific to the S&P 500 and US macroeconomy on bearish phases of the S&P 500, examining the importance of predicting the stock returns for the broader US economy. Using a Markov switching autoregressive model, Chen (2008) identifies bull and bear markets, finding them to be persistent in the S&P 500. Chen finds bull regimes continue for 20 months, while bear market regimes persist for 6.67 months. Furthermore, Chen (2008) finds a lower variance and higher mean of return in bull periods, while finding a higher variance and lower mean during bear periods. After testing for predictive accuracy within the series using in-sample and out-of-sample predictive regression, Chen (2008) finds inflation rates to be the most robust predictor of recessionary periods in the S&P 500.

Chen, Chen, and Chou (2017) follow the work of Chen (2008) in investigating the prediction efficacy of recessionary periods in the market using series specific to the CRSP index and US macroeconomy, a dataset not publicly available. Implementing a regime-switching model like Chen (2008), Chen, Chen, and Chou (2017) find a lower variance and higher mean in bull periods, while finding a higher variance and lower mean during bear periods, conforming with economic intuition and the findings of Chen (2008). Chen, Chen, and Chou (2017) perform in-sample and out-of-sample tests for robustness and find inflation rates as the most important macroeconomic series in the forecasting of bear markets, conforming with the findings of Chen (2008). This research serves as a foundational understanding of market behavior.

Section 3: Economic Framework

The Efficient Market Hypothesis (EMH), introduced by Fama (1963, 1965) and Samuelson (1965), suggests that financial markets promptly and accurately represent all available information into asset prices. Within the US economy, financial markets are observed to exhibit a “semi-strong” form of market efficiency. This suggests that stock prices reflect all information regarding fundamental analysis, historical data, and market sentiment, while also swiftly integrating this new information. This renders endeavors to outperform the market using security selection or market timing largely futile.

Despite the prevailing efficiency in financial markets as suggested by the EMH, forecasting bear markets is not entirely impossible due to deviations from semi-strong market efficiency. These deviations may arise due to cogitative biases or limitations, asymmetric information, misinterpretation of data, errors in algorithm trading, or forecasting with non-public or insider information. Furthermore, lags in how quickly asset prices reflect new information can provide windows of opportunity for investors who can interpret information faster than the market.

Documented instances of market efficiency violations include using metrics such as book-to-market ratios, dividend-price, and earnings-price in the forecasting of future stock prices, indicating nonconformities with the EMH (Campbell and Shiller, 1988, 1989; Fama and French, 1988; Goetzmann and Jorion, 1993; Pontiff and Schall, 1998; Lewellen, 1999; Menzly; Lewell, 2004; Santos and Veronesi, 2004; Lettau and Ludvigson, 2005; Chen, 2008; Ang and Timmermann, 2012; Chen, Chen, and Chou, 2017). Alongside financial variables, macroeconomic variables have also demonstrated significance in forecasting future stock returns (Thorbecke, 1997; Shen, 2003; Rapach, Wohar, and Rangvid, 2005; Chen 2008). Furthermore, Shen (2003) shows that market-timing strategies may outperform "buy-and-hold" strategies by using stock and macroeconomic variables to forecast bear markets. The EMH provides a conceptual basis for understanding market efficiency; however, recent empirical evidence suggests deviations from its principles and facilitates the foundation for further investigation on the predictive power of specific series.

Section 4: Data

The monthly data utilized in this study cover a comprehensive sample period spanning 1960M1 to 2023M12. These data were obtained from the U.S. Stock Markets and CAPE Ratio dataset, compiled by the Yale Department of Economics, under the leadership of renowned economist Robert Schiller. All series have been adjusted for inflation. To identify potential predictors of bear markets, seven key series are selected from this dataset. These predictors encompass a range of financial and macroeconomic indicators, each chosen for its relevance and potential impact on market downturns. Specifically, the predictors include:

Inflation (Consumer Price Index): Mirroring the general trend in prices, inflation serves as a fundamental influence on market sentiment.

