Machine Learning for Recession Prediction and Dynamic Asset Allocation

ALEXANDER JAMES, YASER S. ABU-MOSTAFA, AND XIAO QIAO

ALEXANDER JAMES
is a co-head of research at Paraconic Technologies US, Inc. in New York, NY.
alex@paraconic.com

YASER S. ABU-MOSTAFA
is a professor of electrical engineering and computer science at the California Institute of Technology in Pasadena, CA.
yaser@caltech.edu

XIAO QIAO
is a co-head of research at Paraconic Technologies US, Inc. in New York, NY.
xiao@paraconic.com

ABSTRACT: *The authors introduce a novel application of support vector machines (SVM), an important machine learning algorithm, to determine the beginning and end of recessions in real time. Nowcasting, forecasting a condition in the present time because the full information will not be available until later, is key for recessions, which are only determined months after the fact. The authors show that SVM has excellent predictive performance for this task, capturing all six recessions from 1973 to 2018 and providing the signal with minimal delay. The authors take advantage of the timeliness of SVM signals to test dynamic asset allocation between stocks and bonds. A dynamic risk budgeting approach using SVM outputs appears superior to an equal-risk contribution portfolio, improving the average returns by 85 bps per annum without increased tail risk.*

TOPICS: ***Big data/machine learning, financial crises and financial market history, portfolio construction, tail risks****

*All articles are now categorized by topics and subtopics. **View at PM-Research.com.**

Real-time business cycle dating is of central importance in modern macroeconomics. Recessions reflect great dislocation in the economy and are often the source of societal anxiety. Accurately identifying turning points from expansions to recessions has broad use for policymakers, business executives, academics, and individuals. Investors with enough resources to use this information in their investment process may change their portfolios as the economy turns from growth to contraction.

The National Bureau of Economic Research (NBER) provides the official dating of expansions and recessions. The NBER's Business Cycle Dating Committee periodically assesses the prevailing conditions in the macroeconomy and determines if the economy is in an expansion or a recession.[1,2] The committee releases announcements about the dates of the turning points. Because of the reliance on macroeconomic data, which may be revised or released with a lag, and the committee's preference for accuracy over timeliness, the NBER has historically announced turning points with a delay of 4 to 21 months (Giusto and Piger 2017).

Is it possible to identify business cycle turning points in a more timely manner? This is the focus of a large body of literature going back to Burns and Mitchell (1946) and greatly expanded in a series of papers

[1] NBER uses *expansion* and *contraction* to describe business cycles. In this article, we use *recession* and *contraction* interchangeably.

[2] NBER defines a recession as a "significant decline in economic activity spread across the economy, lasting more than a few months." The committee does not follow a fixed rule of labeling a recession as at least two consecutive quarters of negative GDP growth (NBER 2008).

by Stock and Watson (1989, 2002). This question is also the first question we try to address in our article. We use support vector machines (SVM), a powerful machine learning algorithm, to identify turning points in the macroeconomy.

Machine learning algorithms have been shown to be useful tools in many settings outside of the social sciences and only more recently have been adopted more extensively in economics. It is straightforward to set up SVM for classification, and computation is relatively cheap compared to prevailing models in the macroeconomic literature. For example, the dynamic factor Markov-switching (DFMS) (Diebold and Rudebusch 1996; Chauvet 1998; Chauvet and Piger 2008) and dynamic probit (Fossati 2016) models, which have been used extensively for recession prediction, are significantly more computationally intensive to estimate.

Could recession prediction help investors navigate the changing macroeconomic environment? This is the second question of our article. An investor with knowledge of how stocks and bonds perform in recessions and expansions may want to dynamically adjust his or her portfolio to reflect the prevailing macroeconomic conditions. Typically, stocks are thought to be risky investments that perform well in expansions and perform poorly in recessions, whereas bonds are thought to be more defensive instruments that exhibit some stability across regimes. We explore the possibility of using SVM recession prediction as the basis of a dynamic portfolio of stocks and bonds.

We start the article with a brief description of the SVM algorithm and a discussion of its advantages and disadvantages. We then provide details of our implementation of this algorithm. There are two main empirical choices: the kernel, which models the nonlinearity implemented by the algorithm, and the soft margin cost, which controls the trade-off between model stability and penalty for misclassified data. We use a radial basis kernel (Scholkopf et al. 1997) and select the soft margin cost parameter using 10-fold cross validation.

Inputs to the SVM model includes four variables: the monthly log difference in nonfarm payrolls, the log difference in the average monthly S&P 500 price level, the level of the production index from the Manufacturing ISM Report on Business, and the 10-year US Treasury yield minus the federal funds rate. These four variables represent information from four important and distinct markets: the labor market, the stock market, the goods market, and the bond market.

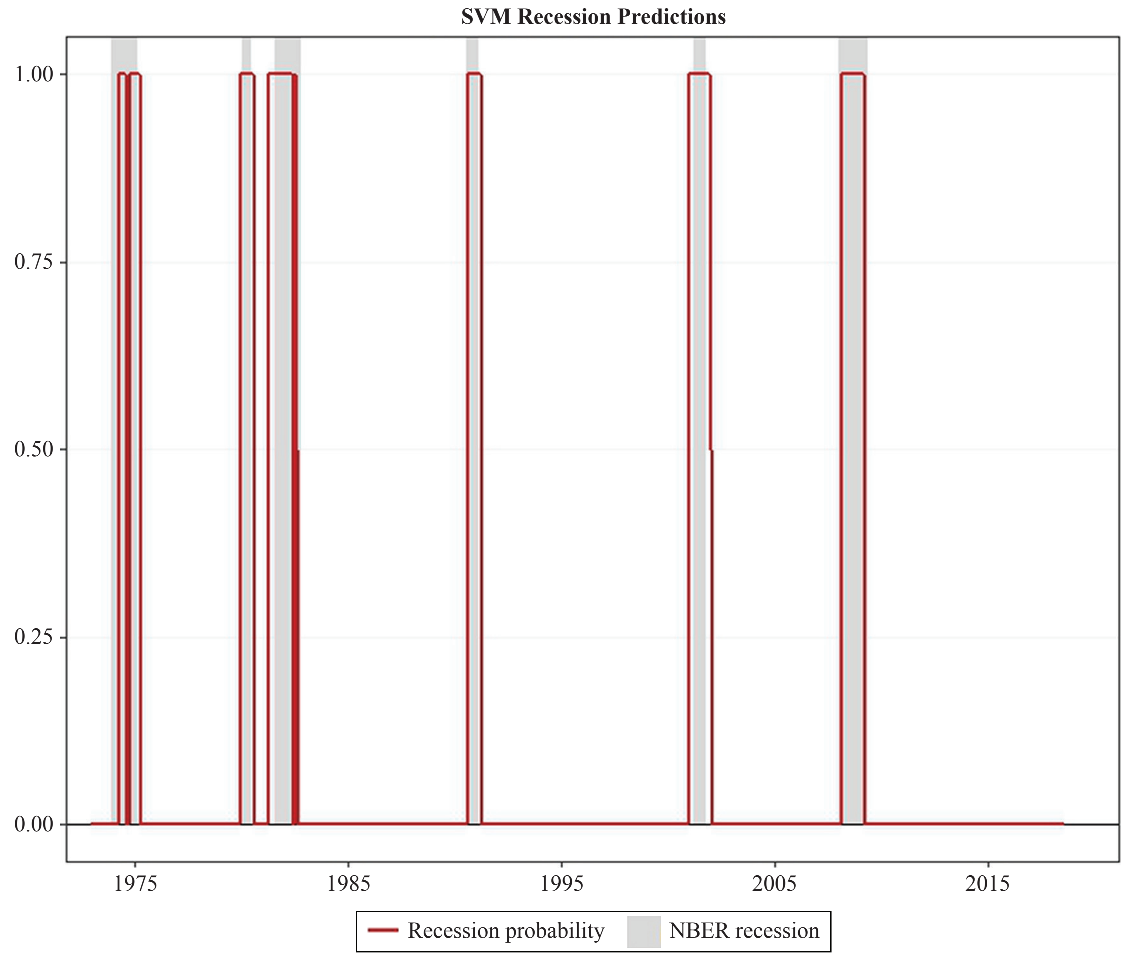
We fit the model once a month to form nowcasts and forecasts. Model estimation occurs once all data referring to the previous calendar month become available, usually within the first 10 days of the month. We are careful not to use variables that are not available in real time; only information available at the time of each nowcast or forecast is included in the model. Our sample is from 1959 to 2018. We use the first 14 years to train the model and make the first nowcast in 1973. From 1959 to 1973, two recessions are identified by the NBER: 1960 to 1961 and 1969 to 1970. These episodes provide targets to train the model.

