Predicting Economic Recessions Using Technical Indicators

Cheuk Hei, Yip

8 August 2019

Abstract: This paper examines the predictive power of two technical indicators – Relative Strength Index and Directional Movement Index – for economic recessions. Complementing the treasury yield spread, these two technical indicators significantly improve the yield-spread only model for out-of-sample recession forecasts at two-quarter and one-quarter forecast horizons. While daily predictions are too noisy for practical applications, the quarterly and monthly smoothed versions of technical indicator-models can predict 90% - 100% recessions, with less than 20% being "false alarms" at the one-quarter horizon. Given their strong predictive power and easy computation, the large family of technical indicators can benefit every economist and professional forecaster who need reliable predictions ahead of recessions.

---

1. Introduction

Predicting economic recessions is a *practical* issue faced by all firms, households and policy makers. Having reliable signals *ahead* of upcoming economic downturns is crucial yet challenging. Due to the intimacy of macroeconomics and finance, there is a huge literature on the predictability of economic recessions using financial data. Estrella and Hardouvelis (1991) laid down the foundation by demonstrating the strong predictive power of treasury yield spreads. Based on their work, a wide variety of financial variables have then been examined, including stock prices and returns (Estrella and Mishkin, 1998; Stock and Watson, 2003, Nyberg, 2010), and market depth and liquidity (Erdogan, Bennett and Ozyildirim, 2015). Liu and Moench (2016) document the predictive power of a list of common financial variables together with economic leading indicators at different forecast horizons.

While previous works have included most of the financial variables, technical indicators are completely omitted in the field of recession prediction. Indeed, technical analysis has been controversial among academic economists at least since Fama and Blume (1966). However, there are plenty of empirical researches revealing that technical analysis can in fact "beat the market" with statistical significance (Lo, Mamaysky, and Wang, 2000; Hsu and Kuan, 2005). The economic contents of technical indicators give them the potential to perform other tasks, for example, recession prediction in this paper.

Intuitively, the predictive power of technical indicators should not be surprising. Market sediment has long been an informal yet essential leading signal for business cycle fluctuations. When the real economy is doing well, investors are more willing to invest in the stock market. Stock prices exhibit strong uptrends, which often persist even though the stocks have been overbought. Nevertheless, on the eve of recessions, investors start to anticipate for upcoming crises, and thus begin to reduce investments in the stock market. Uptrends are thus weakened, and overbought condition becomes less common. Consequently, stock prices usually peak and start to decline *before* recessions come.

Technical indicators which track the above described market behavior can thus give accurate predictions of upcoming recessions. Within the large family of technical indicators, Relative Strength Index (RSI) and Directional Movement Index (DMI) exactly identify the characteristics of this sediment – RSI reflects overbought and oversold conditions and DMI reflects trend. Moreover, since all technical indicators are available daily, the richness of data unveils much more information than months and quarterly data used in the literature. Therefore, in terms of both the economic contents and data availability, it is natural to use technical indicators as predictors of economic recessions.

With this in mind, this paper uses RSI and DMI of stock indexes (S&P 500, NYSE and Dow Jones Industrial Average) to give *out-of-sample* forecasts of the binary recession indicator defined by National Bureau of Economic Research (NBER) in the period 2000 – 2018. Daily predictions are first reported for overall evaluation. Quarterly and monthly averaging are then performed to give smoother forecasts with practical meanings. It is shown that these technical indicators can complement the baseline yield-spread-only model at the two-quarter-ahead horizon, and give very accurate predictions at the one-quarter-ahead horizon.

The paper is organized as follows. Section 2 details data, methodologies and model evaluation metrics for recession prediction. Section 3 first examines the raw daily predictions, and then turn to practically meaningful quarterly and monthly forecasts. Section 4 summarizes and concludes.

2. Data and Methodology

2.1. Data

The predicted variable is the NBER defined binary variable indicating recession. It takes the value of one during recessions and zero otherwise. For predictors, the only predictor involving treasury yields is yield spread, defined as the difference between 10-year Treasury bonds and 3-month Treasury bills. This specific spread is adopted since it is found to provide the best results among a variety of spreads (Moneta, 2005). Other predictors are all technical indicators of indexes of S&P 500, NYSE and Dow Jones Industrial Average. Technical indicators can be divided into two categories: RSI and DMI.

