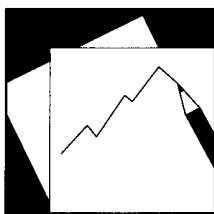


Working Paper

INTERNATIONAL MONETARY FUND



IMF Working Paper

Predicting Recessions: A New Approach For Identifying Leading Indicators and Forecast Combinations

Chikako Baba and Turgut Kışınbay

IMF Working Paper

Monetary and Capital Markets Department

Predicting Recessions: A New Approach for Identifying Leading Indicators and Forecast Combinations

Prepared by Chikako Baba and Turgut Kışınbay¹

Authorized for distribution by Aditya Narain

October 2011

Abstract

This study proposes a data-based algorithm to select a subset of indicators from a large data set with a focus on forecasting recessions. The algorithm selects leading indicators of recessions based on the forecast encompassing principle and combines the forecasts. An application to U.S. data shows that forecasts obtained from the algorithm are consistently among the best in a large comparative forecasting exercise at various forecasting horizons. In addition, the selected indicators are reasonable and consistent with the standard leading indicators followed by many observers of business cycles. The suggested algorithm has several advantages, including wide applicability and objective variable selection.

JEL Classification Numbers: C30, C32, C52, C53, E32, E37

Keywords: business cycles, leading indicators, forecast encompassing, forecast combination

Author's E-Mail Address: cbaba@imf.org, tkisinbay@oppenheimerfunds.com

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

¹ John Bluedorn, David Harvey, Ayhan Kose, Grace Li, Tara Sinclair, Herman Stekler and seminar audiences at George Washington University, and Goldman Sachs provided valuable comments that improved this work. At the time of writing this paper, the second author was with the Monetary and Capital Markets Department of the IMF.

Contents	Page
I. Introduction	<u>3</u>
II. Forecasting Models and Forecast Evaluation Methods.....	<u>5</u>
A. Forecasting Models	<u>5</u>
B. Forecast Evaluation	<u>6</u>
C. Forecast Encompassing Tests for Probability Forecasts	<u>6</u>
III. The Encompassing Algorithm	<u>7</u>
IV. Empirical Application: Predicting U.S. Recessions	<u>8</u>
A. Data and Estimation Set-up.....	<u>8</u>
B. Choice of Algorithm Parameters.....	<u>10</u>
C. Performance of the Algorithm Compared to Single-Indicator Models and Indices.....	<u>11</u>
D. Analysis of Variables Chosen by the Encompassing Algorithm	<u>12</u>
E. Observations on the Performance of Composite Indices and Individual Indicators.....	<u>17</u>
F. Discussion and Methodological Postscript.....	<u>18</u>
V. Conclusion	<u>19</u>
Figure 1. Combined Forecasts from the EAL Algorithm	<u>10</u>
Tables	
1. Forecast Loss at Different Significance Levels	<u>21</u>
2. Performance of the Algorithm Relative to Simple Forecast Averaging	<u>21</u>
3a. Best 25 Indicators Based on QPS Ranking	<u>22</u>
3b. Best 25 Indicators Based on LPS Ranking	<u>23</u>
4. List of Variables Chosen by the EAL Algorithm	<u>24</u>
5. Components of the Conference Board Leading and Coincident Indicators	<u>24</u>
References.....	<u>25</u>
Appendix:Description of Data	<u>27</u>

I. INTRODUCTION

Forecasting phases of the business cycle has been of interest to economists and policy makers for decades.² Economic theory provides broad guidance to economists about which indicators to monitor, but the details are left to applied economists and practitioners. Often there are numerous candidate indicators to forecast business cycles, but it is not always obvious which of these are the most important for forecasting. The abundance of data that are now available due to advances in computing power, data collection and dissemination, has made this problem more acute. For analysis and decision making, it is desirable to develop tools that allow us to identify the most relevant indicators.

This paper proposes a formal method to identify leading indicators from a large data set and advocates a forecast combination based method to forecast recessions. The method aims to identify all the indicators that provide useful marginal information that is not contained in other variables. In order to do that, the algorithm relies on the forecast encompassing principle that determines whether one of the two forecasts in a pair encompasses, i.e., contains all of the useful information of, the other forecast. If one forecast does not encompass the other but rather both models contain some incremental information, then there is the potential to form a combined forecast that incorporates the useful information from both models.

Encompassing tests generally compare two set of forecasts from two different models, while the selection of leading indicators from all possible models or data sets is a multi-dimensional problem. The algorithm proposed in Kışınbay (2010), and adapted here, allows for multiple comparisons in order to eliminate all the encompassed variables from the data set. The method is applied to a large data set to select a subset of indicators that provide different and complementary pieces of information. The collection of such variables forms the leading indicators.

Recent literature suggests that often more data is better than less for forecasting.³ The advantage of using a large data set is evidenced by the popularity of forecasting with dynamic factor models in recent years, with some success in forecasting performance. In empirical applications with these models, typically all available data are used to obtain the factors; a preliminary stage to filter the large data set is not common. In the forecast combinations literature as well, typically all forecasts enter the combination.

However, there is a case for using a subset of a large data set or a subset of forecasts to achieve the best results. Boivin and Ng (2006) provide a number of theoretical reasons for

² See Diebold (1998) for a history of macro-econometric modeling and Hamilton (2011) for the recent literature on various approaches to dating and forecasting business cycles.

³ See Stock and Watson (2011) for a review of forecasting with factor models; and Timmerman (2006) for a review of forecast combinations.

using a subset of a large data set in factor analysis; for example, if forecasting power is provided by a factor that is dominant in a small dataset but is dominated in a larger dataset, then forecasting performance may deteriorate. In the forecast combination literature too, more often than not all forecasts enter the combination, but there may be advantages in using a sub-sample. Aiolfi and Timmerman (2006) propose a method where forecasts are sorted into quartiles based on past performance, and transition probabilities are used to select a subset of quartiles that enter the final forecast combination. A potential advantage of using a subset of forecasts is reduced parameter estimation error in the combination weights.

Beyond the statistical rationale, a smaller number of indicators may be preferred for macroeconomic interpretations. Despite the existence of much more data, most commentators usually focus on a core set of indicators as it is not easy to understand the dynamics of all series, and their relationships to aggregate activity. Moreover, forecasts are generally presented as part of a macroeconomic story. As Leamer (2009) argues in his book *Macroeconomic Patterns and Stories*, human beings are pattern-seeking animals. The forecasting process involves not only seeking patterns and correlations in data, but also telling causal stories behind the forecasts. The latter cannot be done with hundreds of indicators; at the same time a forecaster would still want to benefit from the large data set. Methods that allow one to represent the larger data set with a core subset help balance both these needs.

Our method adds a new tool to the leading indicator literature with an additional layer of objectivity compared to some of the existing methods. Marcellino (2006) provides a comprehensive review of the leading indicator literature, where he shows that data selection methodologies typically use both subjective and objective criteria. For example, the OECD (2002) approach has a variety of criteria for model selection, ranging from the use of economic theory to the timeliness and smoothness of the series. Often, quite a bit of expert judgment goes into the indicator selection. Similarly, Conference Board indicators are selected based on economic and statistical criteria as well as judgment.⁴ While the use of expert knowledge in the selection of indicators may have merit, we argue that pure statistical methods can certainly add value. At the very least, an automated approach can be a useful starting point for further analysis. In many cases there may not be enough prior knowledge or sound theory about how to monitor a variable.⁵

This paper aims to make contributions in three areas. First, it proposes a formal econometric method to identify leading indicators from a large data set. Second, it adapts the new forecast combination approach proposed in Kışınbay (2010) for probabilistic forecasts. Third, using

⁴ Conference Board (2001), Section II describes the methodology. The criteria include concepts such as conformity (the series must conform well to the business cycle), consistent timing (the series must exhibit a consistent timing pattern over time as a leading, coincident or lagging indicator), and economic significance (cyclical timing must be economically logical), among others.

⁵ Section IV F provides a more detailed discussion of the benefits of a purely econometric approach.

U.S. data, it assesses the forecasting performance of the three well-known business cycle indices as well as a large set of individual indicators.⁶ Thus, the paper provides a comprehensive empirical forecast evaluation exercise.

II. FORECASTING MODELS AND FORECAST EVALUATION METHODS

A. Forecasting Models

We first obtain a pseudo out-of-sample probability forecast f_t^i of a binary event y_t based on each indicator series $x_t^i, i = 1, \dots, N$, by applying a standard probit model. y_t equals one if the economy is in a recession at time t and zero otherwise. Let y_t^* be an unobserved dependent variable that determines the occurrence of the event in a way that $y_t = 1$ if $y_t^* > 0$ and $y_t = 0$ otherwise. Let $X_t^i = [1, x_{t-h}^i, x_{t-h-1}^i, \dots, x_{t-h-k}^i]'$ be a vector of an observed leading indicator that includes a constant and own lags and h is the forecast horizon. The following model is fitted to the data:

$$y_t^* = \beta' X_t^i + \varepsilon_t \quad (1)$$

where ε_t is distributed normally. The probability of the event is expressed by the cumulative normal distribution function F , that is:

$$\Pr(y_t = 1 | X_t^i) = \Pr(-\varepsilon_t < \beta' X_t^i) = F(\beta' X_t^i) \quad (2)$$

And β is obtained by maximizing the following log-likelihood function:

$$\log L = \sum_{y_{t+k}=1} \log F(\beta' X_t^i) + \sum_{y_{t+k}=0} \log [1 - F(\beta' X_t^i)] \quad (3)$$

The lag length k can be optimally chosen for each of the single-indicator models by a standard information criterion. In the following application to the US recession forecasting, we employ the Bayesian Information Criterion (BIC).⁷

⁶ The indices are the Conference Board Coincident and Leading Indicators (CBCI and CBLI, respectively), OECD Leading Indicator for the U.S. (OECD-LI), and Chicago Fed's National Activity Index (CFNAI).