There are six NBER-defined recessions from 1973 to 2018, shown in Exhibit 1. The SVM model successfully captures all six events. The model identifies the transition from expansion to recession (and back) typically within one to three months of the NBER definition. We use classification error to formally evaluate the model performance. The SVM model misclassifies 5.3% of the months from 1973 to 2018.

The SVM model is competitive with other recession-prediction models proposed in the literature. The DFMS model proposed by Diebold and Rudebusch (1996) and used by Chauvet (1998) has a classification error of 4.7% or 9.3% depending on implementation. The gross domestic product (GDP)-based model proposed by Chauvet and Hamilton (2006) has a classification error of 7.3%. Although the SVM model slightly underperforms one implementation of the DFMS model, its results are available with minimal delay, and it is computationally simpler than the DFMS model. The GDP-based model only uses one variable and makes quarterly predictions. Both elements may contribute to its higher classification error compared to SVM.

In our implementation of the SVM model, we attempt to provide the most timely nowcast. If we can delay our signal to use revised data, the model accuracy may be improved. Giusto and Piger (2017) used learning vector quantization (LVQ) to identify business cycle turning points, allowing for delays in the revised data. As a result, their model has a very low classification error (1.8%) from 1976 to 2013. In the same period, SVM has a classification error of 4.8%. Although the LVQ model identifies recessions more accurately than SVM, it requires significant delays to do so: an average lag of 184 days. In contrast, the SVM model has a delay

Exhibit 1
SVM Recession Predictions



Notes: The vertical axis shows the recession probability. The line represents the recession predictions from the SVM model. The official NBER recessions are shaded. The sample is from 1973 to 2018.

of less than 10 days. If the statistician's objective is to minimize classification error, the LVQ model may be preferred over SVM. If real-time recession identification is the goal, SVM would be preferred.

When the most recent data are simply not available, SVM could still be used to forecast recessions. The one-month-ahead forecasts have a classification error of 6.7%, whereas the two- and three-month-ahead forecasts have classification errors of 6.8% and 8.0%, respectively. The forecast errors grow monotonically with forecasting horizon, but the increases are modest. This is perhaps helped by the persistent nature of recessions and expansions.

Stock and bond markets behave differently in expansions and contractions. Intuitively, stocks are riskier than bonds and react more sensitively to changes in the underlying macroeconomy. When stock markets suffer large downturns, investors often turn to the bond markets in a flight to quality. Therefore, stocks tend to have better performance and Sharpe ratios in expansions compared to recessions, whereas bonds have smaller differences across regimes. We can exploit the differential behavior of stocks and bonds through a dynamic allocation approach. We start from an equal-risk contribution (ERC) allocation in stock and bonds, with 50% risk contribution coming from each. If in a recession, set the bond risk budget to 75% and stock risk budget to 25%. If in an expansion, set the bond risk budget to 25% and the stock risk budget to 75%. Whether we are in a recession or an expansion is determined from the SVM model.

On each date we estimate the SVM model, we solve an optimization problem including the recession-dependent risk budgets and the covariance matrix using the past 126 days to obtain weights on stocks and bonds. We hold this portfolio until the next estimation of the SVM model. We compare this dynamic risk budgeting approach with ERC.

From 1995 to 2019, the annual return is 7.3% for the dynamic risk budgeting strategy and 6.4% for ERC, a difference of 85 bps. Volatility is also higher for the dynamic strategy at 5.1%, whereas ERC has a lower volatility at 4.4%. The Sharpe ratio for the dynamic strategy is a bit higher than that for ERC (0.94 and 0.91, respectively). Dynamic risk budgeting does not lead to higher tail risk; its max drawdown is similar to that for ERC. Higher returns and lower drawdown lead to a higher Calmar ratio for dynamic risk budgeting (0.94) than for ERC (0.81). Our simple dynamic allocation strategy illustrates how the investor may use the SVM model to improve his or her portfolio through business cycle fluctuations.

Our article makes two contributions. First, we demonstrate the merit of a new methodology applied to a large existing macroeconomics literature on classifying recessions. SVM is a novel technique for time-series analysis in macroeconomics. We show it works well in classifying NBER-defined recessions. In an attempt to establish a link between a machine learning methodology and the macroeconomic forecasting literature, we hope to expand the macroeconomist's toolkit to form superior nowcasts and forecasts.

The second contribution of the article is to incorporate recession prediction in portfolio management, thereby establishing a connection between the macroeconomics literature on recession prediction and the asset pricing literature on asset allocation. Historically, recession prediction and asset allocation developed independently, and it was not clear whether the two areas shared useful insight. We demonstrate that recession prediction can be useful to dynamically allocate between stocks and bonds because the behavior of these assets is sensitive to the prevailing macroeconomic conditions.

Our article fits into the large literature on forecasting and nowcasting recessions. Existing approaches include the probit model, dynamic probit, and other linear probability models. Stock and Watson (1989) proposed a *recession index*, a time series of the probability that the economy will be in a recession in six months. Estrella and Mishkin (1998) used a static probit model to show that financial variables contain information about future recessions. Wright (2006) also used a probit model in investigating the usefulness of the Treasury yield curve in recession prediction. Kauppi and Saikkonen (2008) argued dynamic probit models are superior to static probit models in forecasting recessions. Compared to these studies, our work allows for more complex nonlinearities, which expands the richness in dynamics that may be captured by our model.

There is also an emerging literature that uses statistical and machine learning methods in macroeconomic time-series analysis. Qi (2001) applied neural networks to forecast recessions and identified several useful variables. Berge (2015) examined use of boosting in recession prediction and found improvement to a Bayesian model averaging benchmark. Giusto and Piger (2017) used LVQ to identify recessions, which resulted in a shorter time lag compared to the NBER announcements. Our article complements these papers in demonstrating the efficacy of machine learning algorithms for macroeconomic applications.

Our article is also related to the literature on dynamic asset allocation. Campbell and Viceira (1999) studied dynamic portfolio allocation between a riskless asset and a risky asset with time-varying expected returns. Campbell and Viceira (2001) included a section on dynamic allocation between stocks and bonds with changing real interest rates. Brennan, Schwartz, and Lagnado (1997) numerically analyzed the portfolio problem of an investor who can invest in bonds, stocks, or cash under time-varying interest rates and equity returns. Compared to these papers, our article considers changing investment opportunity sets as a function of the macroeconomic environment. We illustrate how information about recessions could be incorporated in the investor's portfolio choice between stocks and bonds.

SUPPORT VECTOR MACHINE

SVM is a popular and flexible method for classification. It is widely used in many areas, including text and hypertext categorization, hand-written character recognition, image classification, and numerous others.