RSI can identify overbought and oversold conditions of stock markets. According to Murphy's (1999) formula, the x-day RSI are be calculated as follows:

$$RSI_x = 100 - \frac{100}{1 + RS},$$

where the Relative Strength (RS) is

$$RS_x = \frac{Average\ of\ x\ days'up\ closes}{Average\ of\ x\ days'down\ closes}.$$

For trading purpose, usually 9-day or 14-day RSI are used. Yet, for recession prediction purpose, it is necessary to take a longer time frame to smooth out excessive noise. In the paper, I adopt 100-day, 200-day and 300-day RSI of stock indexes as predictors.

In addition, DMI, which reflects directional movements of stock markets, is another category of technical indicators used as recession predictors. Since intraday movements are involved in the calculation of DMI, for convenience I obtain 100-day, 200-day and 300-day +DMI and -DMI from Bloomberg directly. The formulas are omitted here but the computation follows Bloomberg (2019). For each +DMI and -DMI, their differences are also included as predictors to indicate trends.

The dataset covers the period 1 January 1970 – 31 December 2018. The first few days of data are used to compute RSI and thus some data points are not included in the analysis. All data are recorded daily and standardized. Holidays and days with any missing values are omitted from the analysis.

2.2. Methodology

To give forecasts ahead of recessions, *lag* terms of predictors are used. In the literature, where monthly or quarterly data are used, this can simply be done by lagging variables by certain numbers of months or quarters. Since daily data are used in this paper, the terms "month" and "quarter" have to be carefully defined. Taking average across 1 January 1970 to 31 December 2018, I define 21 days – the average number of days in a month across the sample period after removing days with missing data– as an effective month, and similarly 63 days as an effective quarter. Months and quarters are referred to these effective definitions throughout this paper.

Focusing on the practical application of recession prediction, this paper gives out-of-sample forecasts using data available when one is making the prediction. The first prediction is the first data point in year 2000. Standard support vector machine is trained to give forecasts. Validation is performed to choose the hyperparameter C, the penalty parameter, from the set $\{0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1, 10\}$. The hyperparameter with the best performance for the validation set will be chosen as the hyperparameter.[1] Note that training set and validation set of data are separated. Depending on the forecast horizon, the training set, validation set and out-of-sample set differ, which will be explained in details one by one:

<u>A) Quarterly-updating daily model</u>

Define $D_t \in \{0,1\}$ as the *daily* binary indicator of recessions. The objective is to forecast D_{t+h} using data available at the end of day t, where h is the forecast horizon. The forecast horizon is determined by when the last predictors are realized. For instance, if 2-quarter lagged RSI, ie. RSI (-2Q) and RSI (-1Q) are used as predictors, the forecast horizon will be 1 quarter = 63 days.

The model is updated quarterly, whose implication can be best illustrated using the above example. At the beginning, the in-sample data lasts from the very start of the dataset to 1999

[1] Model evaluation metrics will be discussed in the next subsection.

Q3 [2]. The last 48 quarters of the in-sample data are used as validation set. Such a relatively large proportion of data are used for validation because recession is a rare event and the validation set has to be large enough to sample recession days. After the best hyperparameter is selected, the model is trained and give predictions to the first 63 days starting from the 2 January 2000. The model will then be updated by using the in-sample data ranging from the start of the dataset to 1999 Q4 (added 63 data points), within which the last 48 quarters of the in-sample data are used as validation set to select hyperparameter, and then give predictions for the 64^{th} to 126^{th} data points (the subsequent 63 data points). The process continues until no data are left.

B) Quarterly-updating quarterly model

While the above model gives daily forecasts, length of recessions should be measured in months or quarters instead of days. Moreover, given the nature of technical indicators, daily recession forecasts are susceptible to sudden and drastic changes in stock indexes which are not related to the real economy. It is necessary to have quarterly or monthly forecasts so that meaningful economic interpretations can be made.

The quarterly-updating quarterly model is simply a smoothed version of the quarterly-updating daily model. Define $Q_t \in \{0,1\}$ as the *quarterly* binary indicator of recessions. The objective is to forecast Q_{t+h} using data available at the end of effective quarter t, where h is the forecast horizon. The training set, validation set and prediction set are exactly identical as above. However, after each batch of predictions are made, an average is taken over the predictions. Using the conventional 0.5 threshold, if the average of a batch of predictions is greater or equal to 0.5, then recession warning will be given for the whole effect quarter. Similarly, realized recessions are also taken average each effective quarter for evaluation purpose.