⁷ In the application, we also experimented with a single regression using a moving average of an indicator

variable, i.e. $X_t^i = \left[1, \frac{1}{k} \sum_{j=0}^k x_{t-j}^i \right]'$ where k is up to 6 lags. The forecasts generally perform better in the multiple-regression with the BIC based lags.

With N indicators at hand, we obtain N sequences of probability forecasts from the single-indicator models $\{f_t^1, \dots, f_t^N\}_{t=1}^T$ and a series of realizations $\{y_t\}_{t=1}^T$. Clearly,

$$f_t^i \equiv F(\hat{\beta}' X_t^i) \quad (4)$$

In practice, the algorithm utilizes out-of-sample forecasts, so $\hat{\beta}$ is re-estimated for each indicator and each time.

B. Forecast Evaluation

To evaluate the accuracy of forecasts, we use two measures: the Quadratic Probability Score (QPS) and the Log Probability Score (LPS) (see Diebold and Rudebusch, 1989). These measures indicate the average closeness of the predicted probabilities and the observed realizations, the latter a zero-one dummy variable.⁸ For each time series of probability forecasts based on indicator i , the two probability forecast loss functions are defined as follows:

$$QPS^i = \frac{1}{T} \sum_{t=1}^T 2(f_t^i - y_t)^2 \quad (5)$$

$$LPS^i = -\frac{1}{T} \sum_{t=1}^T ((1 - y_t) \log(1 - f_t^i) + y_t \log f_t^i) \quad (6)$$

where $QPS \in [0, 2]$, and $LPS \in [0, \infty)$. In both cases, 0 indicates perfect accuracy, and lower numbers indicate more accuracy. The LPS penalizes large mistakes more heavily than the QPS.

C. Forecast Encompassing Tests for Probability Forecasts

Encompassing tests assess whether a combination of forecasts provides a statistically significant reduction in forecast loss compared to the individual forecasts that are under consideration. If not, there is potential to improve forecasts by combining them. Most of the tests proposed in the literature are designed for testing for encompassing in the case of point forecasts, but often the object of interest is a probability forecast. Fortunately, a recent study by Clements and Harvey (2010) extends the results for encompassing tests to probability forecasts, henceforth called the CH test.⁹ Their test builds on the Harvey *et al.* (1998) version of the Diebold and Mariano (1995) test, which was developed for the linear regression

⁸ See Galbraith and van Norden (2011) for a new approach where probability forecasts are evaluated using kernel estimators, instead of binary or other discrete groupings.

⁹ Clements and Harvey offer three alternative tests, and note that there is little to suggest the use of one formulation over another in the literature. We choose to use the Harvey *et al.* version, which is FE(2) in their notation, as it is most commonly used in recent empirical studies in the literature.

framework. Clements and Harvey show the applicability of the Harvey *et al.* test in the context of probability forecasts.

Let f_t^1 and f_t^2 be two alternative probability forecasts of a binary variable y_t , and forecast errors are defined as $e_t^i = y_t - f_t^i$, $i = 1, 2$.

$$CH = \left(T^{-1/2} \left[T + 1 - 2h + T^{-1}h(h-1) \right]^{1/2} \right) \times \frac{T\bar{d}}{\sqrt{\sum_{\tau=-(h-1)}^{h-1} \sum_{t=|\tau|+1}^T (d_t - \bar{d})(d_{t-|\tau|} - \bar{d})}} \quad (7)$$

Where $d_t = (e_t^1 - \bar{e}^1) \left[(e_t^1 - \bar{e}^1) - (e_t^2 - \bar{e}^2) \right]$ and $\bar{d} = T^{-1} \sum_{t=1}^T d_t$, and h denotes the forecast horizon in a multi-period framework such that f_t refers to a forecast of period t , made at period $t-h$. The test statistic is compared to a Student t distribution with $T-1$ degrees of freedom. The authors show that the CH test has the asymptotic standard normal distribution in this context, and their simulations show that the test has good finite sample properties.¹⁰

III. THE ENCOMPASSING ALGORITHM

The algorithm that we utilize here selects a set of indicators based on the out-of-sample performance of single-indicator probability models. It is a variant of the one proposed for linear regression forecasts in Kışınbay (2010), and is called the Encompassing Algorithm, henceforth EAL. The idea is to compare indicator forecasts with each other using encompassing tests, eliminate those that are encompassed by others, and combine the remaining forecasts. The comparisons are done bilaterally using the CH test.¹¹ Once the pseudo out-of-sample forecasts are obtained,¹² the following steps are taken to eliminate forecasts encompassed by others, and obtain the forecast combination.

Step 1. Calculate the out-of-sample forecasts. Rank the models according to their past performance based on QPS.

Step 2. Pick the best model (i.e., the model with the lowest QPS), and test whether the best model encompasses other models, using the CH test. If the best model encompasses the alternative model, delete the alternative model from the list of models.

¹⁰ Note that the encompassing tests are based on estimated probit models but do not account for the parameter estimation uncertainty.

¹¹ One can use alternative encompassing tests and different loss functions, The current version of the algorithm is based on QPS only.

¹² Forecasts are called ‘pseudo’ out-of-sample as we use revised data, as opposed to real time. In that sense, the analysis here differs from a complete real-time out-of-sample forecasting exercise.

Step 3. Repeat Step 3, with the second best model. Note that the list now contains only models that are not encompassed by the best model, and the best model.

Steps 4 and 4+. Continue with the third best model, and so on, until no encompassed model remains in the list.

Last Step. Calculate the combined forecast by taking the average of all remaining forecasts.

There are two issues to consider when using the EAL in applications. First, an initial set of out-of-sample forecasts with appropriate modeling specifications is required to apply the CH test. One option is to use all the available forecasts prior to the date on which the forecast is being produced by recursively extending a window, which arguably generates more robust estimates by extending observations. An alternative is to choose a rolling window of a fixed number of observations, which allows for structural changes in parameters. In either case, a window must be long enough to cover some occurrence of the events to successfully estimate a probit model.

Secondly, the significance level for the encompassing test needs to be specified. A choice of the significance level affects the number of indicators that remain in the final combined forecast and hence affects the forecast performance. At higher significance levels, fewer variables are eliminated from the data set, and the resulting combined forecasts are closer to the simple average forecast. At low significance levels, very few variables remain in the combination and the forecast benefits less from the advantages of combining. Results in Kışınbay (2010) suggest that significance levels from 0.20 to 0.35 give the best results for the algorithm in a linear regression context.

The algorithm aims to be comprehensive and to capture, to the extent possible, all the useful and separate pieces of information relevant for forecasting the target variable. The key feature of this method is to select indicators that provide complementary information. The selected indicators are most useful as a group since they are selected based on their relationship to other variables within the group. In this way it is a holistic approach, as opposed to considering each individual series in isolation, and then choosing a subset of series based on some subjective criteria.

IV. EMPIRICAL APPLICATION: PREDICTING U.S. RECESSIONS

A. Data and Estimation Set-up

We apply the proposed selection algorithm for forecasting recessions in the U.S.¹³ The analysis covers the monthly U.S. economic and financial data from January 1959 to December 2008. Standard probit regressions are used to forecast the probability of recessions where the independent variable in the regression is a binary variable that takes value one to

¹³ Recessions are particularly difficult to forecast, as shown by Loungani (2001).

indicate that the U.S. economy was in recession at period t and zero otherwise. We use the dates identified by the National Bureau of Economic Research (NBER) as recessions.

A set of leading indicators includes 166 monthly macroeconomic time series. All of them are either directly taken from or calculated by using Global Insight's DRI Basic Economics database. We follow Marcellino *et al.* (2006) for the selection of variables and for the transformation criteria. All the series are transformed to stationarity either by taking logarithm, difference, log-difference or used in level. Detailed data descriptions, and transformations applied to individual series are explained in the Data Appendix.

The data series are classified into five categories: (i) income, output, capacity utilization, and expectations; (ii) employment and unemployment; (iii) construction, inventories, and orders; (iv) interest rates and asset prices; and (v) nominal prices, wages, and money. Besides these leading indicators, we also use well-known composite leading indicators to compare the performance of our method. They include the Conference Board (CB)'s Coincident and Leading Indicators, denoted as CBCI and CBLI respectively, the OECD Leading Index (denoted as OECD-LI), and Chicago Fed's National Activity Indicator (CFNAI).