Its theoretical development is mainly due to Vapnik and can be found in Vapnik (2013). We provide an overview of SVM in the following sections. For a more in-depth introduction, please see Friedman, Hastie, and Tibshirani (2009).

Linear SVM

We start with a discussion of the linear SVM, the building block of more sophisticated SVM algorithms. Consider a classification problem with two classes, labeled as $y_i \in \{-1, 1\}$. We have predictor variables x_i paired with y_i. Suppose the positive and negative classes are linearly separable; then we can construct a hyperplane to completely separate the two classes:

$$\{x : x^T\beta + \beta_0 = 0\}$$

The normal vector β and constant β_0 could be rescaled by a constant and have no effect on the hyperplane (2β, $2\beta_0$ give the same hyperplane). To uniquely identify the parameters, let us impose the unit norm on β, $\|\beta\|=1$, where $\|\cdot\|$ is the Euclidean norm. We may classify the data points according to the following rule:

$$y_i = 1 \; if \; x_i^T\beta + \beta_0 > 0$$

$$y_i = -1 \; if \; x_i^T\beta + \beta_0 < 0$$

This setup gives us a hyperplane that separates the two classes such that $y_i(x_i^T\beta + \beta_0) > 0$ for all i. To induce the largest separation between the two classes, we want to maximize the margin $y_i(x_i^T\beta + \beta_0)$. This leads to the formal optimization problem:

$$\underset{\beta,\beta_0}{\text{Maximize}}\, M \quad s.t. \quad y_i(x_i^T\beta + \beta_0) \geq M, \forall i \tag{1}$$

$$\|\beta\| = 1$$

where M is the margin, or the closest distance between each class and the separating hyperplane. The two classes are separated by a distance of $2M$. This setup is known as a *maximum margin classifier* because it picks the hyperplane with the largest separation between the positive and negative classes. The preceding problem cannot be solved using standard optimization methods because it is not convex. Friedman, Hastie, and Tibshirani (2009) demonstrated that the preceding optimization problem can be transformed into the following convex problem:

$$\underset{\beta,\beta_0}{\text{Minimize}}\, \|\beta\| \quad s.t. \quad y_i(x_i^T\beta + \beta_0) \geq 1, \forall i \tag{2}$$

This optimization problem is convex and can be solved using standard optimization packages.

Now suppose the two classes are not perfectly separable by a hyperplane. We still want to provide the largest separation between the two classes, but now possibly allowing for some data points to be misclassified. We can introduce slack variables ξ_i to allow for misclassification:

$$\underset{\beta,\beta_0}{\text{Minimize}}\, \|\beta\| \quad s.t. \quad y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i, \forall i \tag{3}$$

$$\xi_i \geq 0$$

$$\sum_i \xi_i \leq C$$

Note that Equation 3 is just Equation 2 with slack variables to allow for potential misclassification. If $\xi_i < 1$, i is still classified correctly, although it violates the margin. Misclassification occurs if $\xi_i > 1$. The last constraint limits the total number of misclassifications by restricting the sum of all ξ_i to be smaller than some constant C, called the soft margin cost.

Nonlinear SVM

Nonlinear SVM (referred to simply as SVM) extends the idea of the linear version. If the original data can be easily separated by a hyperplane, linear SVM performs well. However, in many applications, the original data may not be linearly separable even if we allow for misclassification. SVM solves this problem by mapping the original feature space into a higher-dimensional space, then looking for a separating hyperplane in the higher-dimensional space. If a separating hyperplane is found, it can be mapped back into the original feature space. SVM performs this mapping using a kernel function. The kernel function is used to compute inner products in the transformed, higher-dimensional space and can be viewed as a measure of similarity.

SVM has several advantages compared to other classification methods. Different choices of kernel functions make this method versatile and effective in high-dimensional space. It is even effective when the number of dimensions is greater than the number of samples, owing to its self-regularizing nature. SVM can also be used for unsupervised learning (Ben-Hur et al. 2001) and regressions (Drucker et al. 1997).

There are also some disadvantages in using SVM. The choice of kernel function can be a double-edged sword, and a poor choice could lead to undesirable results. SVM does not directly produce probability estimates but simply the final classification outcome. Inference for SVM can be a challenge because the parameters may be difficult to interpret. For our purposes, prediction is the ultimate goal, and we relegate inference to a lower priority.

SVM Implementation

We implement a standard SVM as outlined, using a radial basis kernel. The radial basis kernel is a common choice for SVM implementation. It can be represented as follows:

$$K(u,v) = \exp(-\gamma \| u - v \|^2)$$

where $\|u - v\|^2$ is the squared Euclidean distance between the two vectors u and v, and γ controls the size of the kernel. Small γ is associated with large variance for the Gaussian kernels, so if v is a support vector, it will influence the classification of u even if the distance between them is large. Large γ is associated with small variance for the Gaussian kernels, so the support vectors have little influence on deciding the class. The size of γ also relates to how much the Gaussian kernels will overlap with one another. As such, the choice of γ is related to the dimensionality of the input space, which depends on the number of input variables. The ideal choice of γ scales each of the Gaussian distributions to have partial, but not excessive, overlaps with others.

If in each coordinate the variable is Gaussian with the same variance, the total variance of the resulting high-dimensional Gaussian would be proportional to the number of dimensions. Because γ is inversely proportional to the variance, it would be inversely proportional to the number of dimensions of the input space. This consideration leads to a heuristic choice of γ that equals the inverse of the number of predictors.

The other important parameter choice is the soft margin cost function C, the cost of misclassification. This parameter controls the trade-off between model stability and penalties for misclassified data points. We select C with the following procedure. We start with a set of candidate values for C. On each of the first 100 backtest dates, we run each candidate model (with a different C) 100 times using 10-fold cross validation. We compute a regret measure for each date by taking the difference between each cross-validation error and the lowest cross-validation error of that date. This measure allows us to compare models across different samples. We then select the value of C that minimizes the total regret over the 100 backtest dates.

CLASSIFYING RECESSIONS

Data

The NBER announces a set of dates for the US business cycle expansions and contractions.[3] The NBER's Business Cycle Dating Committee uses data from the National Income and Product Accounts (NIPA) to determine the dates for economic expansions and recessions. Expansions are defined as the previous trough to the current peak in economic activity; recessions are defined as previous peak to the current trough. Additionally, economic cycles defined as trough to trough and peak to peak are available.

Our model is monthly, the same frequency as the NBER recession indicator variable. NBER-defined recessions are marked as the positive class, and expansions are marked as the negative class. Our goal is to perform classification on the positive and negative classes using SVM.

We use four monthly series that contain information about the prevailing state of the macroeconomy. These four variables broadly capture information from four distinct markets:

1. Monthly log difference in nonfarm payrolls from the Employment Situation Report published by the Bureau of Labor Statistics (BLS): We use the same transformation as the datasets compiled by

[3] See: https://www.nber.org/cycles.html.

the Federal Reserve Bank of St. Louis. This variable has been used in multiple past studies (see Camacho, Perez-Quiros, and Poncela 2012) and captures the conditions in the labor market.

2. Log difference in average monthly price of the S&P 500 (Estrella and Mishkin 1998; Qi 2001). This series captures information from the stock market.
3. Production index from the Manufacturing ISM Report on Business published by the Institute for Supply Management: Lahiri and Monokroussos (2013) showed that this variable contains useful information for forecasting GDP. Bok et al. (2018) used this variable for macroeconomic nowcasting. We extend this idea to nowcasting recessions. This variable provides information about the goods market.
4. The 10-year Treasury yield minus the federal funds rate: Numerous academic papers have shown the usefulness of this variable in forecasting recessions (Stock and Watson 1989; Wright 2006). Also known as the term spread, the slope of the Treasury yield curve reflects conditions in the bond market.