C) Monthly-updating monthly model

While quarterly forecasts may be too coarse for business cycle, monthly signals are able to give users recession warnings in a timely fashion. To compensate the reduction of smoothing effect, the model is updated monthly to include more data in the training set. Define $M_t \in \{0,1\}$ as the

[2] The procedure is done in a daily basis using the effective quarter definition. Therefore, it may be possible that some October data are included while it may also be possible that some September data are excluded.

monthly binary indicator of recessions. The objective is to forecast M_{t+h} using data available at the end of effective month t, where h is the forecast horizon.

Using the same example, at the beginning the in-sample data lasts from the start of the dataset to 1999 Q3. The validation set is still the last 48 quarters of the in-sample data. After the best hyperparameter is selected, the model is trained and give predictions to the first *21* days starting from the 2 January 2000. After that, 21 data points are added to the in-sample. The model then goes through same validation procedure, and gives another 21 predictions. This process continues until no data are left.

2.3. Evaluation Metrics

Various evaluation metrics have been used in the literature. For example, Estrella and Mishkin (1996, 1998) and Erdogan, Bennett and Ozyildirim (2015) use pseudo-R^2 to evaluate their models. However, pseudo-R^2 is more about model fitness which may not be appropriate for the applied propose. Meanwhile, Liu and Moench (2016) use the area under the receiver operating characteristic curve (AUROC) to evaluate their models. However, AUROC has two major disadvantages:

(1) It may not be a good measure for imbalanced classes such as recessions and expansions. Table 1 clearly illustrates the imbalance of classes for the whole dataset and the out-of-sample period.

Table 1: Class Imbalance

Period	Number of recession days	Number of expansion days
Whole dataset	1778 (14%)	10877 (86%)
Out-of-sample	566 (11%)	4388 (89%)

From the perspective of both economics and forecasting, missing a recession day (month or quarter) is much more costly than missing an expansion day (month or quarter). Yet, AUROC cannot address this concern properly (Davis, Goadrich, 2006). Consider the following confusion matrix:

	Predicted Recession		
Actual Recession		Yes	No
	Yes	True Positive (TP)	False Negative (FN)
	No	False Positive (FP)	Ture Negative (TN)

AUROC evaluates how well a model balances $Recall$ and $Specificity$, where

$$Recall = \frac{TP}{TP + FN}, \qquad Specificity = \frac{TN}{TN + FP}.$$

For recession prediction, where recessions are kind of rare events, a model always giving recession warning can still receive a high $Specificity$ and thus a decent AUROC. In other words, AUROC is not a suitable metric since it values $Recall$ and $Specificity$ equally.

(2) For classification tasks, one needs a threshold to determine to which class a prediction belongs. While AUROC evaluates performance across a spectrum of different thresholds, this may not be practically meaningful – a good AUROC can be obtained if the model performs somewhat good across thresholds. However, for applied purposes, one actually needs a model to have excellent performance at a single threshold rather than being just average over different thresholds.

Emphasizing the practical nature of recession prediction, this paper uses $F1$ score as the evaluation metric. $F1$ can be calculated as follow:

$$F1 = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision},$$

where

$$Recall = \frac{TP}{TP + FN}, \qquad Precision = \frac{TP}{TP + FP}.$$

It is obvious that $F1$ score is a point-wise measure of both $Recall$ and $Precision$. In simple words, it penalizes the over-pessimism and the over-optimism, and thus is suitable for classifying imbalanced classes.

3. Results

3.1. Quarterly-updating Daily Models

First, I estimate the yield-spread model as the baseline. Following the literature, the fourth quarter lag of yield spread is used as the predictor. Table 2 gives the categorical descriptive statistics of it. As documented in the literature, the yield curve is usually upwards sloping, but prior to recessions, it is flattened. This suggests that the fourth quarter lag of yield spread is a valid predictor.

Table 2.2 gives the prediction results of this baseline model. Recall and Precision are around 0.4 to 0.5, resulting a mediocre F1 score of 0.4783. This shows that, although yield spread is arguably the most well-established recession predictor, it is far from satisfactory when used to give out-of-sample-forecasts which are essential for practical purposes.