The forecasts of recessions are made using the h -step-ahead out-of-sample regression model for $h = 1, 3, 6, \text{ and } 12$ -month horizons. Each of the individual indicators includes own lags that are initially specified as $X_t^i = [1, x_{t-1}^i, x_{t-3}^i, x_{t-6}^i, x_{t-9}^i]'$ for $h = 1$, $X_t^i = [1, x_{t-3}^i, x_{t-6}^i, x_{t-9}^i, x_{t-12}^i]'$ for $h = 3$, $X_t^i = [1, x_{t-6}^i, x_{t-9}^i, x_{t-12}^i, x_{t-15}^i]'$ for $h = 6$, and $X_t^i = [1, x_{t-12}^i, x_{t-15}^i, x_{t-18}^i, x_{t-21}^i]'$ for $h = 12$. A lag length for each model is optimally chosen to minimize the BIC.

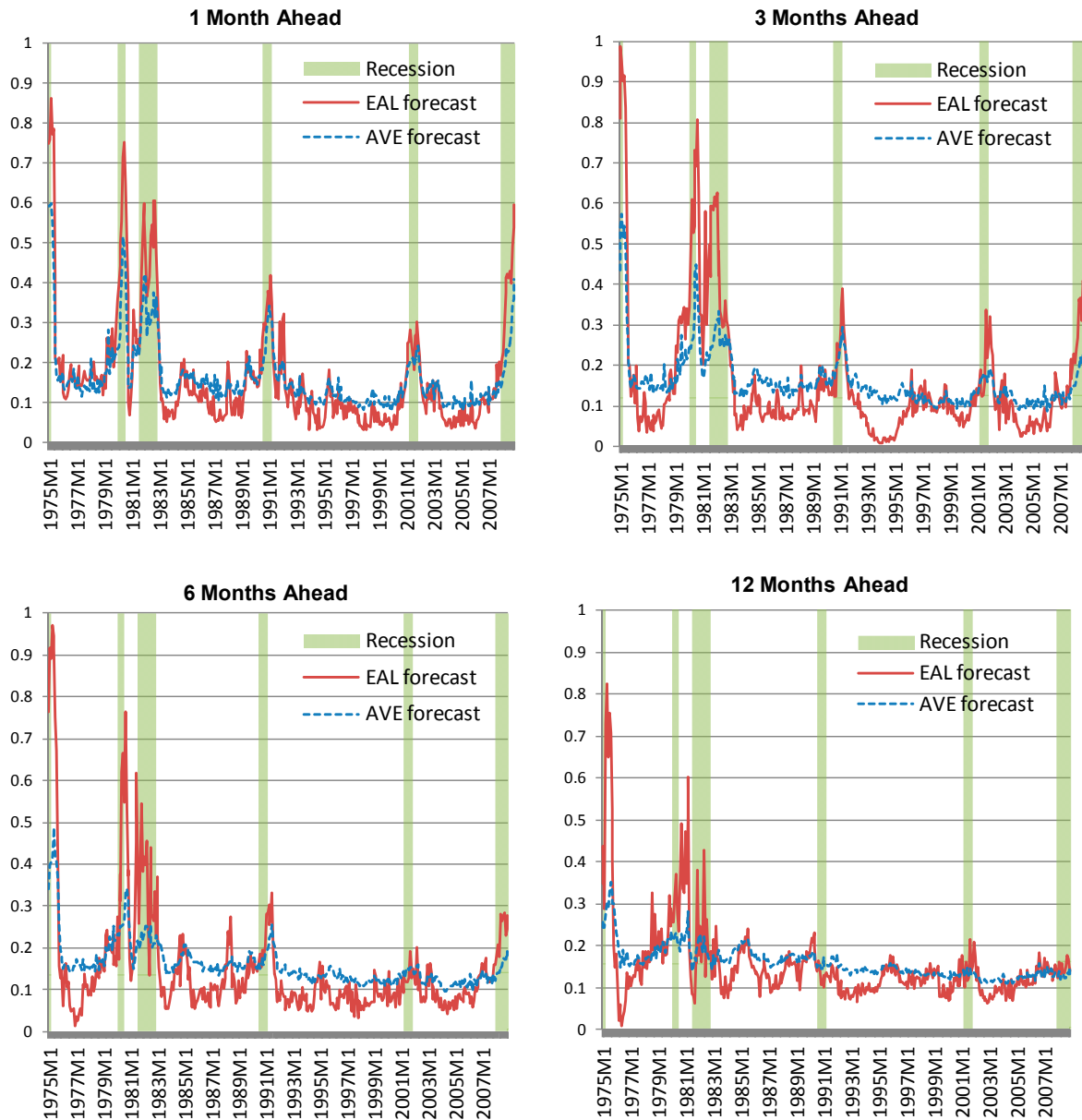
The sample period is divided into three parts. The observations prior to the date t_0 are used to estimate probit models and construct h -step-ahead out-of-sample forecasts. A second window of out-of-sample forecasts is needed to implement the algorithm. The third window is used to assess forecasts. The first set of probit models is estimated using data from 1959:01 to 1969:12. Then one-month-ahead forecasts are produced for 1970:01, using all the leading indicators. The sample period is expanded by one period (recursive window), and an estimate is obtained with data spanning 1959:01 to 1970:01, and forecasts are produced for 1970:02. And so on.

While many series date back to 1959, there are some that start in later years. The series that start after January 1959 are kept out from the sample until they pool a sufficient number of observations to compare the scores. The criterion for including new series is that at least 10 years (120 months) of out-of-sample forecasts are accumulated before including them in the algorithm. We pick 10 years to assure that the window for the out-of-sample forecasts includes at least one business cycle.

B. Choice of Algorithm Parameters

As discussed, one of the key parameters of the EAL is the significance level of the encompassing test. With a higher significance level, the algorithm allows more variables to remain in the final combined forecast. Table 1 shows the QPS and LPS for different significance levels for various forecast horizons. The first column, which reports the QPS and LPS results for the benchmark model, provides an average of all available forecasts. This average is the standard un-weighted forecast combination applied to probit forecasts, and is denoted by AVE. The remaining columns show the results for the EAL forecast at various significance levels. Figure 1 provides a visual presentation. Two results stand out.

Figure 1. Combined Forecasts from the EAL Algorithm



Source: Authors' calculations. The EAL forecasts are chosen by the encompassing test with $\alpha=0.25$.

First, in most of the cases, the EAL forecasts provide lower forecast loss than the forecast combination obtained by the AVE forecasts, i.e., simple averaging without any selection algorithm. The gains are higher especially at horizons of 1, 3, and 6-months-ahead, but AVE seems to perform better at 12 months-ahead for low significance levels. The results are broadly similar for the two loss functions, but there are a few exceptions. For example, at the 3-month-ahead horizon and when the significance level is 0.01 and 0.05, the EAL forecasts have lower QPS than the AVE forecasts, but higher LPS. For the other significance levels, however, the EAL outperforms AVE based on both loss functions. Note that the poorer performance of EAL forecasts generally occurs when the significance levels are low; that is, when there are only a few forecasts in the combination and hence diversification gains are minimal or non-existent. For a significance level greater than 15 percent, an EAL forecast generally outperforms the AVE forecast.

Second, there is no noticeable difference in the performance of the algorithm across different significance levels. This result differs from Kışınbay (2010), where it is found that significance levels from 0.20 to 0.35 give the best results for the algorithm in a linear regression context. The study there, however, is more comprehensive, as about 110 target variables are examined as opposed to the single case here that focuses on recessions. As there is no noticeable difference in algorithm performance across significance levels, we pick the 25 percent level because, based on the earlier results, that level performs well.

C. Performance of the Algorithm Compared to Single-Indicator Models and Indices

This section presents results that compare the performance of algorithm forecasts with other leading indicators in our data set, as well as with the benchmark AVE forecasts. Table 2 suggests that there is about 10 to 20 percent reduction in QPS and LPS that can be obtained from using the algorithm as compared to the benchmark model at horizons 1-, 3- and 6-month-ahead. Tables 3A and 3B show the best 25 indicators based on QPS and LPS scores for the four horizons. The tables rank results for the 166 series in the data set, as well as for the CBCI and CBLI, the OECD-LI, CFNAI, and the EAL algorithm.

The comparisons in Tables 3A and 3B show that the EAL forecasts are consistently among the top 20 indicators at all horizons. With QPS loss, the EAL forecast is ranked 18th for the 1-month-ahead forecast, 10th for 3-months-ahead, 4th for 6-months ahead, and 21st for 12-months-ahead. The relative performance is better with LPS loss, which penalizes outliers more than QPS indicating that EAL forecasts generally contain fewer outliers. Based on LPS, the EAL forecasts are ranked 13th for 1-month-ahead, 5th for 3-months-ahead, 1st for 6-months ahead, and 15th for 12-months-ahead. It needs to be noted that these ranks are affected by several spread variables. These spread variables provide very similar information content. Netting them out, ranks for all other variables would be better.

