Forecasting the probability of US recessions: a Probit and dynamic factor modelling approach

Author(s): Zhihong Chen, Azhar Iqbal and Huiwen Lai

Source: *The Canadian Journal of Economics / Revue canadienne d'Economique*, May / mai 2011, Vol. 44, No. 2 (May / mai 2011), pp. 651-672

Published by: Wiley on behalf of the Canadian Economics Association

Stable URL: https://www.jstor.org/stable/41336378

Forecasting the probability of US recessions: a Probit and dynamic factor modelling approach

Zhihong Chen *School of International Trade and Economics, University of International Business and Economics, Beijing*
Azhar Iqbal *Wells Fargo Securities, Charlotte, NC*
Huiwen Lai *School of Accounting and Finance, Hong Kong Polytechnic University*

Abstract. Quantifying the probability of U.S. recessions has become increasingly important since August 2007. In a data-rich environment, this paper is the first to apply a Probit model to common factors extracted from a large set of explanatory variables to model and forecast recession probability. The results show the advantages of the proposed approach over many existing models. Simulated real-time analysis captures all recessions since 1980. The proposed model also detects a significant jump in the next six-month recession probability based on data up to November 2007, one year before the formal declaration of the recent recession by the NBER. JEL classification: C53, E17

Prédiction de la probabilité des récessions américaines : une approche utilisant la technique probit et une modélisation basée sur les facteurs dynamiques. Quantifier la probabilité des récessions américaines est devenu de plus en plus important depuis août 2007. Dans un environnement où l'information foisonne, ce texte est le premier à appliquer la technique probit à des facteurs communs extraits d'un vaste ensemble de variables explicatives pour modéliser et prédire la probabilité de récession. Les résultats montrent les avantages de l'approche utilisée sur plusieurs des modèles en vogue. Une analyse de simulation en temps réel saisit toutes les récessions depuis 1980. Le modèle proposé détecte aussi un saut significatif dans la probabilité de récession dans les prochains six mois à partir des données disponibles jusqu'à novembre 2007 – un an avant que le NBER n'annonce formellement le commencement de la récente récession.

We are grateful to the anonymous referees for careful reading and constructive comments. The usual disclaimer applies. Zhihong Chen acknowledges Project 211 research fund support from the University of International Business and Economics (Grant 73200001 and the host of Columbia University in spring 2010). Huiwen Lai acknowledges research fund (code 4-ZZ6N) support from the Hong Kong Polytechnic University, as well as support from former colleagues at the Economics Group of Wachovia Bank (now part of Wells Fargo Bank).
Email: afhlai@inet.polyu.edu.hk

Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 44, No. 2
May / mai 2011. Printed in Canada / Imprimé au Canada

0008-4085 / 11 / 651–672 / © Canadian Economics Association

1. Introduction

After roughly doubling in value from 2000 to 2005, home prices in the United States began to fall in the summer of 2006, according to the widely watched Case-Shiller index. Concerns about a U.S. recession surfaced when the downturn in the housing market became apparent and intensified with the credit crunch that started in August 2007. Officially, an economy is not in recession until the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) makes the recession call. The NBER defines a recession as 'a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in production, employment, real income, and other indicators.' Extracting information from abundant economic variables in a data-rich environment to quantify recession probability becomes a topic of increasing interest.

Using a Markov switching approach, Hamilton (1989) modelled turning points in business cycles as regime switches. Birchenhall et al. (1999) generated explanatory variables (composite coincident indicators) and used logistic classification methods to predict business cycle regimes. Estrella and Mishkin (1998) incorporated financial variables directly into a time series Probit model, an approach that has since been used by a number of researchers, including Filardo (1999), Chauvet and Potter (2005), and Wright (2006). Arguing that a typical Probit model specification derived from a small number of explanatory variables limits model performance, Silvia, Bullard, and Lai (2008) applied a stepwise regression procedure to a large number of variables, and the resulting Probit specification improved the forecast performance of the models used by the above mentioned authors.

Although successful to varying degrees, the existing Probit models share serious unresolved deficiencies. First, each model can contain only a very few explanatory variables because of the relatively short time series and rare recessionary periods. The final specification of each existing model typically includes only one transformation of each variable, and thus fails to fully capture the predictive content of the lag dynamics.[1] Useful information embedded in a large number of available but excluded variables and their lag dynamics may not be fully utilized. Second, structural changes and data revisions related to the chosen explanatory variables can significantly erode model performance over time. Many economic variables suffer significant bias, owing to revision and sometimes even redefinition. In such cases, a model that performs very well in simulating real-time forecasts of past recessions may fail woefully when used to forecast future recessions.

1 For example, in Wright (2006) and Silvia et al. (2008), the three-month moving average of each variable is used as the quarterly representation of the variable. In a model of monthly frequency, a model that includes three lags of the variable separately will have better forecast ability than one that includes the moving average of the variable because the former captures more flexibility.

