IAQF Annual Academic Competition 2022

What State Are We In?

Team QuAntiVenom

Abstract

Changing *animal spirits* often govern market behavior over time, with bullish confidence dictating market rises and bearish sentiment bringing about corrections, while oftentimes these forces counteract each other leading to a static regime. Recognizing the market regime can be of paramount importance for investors and traders for implementing market-timing-based strategies. We train a three state Hidden Markov Model (HMM) to identify market regimes using only past Russell 3000 index returns. We find that the HMM trained on this purely technical data performs better than a buy and hold model. However, it often mis-classifies states compared to a "true" Bull-Bear-Static classification method based on standard heuristic methods used by economists. We propose a novel approach of combining the results of the Hidden Markov Model with the states detected by some leading indicators of market regime changes. The profit of the trading strategy obtained using this approach exceeds the gains using the Hidden Markov Model from 2018 to 2021.

1 Introduction

Financial markets are ever-changing like the sea. While some changes may be transitory, the changed behavior often persists for many periods. The recent turmoil spurred by the pandemic-induced financial crisis presents an exhibit of the impact of a paradigm shift in investors' assessment of risky assets. The mean, volatility, and correlation patterns in stock returns, for instance, changed drastically at the start of and continued through the global financial crisis of 2007-2009. Similar regime changes, some of which can be recurring like recessions vs expansions, and some of which can be permanent (structural breaks), are prevalent across a wide range of financial markets and in the behavior of many macroeconomic variables. The observed regimes in financial markets are closely related to the phases of the business cycle as well as other factors. Regime changes present a big challenge to traditional trading strategies, demanding a more adaptive approach. Moreover, the time-varying behavior of risk premiums, volatility, and correlation has important implications for risk management.

It is generally accepted that one cannot predict the price of individual security, but it may be possible to determine the “hidden” state of the market. To that end, market regime detection models provide the ability to identify these hidden states to adjust one's portfolio efficiently. This study aims to develop a technique to determine the state of the market - bull, bear or static. Bull markets are characterized by expansions, while bear markets are a consequence of sudden market crashes. While static markets lack a general consensus in definition, in essence, they are periods which although possibly turbulent, give an investor no incentive to stay in the market, as there is no long-term expansion or contraction.

This study has been divided into several segments. We start with defining benchmarks for our predictive models, against which we measure the validity of our proposed model. Next, we briefly explain the theory behind the Hidden Markov Model and how the Expectation-Maximization algorithm and its variations can be effectively used to identify different states of the market (regimes). We also run through various macro-economic and financial indicators that are expected to be useful in identifying the market state, both in-sample and out of sample. Finally, we propose our market state detection model and examine how it measures up against the benchmarks we defined as well as the standard buy-and-hold strategy. We conclude the study by depicting the results of various models and discussing their future implications.

2 Laying the foundation

We need to define a benchmark of market states to examine the accuracy of the state prediction algorithm that we will use. This will allow us to validate our model by comparing the PnLs generated by the classified market states to the PnLs generated by the benchmark market states. The following subsection presents how we decide these "true states". (Note: The actual inherent market state is unknown, so throughout the study, whenever we refer to "true states", we mean the states that will be assigned using the following algorithm.)

Figure 1: True States ● Bull ● Bear ● Static

2.1 Defining the "True States"

Here, as we attempt to assign the "true state" in historical data, we have the luxury of retrospectively identifying the prevalent state. The steps that we undertake[1] to define these "true states" of the market are as under:

- If the Russell 3000 index rises by at least 10% from the previous market low till the next market high, label it as true market state BULL.
- If the Russell 3000 index drops by at least 10% from the previous market high till the next market low, label it as true market state BEAR.
- For identification of static states, we look at the change in index levels during a 3 month period.
  - If change > 20%, call it true market state BULL
  - Else if change < -15%, call it true market state BEAR
  - Else, call it true market state STATIC

We also combine consecutive high change states into one state and check the change in index value in all resultant states as well.

Following are the identified true states according to the above algorithm.

2.2 Basing a trading strategy

If the market state can be accurately identified, a trading strategy can be based as follows. On the first occurrence of a bull state in the sequence, a long position is taken in the market, and that position is held as long as the market remains in the same state. Similarly, a short position is taken on the first occurrence of a bear state and is held as long as the market remains in a bear state. If the identified market state is static, no active position is taken in the market, and the entire capital is invested in 3-month T-Bills (assumed as risk-free). As soon as a new state is identified in the sequence, the existing position is closed out and a position based on the next identified state is taken. Such a strategy ensures the investor benefits from market movements in both directions and opts to take a risk-off position in cases where there is no clear direction of the market.