Long-Term Interest Rate: This series represents 10-year US Treasury yield, tracking changes in borrowing costs and providing valuable insights into investor expectations regarding future economic conditions.

Average Bond Price: Derived from discounting expected cash flows, this series offers valuable insights into bond market dynamics, future economic conditions, and interest rate movements.

Dividend: Dividend disbursement by S&P 500 corporations represents the profitability and financial well-being of firms.

Total Return: Representing the overall change in the value of the S&P 500 index, total return serves as a broad measure of market performance and investor returns.

Total Return Scaled Earnings: This series evaluates the valuation of stocks relative to their total return, providing insights into market valuation dynamics and investor expectations.

Total Return Scaled Cyclically Adjusted Price Earnings Ratio (TRCAPE Ratio): Offering an assessment of companies' valuation scaled to the total return of holding the stock over time, the TRCAPE ratio is an indicator of market valuation and potential overvaluation.

To adhere to the assumptions of time series analysis, stationarity testing is performed using both Augmented Dicky-Fuller and Phillips-Perron unit root tests. The results indicated all series fail to reject the null hypothesis of possessing a unit root, implying non-stationarity. In response, methodologies analogous to Chen (2008) and Chen, Chen, and Chou (2017) such as moving averages and financial ratios were considered. However, to enhance interpretability, simplify the model, and maintain data integrity, the first differences were implimented. Following this adjustment, all predictor series exhibited stationarity, with p-values < 0.01, affirming them suitable for time-series analysis. Additionally gathered from the dataset is the S&P 500 price, which is used to estimate transition and filtered probabilities.

Section 5: Markov Switching Model

Drawing upon approaches from Maheu and McCurdy (2000), Frauendorfer, Jacoby, and Schwendener (2007), Chen (2008), Chen, Chen, and Chou (2017), and Hammerschmidt and Lohre (2018), the identification of bear and bull periods in the stock market is best modeled by a Markov-switching model. Using a Bayesian approach, a two-regime Markov switching autoregressive (MS-AR) model is implemented to identify bear and bull markets. The possibility of a three regime model that includes extreme bull or extreme bear periods may be possible, however for interpretation and simplicity a two-regime model is chosen. The two-regime MS-AR(p) model of stock returns with autoregressive lag length p is described as follows:

Two Regime Markov-switching autoregressive model:

$r_t = \alpha_{S_t} + \mu_{S_t} + \in_{S_t}$

Where $r_t = 100 * \log(S\&P\ 500\ Price)$, α_{S_t} and μ_{S_t} represent the regime-dependent intercept and mean, while $\in_{S_t}$ represents the regime-dependent error term.

The stock returns are assumed to follow a two-regime Markov process with fixed transition probabilities, these probabilities are determined by logistic functions and are represented in a probability matrix as seen in figure 5.1.

Figure 5.1:

$$P= \begin{bmatrix} P\ (s_t = 0\ |s_{t-1} = 0) & P\ (s_t = 0\ |s_{t-1} = 1) \\ P\ (s_t = 1\ |s_{t-1} = 0) & P\ (s_t = 1\ |s_{t-1} = 1) \end{bmatrix}$$

$$= \begin{bmatrix} p^{00} & p^{01} \\ p^{10} & p^{11} \end{bmatrix}$$

s_t represents the state (bull or bear market), p^{00} represents the probability of staying in regime 0 (bear market), p^{01} represents the probability of transitioning from regime 0 to 1 (bear market to bull market), p^{11} represents the probability of staying in regime 1(bull market), and p^{10} represents the probability of transitioning from regime 1 to 0 (bull market to bear market).

Additionally, the expected duration of regimes can be calculated by:

Bear Market (Regime 0): $\frac{1}{1-p^{00}}$

Bull Market (Regime 1): $\frac{1}{1-p^{11}}$

After the regimes have been specified, filtered probabilities for each regime are estimated by:

$Q_{j,t} = P(S_t = j|y^t)$, j = {0,1}

Where y^t represents all available information up to time t; j signifies bounded values between 0 and 1.