To address the above mentioned inadequacies, we construct a Probit model using common factors estimated by the dynamic factor modelling method (hereafter, Probit-dynamic factor model; Probit-DFM) to forecast recession probability. In a data-rich environment, many more explanatory variables can be used in the factor model framework. The fundamental assumption of this approach is that each economic variable can be decomposed into a common component (made up of a number of common factors) and an idiosyncratic component. The former is driven by only a few dynamic factors that underlie the whole economy. It has been shown that, with fairly reasonable assumptions, principal component analysis (PCA) can be used to estimate common factors consistently (Stock and Watson 2002a; Bai and Ng 2006). Factors estimated using PCA have been proven successful in forecasting individual economic variables, including employment, wage, industrial production, retail sales, and inflation measures, in a linear framework (Stock and Watson 1999; Stock and Watson 2002a; Bernanke and Boivin 2003).[2] Chauvet (1999) characterized business cycle dynamics using a factor structure and regime switching framework. Bai and Ng (2008) provided an analysis for non-linear estimation and established the conditions under which the estimated factors can be treated as though they were observable. We use this approach to develop our Probit framework to forecast recession. To the best of our knowledge, we are the first to do so to forecast U.S. recessions.

First, following Bai and Ng (2002) and Stock and Watson (2002a), we perform PCA to extract several common factors from 141 economic variables. We then incorporate these factors into the Probit framework. The resulting specification is our Probit-DFM. We compare the performance of the proposed approach with that of many existing models widely used by economists. The Probit-DFM is found to outperform the existing models in terms of both in-sample and out-of-sample statistical criteria. Simulated real-time analysis shows that our model captures all the five recessions since 1980. In addition, it gives fewer false alerts than the Silvia, Bullard, and Lai (2008) specification. Interestingly, our model detects a significant jump in the next six-month recession probability based on data up to November 2007, which is consistent with the NBER's determination that the recent recession started in December 2007. This finding is striking, because during the one-year period from December 2007 to the time when the NBER made the formal call, economists debated the existence and timing of a possible recession. Our model's forecast was right on target. (To stress the truly real time forecast of this work, it is interesting to recall that we had presented the early draft of this paper, including the same figure 1, on different occasions long before NBER's formal declaration. The presentations include the following: Wachovia Bank in January 2008, University of International Business and Economics in February 2008, Hong Kong Polytechnic University in June

2 Since both recessions and the common factors identified through PCA reflect overall economic activity, using these factors to forecast recessions may be more relevant than using a single variable.

2008, and Year 2008 Annual Meeting of Econometric Society in Singapore in July 2008.)

The superior performance of the proposed approach may be due to its ability to address the inadequacies we outlined earlier. First, the Probit-DFM has the potential to capture the fundamental drivers of the economy shared by all economic variables. Second, the lagged effects of the dynamic factors may already be captured in the construction of the common component in the static representation of the dynamic factor model.[3] As a result, the factors estimated by PCA automatically capture the lag dynamics of the underlying macroeconomic factors in forecasting. The above mentioned two advantages of our Probit-DFM resolve the problems identified in the existing literature: too few explanatory variables and too restrictive specifications. Finally, although structural change can cause instability in an individual economic variable, the instabilities embedded in a large number of economic variables tend to cancel out each other. That is, it is much less likely that the whole economy will experience a great structural change within the period covered.

The rest of the paper is organized as follows. Section 2 introduces the Probit-DFM and discusses its advantages. Section 3 describes the data and the implementation strategy. Section 4 presents the empirical results, and section 5 gives a caveat. Section 6 concludes the paper.

2. Probit-DFM specification

2.1. Modelling recession probability

We start with the following Probit model:

$$y^*_{t+h|t} = \beta' x_t + \varepsilon_t \tag{1}$$

$$y_{t+h|t} = 1(\text{if} y^*_{t+h|t} > 0), \tag{2}$$

where $y^*_{t+h|t}$ is an unobserved variable that determines, at time t, if a recession will occur within the next h periods; $y_{t+h|t}$ is an observable dummy variable determined by $y^*_{t+h|t}$; x_t is a vector of the independent variables at time t; β is a vector representing the coefficients, including an intercept; and ε_t is a normally distributed error term.

Given the historical data of the occurrence of recessions and the chosen explanatory variables x_t, the task is to estimate a parameter vector β and forecast

3 Stock and Watson (2002b) found that the lags of estimated common factors do not contribute to forecasting. We believe this provides empirical validation of the static representation of dynamic factor models. While estimated factors reflect the common component in static form, the lags of the underlying dynamic factors have already been fully captured.

the probability that a recession will fall in the next h periods:

$$P(y_{t+h|t} = 1) = P(y^*_{t+h|t} > 0) = \Phi(\beta' x_t), \tag{3}$$

where function $\Phi(.)$ is the cumulative density function of the standard normal distribution.

2.2. Probit-dynamic factor model (Probit-DFM)

To capture the predictability embedded in a large number of variables, we consider a dynamic factor model (DFM), in which the co-movements among these economic variables are treated as arising from a small number of unobserved sources, or dynamic factors. The estimates of those factors are incorporated into the Probit framework to model the probability of a recession. We denote this model as the Probit-DFM.