[1]Bull and bear rules referred from here

3 Hidden Markov Model

3.1 Motivation

The data set used in our analysis[2] is the Russell 3000 index starting from 1995 to 2018, which covers different market regimes as well as market crashes. Such a time series data consists of many fluctuations of price movements over time. If we disregard short-term fluctuations, we can see that the price movements can be classified into regimes of bull, bear, and static accompanied by periods of high and low volatility. One can leverage these market conditions into constructing a market-timing-based trading strategy to maximize profits. Therefore, classifying such regimes becomes important for investors to take trading positions in the market.

Modeling time series is used to forecast future events based on past observations. In reality, financial markets depend on various factors which contribute to the price changes, that are not explained using a single model. Instead, we attempt to identify market regimes. This is built on the intuition that for different time windows, we may get different market states.

[2]Sourced from YahooFinance

Modeling and forecasting next-day returns is difficult for the reasons mentioned above, but for traders, identifying current regimes can be useful so that they can make informed decisions. Therefore, we need to consider several factors in a silo as well as in conjunction which essentially govern the changes in the market regime. For modeling the market states, we first consider the Hidden Markov Model (HMM).

HMM assumes that the time series consists of an unobservant (hidden) state which is a Markov process, making the future states independent of the past states, given the present state (essentially making the model memoryless). This will help us capture the stochastic nature of the time series data using observable states. These hidden states will determine the behavior of the index prices which is hidden from the trader/investor. Therefore, we can model HMM on our historical data of the Russell 3000 index and identify the hidden state. This will classify the market regime although will not directly forecast values

3.2 Identifying market state using Russell 3000 index price

Put in simpler terms, Hidden Markov Model is a probabilistic model that gives an idea about the hidden truth, given the observed data of the model. These models are useful in finance to obtain information about a non-observable metric like the market state (i.e. Bear, Bull, or Static) given the observed values like stock prices, index values, returns, etc. In technical terms, the non-observable metric is referred to as the state variable whereas the observed data is referred to as the observed variable.

Here, the state variable is the hidden state of the market. Let us denote this state variable at time t by S_t. The model accepts a hyper-parameter in the form of the number of distinct states to be identified in the data. For our model, we attempt to identify 3 distinct states. The observed variable at time t is the returns on Russell 3000 index at time t, denoted by R_t. (Note that the state of the market is not observable and hence is a latent variable for the model). Given a hidden state, there will be a fixed set of probabilities of transition to the next hidden state. This probability information is captured in state transition matrix A. Also, given the hidden state of the market, there is again a fixed set of probabilities of it corresponding to a specific value or returns. This probability information is captured in state emission matrix B. Finally, the initial state of the market is denoted by π_0. The idea behind HMM is to tune these three parameters (i.e. A, B and π_0) using a modified version of expectation maximization (which is generally used for parameter estimation of models with latent variables) and then use the tuned parameters to predict the hidden state of the market. The process is naturally divided into two parts - Learning and Decoding[3].

3.2.1 Learning

The training data used here range from 1995 to 2017. For this part, the Baum-Welch algorithm is used, which is a special case of expectation maximization (EM) algorithm, building on the basic EM algorithm using a forward-backward algorithm as follows:

Step 1: Initialize the state transition matrix A, state emission matrix B and the initial state of the market as π_0 to some random values.

Step 2: Compute the following probabilities using forward - backward technique:

- $\alpha_i(t)$: The forward probability α of being in a particular market state at time t is calculated during the forward pass using observed variables up to time t.
- $\beta_i(t)$: The backward probability β of observing T-t observable values from time t onward given a particular market state at time t is calculated through the backward pass.

Step 3: Use these computed probabilities to find a better estimate for the parameters (A, B, π_0).

- Define temporary probability variables:
  (i) $\gamma_i(t)$ is probability of being in market state i at time t given the observed sequence of returns and the parameters (A, B, π_0)
  (ii) $\xi_{ij}(t)$ is the probability of being in market states i and j at times t and t+1 respectively given the observed sequence of returns and parameters (A, B, π_0) for Russell 3000 index.
- Update the parameters (A, B, π_0) using these temporary probability variables:
  (i) π_{0i}, which is the expected frequency spent in market state i at time 0
  (ii) a_{ij}, which is the expected number of transitions from market state i to market state j compared to the expected total number of transitions away from market state i
  (iii) b_{ij}, which is the expected number of times the returns have been equal to R_i while in market state j over the expected total number of times in market state j.