Accordingly: $Q_{0,t} = P(S_t = 0|y^t)$ are estimates of probabilities of the bear market at time t.

Transition probabilities represent the probability of transitioning between regimes with a defined timeframe. These probabilities remain fixed parameters that are unchanged by temporal shifts and are based on historical patterns of r_t. Conversely, filtered probabilities indicate the probability of existing within a particular regime at a specific time, based on observed data. Unlike transition probabilities, filtered probabilities are dynamic values and evolve with new data. These probabilities are extracted using the Bayesian filtering technique Kalman Filter to identify discerning trends and patterns amidst market volatility.

Optimal lag length is determined through the Psaradakis–Spagnolo Akaike information criterion (AIC). Yielding values of MS-AR(1): 4015.384, MS-AR(2): 4654.216, MS-AR(3): 4975.541, MS-AR(4): 5200.042. In this case, MS-AR(1) is determined to be the most suitable model. The MS-AR(1) reveals distinct characteristics for bear (regime 0) and bull (regime 1) markets. For bear markets, the mean is estimated at 0.1209 with a residual standard error of

5.4568, while bull markets show a mean of 2.2418 and a residual standard error of 2.4655. Lower mean and higher variance of returns for bear market periods is consistent with economic theory and the findings of Chen (2008) and Chen, Chen, and Chou (2017). The transition probability matrix of MS-AR(1) is shown in figure 5.2.

Figure 5.2: Transition Probabilities within the S&P 500

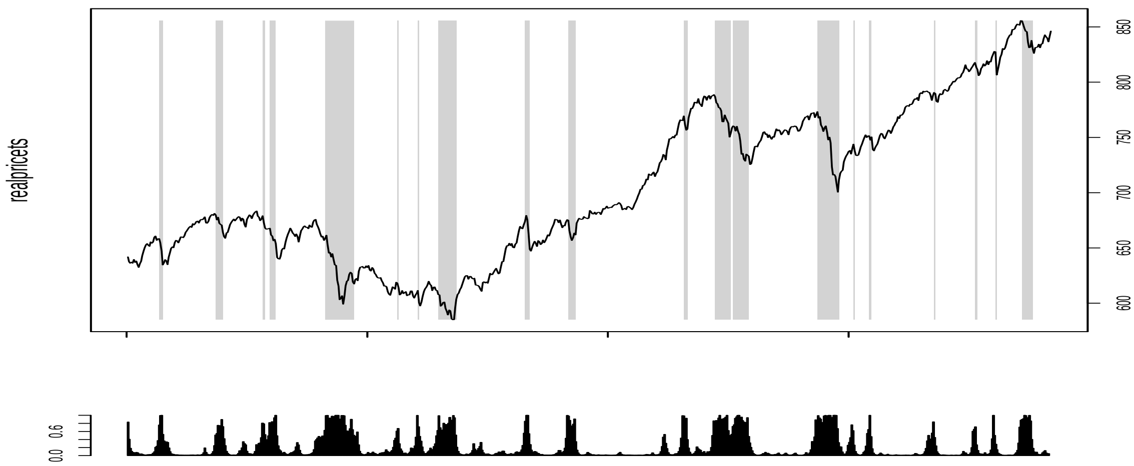
	$Bear\ (0)$	$Bull\ (1)$
$Bear\ (0)$	0.833	0.167
$Bull\ (1)$	0.055	0.945

(94.5% chance of being in a bull market and staying in a bull market, 5.5% chance of being in a bull market and transitioning into a bear market. 83.3% chance of being in a bear market, 16.7% chance of being in a bear market and transitioning into a bull market.)

The transition probabilities show that both bear markets and bull markets are persistent in the S&P 500. Bear markets ($\frac{1}{1-0.833}$) last an estimated 6 months, and bull markets ($\frac{1}{1-0.945}$) last an estimated 18.2 months, consistent with findings of Chen (2008).