Table 2.1: Categorical Mean of Yield Spread (-4Q)

Variable	Expansion Days	Recession Days
Yield Spread (-4Q)	1.8765	0.0950

Table 2.2: Performance – Yield Spread (-4Q)

Predictor(s)	Yield Spread (-4Q)
Forecast Horizon	4Q
Model	Quarterly-updating Daily Model
Recall	0.4488
Precision	0.5121
F1	0.4784

Next, I examine the predictive power of RSI and DMI with two-quarter lag. Table 3.1 justifies the intuition of using RSI and DMI as predictors. Prior to recessions, RSI are lower; +DMI drops while -DMI increases, and the trend is weakened. This matches the observation that stock markets start to decline before recessions.

Table 3.2 reports the mixed results on their predictive power at the two-quarter horizon. On the one hand, RSI (-2Q) complements and improves the baseline model considerably, with both Recall and Precision over 0.6; on the other hand, DMI (-2Q) only improves the baseline modestly in terms of F1 score (0.5261). In fact, the DMI (-2Q) model scores less in terms of Recall (0.4982), meaning that it misses even more recessions than the baseline. The third panel of Table 3.2 reports the performance of both RSI (-2Q) and DMI (-2Q). Similar to the second panel, DMI (-2Q) does not work well and it actually deteriorates the performance of RSI (-2Q), with a $F1$ score of 0.5166.

Table 3.1: Categorical Mean of RSI (-2Q), DMI (-2Q)		
Variable	Expansion Days	Recession Days
NYSE_ RSI_{100} (-2Q)	53.7428	47.4509
NYSE_ RSI_{200} (-2Q)	53.4876	48.7720
NYSE_ RSI_{300} (-2Q)	53.1616	49.9148
SPX_ RSI_{100} (-2Q)	53.9819	47.1491
SPX_ RSI_{200} (-2Q)	53.5729	48.2869
SPX_ RSI_{300} (-2Q)	53.2172	49.5735
INDU_ RSI_{100} (-2Q)	53.7606	47.9236
INDU_ RSI_{200} (-2Q)	53.3873	47.5052
INDU_ RSI_{300} (-2Q)	53.0434	49.7648
NYSE_ $+DMI_{100}$ (-2Q)	53.3790	48.1034
NYSE_ $-DMI_{100}$ (-2Q)	46.6210	51.8967
NYSE_ $Trend_{100}$ (-2Q)	6.7581	-3.7933
NYSE_ $+DMI_{200}$ (-2Q)	52.8526	49.3304
NYSE_ $-DMI_{200}$ (-2Q)	47.1474	50.6696
NYSE_ $Trend_{200}$ (-2Q)	5.7053	-1.3392
NYSE_ $+DMI_{300}$ (-2Q)	52.5523	49.9526
NYSE_ $-DMI_{300}$ (-2Q)	47.4477	50.0473
NYSE_ $Trend_{300}$ (-2Q)	5.1046	-0.0947
SPX_ $+DMI_{100}$ (-2Q)	53.4155	47.8760
SPX _ $-DMI_{100}$ (-2Q)	46.5845	52.1240
SPX _ $Trend_{100}$ (-2Q)	6.8310	-4.2479
SPX _ $+DMI_{200}$ (-2Q)	52.8864	49.1274
SPX_ $-DMI_{200}$ (-2Q)	47.1136	50.8726
SPX _ $Trend_{200}$ (-2Q)	5.7729	-1.7451
SPX _ $+DMI_{300}$ (-2Q)	52.5803	49.7428
SPX_ $-DMI_{300}$ (-2Q)	47.4197	50.2572
SPX _ $Trend_{300}$ (-2Q)	5.1605	-0.5145
INDU_ $+DMI_{100}$ (-2Q)	53.2500	48.4154
INDU_ $-DMI_{100}$ (-2Q)	46.7500	51.5846
INDU_ $Trend_{100}$ (-2Q)	6.4999	-3.1691
INDU_ $+DMI_{200}$ (-2Q)	52.7582	49.3153
INDU_ $-DMI_{200}$ (-2Q)	47.2418	50.6847
INDU_ $Trend_{200}$ (-2Q)	5.5165	-1.3693
INDU_ $+DMI_{300}$ (-2Q)	52.4740	49.7269
INDU_ $-DMI_{300}$ (-2Q)	47.5260	50.2731
INDU_ $Trend_{300}$ (-2Q)	4.9479	-0.5462