Step 4: Iterate through steps 2 and 3 until parameters (A, B, π_0) converge.

[3]Refer here for detailed formulae on learning and decoding

3.2.2 Decoding

Once the trained parameters (state transition matrix A, state emission matrix B, and initial state of the market π_0) are obtained, inferences can be made based on a trained model and observed returns. These inferences are made using the Viterbi algorithm that divides the observed state into one of 3 states at time t, given the model parameters (A, B, π_0) and observed return upto time t. The algorithm is as follows:

Step 1: Define a matrix *trellis* of dimension (3 x t) to hold probability p of each market state upto time t, given each observation of returns up to time t

Step 2: Define another matrix *pointers* of dimension (3 x t) to hold backpointer to best prior state

Step 3: Determine each market state's probability at time 0

for s in range(3):
  trellis[s, 0] $\leftarrow$ *π[s] * B[s, 0]*

Step 4: Track each market state's most likely prior market state (call it state k)

for r in range(1, length(R)):
  for s in range(3):
    k $\leftarrow$ *argmax(k in trellis[k,r-1]*A[k,s]*B[s,o])*
    trellis[s,r] $\leftarrow$ *trellis[k,r-1]*A[k,s]*B[s,o]*
    pointers[s,r] $\leftarrow$ *k*

Step 5: Collect the sequence of most likely hidden states in *bestPath*

k $\leftarrow$ *argmax(k in trellis[k, length(R)-1])*
for r in range(length(R)-1, -1, -1):
  bestPath.insert(0, S[k])
  k $\leftarrow$ *pointers[k, r]*

Step 6: Return sequence of states found in *bestPath*

Thus, the hidden market states for each timepoint up to time t are obtained using the Hidden Markov Model. However, HMM does not label these states as Bull, Bear, or Static, rather it merely identifies distinct states (0, 1, 2) within the data without attaching any inherent characteristic to each state. In order to identify which labeled state coincides with our definition of Bull, Bear, Static states, we try across all possible 3! permutations of state assignments to the labels given by HMM, and choose the assignment that maximizes the return on the training period based on our trading strategy.

The following state transition diagram depicts the initial state probabilities and the state transition probabilities for all market states.

We can now use this decoding algorithm to identify market states in the test period from 2018 to 2021 by mapping the labels given by HMM into actual Bull, Bear or Neutral state.

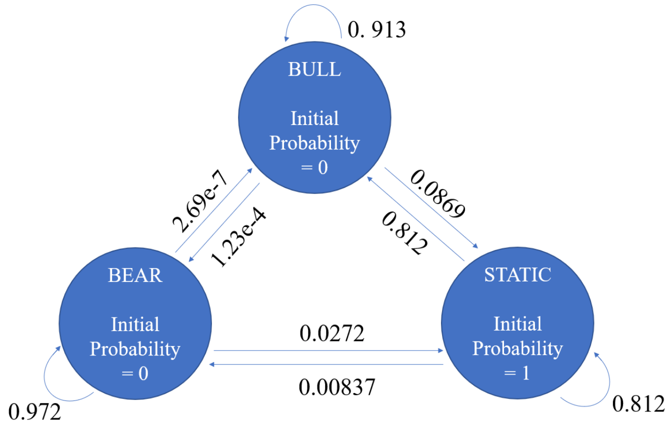

Figure 2: State Transition

4 Macroeconomic & Financial Indicators

4.1 Motivation

HMM predicted states present a good starting point for identifying 'true states', but the model tends to lag in identifying state changes, which would result in missed opportunities of entering into trading positions during dips and spikes. HMM predictions also result in some very short-lived market regimes (of 5 days or less), which is less consistent with the economic significance of market states.

In order to build upon the predictions made by the HMM, we attempt to incorporate leading indicators into our model by analyzing macro-economic and financial indicators that are expected to hold predictive power for identifying market movements.

4.2 Exploring additional features

Following are the new features that were explored to be added to the model, in addition to the Russell 3000 index price.