The filtered probabilities for bear markets (Q_0) are shown in figure 5.3

Figure 5.3: Filtered probabilities of bear markets in the S&P 500



Section 6: Predictive Regression and In-sample tests

Following the attainment of filtered probabilities from MS-AR(1), predictive regression based on the methods of Chen (2008) and Chen, Chen, and Chou (2017) is modeled as:

$$Q_{0,t+k} = \alpha + \beta x_t + e_{t+k}$$

where $Q_{0,t+k}$ represents the bear market probability at a time t plus horizon k (1, 2, 3, 6, 12, 18, 24) months ahead, and x_t denotes the predictive series under investigation.

In-sample OLS and Beta Regression tests are conducted to evaluate predictability, testing the null hypothesis of no predictive power ($\beta = 0$) against the alternative hypothesis ($\beta \neq 0$). The null hypothesis being consistent with the EMH, while the alternative is a deviation from the EMH.

OLS Regression revealed that inflation demonstrated predictive power at k=2, long-range interest rates at k= 1, 2, and 3, bond price at k=1, 2, and 3, dividend at k= 1, 2, 3, 6, 12, 18, and

24, total return at k=1, total return scaled earnings at k=1, 2, 3, 6, 12, 18, and 24, and TRCAPE at k= 1, 12, and 18. See figure 6.1. Notably, total revenue-scaled earnings exhibited the most predictive power, with p-values <0.01 across all horizons, followed closely by dividend with similar significance, excluding k=24 ($p < 0.10$). This represents noteworthy deviations from the semi-strong EMH.

Table 6.1: OLS In-Sample Testing

	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	-0.005	-0.034**	-0.025	-0.011	0.001	-0.013	0.016
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Observations	767	766	765	762	756	750	744
R^2	0.0001	0.005	0.003	0.001	0.00001	0.001	0.001
Adjusted R^2	-0.001	0.004	0.001	-0.001	-0.001	-0.001	-0.0002
LRIR10	-0.111***	-0.152***	-0.103***	0.004	0.011	-0.017	-0.016
	(0.036)	(0.036)	(0.037)	(0.037)	(0.037)	(0.037)	(0.036)
Observations	767	766	765	762	756	750	744
R^2	0.012	0.022	0.010	0.00001	0.0001	0.0003	0.0003
Adjusted R^2	0.011	0.021	0.009	-0.001	-0.001	-0.001	-0.001
Real Bond	0.063***	0.054***	0.030*	-0.0003	-0.007	0.008	-0.008
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Observations	767	766	765	762	756	750	744
R^2	0.017	0.013	0.004	0.00000	0.0002	0.0003	0.0003
Adjusted R^2	0.016	0.011	0.003	-0.001	-0.001	-0.001	-0.001

Dividend	-0.195***	-0.206***	-0.258***	-0.325***	-0.335***	-0.189***	-0.078*
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.043)	(0.043)
Observations	767	766	765	762	756	750	744
R^2	0.027	0.030	0.047	0.075	0.079	0.026	0.005
Adjusted R^2	0.026	0.029	0.046	0.073	0.078	0.025	0.003
Total Return	-0.00000***	-0.00000	0.00000	0.00000	0.00000	0.00000	-0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Observations	767	766	765	762	756	750	744
R^2	0.108	0.002	0.0001	0.00000	0.001	0.002	0.003
Adjusted R^2	0.107	0.001	-0.001	-0.001	-0.0003	0.001	0.001
TR Scaled Earnings	-0.0001***	-0.0001***	-0.00005***	-0.00003***	0.00003***	0.00003***	0.00002***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Observations	767	766	765	762	756	750	744
R^2	0.085	0.076	0.056	0.016	0.015	0.025	0.010
Adjusted R^2	0.084	0.074	0.055	0.015	0.013	0.024	0.009
TRCAPE	-0.128***	-0.017	0.012	0.007	0.019*	0.019*	-0.010
	(0.010)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)
Observations	767	766	765	762	756	750	744
R^2	0.163	0.003	0.002	0.0004	0.004	0.004	0.001
Adjusted R^2	0.162	0.002	0.0002	-0.001	0.002	0.002	-0.0003