Table 3.2: Performance – RSI (-2Q), DMI (-2Q)	
Predictor(s)	Yield Spread (-4Q), RSI (-2Q)
Forecast Horizon	2Q
Model	Quarterly-updating Daily Model
Recall	0.6360
Precision	0.6060
F1	0.6207
Predictor(s)	Yield Spread (-4Q), DMI (-2Q)
Forecast Horizon	2Q
Model	Quarterly-updating Daily Model
Recall	0.4982
Precision	0.5573
F1	0. 5261
Predictor(s)	Yield Spread (-4Q), RSI (-2Q), DMI (-2Q)
Forecast Horizon	2Q
Model	Quarterly-updating Daily Model
Recall	0.5901
Precision	0.4594
F1	0.5166

While technical indicators have somewhat mixed results at the two-quarter horizon, they significantly improve the baseline model at the one-quarter horizon. Table 4.1 gives the categorical mean of RSI (-1Q) and DMI (-1Q). The interpretation of Table 4.1 is identical to that of its counterpart Table 3.1 at the two-quarter horizon, which again justifies the use of RSI (-1Q) and DMI (-1Q) as discriminating factors of recessions.

Table 4.2 shows that RSI (-1Q) and DMI (-1Q) have greatly improved baseline yield-spread-only model. RSI (-1Q) complements with Yield Spread (-4Q) and gives a F1 score of 0.7678. The predictive power of DMI (-1) is weaker than RSI (-1), but still it hugely enhances the baseline model with a F1 score of 0.6801. Combining RSI (-1Q) and DMI (-1Q) gives the best prediction than just using either one, where the F1 score is 0.7889.

Table 4.1: Categorical Mean of RSI (-1Q), DMI (-1Q)		
Variable	Expansion Days	Recession Days
NYSE_ RSI_{100} (-1Q)	54.1680	46.1549
NYSE_ RSI_{200} (-1Q)	53.6967	47.2473
NYSE_ RSI_{300} (-1Q)	53.3641	48.5277
SPX_ RSI_{100} (-1Q)	54.1964	46.1451
SPX_ RSI_{200} (-1Q)	53.7730	46.9613
SPX_ RSI_{300} (-1Q)	53.4430	48.1546
INDU_ RSI_{100} (-1Q)	53.9835	46..7603
INDU_ RSI_{200} (-1Q)	53.5693	47.4823
INDU_ RSI_{300} (-1Q)	53.2527	48.4835
NYSE_$+DMI_{100}$ (-1Q)	53.6062	46.8401
NYSE_$-DMI_{100}$ (-1Q)	46.3938	53.1599
NYSE_$Trend_{100}$ (-1Q)	7.2124	-6.3198
NYSE_$+DMI_{200}$ (-1Q)	53.0691	48.0964
NYSE_$-DMI_{200}$ (-1Q)	46.9308	51.9036
NYSE_$Trend_{200}$ (-1Q)	6.1383	-3.9072
NYSE_$+DMI_{300}$ (-1Q)	52.7462	48.8366
NYSE_$-DMI_{300}$ (-1Q)	47.2538	51.1634
NYSE_$Trend_{300}$ (-1Q)	5.4925	-2.3267
SPX_$+DMI_{100}$ (-1Q)	53.6401	46.7355
SPX _$-DMI_{100}$ (-1Q)	46.3599	53.2645
SPX _$Trend_{100}$ (-1Q)	7.2802	-6.5290
SPX _$+DMI_{200}$ (-1Q)	53.1053	47.9669
SPX_$-DMI_{200}$ (-1Q)	46.8946	52.0331
SPX _$Trend_{200}$ (-1Q)	6.2107	-4.0662
SPX _$+DMI_{300}$ (-1Q)	52.7774	48.6891
SPX_$-DMI_{300}$ (-1Q)	47.2226	51.2110
SPX _$Trend_{300}$ (-1Q)	5.5548	-2.6219
INDU_$+DMI_{100}$ (-1Q)	53.4626	47.2976
INDU_$-DMI_{100}$ (-1Q)	46.5374	52.7024
INDU_$Trend_{100}$ (-1Q)	6.9252	-5.4048
INDU_$+DMI_{200}$ (-1Q)	52.9588	48.2523
INDU_$-DMI_{200}$ (-1Q)	47.0412	51.7477
INDU_$Trend_{200}$ (-1Q)	5.9177	-3.4954
INDU_$+DMI_{300}$ (-1Q)	52.6547	48.7765
INDU_$-DMI_{300}$ (-1Q)	47.3453	51.2235
INDU_$Trend_{300}$ (-1Q)	5.3093	-2.4470