1. **Realized Volatility:** The biweekly realized volatility is computed using the standard deviation of the Russell 3000 index over a 10 trading day period. Using scatter plots, we see a marked correlation between the biweekly realized volatility and the change of the market states. We define a lower threshold, such that all points with realized volatility below that threshold will be classified as bull and we define an upper threshold, such

that all points with realized volatility above that threshold will be classified as bear, with anything in between these thresholds being static. We now compute the optimal values of these thresholds based on values of realized volatility, such that the sum of the proportion of correctly classified bulls below the lower threshold and the proportion of correctly classified bears above the upper threshold is maximized. The prediction of the change of state of the market generated by the HMM and the realized volatility signal was a better predictor than the signal generated by the implied volatility and the HMM. Its performance (on the validation set) was however worse than the signal generated by using vol of vol signal and HMM, therefore we chose to drop realized volatility of Russell from our selected indicators.

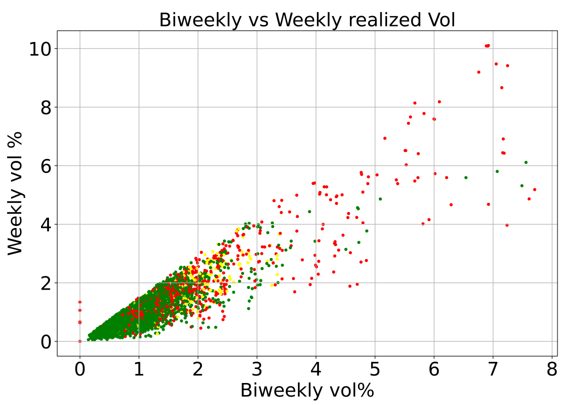

Figure 3: Realized Volatility ● Bull ● Bear ● Static

2. **VIX[4]:** VIX or the Fear Index was used with the motivation of being able to predict the bear regimes better. The VIX index is based on the implied volatility of the near-term S&P 500 index options. This failed to give any promising results for state changes. The predictions of the change of state were lagging and failed to capture the "true state" of the market at the right time leading to a prediction model which was mistimed. The state changes using realized volatility of VIX were better in terms of timing and we eventually used them in our prediction model.

3. **Realized Volatility of VIX Index:** We will refer to this as vol of vol henceforth. The biweekly vol of vol is computed using the standard deviation of the VIX index over a 10 trading day period. Using scatter plots, we see a marked correlation between the biweekly realized vol of vol and the change of the market states. A similar optimization algorithm was used to calculate thresholds for vol of vol as described above. The prediction of the change of state of the market generated by the HMM and the realized vol of vol was a better predictor than the signal generated by the implied vol of vol (VVIX) and the HMM. Therefore, we expect it to contribute further information in identifying state changes and incorporate this signal in our final model.

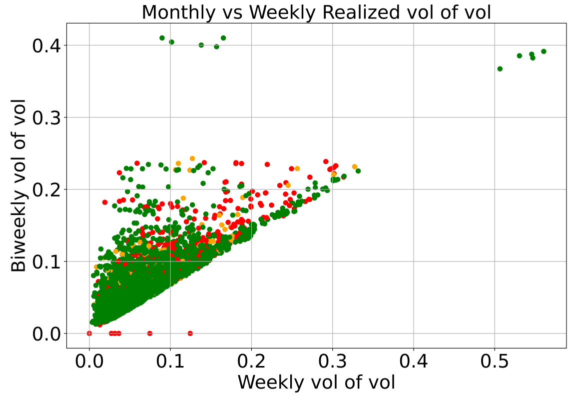

Figure 4: Realized Volatility of VIX ● Bull ● Bear ● Static

4. **VVIX[5]:** Implied volatility of the VIX index (implied volatility of volatility) was also analyzed for the prediction model. As with the VIX feature, the state changes predicted using the VVIX index and the HMM (which used the Russell 3000 index as input) also were lagging, resulting in not efficiently capturing the spikes and the plummets of the Russell 3000 index for efficient money-making trading strategy. The state changes using realized volatility of VIX were better predictors of change of state and hence captured the peaks and troughs better, leading to better cumulative returns on the training data. To capture sudden movements in VIX from persisting trends, we defined a feature as the deviation of the daily levels from the 15-day moving average of the VIX index.

5. **Shiller Housing Prices Index[6]:** Since housing bubbles are known to precede bear markets, we tried using the Shiller housing prices index as a feature of our model. The issue that was a hindrance to the model was the frequency.