Numerous papers have investigated predictive variables for recessions. We use four variables that have been proposed in the literature. There may be a potential selection bias among these variables stemming from a multiple comparisons problem, as failed predictors are rarely published. To the extent there is a selection bias among these variables, our model will inherit that bias.

We restrict the number of variables to keep the model parsimonious. The small number of variables is comparable to that in other studies using machine learning techniques to produce reduced-form forecasts in macroeconomics. For example, Davig and Hall (2019) used the same four variables in a naïve Bayes classifier to forecast recessions. Giusto and Piger (2017) also used four variables, albeit somewhat different from our choices, as inputs for real-time identification of recessions using an algorithm called LVQ.

In modern macroeconomics, the economy is commonly thought to be driven by a small set of unobserved state variables. Observed macroeconomic variables are combinations of these latent state variables plus some measurement error. The latent state variables are slow moving and contain some lead–lag structure. Given this consideration, we not only include the current value of the aforementioned four variables but also 11 lags to allow for any persistent effects in the underlying macroeconomy. In effect, we are using one year of information of each of the four variables as input to our SVM model, for a total of 48 input variables.

Macroeconomic data are often revised. At a given point in time, the data available in real time up to that point are known as the vintage series. If there are revisions to the historical data in the next period, those vintage data are overwritten with the revised series. Using the revised series may build in a look-ahead bias in forecasting because we are using information we would not have had in real time. To alleviate this problem, we are careful in considering vintage variables in our forecasts. At each point in time, we only use real-time information that the econometrician could have accessed. We do not use revised data until the time they are known.

Modeling Choices

We estimate our model monthly when all the data become available for that month. S&P 500 and Treasury yields are available daily. The Manufacturing ISM Report on Business typically comes out on the first business day of each month at 10 a.m. Eastern time.[4] The BLS releases the Employment Situation Report typically on the first Friday of each month.[5] This report contains nonfarm payroll data and other macroeconomic data, including the unemployment rate, average hourly earnings, and so on. We run our model after nonfarm payroll is released by the BLS. The SVM nowcasts produced at the beginning of month t are evaluated against the NBER definition in month $t - 1$. Therefore, the SVM model produces real-time forecasts with a delay equal to the time between the end of one month and the first Friday of the following month, typically fewer than 10 days. The SVM forecasts are compared to the NBER definition in month t, $t + 1$, or $t + 2$, depending on the forecast horizon.

The NBER's Business Cycle Dating Committee uses NIPA data to determine whether the economy is in an expansion or a recession. NIPA data are often released with a lag and are sometimes revised. As a result, the NBER business cycle dates have historically

[4] See: https://www.instituteforsupplymanagement.org/ISMReport/content.cfm?ItemNumber=10745&SSO=1.

[5] See: https://www.bls.gov/news.release/empsit.nr0.htm.

been reported with a delay of between 4 and 21 months (Giusto and Piger 2017). We are mindful of this lag in reporting, and we are careful in our modeling choices to account for it. When we fit the model, we stop the training period 12 months before the date of estimation.

We use an expanding window when estimating our model. The start date of our estimation is January 1959, comparable to other studies (e.g., Davig and Hall 2019 also started in January 1959). We evaluate the model nowcasts and forecasts starting in 1973, to allow for a 14-year training period from 1959 to 1973. In this period, there were two recessions, one from April 1960 to February 1961 and another from December 1969 to November 1970. These events provide the positive examples on which our model is initially trained.

Our model produces nowcasts and forecasts of the macroeconomy and classifies the prevailing conditions as either expansion or recession. We compare our model predictions with the NBER business-cycle chronology to evaluate model performance.

SVM Nowcasting

How does the SVM model compare to other recession-prediction models? We compare the SVM model to two models published by the Federal Reserve Bank of St. Louis: the DFMS model and the GDP-based recession prediction model.

The original work on the DFMS model is due to Diebold and Rudebusch (1996), and Chauvet (1998) used four monthly variables to identify business cycle turning points. Chauvet and Piger (2008) expanded the work of Chauvet (1998) to analyze the performance of the DFMS model. The St. Louis Fed releases the smoothed US recession probabilities from the DFMS model. We take these probabilities and impose a threshold of 50% for classification. If the probability of recession is greater than or equal to 50%, we classify that month as being in a recession. If the probability is lower than 50%, it is classified as an expansionary month.

Chauvet and Hamilton (2006) proposed a GDP-based recession prediction model. The model identifies recessions at a quarterly frequency for the quarter just preceding the most recently available GDP numbers. The St. Louis Fed publishes both the probability of recession and the time series of a GDP-based recession indicator. We compare the SVM model to the latter, transforming the quarterly series by assuming the three months within the quarter are all in the same class, either recession or expansion. Admittedly, comparing a quarterly series of recession predictions to monthly models may put the quarterly model at a disadvantage because of its coarser granularity.

We compare SVM, DFMS, and GDP-based models on the metric of classification error. Classification error is a simple measure of how classification models perform when compared to the actual outcomes. It is defined as follows:

$$Error = \frac{\#\ Incorrect\ observations}{\#\ Total\ observations}$$

where the classification *Error* for a model is the ratio of the number of incorrectly classified data points to the total number of observations. Incorrectly classified data are a combination of Type I and Type II errors. In our setting, an incorrect observation corresponds to the model classifying a month as a recession but NBER classifying it as an expansion, or the model classifying a month as an expansion but the NBER classifying it as a recession. We present the classification errors for the three models in Exhibit 2.

The Federal Reserve Bank of St. Louis publishes the smoothed recession probabilities from the DFMS model. We transform the probabilities into recession predictions in two ways. First, we use a simple cutoff rule: If the probability exceeds 50% in a month, that month is classified as a recession. Second, we use the

Exhibit 2

Classification Errors for SVM, DFMS, and GDP-Based Models

Classification Error 1973–2018			
SVM	**DFMS (50% cutoff)**	**DFMS (CP)**	**GDP Based**
5.3%	4.7%	9.3%	7.3%

Notes: Classification error is calculated as $Error = \frac{\#\ Incorrect\ observations}{\#\ Total\ observations}$, *where* # Incorrect observations *is the total number of observations that are not classified correctly, and* # Total observations *is the total number of data points for which the models are evaluated. We transform the DFMS smoothed recession probabilities from the St. Louis Fed into recession predictions in two ways. First, we use a 50% cutoff; any month with probability greater than 50% is classified as a recession. Second, we use the methodology from Chauvet and Piger (2008): Three consecutive months of probabilities greater than 80% indicate the start of a new recession; three consecutive months of probabilities below 20% indicate the start of a new expansion.*

methodology outlined by Chauvet and Piger (2008): Three consecutive months of probabilities greater than 80% indicate the beginning of a new recession; three consecutive months of probabilities lower than 20% indicate the beginning of a new expansion. These are marked as DFMS (50% cutoff) and DFMS (CP) in Exhibit 2.

From 1973 to 2018, recessions occurred 13% of the time; thus, if a model always predicts expansion, it would have a classification error of 13%. All models in Exhibit 2 have errors much lower than this naïve model. The SVM model has a classification error of 5.3%. The DFMS model with a 50% cutoff rule has a slightly lower rate of 4.7%, indicating the more sophisticated structure of DFMS may capture some nuances beyond the SVM model. Using the three-month rule put forward by Chauvet and Piger (2008) results in a larger classification error of 9.3% for the DFMS. Compared to SVM, DFMS is much more computationally intensive to estimate. SVM provides a simpler alternative that achieves a similar classification error.