The value of indices and the EAL method is revealed by the observation that none of the individual indicators consistently shows up among the best indicators across horizons. While methods that combine information from different indicators are often among the best indicators at all horizons, individual indicators do not consistently have good indicator properties across different horizons. In other words, the usefulness of individual indicators about the state of the economy is horizon specific. For example, indicators that measure price developments lead recessions with a one-year lag, while employment variables have more coincident and lagging properties. No individual variable can capture all phases of the business cycle. On the other hand, methods that combine information from different sources blend relevant information and so can be useful more consistently across horizons. Similar results have been obtained by Clements and Galvão (2009) and Berge and Jordà (2011).

Comparisons with the CFNAI, OECD-LI, and the CB indices also show encouraging results for the EAL forecasts. No index consistently improves upon the EAL algorithm. CFNAI performs better than EAL at 1-month-ahead but not at other horizons. Similarly, OECD-LI performs better than EAL at short horizons, but not at 6- and 12-months- ahead. CBLI outperforms EAL only at the 12-month-ahead horizon with QPS loss, but not with LPS loss. CBCI does not perform well even at short horizons, at least based on the approach we adopt here.

D. Analysis of Variables Chosen by the Encompassing Algorithm

Examination of indicators chosen by EAL is informative. They are the variables that are not encompassed by other indicators and hence remain in the forecast combination. To facilitate comparison with the CB indices, we focus on the significance levels where the number of chosen variables is close to the number of variables in the CB indices. As mentioned before, there is no noticeable difference in the performance of the algorithm when the significance level of the test varies. This allows for the flexibility to choose a significance level that serves our purpose. The CBCI has four variables, and the CBLI has 10. The four variables in the CBCI are the same that the NBER Business Cycle Committee monitors to identify recessions (see NBER, 2008).¹⁴ In several cases we choose a slightly higher number than the CB variables as the encompassing test may not differentiate variables that are too similar to each other. For example, various spread definitions show up in the chosen variables, but they likely possess similar predictive power. In those cases, we just focus on the best performing variable and ignore the rest in making comparisons.

¹⁴ Note that we do not exactly calculate a coincident index as doing that would require using in-sample encompassing tests, as opposed to forecast encompassing tests. Yet, the difference between one-month-ahead versus current estimates of a recession should not be very significant.

The 1-month to 6-months-ahead horizons

At the short horizons, the indicators chosen by the EAL are primarily related to the labor market, housing, and consumption. At the 1-month-horizon, which is close to a coincident index, consumers are at the center (see Box 1). Several housing sector variables are chosen at this horizon, including those measuring sales and housing starts. In addition, at this horizon labor market indicators are among the best, including the Ratio of Index of Help-Wanted Advertising to Number of Persons, and Unemployment Insurance Claims. Finally, consumer confidence and interest rate spread variables are chosen at short horizons. A key difference between 1-month-ahead and 3-month-ahead forecasts is that at the latter horizon, supply side and production related variables are also chosen. In particular, several indices of the Institute for Supply Management (ISM) as well as the New Orders Index of the Philadelphia Fed are not encompassed by other variables.

A similar set of variables is highly ranked at 3- and 6-months ahead. The order of ranking varies, however. At the 6-month-ahead horizon, financial spread variables dominate the top ranks, followed by housing sector variables. We get complementary information from the stock market (three variables), capacity utilization (two variables), and the labor market. Unlike the 3-month-ahead horizon, ISM variables are not very prominent at the 6-month-ahead horizon. Another difference from the 3-month-ahead horizon is that capacity utilization variables show up at the 6-month-ahead horizon.

The 12-month-ahead horizon

At the 12-month-ahead horizon, the chosen variables have a somewhat different character. An important difference with other horizons is that several price variables are chosen by the algorithm at this horizon. The yield spread (various definitions) and the ISM Slower Deliveries Diffusion Index are the only two typical leading indicators chosen by the algorithm at this horizon. Interestingly, interest rates in levels are also chosen, suggesting that such variables may also provide predictive information content that is not contained in the spread variables. The level of the Effective Interest Rate on Conventional Mortgages (FYMCLE), the Federal Funds Rate (FYFF), and the Prime Rate charged by banks (RM1) are all among the variables chosen by the EAL algorithm. The only labor market variable chosen is the Average Duration of Unemployment (LHU680).

This set of variables could be consistent with a view that makes monetary policy a cause of recessions. Signs of overheating emerge a year ahead, as signaled by several inflation measures and the levels of interest rates picked up by the algorithm. Typically, the Fed responds to rising inflation by raising interest rates, which inverts the yield curve. The literature on monetary policy transmission suggests that the major impact of a tightening is felt on output in about a year, and thus it is expected that recessions may occur a year after the emergence of inflation, see Christiano *et al.* (1999). If these signs are followed six months later by sustained high spreads, increasing strains in the housing sector, capacity

utilization, and indices of stock prices, the likelihood of a recession increases. Finally, to monitor short-term signals, in addition to the housing sector, one might look at the indicators selected at 1- month and 3-months ahead, such as labor market indicators, ISM indices, and consumer expectations.

Similarities and differences between variables chosen by the EAL and the CB indices

Many of the CBLI components or their close substitutes are also chosen by the EAL algorithm. The two approaches differ in that the EAL approach is tailored to different horizons and, as a result, picks different indicators at each horizon, while the CB seeks more general leading indicator properties without an explicit emphasis on the next few quarters. The similarities or overlaps are more relevant at the shorter horizons. Five of the 10 variables of the CBLI are also selected by the EAL at the 1-month and 3-month horizons. These variables are Consumer Expectations, Unemployment Insurance Claims, the Yield Spread, Building Permits, and the S&P 500 Composite Index. The EAL also chooses variables that are quite similar to CBLI components; in those cases, the definitions are similar but the indicator in question is a different one. For example, instead of the CBLI's New Manufacturing Orders, and New Orders, Nondefense Capital Goods, the EAL approach favors two alternative indices: Philadelphia Fed's Diffusion Index of New Orders (JFIFFO), and the ISM New Orders Index (PMNO). Interestingly, at shorter horizons, the EAL-chosen indicators have little in common with the CBLI but more with the CBLI.

The key difference between the EAL and CBLI variables is related to the consumer side of the economy, especially to the housing sector. In Table 4, variables that are identical or similar between the EAL and CBLI are indicated by italics. The remaining variables, which highlight the difference between the two approaches, are mostly related to the housing sector and the labor market, a pattern especially pronounced at shorter horizons. For example, in addition to CBLI's Building Permits variable (A0M029), the EAL chooses four other housing variables. The best performing one is New 1-Family Houses, Month's Supply at Current Sales Ratio (HNR). The others are HUSTS1, HNS, HNIV and CONDO9; showing that the housing sector variables play a key role based on the results of the EAL approach. For labor markets, EAL selected variables include LHELX, which is the overall best indicator at the 1-month horizon, and others such as PMEMP and LHEL, which also capture developments in the labor market. Of course, CBLI also contains variables related to these sectors, but it is not as consumer heavy as the EAL chosen variables; the difference is the degree of emphasis.

Importance of consumer-related information

The important role of the consumer side of the economy for forecasting recessions is highlighted by prominent business cycle researchers. Leamer (2007) argues that eight of ten recessions he analyzed were preceded by sustained and substantial problems in housing. He shows that all major consumer aggregates contribute to weakness before the recession. The

most important one is residential investment, which is the major contributor to weakness in the year before the recession, followed by consumer durables, consumer services, and consumer nondurables. Business investment in equipment and software is the primary source of weakness during the recession, but not a leading indicator. In Leamer's words, "*It's a Consumer Cycle, not a Business Cycle.*" His results also suggest that housing is an accurate predictor of recessions; the next best predictor is consumer durables. Our results are largely in agreement with Leamer's and complement them. His analysis is based on components of GDP and on quarterly data. Our study offers numerous alternative and higher frequency (monthly) tools to monitor the cycle.

Sinai (2010) argues that the changing economy requires that the conceptualization and measurement of the business cycle be revisited. Sinai argues that the current approach to business cycle dating and measurement is tilted towards the goods side of the economy relative to services. This is a reflection of the tradition of business cycle analysis of the NBER in the first half of the 20th century, at a time when the U.S. economy was fundamentally manufacturing and industrial based. Now it is services centered, but this aspect may not be adequately captured by the current business cycle measurement approaches. Second, Sinai argues that business, consumer, and financial sector surveys are ample now, and could play a more prominent role in business cycle analysis, especially for analyzing turning points. Although CB's four coincident indicators currently receive a lot of attention, they may actually have a diminished role in monitoring recessions. For example, industrial production is now a smaller share of the economy, and the correlation between the overall level of activity and payroll employment has changed, with the latter becoming a lagging indicator in the recent cycles. Our results are in line with these observations, as the selected variables are not heavy on the manufacturing side of the economy, and ISM surveys play an important role at short horizons.

Summary

To conclude this section, our results suggest the EAL approach is a good candidate for forecasting recessions and identifying leading indicators. First, the forecasting performance is on par with or better compared to other well-established approaches to the monitoring of business cycles. Second, a closer examination of the individual variables suggests that the EAL selects variables that are similar to other well-known approaches. Naturally there are differences; notably the variables selected by the EAL put more weight on housing and consumers. These differences may point to an advantage of the proposed approach, since it is able to identify key variables that may not be given enough attention in other approaches.