[4]Sourced from YahooFinance

[5]Sourced from YahooFinance

[6]Data sourced from here

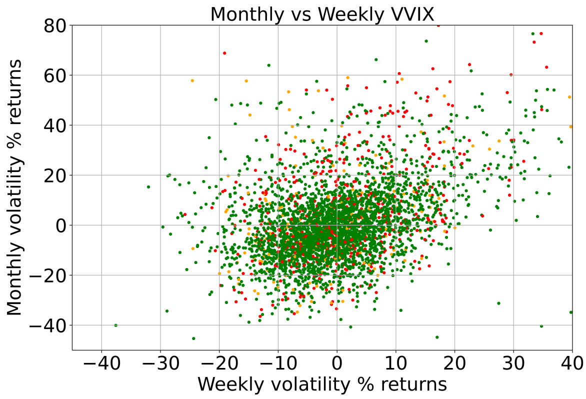

Figure 5: VVIX ● Bull ● Bear ● Static

The Shiller housing index is published monthly whereas the other features we use are of daily frequency. Given the frequency mismatch, the Shiller housing price index was not able to provide any significant predictive power to our baseline HMM. Hence, we eventually ended up dropping this feature.

6. **Treasury yields**[7]**:** Treasury yields are typically a good indicator of investor confidence. Yields typically narrow during periods of low market confidence, indicating investors prefer moving out of risky assets and widen during periods of high confidence. Based on this theory, we expected Treasury yields to be a good leading indicator of the market state. However, on analysis, it was observed that this signal does not hold significant predictive power over and above the baseline HMM. This is possibly due to the fact that there may be a variety of other reasons for the movement of Treasury yields, such as changing interest rate expectations, and yield movements are also regulated by Fed. Hence, we eventually ended up excluding this signal from our model.

7. **High Yield Spreads**[8]**:** The spread between the yield-to-worst of High Yield corporate bonds and the 10-Y treasury rate was also analyzed as a potential indicator for the prediction model. We expect corporate bond spreads to widen preemptively as markets enter a bear phase, and tighten during a strong bull run this is also because of investor confidence. Due to the hybrid characteristics of High Yield bonds (credit products with high risk), this effect is expected to be more prominent for High Yield bonds as compared to Investment Grade bonds. This was verified in the analysis, as High Yield spreads proved to capture the market state better than Investment Grade spreads. We use the "BarCap US Corp HY YTW - 10 Year Spread" Index to proxy for High Yield spreads. We include this indicator to extract an alternative signal within our model.

8. **Daily Price Fluctuation:** Intra-day price fluctuations provide information regarding volatility and sudden market movements on a daily basis, which led us to believe such a signal may help capture the state of the market. We define the formula as follows:

$$\text{T-1 Daily Fluc}^{9} = \frac{high_{t-1} - low_{t-1}}{open_{t-1}}$$

Sudden market crashes typically involve a large amount of fluctuation of the index price within a day, whereas Bull runs often have a more steady intra-day trading pattern. We thus, expected such a signal would be useful in capturing bear states accurately, which was verified through analysis as well. We include this feature to obtain a signal for our model.

5 Final Model

5.1 Identifying market states using Signal Extraction

This section elaborates on the final indicators chosen to incorporate in our model to identify market states, and the methodology to extract signals from these indicators.

1. **Signal 1:** For this specific signal, we use the features T-1 daily returns of Russell 3000 index, T-1 daily fluctuations, T-1 High Yield spreads, T-1 Open - Close prices of Russell 3000, and T-1 VIX deviations from 15-day moving average. These features are trained on a Random Forest Classifier to classify the likely market state based on our "true states". Since machine learning models are unable to incorporate the concept of predicting "contiguous targets", there is expected to be some noise in identified states, i.e the odd daily state identified as bear/static within a broader bull run and vice versa. To gain more confidence from the signal extracted from our machine learning models, our predicted state for each data point is

[7] Sourced from Bloomberg
[8] Sourced from Bloomberg
[9] Open, High and Low data sourced from Bloomberg

the most frequently occurring state in the set of states predicted for the current and trailing four data points. This smoothing ensures we reduce the noise in the predicted state sequence, which we feel is essential based on the spirit of the definition of a 'market state'.

2. **Signal 2:** For the second signal, we use the bi-weekly realized volatility of the VIX index as well as the daily fluctuation in intra-day prices of Russell 3000 as defined above. These features along with the "true states" are trained using a Random Forest Classifier to predict market states. We apply the same methodology to smooth the state sequence as signal 1. We expect this signal to hold information crucial for predicting change in states, and we attempt to validate this in our validation runs.

3. **Signal 3:** After looking at the realized volatility results, we decided to identify the states of the market through realized vol of vol as well. For defining the states using the vol of vol model, we used the thresholds obtained using the optimization approach to classify bear, bull, and static states of the market. The results of this signal have been presented in the tables under the "Results" section. It was observed that the states being identified by the vol of vol model were able to predict the change in the states better than any of the other signals. The timing was the best for this vol of vol signal when compared to the other signals leading to the highest PnL amounts. This model is also expected to result in some noise in the final state and to deal with this issue, we use the same smoothing approach as earlier.