Another attractive aspect of SVM is its built-in, in-sample metric that can be used to characterize out-of-sample performance. This generalization result can be stated as

$$E(e_{out}) \leq \frac{E[\#\ of\ support\ vectors]}{N}$$

where e_{out} is the out-of-sample classification error estimated by cross validation, and the expected value is taken with respect to different datasets. Intuitively, support vectors correspond to the effective parameters used in an SVM model, so this generalization result relates the expected out-of-sample error to the ratio of number of parameters to the number of observations. This result allows us to place at least a somewhat loose upper bound on the out-of-sample performance of the model. In our SVM model, the ratio of support vectors to observations declines steadily over the sample period—starting from a high of 43% and decreasing to a low of 26% on the most recent backtest date. Although the bound does not pin down e_{out} exactly, its value is consistent with the fact that the out-of-sample performance has improved over time as the expanding training window includes more recession examples from which to learn.

The DFMS model presents a smooth recession probability, which uses the most recent data available and is potentially influenced by data that were not available the first time a recession probability for a given month was calculated. In contrast, the SVM model produces nowcasts in real time, using the most recent data and never revising the past predictions. The DFMS model also has a two-month lag in reporting the recession probabilities because one of the inputs is real manufacturing and trade sales produced by the US Census Bureau, which is only available after a two-month lag. The SVM model remains competitive with the DFMS model even without any revisions or a two-month lag.

The GDP-based model has a higher classification error of 7.3% in this period. We conjecture that this result is due to the combination of using only GDP as the input and the coarser output frequency. The SVM model uses a broader dataset and produces a monthly series, allowing greater flexibility compared to the GDP-based model.

Giusto and Piger (2017) proposed using the LVQ model to identify business cycle turning points. Their sample was from 1976 through 2013. We compare our results to theirs in this sample period. The classification errors of these two models are reported in Exhibit 3.

EXHIBIT 3

Classification Errors for SVM and LVQ Models, 1976–2013

Classification Error 1976–2013	
SVM	**LVQ**
4.8%	1.8%

Notes: Classification error is calculated as $Error = \frac{\#\ Incorrect\ observations}{\#\ Total\ observations}$*, where # Incorrect observations is the total number of observations that are not classified correctly, and # Total observations is the total number of data points for which the models are evaluated.*

The LVQ model has a lower classification error compared to SVM. As indicated in their paper, Giusto and Piger (2017) demonstrated that the LVQ model typically identifies business cycle peaks and troughs within one month of the NBER definitions. Although the LVQ identifies recessions more effectively than SVM, it requires a significant delay to do so. Giusto and Piger (2017) were on average 134 days late in identifying NBER peaks (compared to the 224-day delay from the NBER) and 234 days late in identifying NBER troughs (a 446-day delay from the NBER). In contrast, SVM

EXHIBIT 4
Classification Errors for SVM Forecasts

SVM Forecast Classification Error		
1 Month	**2 Month**	**3 Month**
6.7%	6.8%	8.0%

Notes: We compute classification error for forecasts made using the SVM model. Forecast horizons include one, two, and three months. The sample is 1973 to 2018.

can identify recessions with a much shorter delay of under 10 days. If the econometrician's objective is to minimize classification error, then LVQ may be preferred over SVM. If the econometrician values real-time identification, SVM is the preferred model.

SVM Forecasting

We investigate the efficacy of SVM as a forecasting model, using the information set available in real time to predict whether the economy will be in a recession in the future. We use the same set of four variables and their lags in our model and maintain our model parameter choices from the previous section. The classification error from SVM forecasts is shown in Exhibit 4.

From 1973 to 2018, the SVM model produced nowcasts with a classification error of 5.3%. At the one-month forecasting horizon, this figure increases to 6.7%. Perhaps not surprisingly, as the forecast horizon increases, the classification error also increases. At the two-month horizon, the classification error is 6.8%. At the three-month horizon, it is 8.0%. Forecasts made at longer horizons do not use the most recent information, which results in noisier identification of the macroeconomic conditions.

TACTICAL ASSET ALLOCATION

Behavior of Stocks and Bonds

SVM provides a novel approach to classify the macroeconomy into expansions and recessions. At the beginning of month t, the SVM model produces a nowcast of the macroeconomic condition of $t-1$, the previous month. If we use the SVM model to forecast recessions, setting the forecast horizon to one month gives the prevailing economic conditions for t, the current month. Compared to alternative methods, an important advantage of SVM is its timeliness. We use this advantage in an investment strategy that tactically allocates to stock and bond markets. The S&P 500 is used to represent the aggregate stock market. Bloomberg Barclays US Treasury Index[6] is used to represent the aggregate bond market.

Stocks and bonds behave differently across expansions and contractions. Exhibit 5 presents summary statistics for the two asset classes. Average returns are slightly higher in recessions than in expansions, which reflects bonds being safe assets that investors flock to when the economy is in a downturn. Average returns are large and positive for stocks in expansions and large and negative in recessions, corresponding to the risk-on, risk-off behavior of the stock market. For both stocks and bonds, volatility is higher in recessions compared to in expansions. Whether in expansions or recessions, stock market volatility is three to four times higher than bond market volatility. From 1995 to 2019, bonds appear to have approximately equal Sharpe ratios in expansions and recessions. In contrast, stocks have a positive and high Sharpe ratio in expansions but a negative Sharpe ratio in recessions.

Dynamic Risk Budgeting

We can exploit the differential behavior of stocks and bonds across recessions and expansions, given a signal about the prevailing economic conditions. SVM offers such a signal. With a forecast horizon of one month, the SVM model provides the recession prediction for the upcoming month. We use this prediction to allocate dynamically to stocks and bonds.