Box 1. Which Variables are Chosen by the Algorithm?

At the 1-month-ahead horizon, using a 0.01 significance level, only three variables are chosen. These are the Ratio of Index of Help-Wanted Advertising to Number of Persons (LHELX); New 1-Family Houses, Month's Supply at Current Sales Ratio (HNR); and another housing variable, Housing Starts, Private Including Farm, One Unit (HUSTS1). These are not the typical variables used by the CB. Using a 0.05 significance level, six additional variables are chosen. In addition to the three chosen at the 0.01 significance level, we have Consumer Expectations (AWM123); Unemployment Insurance Claims (A0M005); the spread between 6-month T-bill yields and the Federal Funds Rate (SFYGM6); two more housing sector variables—New One Family Houses Sold (HNS) and New One Family Houses for Sale at the end of the Month (HNIV); and, finally, Real Personal Income, Excluding Transfers (A0M051).

At the 3-month-ahead horizon, using a 0.30 significance level, EAL-chosen variables include: New 1-Family Houses, Month's Supply at Current Sales Ratio (HNR), the best variable; several interest rate spreads, the best being the one derived from 3-month T-bills (SFYGM3); Housing Starts, Private Including Farm, One Unit (HUSTS1); Building Permits (A0M029); the Ratio of Index of Help-Wanted Advertising to Number of Persons (LHELX); Philadelphia Fed's Expected New Orders Index (JDIFFO); Unemployment Insurance Claims (A0M005); Institute of Supply Management (ISM) Employment Index (PMEMP); ISM New Orders Index (PMNO); ISM Production Index (PMP); S&P 500 Composite Index (U0M019); Index of Help Wanted Advertisements in Newspapers (LHEL); and Construction Contracts, Commercial and Industrial Buildings (CONDO9).

At the 6-month-ahead horizon, using a 0.30 significance level, EAL-chosen variables include: several yield spreads, the one derived from the 10-year-bond yield spread (SFYGT10) being the overall best indicator; New 1-Family Houses, Month's Supply at Current Sales Ratio (HNR); New Single Family Private Housing Units (HUATZC1); Capacity Utilization in Mining (UTL35); Building Permits (A0M029); the Average Duration of Unemployment (LHU680); Housing Starts – West (HSWST); ISM Employment Index (PMEMP); S&P's Common Stock Price Index-Industrials (FSPIN); Capacity Utilization—Electricity and Gas Utilities (UTL36); Housing Starts –Midwest (HSMW); Share Price Index—Dow Jones Industrial (JPSHARE); and S&P Composite Index: Dividend Yield (FSDXP).

At the 12-month-ahead horizon, using a 0.25 significance level, EAL-chosen variables include: several yield spreads, the one derived from the 10-year-bond yield spread (SFYGT10) being the overall best indicator; the level of the effective interest rate on conventional mortgages (FYMCLE); the Federal Funds Rate (FYFF); ISM Vendor Deliveries Index (U0M032); the Average Duration of Unemployment (LHU680); CPI-Less Medical Care (PUXM); CPI-U, Non-durables (PU882); Prime Rate charged by banks (RM1); PPI, Crude Materials (PWCMSA); and Construction Contracts, Commercial and Industrial Buildings (CONDO9).

E. Observations on the Performance of Composite Indices and Individual Indicators

The leading and coincident indicators provided by the Chicago Fed, the Conference Board, and the OECD-LI can be used in a complementary way. At the short horizon of 1-month-ahead, Chicago Fed's CFNAI is the best composite indicator. The CFNAI has a broad coverage of real sector variables, and based on our assessment framework, it is the best index at achieving that aim. Interestingly, the CBCI does not perform well at short-horizons relative to other indices; CBCI ranks 24th with QPS loss, and 20th with LPS loss. The relatively weak performance of the CBCI is important, since its components are closely monitored by the NBER to date business cycles. The OECD-LI comes a close second at short horizons, whereas the CBLI's relative performance improves at longer horizons. This result suggests that these two well-established indicators can be used in a complementary way. CBLI may give early warning signs about a looming recession about a year or six months ahead. After getting such signals, one might turn to the OECD-LI leading indicator for more accurate short-term signals, and finally to the CFNAI for a coincident indicator. Although CBCI does not perform as well as other indices, it has the advantage of being composed of four important variables that can be easily monitored and quickly updated.

Turning back to individual indicators, at short horizons several labor market indicators perform very well and deserve close monitoring. Notably, the LHEL and LHELX indicators of newspaper job-wanted advertisements are among the best indicators at short horizons. Baumol (2008) and Goldman Sachs (2008) argue that these variables, while important in the past, may have lost their impact on financial markets, partly due to a lag in their publication relative to the employment situation report and also because newspaper ads have lost share to internet job-search engines and companies' own web pages. Recently, the Conference Board discontinued this data series, and instead created a new index, called Help Wanted OnLine. Given the good performance of such indicators, it would be useful to blend the information in the old and new series to create a continuous index that would cover past business cycles.

The behavior of traditional labor market indicators may have changed. Sinai (2010) argues that payroll data were largely coincident with the business cycle in the post-World War II era, but this pattern changed with the recessions of 1990-1991 and 2001. Job creation lagged the recovery after the 1990-1991 recession, and did not show a sustained improvement until about a year after the trough. A similar pattern occurred after the 2001 recession. The payroll data may have become a lagging indicator in modern business cycles. Gordon (2010) argues that in the current economy, managers find it easier to shed workers and cut back hours in recessions, and are more likely to rely on productivity gains in an upturn by offering longer hours instead of new hiring, especially at the early stages of a recovery. The result is jobless recoveries, at least at the early stages, which break the coincident indicator property of the payroll data. On the other hand, variables measuring longer hours, or new temporary job adds, may be better correlated with output and therefore have good coincident properties.

The Real Money Supply, a traditional indicator of business cycles and a component of CBLI, does not seem to have good coincident or leading indicator properties. It is poorly correlated with recessions and ranks low among individual variables. The variable is not selected by the EAL. The marginal information coming from this variable may be limited and it could be replaced with better variables that capture financial market conditions that are relevant for business cycles. In fact, Conference Board researchers also observed this point, and documented the poor performance of the indicator in the last two recessions (Conference Board, 2010). They argue that the Real Money Supply may be removed from the CBLI and replaced with a ‘suitable indicator of monetary and credit conditions’.

F. Discussion and Methodological Postscript

A formal statistical approach to choose indicators from a large data set has a number of advantages. In contrast to approaches that benefit from judgment or historical precedent, a formal statistical approach (i) can be applied in different contexts; (ii) is objective; easily replicable, and time-consistent; (iii) has advantages in analyzing large data sets; and (iv) is flexible enough so that it can be tailored to meet the objectives of different users. The following discussion elaborates on these points.

A formal statistical approach is more applicable in different contexts compared to those that rely on judgment and historical accumulation of knowledge. In the U.S., where work on leading indicators and business cycles in general dates back at least to the 1920s, one can rely on the accumulated wisdom and build on it. In many countries outside the U.S., however, a tradition of business cycle forecasting is less developed or even non-existent. In the absence of a historical knowledge, a formal model that identifies leading indicators from a large data set could be a useful tool. Even within the U.S., leading indicator literature is primarily concerned with forecasting GDP and to a lesser extent, inflation. For many other variables of interest, there is much less accumulated knowledge on standard leading indicators. Again, a formal statistical approach can be adopted in various contexts; for example, formal methods can be used to develop leading indicators of consumption, the housing market, and any other series of interest. The variables selected by an algorithm can then be fine-tuned by judgment, if there is scope to do so. But even then formal methods provide a good starting point for the analysis in the absence of prior knowledge.

A formal statistical approach is objective, easily replicable, and time-consistent. Variable selection in current leading indicator literature still largely relies on judgment with certain pros and cons attached to it. A level of judgment is unavoidable, in fact desirable in forecasting. Models cannot capture all key aspects and subtleties of an economy, nor can they always be applied mechanically. Yet there is a case for using formal methods to inform judgment, or sometimes supersede it. Algorithms can bring in an additional layer of objectivity to leading indicator literature. They are transparent; work done using an algorithm can easily be replicated, which may not be possible with judgment based approaches. Algorithms are time-consistent; subjective methods need not be and likely often they are not.

Revisions to the methods may seemingly improve performance, but they may also be subject to data snooping biases. There may be conservatism and inertia in human thinking. Human beings may be attached to stories, and overemphasize data that become less informative through time, or put insufficient emphasis on new data series. Recall the arguments by Sinai, and the debate on the role of the payroll data. Formal methods can help to reduce such biases, and provide a way to cross-check the traditional analyses.

In a world of large data sets, formal models are required to read, filter, and analyze numerous available data series. Advances in computer technology have brought a striking increase in the availability of data. With more data, new tools are needed to utilize them efficiently; as data sets improve and expand, analyzing them just by intuition may not be straightforward or feasible. Formal methods will be needed to facilitate analyses. For example, one might need tools to read data from a large data set; to make the necessary filtering such as a seasonal adjustment or conversion to stationary; to produce numerous charts, run numerous models; and to summarize results and present them.