Note: *p**p***p<0.01

Beta Regression analysis revealed similar results as seen in figure 6.2. The model shows inflation, long-range interest rates, bond prices, dividends, total returns, total return scaled earnings, and TRCAPE showing predictive power across various horizons. Notably, total revenue-scaled earnings displayed consistent predictive power across all horizons ($p < 0.01$, excluding k=24). These results exhibit further deviations from the EMH.

Table 6.2: Beta Regression In-Sample Testing

	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	-0.217***	-0.444***	-0.268***	0.003	0.035	-0.002	0.158**
	(0.072)	(0.070)	(0.071)	(0.072)	(0.072)	(0.072)	(0.072)
Observations	767	766	765	762	756	750	744
R^2	0.006	0.039	0.013	0.00000	0.0003	0.00000	0.006
LRIR10	-0.475***	-0.666***	-0.748***	0.017	-0.094	-0.040	0.021
	(0.157)	(0.156)	(0.156)	(0.158)	(0.158)	(0.158)	(0.157)
Observations	767	766	765	762	756	750	744
R^2	0.008	0.017	0.020	0.00001	0.0004	0.0001	0.00002
Real Bond	0.413***	0.426***	0.354***	-0.006	0.037	0.013	-0.103
	(0.073)	(0.073)	(0.074)	(0.074)	(0.075)	(0.074)	(0.074)
Observations	767	766	765	762	756	750	744
R^2	0.027	0.028	0.027	0.00001	0.0003	0.00004	0.002
Dividend	-0.512***	-0.480***	-0.874***	-1.642***	-1.742***	-0.686***	-0.219
	(0.184)	(0.185)	(0.184)	(0.182)	(0.181)	(0.185)	(0.185)
Observations	767	766	765	762	756	750	744
R^2	0.008	0.006	0.023	0.074	0.086	0.014	0.002
Total Return	-0.00002***	-0.00000	0.00000**	0.00000	0.00000	0.00000*	-0.00000**
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Observations	767	766	765	762	756	750	744
R^2	0.249	0.001	0.004	0.0002	0.003	0.003	0.006

TR Scaled Earnings	-0.0004***	-0.0003***	-0.0003***	-0.0001***	0.0002***	0.0002***	0.0001*
	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Observations	767	766	765	762	756	750	744
R^2	0.107	0.087	0.075	0.015	0.028	0.030	0.004
TRCAPE	-0.854***	-0.072	0.101**	0.029	0.108**	0.095*	-0.058
	(0.046)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)
Observations	767	766	765	762	756	750	744
R^2	0.291	0.002	0.004	0.0004	0.005	0.004	0.002
Note:							*p**p***p<0.01

The estimates of β align with economic intuition: inflation decreases short-term bear market probabilities while increasing them in the long run; long-range interest rates exhibit a negative relationship with bear markets at k =1,2, and 3 due to their association with economic prosperity; bond prices demonstrate a positive relationship with bear markets at short horizons likely attributed to investors seeking refuge in bonds during bear markets; dividends exhibit a consistent negative relationship with bear market probability across all horizons, indicating flourishing firms and market confidence. Total returns and total revenue-scaled earnings show nuanced relationships, reflecting the cyclical nature of market dynamics as shown by the MS-AR model.