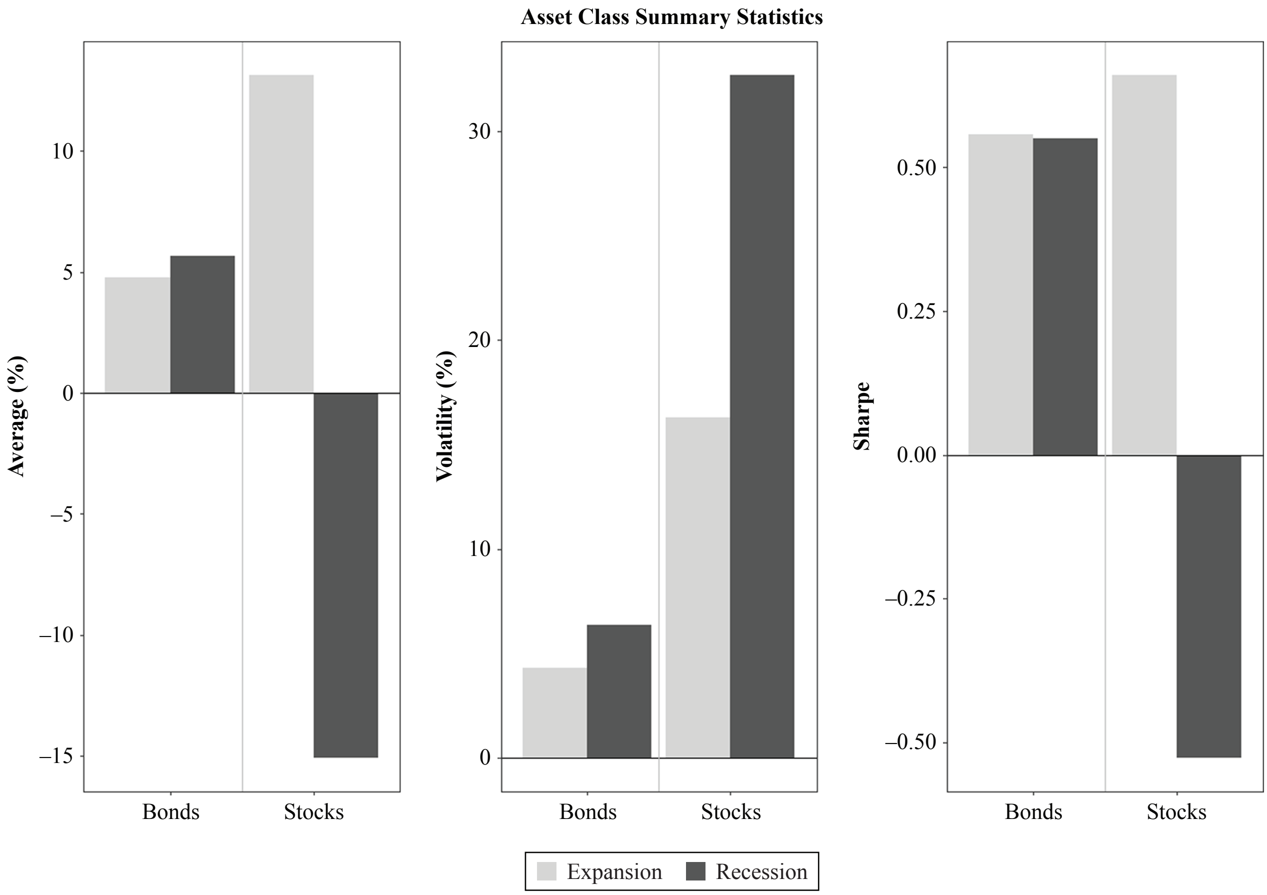
One popular method to allocate across assets is ERC, for which each N assets would contribute a fraction $1/N$ to the total portfolio risk (Maillard, Roncalli, and Teiletche 2008). This methodology provides economically motivated allocations and avoids estimation error in mean returns. We consider ERC as a static benchmark for our dynamic allocation.

We use dynamic risk budgeting to alter the portfolio risk contribution coming from stocks and bonds through different economic conditions. Starting from the benchmark of 50% risk contribution from stocks and 50% from bonds, we increase the risk budget for stocks

[6] BBG ticker: LUATTRUU Index.

EXHIBIT 5

Summary Statistics for Stocks and Bonds across Phases



Notes: We plot the average annual returns, volatility, and sharpe ratios for stocks and bonds conditional on expansions and recessions. The sample is from 1995 to 2019.

EXHIBIT 6

Risk Budgets across Expansions and Recessions

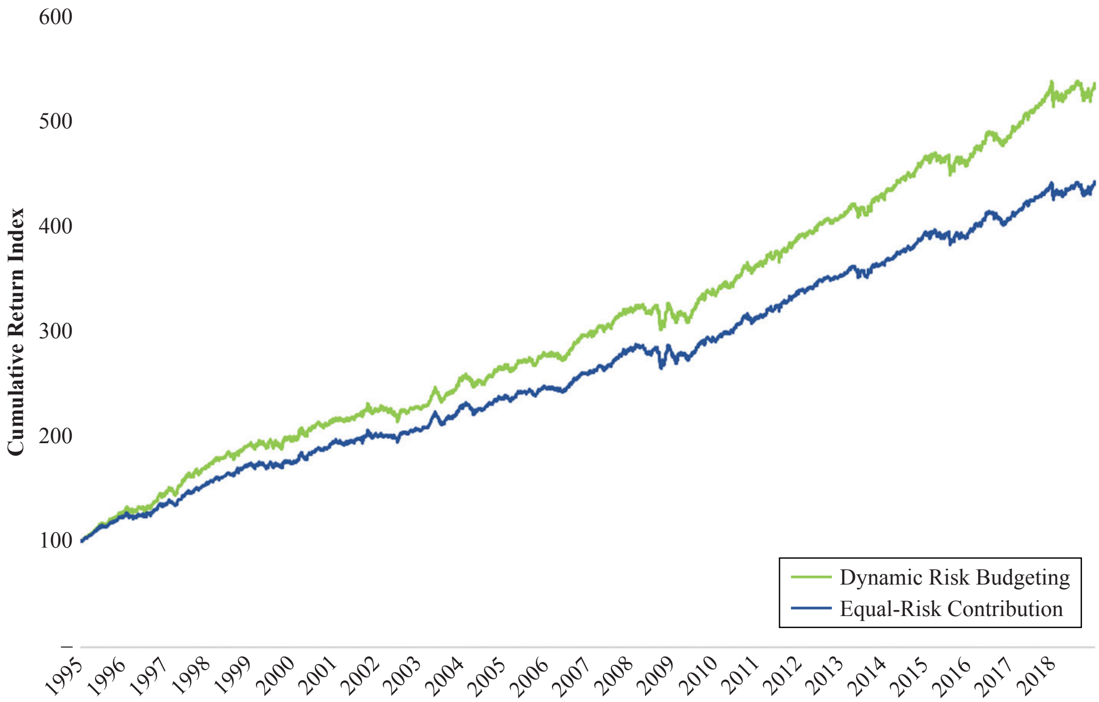
	Expansion	Recession
Panel A: Equal-Risk Contribution		
Stocks	50%	50%
Bonds	50%	50%
Panel B: Dynamic Risk Budgeting		
Stocks	75%	25%
Bonds	25%	75%

to 75% and decrease the risk budget for bonds to 25% in expansions, and we decrease the risk budget for stocks to 25% and increase the risk budget for bonds to 75% in recessions. Exhibit 6 summarizes the risk allocations.

At the beginning of each month, we estimate the SVM model after nonfarm payroll is released by the BLS. The model outputs the recession prediction for the upcoming month. Depending on this prediction, we choose the appropriate risk budgets for our dynamic model. We then solve an optimization problem, which takes in the risk budgets and the covariance matrix estimated using the past 126 days (six months) to obtain the weights on stocks and bonds. We hold this portfolio until the next estimation day for the SVM model. For ERC, we allocate a 50% risk budget to stocks and bonds independent of the recession prediction, and we solve a similar optimization problem to get portfolio weights. ERC is also rebalanced on each SVM estimation day.

Exhibit 7

Cumulative Returns for Dynamic Risk Budgeting and Equal-Risk Contribution Strategies



Notes: We plot the cumulative returns, starting at 100, for two strategies investing in stocks and bonds. Dynamic risk budgeting adjusts risk budgets assigned to stocks and bonds depending on whether the economy is in an expansion or a recession. Equal-risk contribution always allocates 50% risk contribution to stocks and 50% to bonds.

We compare the dynamic risk budgeting approach to ERC in Exhibit 7. Cumulative returns for both series start at 100 on January 3, 1995. As of January 31, 2019, the dynamic risk budgeting strategy reached an index level of 533, whereas the ERC strategy reached a level of 441. The annualized holding-period return is 7.3% for dynamic risk budgeting and 6.4% for ERC.

Exhibit 8 presents the strategy summary for dynamic risk budgeting and ERC strategies. Dynamic risk budgeting results in higher returns and higher volatility compared to ERC, which translates to a marginally higher Sharpe ratio (0.94) relative to ERC (0.91). The maximum drawdowns are virtually identical for the two strategies. Because dynamic risk budgeting has higher returns and about the same max drawdown as ERC, we see a higher Calmar ratio for dynamic risk budgeting (0.94) than for ERC (0.81).

Exhibit 9 takes a closer look at the risk associated with the two asset allocation strategies. We compare drawdowns over 24 years from January 1995 to January 2019.