Table 4.2: Performance – RSI (-1Q), DMI (-1Q)	
Predictor(s)	Yield Spread (-4Q), RSI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.7155
Precision	0.8282
F1	0.7678
Predictor(s)	Yield Spread (-4Q), DMI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.7456
Precision	0.6252
F1	0. 6801
Predictor(s)	Yield Spread (-4Q), RSI (-1Q), DMI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.8516
Precision	0.7348
F1	0.7889

At the moment, all predictions are made using only the information of a single quarter. It is natural to use two-quarter lag predictors to supplement those with one-quarter lag to see if it is possible to obtain better forecasts. The results of such extension are listed in Table 5. Although the first panel reveals that using RSI (-2Q) slightly undermines the predictive power of RSI (-1Q), the second panel shows that DMI (-2Q), surprisingly, considerably enhances the DMI (-1Q) model. Both models have a F1 score of around 0.75.

The best performance of daily prediction can be obtained when all predictors are used together, which is shown in Table 5. The F1 score is 0.8177, which is the highest among all previous models. Specifically, the Recall score is the almost 0.9, meaning that the model only misses 1 out of 10 recession days. It is worth noticing that the high Recall score is achieved while Precision score is not completely forgone – only around one fourth of the recession predictions are "false alarms". For easier reference, this model is named the "comprehensive model" through the rest of the paper.

Figure 1 plots predictions of the comprehensive model with shaded areas being actual recession periods. It is clear that daily predictions are too noisy for meaningful economic interpretations. It is, therefore, necessary to smooth out predictions by quarter or month.

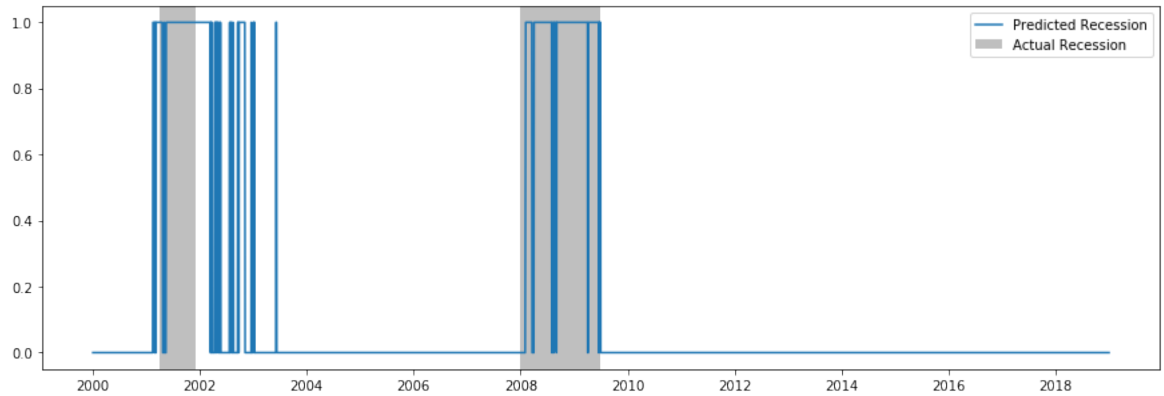

Figure 1: Performance of Type A Comprehensive Model

Table 5: Performance – RSI (-2Q), DMI (-2Q), RSI (-1Q), DMI (-1Q)	
Predictor(s)	Yield Spread (-4Q), RSI (-2Q), RSI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.7614
Precision	0.7655
F1	0.7635
Predictor(s)	Yield Spread (-4Q), DMI (-2Q), DMI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.8693
Precision	0.6482
F1	0. 7426
Predictor(s)	Yield Spread (-4Q), RSI (-2Q), DMI (-2Q), RSI (-1Q), DMI (-1Q)
Forecast Horizon	1Q
Model	Quarterly-updating Daily Model
Recall	0.8993
Precision	0.7496
F1	0.8177