Section 7: Out-of-sample Tests:

Forecast/Back-Cast Comparison:

Restricted Model: $Q_{0,t+k} = \alpha_1 + e_{1t+k}$

Unrestricted Model: $Q_{0,t+k} = \alpha_2 + \beta x_t + e_{2t+k}$

The restricted and unrestricted models are compared for forecasting and back-casting accuracy using the mean squared prediction error (MSPE). The unrestricted model holds the predictor under investigation, while the restricted is an intercept-only model that excludes the predictor. The null hypothesis assumes equality in the forecasting/back-casting power between the unrestricted and restricted models $\frac{MSPE_u}{MSPEr} = 1$, with the alternative hypothesis indicating superior performance of the unrestricted model $\frac{MSPE_u}{MSPEr} < 1$. The formula for MSPE is as follows:

MSPE $= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

where n is the number of observations, y_i is the actual value of the dependent variable for observation i, and $\hat{y}_i$ is the predicted value for of the dependent variable for observation i.

For the forecasting process, data from 1960 to 2010 are used to train the model, after training and prediction using OLS and Beta Regression the predictions are compared to data from 2010 to 2023. For back-casting, data from 1972 to 2023 are used to train the model, after training and prediction using OLS and Beta Regression the predictions are compared to data from 1960 to 1972.

Both OLS Regression and Beta Regression demonstrate the forecasting efficacy of all series at k=1 and 2. Dividends emerge as the most effective series for forecasting, showing remarkable MSPE ratios of less than 0.3 for horizons k= 1, 2, 3, 6, 12, and 18. Other series also showed significant forecasting abilities across various lengths, as seen in table 7.1. Values in bold show the unrestricted model performs better than the restricted model.

See Table 7.1: OLS Forecasting (MSPE ratios)

Horizon	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	**0.917**	**0.658**	**0.738**	**0.860**	**0.901**	1.002	1.091
Dividend	**0.004**	**0.006**	**0.124**	**0.295**	**0.116**	**0.217**	1.013
Long Range Interest Rate	**0.982**	**0.976**	**0.976**	1.006	**0.998**	**0.993**	**0.988**
Bond Price	**0.828**	**0.822**	**0.868**	1.069	**0.952**	**0.994**	**0.963**
Total Return	**0.014**	**0.587**	**0.904**	**0.929**	**0.964**	1.303	**0.912**
TR Scaled Earnings	**0.390**	**0.412**	**0.500**	**0.764**	1.316	1.662	1.454
TR CAPE	**0.826**	**0.968**	1.013	1.012	1.028	1.032	**0.989**

Beta Regression forecasting yields similar results, with all series proving useful for forecasting at k=1 and 2. Again, dividend stands out as the top forecasting series, with remarkable MSPE ratios of less than 0.6 for horizons k= 1, 2, 3, 6, 12, and 18. Other series also showed significant forecasting abilities at various lengths. These out-of-sample findings align well with the in-sample results, indicating the robustness of the models, as seen in table 7.2. Values in bold show the unrestricted model performs better than the restricted model.

Table 7.2: Beta Regression Forecasting (MSPE ratios)

Horizon	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	**0.825**	**0.732**	**0.807**	**0.975**	**0.982**	1.050	1.062
Dividend	**0.495**	**0.556**	**0.389**	**0.206**	**0.244**	**0.591**	1.046
Long Range Interest Rate	**0.994**	**0.981**	**0.958**	1.001	**0.992**	**0.998**	**0.997**
Bond Price	**0.959**	**0.927**	**0.866**	1.004	**0.938**	1.002	**0.998**
Total Return	**0.442**	**0.817**	1.027	**0.888**	1.100	1.237	**0.998**
TR Scaled Earnings	**0.742**	**0.747**	**0.762**	**0.877**	1.190	1.384	1.234
TR CAPE	**0.552**	**0.986**	1.009	**0.997**	1.020	1.020	**0.998**

OLS Regression back-casting reveals starkly different outcomes from forecasting as seen in table 7.3. Only long-range interest rates and bond prices exhibit back-casting power at short horizons (k=1 and 2), while other series such as total return scaled earnings and TRCAPE demonstrate back-casting effectiveness at longer-term horizons. Values in bold show the unrestricted model performs better than the restricted model.