Exhibit 8

Summary Statistics of Dynamic Risk Budgeting and Equal-Risk Contribution Strategies

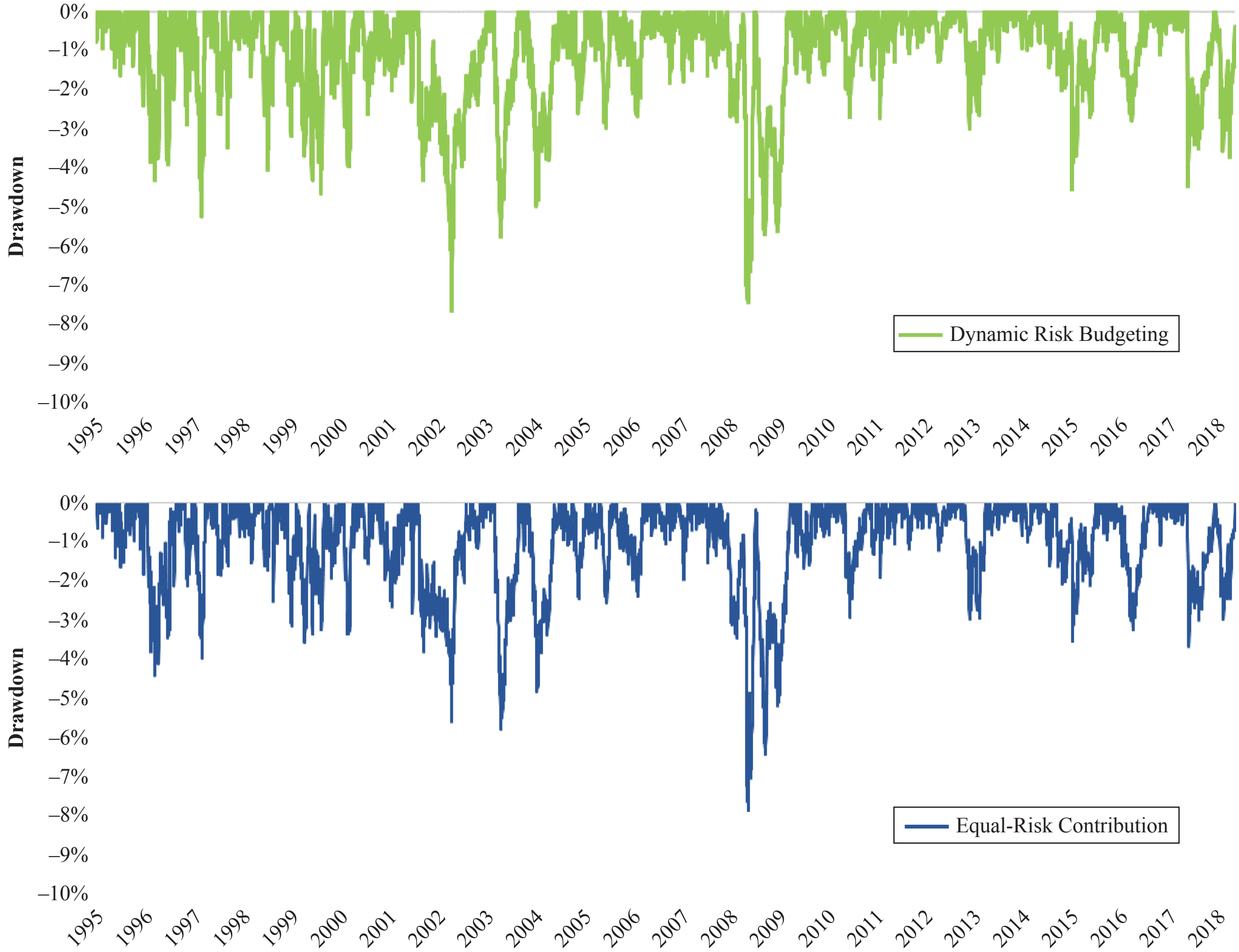
	Dynamic Risk Budgeting	Equal-Risk Contribution
Average Return	7.3%	6.4%
Volatility	5.1%	4.4%
Sharpe Ratio	0.94	0.91
Max Drawdown	–7.6%	–7.8%
Calmar Ratio	0.94	0.81

Notes: We present summary statistics for two strategies investing in stocks and bonds. Dynamic risk budgeting adjusts risk budgets assigned to stocks and bonds depending on whether the economy is in an expansion or a recession. Equal-risk contribution always allocates 50% risk contribution to stocks and 50% to bonds. The sample is from 1995 to 2019.

The worst drawdown periods for both strategies are during the two recessions in this sample: 2001 and 2008. During these periods, the equity market suffered a very large drawdown on the order of 40% to 50%.

E X H I B I T 9

Drawdowns for Dynamic Risk Budgeting and Equal-Risk Contribution Strategies



Notes: We show the drawdowns for two strategies investing in stocks and bonds. Dynamic risk budgeting adjusts risk budgets assigned to stocks and bonds depending on whether the economy is in an expansion or a recession. Equal-risk contribution always allocates 50% risk contribution to stocks and 50% to bonds.

The combined stocks and bonds portfolios, whether dynamic risk budgeting or ERC, suffered their deepest drawdowns in 2001 and 2008. Comparing the two strategies, dynamic risk budgeting had a larger drawdown in 2001 and a slightly smaller drawdown in 2008. The 2001 drawdown for ERC also had a shorter duration than dynamic risk budgeting. Aside from the two drawdowns in recessions, other drawdowns are relatively mild. They are typically smaller than 5% and quickly recover from the troughs.

CONCLUSION

Many problems in economics and finance can be viewed as prediction exercises. Statistical and machine learning techniques, with their impressive predictive power, are becoming increasingly common in economics and finance applications. In this article, we address a classic issue in macroeconomics, recession identification and prediction, using SVM. We find SVM to be a useful tool for this application.

We consider four variables as inputs to SVM: the monthly log difference in nonfarm payrolls, the log difference in the average monthly price of the S&P 500 price level, the level of the production index from the Manufacturing ISM Report on Business, and the 10-year US Treasury yield minus the federal funds rate. These variables reflect prevailing conditions in the labor market, stock market, goods market, and bond market. For SVM, we use a radial basis kernel, and the soft margin cost is chosen using 10-fold cross validation. We train our model from 1959 to 1973 and compare model predictions to NBER chronology from 1973 to 2018.

The SVM model achieves a classification error of 5.3% from 1973 to 2018. In comparison, the DFMS model (Diebold and Rudebusch 1996; Chauvet 1998) has a classification error of 4.7% or 9.3%, depending on how probabilities are transformed into a recession indicator variable; the GDP-based model of Chauvet and Hamilton (2006) has a classification error of 7.3%. Although the DFMS slightly outperforms the SVM model, it is more difficult to set up and computationally intensive to estimate. The SVM model could also be used to forecast recessions. The one-, two-, and three-month forecasts result in classification errors of 6.7%, 6.8%, and 8.0%, respectively.

Recession predictions can be used in dynamic asset allocation. We consider a dynamic risk budgeting approach that allocates a larger risk budget to stocks in expansions and a larger risk budget to bonds in recessions. Compared to a static ERC, the dynamic approach has higher returns (7.3% per year) than ERC (6.4%), higher volatility (5.1% versus 4.4%), and a higher Sharpe ratio (0.94 versus 0.91). Dynamic risk budgeting does not necessarily lead to more tail risk, as the maximum drawdowns for the two strategies are similar (–7.6% and –7.8%). Our dynamic risk budgeting approach illustrates how investors may strategically adjust their portfolio when armed with recession forecasts.