3.2. Quarterly-updating Quarter Models and Monthly-updating Monthly Models

Table 6 gives the smoothed predictions using the procedures described in Section 2.2 for each model reported previously. From the table, it is clear that smoothing does not only contribute to quarterly or monthly signals which are better for applied purposes, it also improves the prediction performance significantly. With only two exceptions, smoothed predictions almost uniformly outperform daily predictions. The best model remains to be the comprehensive model, and indeed its predictions are very accurate – one can *always* foresee an upcoming recession a quarter ahead, while making false warnings only with less than 20% probability. Its monthly counterpart, though less accurate, still records a high F1 score of 0.8814. The comparison between column (b) and (c) is less sharp, because while monthly smoothing is nosier, the model is updated per month to incorporate more data available at the forecast horizon. Indeed, a daily-updated model with monthly or quarterly smoothing should be the best, but due to high computational costs this paper does not include daily-updated models.

Table 6: Model Comparison

	(a)	(b)			(c)		
Predictors	F1	Recall	Precision	F1	Recall	Precision	F1
Yield Spread (-4Q)	0.4783	0.5556	0.6250	0.5882	0.4643	0.5652	0.5098
Yield Spread (-4Q), RSI (-2Q)	0.6207	0.7778	0.6363	0.7000	0.6429	0.6429	0.6429
Yield Spread (-4Q), DMI (-2Q)	0.5261	0.4444	0.5714	0.5000	0.5357	0.6250	0.5769
Yield Spread (-4Q), RSI (-2Q), DMI (-2Q)	0.5166	0.4444	0.4000	0.4211	0.6429	0.5455	0.5902
Yield Spread (-4Q), RSI (-1Q)	0.7678	0.7778	0.8750	0.8235	0.7500	0.8750	0.8077
Yield Spread (-4Q), DMI (-1Q)	0.6801	0.8889	0.7273	0.8000	0.8214	0.6970	0.7541
Yield Spread (-4Q), RSI (-1Q), DMI (-1Q)	0.7889	0.8889	0.8000	0.8421	0.8929	0.7813	0.8333
Yield Spread (-4Q), RSI (-2Q), RSI (-1Q)	0.7635	0.7778	0.7778	0.7778	0.7857	0.7857	0.7857
Yield Spread (-4Q), DMI (-2Q), DMI (-1Q)	0.7426	0.8889	0.6667	0.7619	0.8571	0.6857	0.7619
Comprehensive Model	0.8177	1	0.8182	0.9000	0.9286	0.8387	0.8814

Column (a), (b) and (c) refer to quarter-updating daily models, quarter-updating quarter models, and monthly-updating monthly models respectively. The shaded number is the best F1 score among the three models.

To highlight the predictive power of technical indicators at the one-quarter horizon, I compare the comprehensive model with Anxious Index reported in the Survey of Professional Forecasters, which is to predict upcoming declines in real GDP, but at the same time has a good track record for predicting NBER defined recessions (The Federal Reserve Bank of Philadelphia, 2019). In this comparison, if the one-quarter ahead Anxious Index is higher than a threshold x, it is considered to be a recession warning. Table 7 provides the results of the comparison. This is patent that the comprehensive model performs better than Anxious Index no matter where the threshold is. Since the anxious index takes into account the predictions from a group of economists and professional forecasters, the result here indicates that the majority of professionals have overlooked the predictive power of technical indicators when they form their own recession predictions. Figure 2 plots the recession predictions of the comprehensive model and the Anxious Index model ($x = 40$).

Table 7: Comparison between the Comprehensive Model and Anxious Index (2000– 2018)

	Comprehensive Model	Anxious Index								
x	/	10	20	30	40	50	60	70	80	90
F1	0.900	0.2951	0.6923	0.7500	0.800	0.3636	0.3636	0.3636	N/A[3]	N/A

[3] F1 is undefined when the model does not predict any recession.