Another advantage is that, by design, the EAL can be tailored to the objective of the forecast. Standard indicators may not have this desirable property. A leading indicator may be best at forecasting economic activity for three-months-ahead, while another one can be better at a year ahead—recall the relative performance of OECD-LI versus CBLI indicators at different horizons. Users of indicators may have different objectives. The predictive power at different horizons, for example, can be important for investment decisions. There is some historical evidence suggesting that share prices make considerable gains about six months before the economic trough (Economist, 2001). Investors may want to look for indicators that signal the trough six months ahead, as opposed to an indicator that has some leading indicator properties with no clear horizon-specific properties, or that is documented to be good at certain horizons, but not at the horizons that meet the objective of the decision maker. A tool that can be tailored to the objective of the forecaster, on the other hand, has this kind of flexibility.

V. CONCLUSION

This study proposes a formal method to select a subset of series from a large data set, with a focus on forecasting recessions. By applying the forecast encompassing principle to forecasts based on individual indicators, the method selects indicators that provide complementary information on recessions. In this context, the selected variables can be considered leading indicators of recessions. In parallel, the paper also adapts a forecast combination method.

The empirical application to the U.S. recession episodes confirms the good out-of-sample performance of the methodology. The forecasts obtained through the combination of series selected by the algorithm improve the accuracy of predicting recessions in the sample compared to the averages of all available forecasts, confirming the usefulness of the encompassing principle as a selection criterion. Furthermore, they are consistently among the

best in a large comparative forecasting exercise including those based on well-known composite leading indicators.

The study also shows that the selected variables are reasonable and consistent with the standard leading indicators followed by many observers of business cycles, as well as with narratives of business cycles presented by prominent business cycle researchers. For shorter horizons, key leading indicators of the U.S. recessions are those related to labor market, housing and consumption, while yield spreads are more prominent in forecasting recessions at a longer horizon. The importance of consumer side information is less stressed in the existing composite indicators, however our findings suggest a need for tailoring the set of leading indicators based on the forecast horizons and, possibly, structural changes in an economy. Further, this method would involve putting less emphasis on some traditional indicators, such as the Real Money Supply, which has not performed well as a leading indicator of recessions in recent cycles.

Our findings suggest that more attention should be given to the methods of data selection in business cycle research. Such formal methods have a number of advantages compared to data selection carried out with standard indicators and expert knowledge. The methods can be tailored to the objectives of the forecaster, so different indicators can be selected for different purposes; additionally, the proposed methods have a flexibility that standard indicators and indices do not have. Formal methods can be applied in different, new contexts where there is limited accumulated knowledge. They are more transparent than approaches in which expert knowledge plays a key role. In a world of increasing availability of data, the volume of potentially useful data series can be too large to be feasibly processed without formal methods. We propose a method to do so, with encouraging first set of results.

Table 1. Forecast Loss at Different Significance Levels

		Significance Level									
	AVE	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
1M ahead											
QPS	0.17	0.13	0.13	0.13	0.14	0.14	0.14	0.13	0.14	0.14	0.14
LPS	0.29	0.23	0.25	0.24	0.25	0.25	0.25	0.24	0.25	0.25	0.26
3M ahead											
QPS	0.19	0.18	0.18	0.17	0.17	0.17	0.16	0.17	0.17	0.17	0.18
LPS	0.33	0.60	0.34	0.31	0.31	0.29	0.28	0.28	0.29	0.29	0.30
6M ahead											
QPS	0.21	0.22	0.20	0.18	0.18	0.18	0.19	0.19	0.18	0.19	0.19
LPS	0.36	0.36	0.35	0.33	0.30	0.31	0.31	0.31	0.31	0.31	0.33
12M ahead											
QPS	0.23	0.24	0.23	0.23	0.22	0.23	0.23	0.23	0.22	0.22	0.22
LPS	0.38	0.41	0.40	0.38	0.36	0.38	0.38	0.37	0.37	0.36	0.37

Notes: QPS and LPS of the EAL forecasts are reported for all horizons. The performance is evaluated between 1975M1 and 2008M12.

Table 2. Performance of the Algorithm Relative to Simple Forecast Averaging

	Significance Level									
	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
1M ahead										
QPS	0.75	0.78	0.79	0.82	0.82	0.81	0.80	0.81	0.83	0.85
LPS	0.77	0.84	0.82	0.85	0.84	0.83	0.82	0.83	0.85	0.88
3M ahead										
QPS	0.96	0.92	0.89	0.89	0.86	0.84	0.86	0.88	0.89	0.91
LPS	1.83	1.04	0.94	0.94	0.88	0.83	0.85	0.87	0.89	0.91
6M ahead										
QPS	1.02	0.96	0.86	0.85	0.84	0.88	0.88	0.86	0.87	0.91
LPS	0.99	0.98	0.90	0.83	0.85	0.87	0.87	0.85	0.87	0.91
12M ahead										
QPS	1.06	1.02	1.02	0.97	1.01	1.01	1.00	0.98	0.97	0.97
LPS	1.06	1.04	0.98	0.94	0.99	0.98	0.97	0.96	0.95	0.95

Notes: For each significance level, the table reports the ratio of QPS and LPS of algorithm forecasts to that of the simple averaging. A ratio of less than one indicates better performance by the competing model relative to AVE.

Table 3a. Best 25 Indicators Based on QPS Ranking(QPS scores ($0 \leq \text{QPS} \leq 2$, 0 if perfectly accurate))

Rank	1-month	QPS	3-month	QPS	6-month	QPS	12-month	QPS
1	LHELX	0.0938	HUSTS1	0.1241	HNR	0.1601	SFYGT10	0.1871
2	HUSTS1	0.0987	HNR	0.1373	HUATZC1	0.1796	SFYGT5	0.1897
3	CFNAI	0.1147	HUATZC1	0.1402	SFYGT5	0.1862	SFYAAAC	0.1975
4	HUATZC1	0.1206	OECD-LI	0.1415	EAL-Mean	0.1876	SFYBAAC	0.2000
5	AWM123	0.1232	HNS	0.1502	SFYGT10	0.1881	SFYMCLE	0.2056
6	OECD-LI	0.1236	AWM123	0.1521	HUSTS1	0.1896	HNR	0.2111
7	A0M005	0.1308	A0M029	0.1551	SFYGT1	0.1904	SFYGT1	0.2116
8	HNR	0.1324	HSBR	0.1591	HNS	0.1931	SFYGM6	0.2145
9	EEAP	0.1333	HSFR	0.1597	SFYGM6	0.2005	PU81	0.2208
10	LHUR	0.1336	EAL-Mean	0.1659	CBLI	0.2008	PUXHS	0.2225
11	PMNO	0.1347	HSWST	0.1698	OECD-LI	0.2019	SFYGM3	0.2227
12	EGP	0.1347	CFNAI	0.1753	AWM123	0.2022	PUXM	0.2237
13	HSFR	0.1349	CBLI	0.1781	A0M029	0.2042	CBLI	0.2256
14	A0M041	0.1350	HNIV	0.1786	SFYBAAC	0.2062	PUNEW	0.2257
15	A0M029	0.1372	PMNO	0.1797	SFYAAAC	0.2080	PUXF	0.2260
16	HNS	0.1390	LHELX	0.1799	SFYGM3	0.2081	PUC	0.2262
17	HSBR	0.1393	HSMW	0.1828	HSBR	0.2088	JWSSQMCH	0.2269
18	EAL-Mean	0.1393	PMP	0.1845	HNIV	0.2137	PWFSA	0.2270
19	PMP	0.1416	A0M005	0.1865	HSWST	0.2146	PWFCSA	0.2280
20	HSWST	0.1457	JDIFFONE	0.1914	PMNO	0.2158	JPCSV	0.2285
21	PMI	0.1496	JDIFFO@0	0.1921	PMP	0.2192	EAL-Mean	0.2298
22	EMF	0.1504	PMEMP	0.1924	HSFR	0.2195	PU882	0.2301
23	UTL11	0.1516	HSSOU	0.1950	PUXHS	0.2195	PUS	0.2315
24	CBCI	0.1530	EEAP	0.1954	PUS	0.2201	JPCXFAE	0.2317
25	EMD	0.1534	AWM122	0.1970	A0M021	0.2203	UTL35	0.2327

Notes: The actual evaluation is between 1978M1 to 2008M12. The reported unencompassed forecasts are chosen by the encompassing test with $\alpha=0.25$ and a recursive window.