5.2 Combining with HMM Baseline:

The 'pseudo signal state' extracted from each of the above signals is combined with the 'pseudo HMM state' from the HMM to obtain the final market state prediction by switching the governing signal that predicts the state when either of the pseudo-states indicates a change of state. This method ensures that there is no loss of information from the two signal sources, as we expect each signal to contribute separate crucial information for predicting the market state. To confirm the validity of this method, we run model validation as outlined below. While this method may increase the overall number of state changes within the predictions, validation experiments indicate the increase in state changes is not extreme as compared to the improvement in results.

From here on, we refer Model 1 as the the model combining results from Signal 1 and HMM. We use similar notation for Model 2 and Model 3.

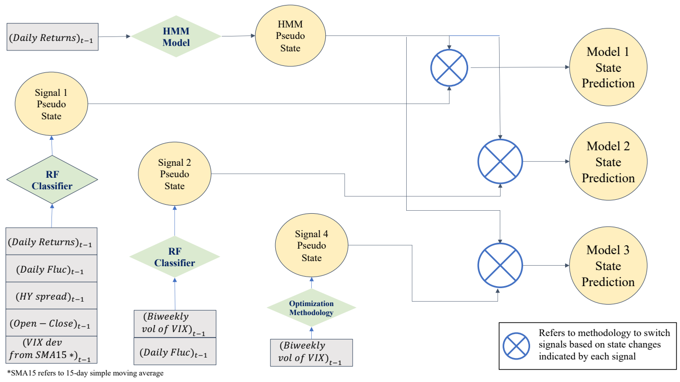

Figure 6: Model Flowchart

5.3 Model Validation and Training

We consider the pseudo-state predicted by HMM as a good starting point to predict the market state, but incorporating signals from the above-identified indicators is expected to improve on the base case HMM prediction by using the above-explained methodology. To validate our final model, we must impose a rolling window validation set, since maintaining the sequence of time series data is crucial. We take a rolling 9-year window for our train/validation sequence. The rationale behind selecting a 9-year window lies in the fact that typical business cycles last for 4.5-5 years on average, so it seems prudent to include a period roughly equal to two business cycles in each train/validation sequence. Within these 9 years, we take a training period of 6 years, and a validation period of 3 years. As we roll this window with a step size of 250 trading days across the entirety of the overall training set (1995 - 2017), the average performance of each model can be evaluated across validation sets using 15 different train and validation sets.

Validation of HMM Pseudo States Here again, we assign states to the labels given by HMM on training data, trying out all 6 combinations and using our trading strategy to select the assignment giving maximum returns. We then apply this assignment to the validation set predictions and pass these assigned market states into our final model methodology for validation.

Validation of Signal Model Pseudo States: Using the same training/validation window, we train each signal model on the true labels from the training data and use this model to obtain validation predictions of the market state. These predictions are passed into our final model methodology for validation.

Validation of Market State Prediction Models: To obtain predictions of market states, we combine 'pseudo states' obtained from the HMM with each of the Signal models as outlined in the previous section. We then compare the returns of our trading strategy based on these predicted market states up against a buy-and-hold strategy to test whether our predictions in each validation set outperforms. Mean returns excess of the buy-and-hold strategy are computed. These results are tabulated below.

Validation Results*	Model 1	Model 2	Model 3
Outperformance in Validation Set	9	8	7
Mean Excess Return	14.06%	15.07%	20.15%

**excess returns signifies returns in excess of buy and hold*

Table 1: Validation Results

Model 1 outperforms in the most number of validation sets, while Model 3 has the highest mean excess returns over the buy and hold strategy. Notably, Model 1 results in far fewer re-balancing instances (state changes) as compared to the other models. Since we aim to maximize our profits in the absence of transaction costs, we believe Model 3 would be the ideal model to choose moving forward. Based on these validation results, we decide Model 2 is not outperforming the other models in either metric, and therefore remove it from final testing. We suggest Model 3 as our final chosen model but include Model 1 in final testing to check performance, as it may prove a good alternative in the presence of transaction costs, as this is an important consideration in practical applications of such a strategy.

6 Results

We report the out-of-sample result of our model below by testing the model's performance on data from 2018 to 2021. We apply a strategy based on the real-time identified market state to buy on a bull signal, short on a bear signal, and invest in 3M t-bills (which is considered as a benchmark risk-free rate) during static phases rather than take an active bet on the market.