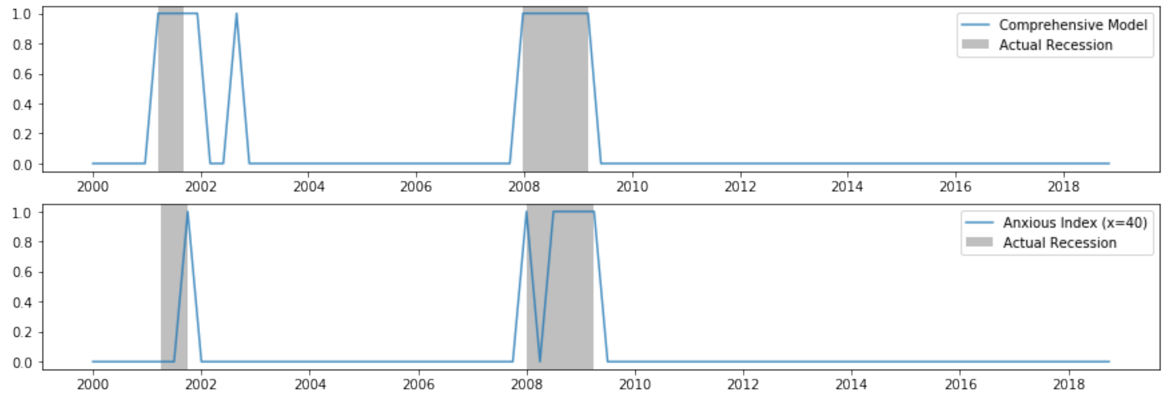

Figure 2: Performance of Type C Comprehensive Model (Top); and Performance of Anxious Index Model ($x = 40$) (Bottom)

4. Conclusion

This paper complements the yield-spread recession prediction model using RSI and DMI with one- and two-quarter lag. While the predictive power of yield spread is well established in the literature, this paper shows that it cannot give reliable out-of-sample predictions, which are essential for countless practical applications. On the other hand, RSI and DMI, though excluded in previous recession prediction literature, are found to have huge predictive power. Not only does it complement well with the yield spread, it also outperforms Anxious Index which pulls together economists' and professional forecasters' viewpoints. Indeed, technical indicators exactly matches the traditional intuitions that stock markets are leading indicators of the real economy and thus should not be completely ignored by researchers and academics. Further research can be dedicated towards the predictive power of other technical indicators and how technical indicators complement with predictors other than the yield spread.

5. References

Bloomberg L.P. (2019) Study Property Manager – Directional Movement Index. Retrieved Aug. 2, 2019 from Bloomberg database.

Brownlee, J. (2016). *Machine learning mastery with Python: Understand your data, create accurate models and work projects end-to-end*. Melbourne, Australia: Jason Brownlee.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning - ICML 06*.

Lo, A., Mamaysky, H., & Wang, J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. *Journal of Finance,55*(4), 1705-1765.

Erdogan, O., Bennett, P., & Ozyildirim, C. (2015). Recession Prediction Using Yield Curve and Stock Market Liquidity Deviation Measures. *Review of Finance*,19(1), 407-422.

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

Estrella, A., & Mishkin, F. S. (1996). The Yield Curve as a Predictor of U.S. Recessions. *Current Issues in Economics and Finance,2*(7), 1-6.

Estrella, A., & Mishkin, F. S. (1998). Predicting U.S. Recessions: Financial Variables as Leading Indicators. *Review of Economics and Statistics,80*(1), 45-61.

Fama, E. F., & Blume, M. E. (1966). Filter Rules and Stock-Market Trading. *The Journal of Business,39*(S1), 226.

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

Hsu, P., & Kuan, C. (2005). Reexamining the Profitability of Technical Analysis with Data Snooping Checks. *Journal of Financial Econometrics,3*(4), 606-628.

Liu, W., & Moench, E. (2016). What predicts US recessions? *International Journal of Forecasting,32*(4), 1138-1150.

Moneta, F. (2005). Does the Yield Spread Predict Recessions in the Euro Area? *International Finance,8*(2), 263-301.

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York: New York Institute of Finance.

Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting,29*(1-2), 215-230.

Stock, J., & Watson, M. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature,41*, 788-829.

The Federal Reserve Bank of Philadelphia. (2019). The Anxious Index. Retrieved from https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/anxious-index/.

Wilder, J. W. (1978). *New concepts in technical trading systems*. Greensboro (N.C.): Trend research.