Table 3b. Best 25 Indicators Based on LPS RankingLPS scores ($0 \leq \text{LPS} < +\infty$, 0 if perfectly accurate)

Rank	1-month	LPS	3-month	LPS	6-month	LPS	12-month	LPS
1	LHELX	0.1994	HUSTS1	0.2499	EAL-Mean	0.3090	SFYGT10	0.2918
2	AWM123	0.2086	OECD-LI	0.2505	SFYGT5	0.3113	SFYGT5	0.3010
3	CFNAI	0.2122	AWM123	0.2657	SFYGT10	0.3137	SFYAAAC	0.3072
4	A0M005	0.2152	HUATZC1	0.2776	CBLI	0.3207	SFYBAAC	0.3095
5	HUSTS1	0.2191	EAL-Mean	0.2778	HUSTS1	0.3251	SFYMCLE	0.3256
6	PMNO	0.2292	A0M029	0.2859	OECD-LI	0.3254	SFYGM6	0.3610
7	EEAP	0.2314	CBLI	0.2864	HUATZC1	0.3298	SFYGT1	0.3640
8	OECD-LI	0.2352	HSWST	0.2926	SFYGT1	0.3354	PU81	0.3715
9	EGP	0.2354	HSFR	0.2928	HNIV	0.3453	SFYGM3	0.3730
10	A0M041	0.2377	HSBR	0.2931	AWM123	0.3461	JWSSQMCH	0.3771
11	PMP	0.2398	HNIV	0.2982	PMNO	0.3509	UTL35	0.3784
12	LHUR	0.2400	A0M005	0.2999	HSWST	0.3564	HNR	0.3788
13	EAL-Mean	0.2461	CFNAI	0.3012	PMP	0.3565	PUXM	0.3790
14	HUATZC1	0.2478	PMNO	0.3024	UTL35	0.3566	PUXHS	0.3814
15	HSFR	0.2503	LHELX	0.3063	HNS	0.3592	EAL-Mean	0.3817
16	A0M029	0.2518	PMP	0.3096	A0M029	0.3598	PUNEW	0.3824
17	HSBR	0.2530	HNS	0.3116	SFYGM6	0.3598	PWFSA	0.3825
18	CBCI	0.2537	JDIFFONE	0.3146	JDIFFONE	0.3618	PUXF	0.3827
19	JDIFFONE	0.2571	A0M021	0.3191	SFYGM3	0.3622	PWFCSA	0.3847
20	HSWST	0.2576	PMEMP	0.3213	PUXHS	0.3630	JPCSV	0.3888
21	UTL11	0.2604	HSMW	0.3234	SFYAAAC	0.3643	PUS	0.3906
22	PMI	0.2638	EEAP	0.3247	JDIFFO@0	0.3670	CES277	0.3908
23	EMF	0.2649	AWM122	0.3265	PUS	0.3695	CES193	0.3924
24	EMD	0.2665	SFYGT1	0.3279	PUXM	0.3704	PUC	0.3924
25	IPS34	0.2690	JDIFFO@0	0.3280	AWM122	0.3705	PUH	0.3934

Notes: The actual evaluation is between 1978M1 to 2008M12. The reported unencompassed forecasts are chosen by the encompassing test with $\alpha=0.25$ and a recursive window.

Table 4. List of Variables Chosen by the EAL Algorithm

1-month	3-month	6-month	12-month
$\alpha=0.01$	$\alpha=0.30$	$\alpha=0.30$	$\alpha=0.25$
LHELX	HNR	SFYGT10	SFYGT10
HNR	SFYGM3	SFYGT5	SFYGT5
HUSTS1	SFYGT10	SFYGT1	FYMCLE
	SFYGT1	HNR	FYFF
$\alpha=0.05$	HUSTS1	HUATZC1	U0M032
LHELX	A0M029	UTL35	LHU680
AWM123	LHELX	A0M029	SFYGM3
HNR	JDIFFO	LHU680	PUXM
A0M005	A0M005	FYAAAC	FYGM3
HUSTS1	PMEMP	HSWST	PU882
SFYGM6	PMNO	PMEMP	RM1
HNS	PMP	FSPIN	PWCMSA
HNIV	U0M019	FYBAAC	FYGT1
A0M051	LHEL	UTL36	CONDO9
	CONDO9	HSMW	
		JPSHARE	
		FSDXP	

Detailed descriptions of the variables are provided in the Data Appendix.

Table 5. Components of the Conference Board Leading and Coincident Indicators**Leading Indicators**

Average weekly hours, manufacturing	A0M001
Average weekly initial claims for unemployment insurance	A0M005
Manufacturers' new orders, consumer goods and materials	OCD&N96
Vendor performance, slower deliveries diffusion index	U0M032
Manufacturers' new orders, nondefense capital goods	A0M027
Building permits, new private housing units	A0M029
Stock prices, 500 common stocks (S&P)	U0M019
Money supply, M2, Real	A0M106
Interest rate spread, 10-year Treasury bonds less Federal funds rate	U0M129
Index of consumer expectations	AWM123

Coincident Indicators

Employees on nonagricultural payrolls	A0M041
Personal income less transfer payments	A0M051
Index of industrial production	A0M047
Manufacturing and trade sales	X56D

Source: IHS Global Insight.

References

- Aiolfi, A., and A. Timmerman, 2006, “Persistence of Forecasting Performance and Combination Strategies,” *Journal of Econometrics*, 135, 31–53.
- Baumohl, B., 2008, *The Secrets of Economic Indicators: Hidden Clues to Future Economic Trends and Investment Opportunities*, New Jersey: Pearson Prentice Hall.
- Berge, T.J., and Ò. Jordà, 2011, "Evaluating the Classification of Economic Activity into Recessions and Expansions," *American Economic Journal: Macroeconomics*, 3, 246–77.
- Boivin, J., and S. Ng, 2006, “Are More Data Always Better for Factor Analysis?,” *Journal of Econometrics*, 132, 169–194.
- Christiano, L., M. Eichenbaum, and C. Evans, 1999, “Monetary Policy Shocks: What Have We Learned and to What End?” in *Handbook of Macroeconomics*, eds. M. Woodford and J. Taylor, North Holland, 65–148.
- Clements, M.P., and A.B. Galvão, 2009, “Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models,” *Journal of Applied Econometrics*, 24, 1187–1206.
- Clements, M.P., and D. Harvey, 2010, “Forecast Encompassing Tests and Probability Forecasts,” *Journal of Applied Econometrics*, 25, 1028–1062.
- Conference Board, 2001, *Business Cycle Indicators Handbook*, New York.
- Conference Board, 2010, “Real M2 and Its Impact on the Conference Board Leading Economic Index for the United States,” *Monthly Report*, 12, No 3.
- Diebold, F.X., 1998, “The Past, Present, and Future of Macroeconomic Forecasting” *Journal of Economic Perspectives*, 12, 175–192.
- Diebold, F.X. and R. Mariano, 1995, “Comparing Predictive Accuracy,” *Journal of Business and Economic Statistics*, 13, 253–263.
- Diebold, F. and G.D. Rudebusch, 1989, “Scoring the Leading Indicators,” *Journal of Business*, 62, 369–391.
- The Economist, 2001, “Waiting for the Midnight Hour,” April 19th.
- Galbraith, J.W., and S. van Norden, 2011, “Kernel-based Calibration Diagnostics For Recession and Inflation Probability Forecasts,” *International Journal of Forecasting*, 27, 1041–1057.
- Goldman Sachs, 2008, *Understanding U.S. Economic Statistics*, New York.
- Gordon, R.J., 2010, “Okun’s Law, Productivity Innovations, and Conundrums in Business Cycle Dating,” *American Economic Review*, 100, 11–15.

- Hamilton, J., 2011, "Calling Recessions In Real Time," *International Journal of Forecasting*, 27, 1006–1026.
- Harvey, D., S. Leybourne, and P. Newbold, 1998, "Tests for Forecast Encompassing," *Journal of Business & Economic Statistics*, 16, 254–259.
- Kışınbay, T., 2010, "The Use of Encompassing Tests for Forecast Combinations," *Journal of Forecasting*, 29, 715–727.
- Leamer, E.L., 2007, "Housing is the Business Cycle" NBER Working Paper No. 13428
- Leamer, E.L., 2010, *Macroeconomic Patterns and Stories*, Berlin: Springer-Verlag.
- Loungani, P., 2001, "How Accurate Are Private Sector Forecasts? Cross-country Evidence From Consensus Forecasts of Output Growth," *International Journal of Forecasting*, 17, 419–432.
- Marcellino, M., 2006, "Leading Indicators," in G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook Of Economic Forecasting*, Amsterdam: Elsevier, 879–960.
- Marcellino, M., J. H. Stock, and M. W. Watson, 2006, "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series," *Journal of Econometrics*, 135, 499–526.
- NBER (National Bureau of Economic Research), 2008, Announcement of December 2007 Business cycle Peak/Beginning of Last Recession. Available at <http://www.nber.org/cycles/main.html>
- OECD, 2002, "An Update of the OECD Composite Leading Indicators," Unclassified Document Available at www.oecd.org/dataoecd/6/2/2410332.pdf.
- Sinai, A., 2010, "The Business Cycle in a Changing Economy: Conceptualization, Measurement, Dating," *American Economic Review*, 100, 25–29.
- Stock, J. H., and M. W. Watson, 2011, "Dynamic Factor Models," in *Oxford Handbook of Forecasting*, M. P. Clements and D. F. Hendry (eds.), Oxford: Oxford University Press.
- Timmerman A., 2006, "Forecast Combinations," in G. Elliott., C.W.J. Granger, and A. Timmerman (eds), *Handbook of Economic Forecasting*, North-Holland: Elsevier, 135–196

Appendix: Description of Data

This appendix lists the time series used in the empirical analysis. Most series are taken from the IHS Global Insight's database; we present the original mnemonics in this appendix. Some series were produced by the author's calculations, in which case the author's calculations and original Global Insight series mnemonics are summarized in the data description field.