Table 7.3: OLS Back-Casting (MSPE ratios)

Horizon	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	1.265	2.026	1.794	1.440	1.101	1.565	**0.356**
Dividend	1.807	1.894	2.199	2.859	2.788	2.183	1.989
Long Range Interest Rate	**0.895**	**0.854**	**0.895**	1.029	1.017	**0.990**	**0.948**
Bond Price	**0.877**	**0.891**	**0.938**	1.011	1.019	**0.979**	1.065
Total Return	1.697	1.091	**0.983**	1.002	**0.936**	**0.872**	1.321
TR Scaled Earnings	1.685	1.662	1.579	1.325	**0.737**	**0.585**	**0.512**
TR CAPE	1.239	1.023	**0.985**	**0.989**	**0.970**	**0.919**	1.141

Beta Regression back-casting demonstrates back-casting power for all series except inflation and dividends at k=1, with other series displaying back-casting effectiveness at various horizons, as seen in Table 6.4. These disparities between back-casting and forecasting results could be attributed to notable differences in financial market dynamics over the past 50 years. Values in bold show the unrestricted model performs better than the restricted model.

See Table 6.4: Beta Regression Back-Casting (MSPE ratios)

Horizon	K=1	K=2	K=3	K=6	K=12	K=18	K=24
CPI	1.261	1.407	1.330	1.022	**0.982**	1.019	**0.823**
Dividend	1.093	1.085	1.160	1.255	1.227	1.133	1.055
Long Range Interest Rate	**0.966**	**0.954**	**0.938**	1.004	**0.993**	1.002	1.003
Bond Price	**0.897**	**0.915**	**0.905**	1.002	**0.995**	1.000	1.015
Total Return	**0.753**	1.014	**0.968**	**0.994**	**0.977**	**0.970**	1.042
TR Scaled Earnings	**0.942**	**0.985**	**0.997**	1.058	**0.857**	**0.889**	**0.963**
TR CAPE	**0.517**	1.005	**0.991**	**0.999**	**0.992**	**0.984**	1.017

Section 7: Concluding Remarks

In this paper, I investigated the effectiveness of forecasting bear markets using pertinent S&P 500 and US macroeconomy series, with a specific focus on series intrinsic to the S&P 500. The aim was to enhance bear market prediction, thereby challenging the semi-strong form of the Efficient Market Hypothesis (EMH). My investigation revealed significant deviations from the semi-strong form of the EMH, specifically by finding multiple series as robust predictors of bull markets through in-sample and out-of-sample testing. Among the predictors, dividend yield and total return scaled earnings emerged as the most consistent at all horizons, while inflation, long-range interest rates, bond price, total return, and total return scaled earnings were also shown to have significant prediction power. These findings not only show the limitations of the EMH, but also offer valuable insights in the forecasting of bear markets. By identifying robust predictors of bear markets and challenging the assumptions of the EMH, this study aids in a deeper understanding of the financial market, providing benefits for policymakers and investors alike.

While my results have generally showed robustness in predictive power, there are still certain limitations and caveats in my study. It is crucial to address the inconsistencies observed in the back-casting results, as they have potential limitations in the reliability and robustness of the predictive model. These discrepancies in the back-casting results compared to actual market outcomes suggest that the model may not entirely capture the complexities of real-world market dynamics. These inconsistencies undermine the confidence in the model's predictive capabilities and further highlight the importance of critical evaluation of different time periods.

Furthermore, despite robust in-sample results and forecasting results, relying solely on in-sample and out-of-sample forecasting results to assess the effectiveness of the predictive model is inherently flawed. Just because the market performed well in retrospect does not guarantee its success in forecasting future bear markets. Financial markets themselves are inherently dynamic and are subject to various unpredictable factors, making this process a tedious and difficult one. Therefore, while these models may provide valuable insights for forecasting bear markets in the S&P 500, they must be complimented with rigorous validation and testing to assess reliability and generalizability before being used to make consequential decisions. By acknowledging the discrepancies in forecasting and back-casting results and emphasizing the need for robust validation, researchers can ensure more thorough and accurate models in the financial domain.