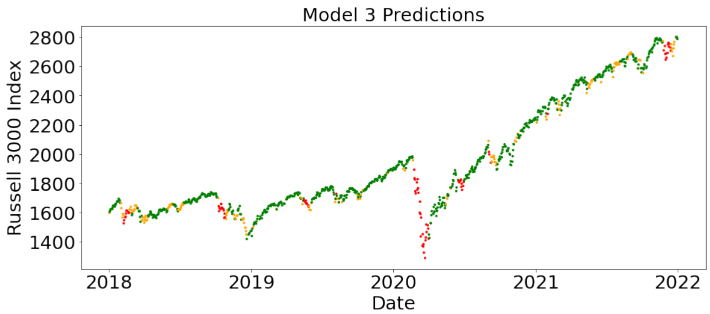

Figure 7: Model 3 ● Bull ● Bear ● Static

The PnL can be compared to the return obtained using a simple buy-and-hold strategy on the Russell 3000 index for the 2018-2021 period. This can be considered a lower bound for returns of a market timing strategy over this period and is around $\sim 74\%$.

If this strategy is implemented on states identified by our true labeling methodology (which requires retrospectively identifying states, therefore not able to be used in real-time), the return obtained is $\sim 415\%$. This can be considered an upper bound to the performance of a regime-based market timing strategy over this period.

The following table shows the comparative analysis of PnL obtained on the Train and Test data for the model.[10].

Cumulative Returns

Data set classification	Time Period	Buy and Hold	Model 3	Trades in Model 3
Train Dataset	2/2/95 - 12/27/17	473.24%	3086.32%	504
Test Dataset	1/3/18 - 12/31/21	73.70%	211.63%	113

Annualised Returns

Data set classification	Time Period	Buy and Hold	Model 3	Annual Trades in Model 3
Train Dataset	2/2/95 - 12/27/17	8.04%	16.56%	23
Test Dataset	1/3/18 - 12/31/21	15.42%	34.35%	28

Table 2: Result Comparison for Model 3

As shown in the table above, implementing our market-timing strategy on states identified by Model 3 returns 212% over 2018-2021, as compared to a buy and hold strategy's return of 74%. Our model dictates

[10]"Trades" refer to no. of times the portfolio is rebalanced

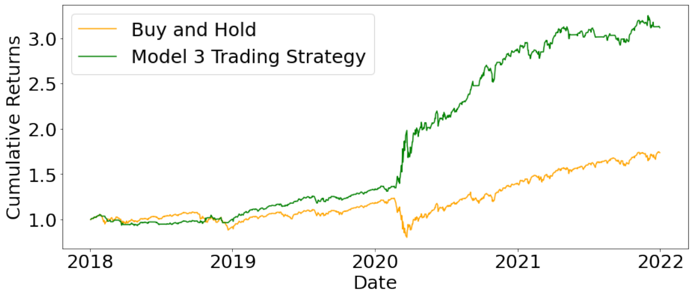

Figure 8: Cumulative Returns: Model 3 vs Buy & Hold

113 re-balances over this period or roughly twice a month.

It is encouraging to see that the PnL of the strategy based on Model 3 is near twice the PnL of the strategy based on the baseline HMM, at the cost of higher re-balancing.

Cumulative Returns

Data set classification	Time Period	Buy and Hold	HMM	Trades in HMM
Train Dataset	2/2/95 - 12/27/17	473.24%	2655.14%	429
Test Dataset	1/3/18 - 12/31/21	73.70%	106.99%	71

Annualised Returns

Data set classification	Time Period	Buy and Hold	HMM	Annual Trades in HMM
Train Dataset	2/2/95 - 12/27/17	8.04%	15.82%	20
Test Dataset	1/3/18 - 12/31/21	15.42%	20.80%	18

Table 3: Result Comparison for HMM

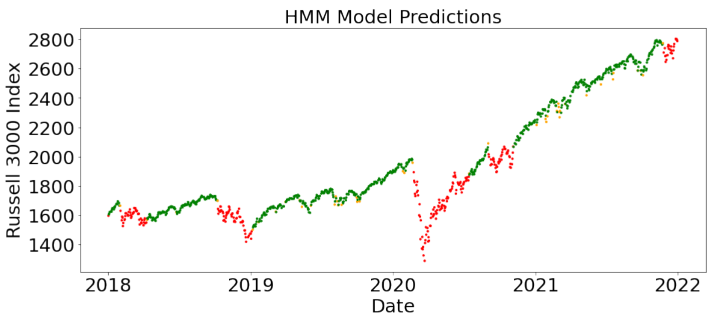

Figure 9: Base HMM • Bull • Bear • Static

Model 1 results are also encouraging, as this may provide a high profit alternative model in the presence of transaction costs.