Following the series name is a transformation code, and a short data description. The transformations are (1) level of the series; (2) first difference; (3) second difference (4) logarithm of the series; (5) first difference of the logarithm. All data are seasonally adjusted. The CBCI, CBLI, and CFNAI are from the same source. OECD Leading Indicators are obtained from OECD Database.

Series	Trans.	Description			
<i>(A) INCOME, OUTPUT, CAPACITY UTILIZATION, EXPECTATIONS</i>					
A0M047	5	Industrial production index - total index			
IPS11	5	Industrial production index - products, total			
IPS12	5	Industrial production index - consumer goods			
IPS13	5	Industrial production index - durable consumer goods			
IPS18	5	Industrial production index - nondurable consumer goods			
IPS25	5	Industrial production index - business equipment			
IPS299	5	Industrial production index - final products			
IPS32	5	Industrial production index - materials			
IPS34	5	Industrial production index - durable goods materials			
IPS38	5	Industrial production index - nondurable goods materials			
IPS43	5	Industrial production index - manufacturing			
UTL10	4	Capacity utilization - total index			
UTL11	4	Capacity utilization - manufacturing			
UTL13	4	Capacity utilization - durable manufacturing			
UTL25	4	Capacity utilization - nondurable manufacturing			
UTL35	4	Capacity utilization - mining			
UTL36	4	Capacity utilization - electric and gas utilities			
IPSB51222	5	Industrial production index, durable manufacturing			
CAPGMFN	5	Industrial production index, nondurable manufacturing			
CAPG21	5	Industrial production index, mining			
IPSB51222	5	Industrial production index, residential utilities			
YPR	5	Personal income, total, units: billions/chained 2000\$			
A0M051	5	Personal income less transfer payments, billions/chained 2000\$			
PCHCR	1	Real personal consumption expenditures			
PCHCDR	1	Real personal consumption expenditures - durable goods			
PCHCNR	1	Real personal consumption expenditures - nondurable goods			
PCHCSVR	1	Real personal consumption expenditures - services			
CDMNA	5	Personal consumption expenditures - durable goods - new autos			
X56D	5	Manufacturing and trade sales, units: millions/chained 1996\$			
A0M059	5	Sales, retail stores, units: millions of chained 2000\$			
A0M092	2	Manufacturers' change in unfilled orders, durable goods.			

(C) CONSTRUCTION, INVENTORIES, ORDERS			
HSFR	4	Housing starts:nonfarm(1947-58);total farm&nonfarm	
HSMW	4	Housing starts:midwest	
HSNE	4	Housing starts:northeast	
HSSOU	4	Housing starts:south	
HSWST	4	Housing starts:west	
HUSTS1	4	Housing starts, private including farm - one unit	
HUATZC1	4	New private housing units authorized by building permits, 1 family	
CONDO9	4	Construct.contracts: comm'l & indus.bldgs(mil.sq.ft.floor sp)	
HNIV	4	New 1-family houses for sale at end of month	
HNR	4	New 1-family houses, month's supply @ current sales rate(ratio)	
HNS	4	New 1-family houses sold during month	
HSBR	4	Housing authorized: total new priv housing units	
HMOB	4	Mobile homes: manufacturers' shipments	
PMI	1	Purchasing managers' index	
PMP	1	NAPM production index	
PMNO	1	NAPM new orders index	
PMDEL	1	NAPM vendor deliveries index	
PMNV	1	NAPM inventories index	
PMEMP	1	NAPM employment index	
PMCP	1	NAPM commodity prices index	
OCD&N96	5	New orders (net) - consumer goods & materials, 1996 dollars	
A0M027	5	Value of manufacturers' new orders/nondefense capital goods ind.	
JDIFFONE	1	ISM's new orders-net,-US	
JWSSQMC	1	Change in labor cost per unit of output in manufacturing, 6 months	
A0M031	1	Change in book value of manufacturing and trade inventories	
INVM&T00C	5	Inventories - manufacturing and trade, chained 2000 dollars (BCI)	
JXDIFFO	1	Philadelphia diffusion index, expected, new orders	
JDIFFO	1	Philadelphia diffusion index - new orders	
JXDIFFINV	1	Philadelphia diffusion index, expected, inventories	
JDIFFINV	1	Philadelphia diffusion index - inventories, diffusion index	
A0M029	4	Building permits for new private housing units	

(D) INTEREST RATES, AND ASSET PRICES				
FYFF	2	Interest rate: federal funds (effective)		
FYGM3	2	Interest rate: U.S.treasury bills,sec mkt,3-mo.		
FYGM6	2	Interest rate: U.S.treasury bills,sec mkt,6-mo.		
FYGT1	2	Interest rate: U.S.treasury const maturities,1-yr.		
FYGT5	2	Interest rate: U.S.treasury const maturities,5-yr.		
FYGT10	2	Interest rate: U.S.treasury const maturities,10-yr.		
FYAAAC	2	Bond yield: Moody's AAA corporate		
FYBAAC	2	Bond yield: Moody's BAA corporate		
FYMCLE	2	Effective interest rate: conventional home mtge loans closed		
RM1	2	Prime rate charged by banks		
U0M019	5	S&P's common stock price index: composite (1941-43=10)		
FSPIN	5	S&P's common stock price index: industrials (1941-43=10)		
FSDXP	4	S&P's composite common stock: dividend yield		
FSPXE	4	S&P's composite common stock: price-earnings ratio		
MF5652	5	Stock prices, New York Stock Exchange, common stock index		
JPSHARE	5	Stock and bond yields, Dow Jones Industrial		
EXRCAN	5	Foreign exchange rate: Canada (canadian \$ per U.S.\$)		
EXRJAN	5	Foreign exchange rate: Japan (yen per U.S.\$)		
EXRSW	5	Foreign exchange rate: Switzerland (swiss franc per U.S.\$)		
EXRUK	5	Foreign exchange rate: United Kingdom (cents per pound)		
EXRUS	5	United states;effective exchange rate		
SFYAAAC	1	(FYAAAC-FYFF)		
SFYBAAC	1	(FYBAAC-FYFF)		
SFYGM3	1	(FYGM3-FYFF)		
SFYGM6	1	(FYGM6-FYFF)		
SFYGT1	1	(FYGT1-FYFF)		
SFYGT10	1	(FYGT10-FYFF)		
SFYGT5	1	(FYGT5-FYFF)		
SFYMCLE	1	(FYMCLE-FYFF)		
MNY2@00	5	Money supply, m2, units: billions/chained 2000\$		
A0M101	5	Commercial and industrial loans outstanding		

(E) NOMINAL PRICES, WAGES, MONEY						
FM1	5	Money stock: M1				
FM2	5	Money stock:M2				
FMFBA	5	Monetary base, adj for reserve requirement changes				
FMRRA	5	Depository inst reserves:total,adj for reserve req chgs				
CDTPI	5	Consumer credit outstanding - total				
CES193	5	Average hourly earnings, private nonfarm payrolls - total private				
CES277	5	Average hourly earnings, private nonfarm payrolls - construction				
CES279	5	Average hourly earnings, private nonfarm payrolls - manufacturing				
AHPTTU	5	Average hourly earnings, prod. workers/trade, transport, util.				
AHPPSP	5	Average hourly earnings, prod. workers/private service providing				
AHPFIN	5	Average hourly earnings, prod. workers: financial activities				
PWCM	5	Producer price index: crude materials (82=100)				
PWFC	5	Producer price index: finished consumer goods (82=100)				
PWF	5	Producer price index: finished goods (82=100)				
PWFX	5	Producer price index: finished goods,excl. foods (82=100)				
PWIM	5	Producer price index: inter.mat.supplies & components(82=100)				
PUNEW	5	CPI-u: all items (82-84=100)				
PU81	5	CPI-u: food & beverages (82-84=100)				
PU83	5	CPI-u: apparel & upkeep (82-84=100)				
PU84	5	CPI-u: transportation (82-84=100)				
PU85	5	CPI-u: medical care (82-84=100)				
PU882	5	CPI-u: nondurables (1982-84=100)				
PUC	5	CPI-u: commodities (82-84=100)				
PUCD	5	CPI-u: durables (82-84=100)				
PUH	5	CPI-u: housing (82-84=100)				
PUS	5	CPI-u: services (82-84=100)				
PUXF	5	CPI-u: all items less food (82-84=100)				
PUXHS	5	CPI-u: all items less shelter (82-84=100)				
PUXM	5	CPI-u: all items less medical care (82-84=100)				
JPCXFAE	5	Personal consumption expenditures less food and energy				
JPCD	5	Personal consumption expenditures - durable goods				
JPCN	5	Personal consumption expenditures - nondurable goods				
JPCSV	5	Personal consumption expenditures - service price index (2005=100)				
ALBBCH	1	Net change in bank loans to businesses				
CDTPI%YP	4	Ratio, consumer installment credit outstanding to personal income				
AOM066	5	Outst debt consumer installment credit outst				
WSSMIMCC	5	Wage&sal income in mining,mfg&constr				