Predicting turning points is inherently interesting, but a focus on recession prediction classifies the economy into just two phases. Within a recession or expansion, the economy does not behave in a uniform way. As such, other models favor alternative approaches that may help uncover intraphase behavior. For example, Aruoba, Diebold, and Scotti (2009) produced cardinal measurements for the prevailing macroeconomic conditions. Researchers may also consider combining distinct models for forecasting recessions and for the level of the economy. Each model has its own value for investors and policymakers.

One interesting future research direction would be to expand the data to include other countries. An out-of-sample test provides the best evaluation of statistical models. In particular, it would be interesting to see if SVM can help identify euro area expansions and recessions, using as training data the chronology maintained by the Business Cycle Dating Committee at the Centre for Economic Policy Research. Moving beyond OECD countries, China and India could provide relevant test cases—although data limitations may hamper model performance. Another potentially interesting research direction is to apply SVM to forecast continuous target variables (support vector regression) in macroeconomics. SVM may be used to forecast GDP growth, unemployment, industrial production, or any other key macroeconomic variables.

Another interesting research direction would be to expand beyond two assets in dynamic allocation. Rather than treating the stock market as a whole, we may consider different stock market factors such as value, momentum, and profitability, or different market segments such as large cap and small cap. These finer divisions may behave in distinct ways across different macroeconomic regimes. Similarly, the bond market can be divided into finer segments, expressing credit risk and duration risk in specific combinations. Those segments could also react to expansions and recessions differently. Lastly, the analysis could be expanded to include additional asset classes such as currencies and commodities.

ACKNOWLEDGMENT

We thank Xavier Botteri, Frank Diebold, Andrea Giusto, Alfredo Hernandez, Jean-Daniel Malan, and Sergiy Verstyuk for very helpful comments.

REFERENCES

Aruoba, S. B., F. X. Diebold, and C. Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27: 417–427.

Ben-Hur, A., D. Horn, H. T. Siegelmann, and V. Vapnik. 2001. "Support Vector Clustering." *Journal of Machine Learning Research* 2: 125–137.

Berge, T. J. 2015. "Predicting Recessions with Leading Indicators: Model Averaging and Selection over the Business Cycle." *Journal of Forecasting* 34: 455–471.

Bok, B., D. Caratelli, D. Giannone, A. M. Sbordone, and A. Tambalotti. 2018. "Macroeconomic Nowcasting and Forecasting with Big Data." *Annual Review of Economics* 10: 615–643.

Brennan, M. J., E. S. Schwartz, and R. Lagnado. 1997. "Strategy Asset Allocation." *Journal of Economic Dynamics and Control* 21: 1377–1403.

Burns, A. F., and W. C. Mitchell. *Measuring Business Cycles*. New York: NBER, 1946.

Camacho, M., G. Perez-Quiros, and P. Poncela. 2012. "Markov-Switching Dynamic Factor Models in Real Time." *International Journal of Forecasting* 34: 598–611.

Campbell, J. Y., and L. M. Viceira. 1999. "Consumption and Portfolio Decisions When Expected Returns Are Time Varying." *Quarterly Journal of Economics* 114: 433–495.

——. 2001. "Who Should Buy Long-Term Bonds?" *American Economic Review* 91: 99–127.

Chauvet, M. 1998. "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching." *International Economic Review* 39: 969–996.

Chauvet, M., and J. D. Hamilton. "Dating Business Cycle Turning Points" In *Nonlinear Time Series Analysis of Business Cycles*, edited by P. R. Costas Milas and D. van Dijk, pp. 1–44. Bingley, UK: Emerald Group Publishing Limited, 2006.

Chauvet, M., and J. Piger. 2008. "A Comparison of the Real-Time Performance of Business Cycle Dating Methods." *Journal of Business and Economic Statistics* 26: 42–49.

Davig, T., and A. S. Hall. 2019. "Recession Forecasting Using Bayesian Classification." *International Journal of Forecasting* 35: 848–867.

Diebold, F. X., and G. Rudebusch. 1996. "Measuring Business Cycles: A Modern Perspective." *Review of Economics and Statistics* 78: 67–77.

Drucker, H., C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik. "Support Vector Regression Machines." In *Advances in Neural Information Processing Systems 9*, pp. 155–161. Cambridge, MA: MIT Press, 1997.

Estrella, A., and F. S. Mishkin. 1998. "Predicting US Recessions: Financial Variables as Leading Indicators." *Review of Economics and Statistics* 80: 45–61.

Fossati, S. 2016. "Dating US Business Cycles with Macro Factors." *Studies in Nonlinear Dynamics & Econometrics* 5: 529–547.

Friedman, J., T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning*, 2nd ed. New York: Springer, 2009.

Giusto, A., and J. Piger. 2017. "Identifying Business Cycle Turning Points in Real Time with Vector Quantization." *International Journal of Forecasting* 33: 174–184.

Kauppi, H., and P. Saikkonen. 2008. "Predicting US Recessions with Dynamic Binary Response Models." *Review of Economics and Statistics* 90: 777–791.

Lahiri, K., and G. Monokroussos. 2013. "Nowcasting US GDP: The Role of ISM Business Surveys." *International Journal of Forecasting* 29: 644–658.

Maillard, S., T. Roncalli, and J. Teiletche. 2008. "The Properties of Equally-Weighted Risk Contributions Portfolios." *The Journal of Portfolio Management* 36: 60–70.

National Bureau of Economic Research. "The NBER's Recession Dating Procedure." 2008, https://www.nber.org/cycles/jan08bcdc_memo.html.

Qi, M. 2001. "Predicting US Recessions with Leading Indicators via Neural Network Models." *International Journal of Forecasting* 17: 383–401.

Scholkopf, B., K. K. Sung, C. J. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik. 1997. "Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers." *IEEE Transactions on Signals Processing* 45: 2758–2765.

Stock, J. H., and M. W. Watson. 1989. "New Indexes of Coincident and Leading Economic Indicators." *NBER Macroeconomics Annual* 4: 351–394.

——. 2002. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20: 147–162.

Vapnik, V. *The Nature of Statistical Learning Theory*. New York: Springer, 2013.

Wright, J. H. "The Yield Curve and Predicting Recessions." In *Finance and Economics Discussion Series*. Federal Reserve Board, 2006.

Disclaimer

The views expressed herein are those of the individual authors and do not necessarily reflect official positions of Paraconic Technologies US, Inc. ("Paraconic"), nor are the views endorsed by Paraconic. This article is not an offer to sell or a solicitation of an offer to buy any investment product or services offered by Paraconic. Paraconic does not guarantee the accuracy or completeness of the information contained herein, and any information provided by third parties has not been independently verified by Paraconic.

ADDITIONAL READING

The Properties of Equally Weighted Risk Contribution Portfolios

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche

The Journal of Portfolio Management

https://jpm.pm-research.com/content/36/4/60

ABSTRACT: *Minimum-variance portfolios and equally weighted portfolios have recently prompted great interest from both academic researchers and market practitioners because their construction does not rely on expected average returns and, therefore, is assumed to be robust. In this article, the authors consider a related approach in which the risk contribution from each portfolio component is made equal, maximizing the diversification of risk, at least, on an ex ante basis. Roughly speaking, the resulting portfolio is similar to a minimum-variance portfolio subject to a diversification constraint on the weights of its components. The authors derive the theoretical properties of such a portfolio and show that its volatility is located between those of minimum-variance and equally weighted portfolios. Empirical applications confirm that ranking. Equally weighted risk contribution portfolios appear to be an attractive alternative to minimum-variance and equally weighted portfolios and, therefore, could be considered a good trade-off between the two approaches in terms of absolute risk level, risk budgeting, and diversification.*