7 Conclusion

To conclude and bring this back to regimes, we can certainly discern patterns of fairly persistent market

Cumulative Returns

Data set classification	Time Period	Buy and Hold	Model 1	Trades in Model 1
Train Dataset	2/2/95 - 12/27/17	473.24%	2655.26%	429
Test Dataset	1/3/18 - 12/31/21	73.70%	182.73%	89

Annualised Returns

Data set classification	Time Period	Buy and Hold	Model 1	Annual Trades in Model 1
Train Dataset	2/2/95 - 12/27/17	8.04%	15.82%	19
Test Dataset	1/3/18 - 12/31/21	15.42%	31.00%	22

Table 4: Result Comparison for Model 1

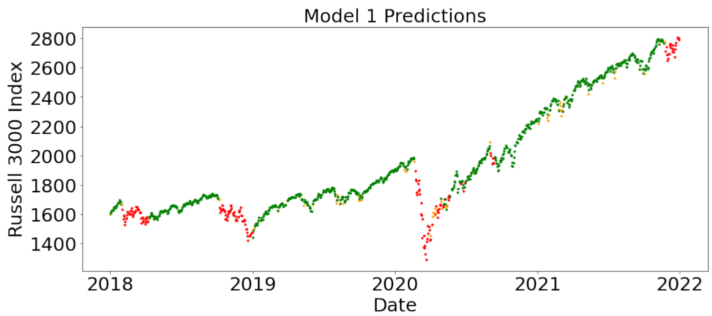

Figure 10: Model 1 • Bull • Bear • Static

conditions using our models. It was, however, rare for one market condition to last prolonged periods without interruption from another. As was evident recently during the Covid-19 crisis, transitory themes may enter on a short-term basis resulting in swift changes in market conditions. Our model is reactive to the Covid-19 changes resulting in abrupt market regime switches. The model was able to capture two high volatility states of the market- the phase just before the bear market and the high volatility neutral states which we labeled for our 'true states' as well.

We have used the economists' consensus to determine our benchmark model and then tried to take a data-driven approach to identify the market regimes. The vol of vol signal is used in addition to the HMM to enhance the performance.

We foresee multiple use cases for the kind of models that we developed. The first and foremost is to take assistance in making appropriate trading decisions according to the state of the market. These models can be used for risk management purposes specifically for scenario and stress testing analysis for portfolios. We can also use the regime-switching model to build various kinds of bull and bear strategies. It can also be used for allocating weights to assets according to the market regime.

As future improvements for these models, we should also consider the impact of transaction costs on these models. The total transaction cost should be optimized in a practical scenario. This constraint may result in choosing another model during the validation stage.

References

[1] Kole, E. and van Dijk, D.J., 2010. How to identify and predict bull and bear markets?. *Econometrics institute, Erasmus University Rotterdam.*

[2] Zegadło, P., 2022. Identifying bull and bear market regimes with a robust rule-based method. *Research in International Business and Finance, 60*, p.101603.

[3] Nyberg, H., 2013. Predicting bear and bull stock markets with dynamic binary time series models. Journal of Banking Finance, 37(9), pp.3351-3363.

[4] Chen, J. and Tsang, E.P., 2020. *Detecting Regime Change in Computational Finance: Data Science, Machine Learning and Algorithmic Trading*. Chapman and Hall/CRC.

[5] Yuan, Y. and Mitra, G., 2016. Market Regime Identification Using Hidden Markov Models. *Available at SSRN 3406068.*

[6] Wang, M., Lin, Y.H. and Mikhelson, I., 2020. Regime-switching factor investing with hidden Markov models. *Journal of Risk and Financial Management, 13*(12), p.311.

[7] Liu, M., Huo, J., Wu, Y. and Wu, J., 2021. Stock Market Trend Analysis Using Hidden Markov Model and Long Short Term Memory. *arXiv preprint arXiv:2104.09700.*

[8] Ryou, H., Bae, H.H., Lee, H.S. and Oh, K.J., 2020. Momentum investment strategy using a hidden Markov model. *Sustainability, 12*(17), p.7031.

[9] Zhu, F. and Nam, C., 2019. Market Regime Classification Using Correlation Networks.