Future research efforts should prioritize the adoption of advanced statistical (machine) learning techniques to uncover further relationships that may elude detection by conventional statistical methods such as OLS and Beta Regression. Embracing innovative approaches to time series data modeling offers researchers the opportunity to contribute to the development of more effective forecasting methodologies. Through ongoing exploration, we can strive toward greater precision in decision-making and risk management within financial markets.

References:

Ang, Andrew, and Allan Timmermann. "Regime changes and financial markets." Annu. Rev. Financ. Econ. 4, no. 1 (2012): 313-337.

Bohl, Martin T., Pierre L. Siklos, and David Sondermann. "Shocking markets: European stock markets and the ECB’s monetary policy surprises." Available at SSRN 1091133 (2007).

Campbell, John Y., and Robert J. Shiller. "The dividend-price ratio and expectations of future dividends and discount factors." The review of financial studies 1, no. 3 (1988): 195-228.

Chen, Nan-Kuang, Shiu-Sheng Chen, and Yu-Hsi Chou. "Further evidence on bear market predictability: The role of the external finance premium." International Review of Economics & Finance 50 (2017): 106-121.

Chen, Shiu-Sheng. "Predicting the bear stock market: Macroeconomic variables as leading indicators." Journal of Banking & Finance 33, no. 2 (2009): 211-223.

Estrella, Arturo, and Frederic S. Mishkin. "Predicting US recessions: Financial variables as leading indicators." Review of Economics and Statistics 80, no. 1 (1998): 45-61.

Fama, Eugene F., and Kenneth R. French. "Dividend yields and expected stock returns." Journal of financial economics 22, no. 1 (1988): 3-25.

Ferrari, Silvia, and Francisco Cribari-Neto. "Beta regression for modelling rates and proportions." Journal of applied statistics 31, no. 7 (2004): 799-815.

Frauendorfer, Karl, Ulrich Jacoby, and Alvin Schwendener. "Regime switching based portfolio selection for pension funds." Journal of Banking & Finance 31, no. 8 (2007): 2265-2280.

Goetzmann, William N., and Philippe Jorion. "Testing the predictive power of dividend yields." The Journal of Finance 48, no. 2 (1993): 663-679.

Hammerschmid, Regina, and Harald Lohre. "Regime shifts and stock return predictability." International Review of Economics & Finance 56 (2018): 138-160.

Lettau, Martin, and Sydney C. Ludvigson. "Expected returns and expected dividend growth." Journal of Financial Economics 76, no. 3 (2005): 583-626.

Lewellen, Jonathan. "The time-series relations among expected return, risk, and book-to-market." Journal of financial Economics 54, no. 1 (1999): 5-43.

Lewellen, Jonathan. "Predicting returns with financial ratios." Journal of Financial Economics 74, no. 2 (2004): 209-235.

Maheu, John M., and Thomas H. McCurdy. "Identifying bull and bear markets in stock returns." Journal of Business & Economic Statistics 18, no. 1 (2000): 100-112.

Menzly, Lior, Tano Santos, and Pietro Veronesi. "Understanding predictability." Journal of Political Economy 112, no. 1 (2004): 1-47.

Pontiff, Jeffrey, and Lawrence D. Schall. "Book-to-market ratios as predictors of market returns." Journal of financial economics 49, no. 2 (1998): 141-160.

Rapach, David E., Mark E. Wohar, and Jesper Rangvid. "Macro variables and international stock return predictability." International journal of forecasting 21, no. 1 (2005): 137-166.

Rigobon, Roberto, and Brian Sack. "Measuring the reaction of monetary policy to the stock market." The quarterly journal of Economics 118, no. 2 (2003): 639-669.

Shen, Pu. "Market-timing strategies that worked." Available at SSRN 445920 (2002)

Thorbecke, Willem. "On stock market returns and monetary policy." The Journal of Finance 52, no. 2 (1997): 635-654.