Let the probability that a recession will fall in the next h periods, $P(y_{t+h|t} = 1)$, be a scalar time series variable to be forecasted, and let X_t be an N-dimensional vector of the economic time series, where N is a large number. Then, the static representation of the dynamic factor model takes the form

$$X_t = \Lambda F_t + e_t. \tag{4}$$

We can rewrite (1) to (3) as the Probit-DFM:

$$y^*_{t+h|t} = \beta'_F F_t + \varepsilon_t \tag{5}$$

$$y_{t+h|t} = 1(\text{if} y^*_{t+h|t} > 0) \tag{6}$$

$$P(y_{t+h|t} = 1) = P(y^*_{t+h|t} > 0) = \Phi(\beta'_F F_t), \tag{7}$$

where e_t is an $N \times 1$ vector of idiosyncratic disturbances, Λ is an $N \times r$ matrix of factor loadings and F_t is an $r \times 1$ vector of the common factors underlying X_t. Data are available for $\{y_{t+h|t}, X_t\}_{t=1}^T$, and our goal is to forecast the recession probability, $P(y_{T+h|T} = 1)$, in the next h periods.

Stock and Watson (2002a) and Bai and Ng (2006) showed that under general conditions, the principal components of X_t are consistent estimators of the true latent factors,F_t. The feasible forecast of an economic variable of interest,z_{t+h}, which is a function of the estimated factors, converges to the infeasible forecast that would be obtained if the factors and coefficients were known. Empirical work in macroeconomics indicates that these results hold when the temporal instability of a dynamic factor model is small.

Bai and Ng (2008) showed that asymptotically the estimated factors have no effect on the estimates of the Probit slope coefficient in equation (7). Therefore, applying the DFM to X_t will give consistent estimations of F and β_F in equation (5). This establishes the theoretical foundation of our Probit application, as reflected in equation (7).

Ideally, a Probit recession model should incorporate the dynamics of the most relevant variables: that is, include multiple lags of each important explanatory variable as predictors, with the number of lags selected by a statistical principle such as the Bayesian information criterion (BIC). Because of the relatively short time series and rare recession events, however, including too many explanatory variables and their lags may cause over-fitting.

According to Stock and Watson (2002a), factors are extracted based on a static representation of the DFM, such as equation (4). The several extracted factors, denoted as $V_{1t}, V_{2t}, \ldots, V_{rt}$, may already incorporate the lags of the several underlying dynamic factors.[4] Thus, a Probit forecasting model that focuses on the current month values of these factors suffices to capture the dynamics and interactions of the extracted factors and underlying economic variables. Stock and Watson (2002b) confirmed empirically that the lags of extracted factors do not contribute to the forecasting of real U.S. macroeconomic variables.[5] Thus, incorporating the current month values of the several extracted factors into the Probit-DFM resolves the above mentioned problems (too few explanatory variables and too restrictive specifications) in the existing literature.

3. Data and implementation strategy

3.1. Data

We obtain most of the data for our empirical study from the IHS Global Insight database. The sampling period is from January 1964 to November 2007. The dependent variable is constructed using the standard NBER dates. We define a

4 According to Stock and Watson (2002a), model (4) is a static representation of a dynamic factor model. Let X_t be the n-dimensional multiple time series of predictors, $t = 1, \ldots, T$. X_t has a dynamic factor model representation with $\bar{r}$ common dynamic factors f_t:
$X_{it} = \lambda_i(L)f_t + e_{it}$ (*) for $i = 1, \ldots, N$, where e_{it} is the idiosyncratic disturbance. $\lambda_i(L)$ is the lag polynomial of L
with a finite order of q; thus,
$$\lambda_i(L) = \sum_{j=0}^{q} \lambda_{ij} L^j.$$
The finite lag assumption allows us to rewrite (*) as
$X_t = \Lambda F_t + e_t$,
where $F_t = (f_t', \ldots f_{t-q}')'$ is an $r \times 1$ vector, where $r \leq (q+1)\bar{r}$.
$e_t = (e_{1t}, e_{2t}, \ldots, e_{Nt})'$. The ith row of Λ is $(\lambda_{i0}, \lambda_{i1}, \ldots, \lambda_{iq})$. Thus, the PCA estimation of F_t, $V_{1t}, V_{2t}, \ldots, V_{rt}$, may include both the current level and lags of the few dynamic factors, f_t.

5 Stock and Watson (2002b) note: 'In most cases the performance of the simpler DI forecasts, which exclude lags of F_t and y_t, is comparable to or even better than that of the DI-AR, Lag forecasts. This is rather surprising, because it implies that essentially all the predictable dynamics of these series are accounted for by the estimated factors' (151).

dummy variable $y_{t+6|t}$, which is equal to one if a recession occurs in any of the next six months after month t, and zero otherwise. We define another dummy variable y_t, which is equal to one if a recession occurs at month t, and zero otherwise.

Following Stock and Watson (2002b), we use monthly macroeconomic time series for factor extraction.[6] The series include 15 main categories: real output; real income; employment and hours; real retail, manufacturing, and trade sales; consumption; housing starts and sales; real inventories and inventory-sale ratios; orders and unfilled orders; stock prices; exchange rates; interest rates; money and credit quantity aggregates; prices indexes; average hourly earnings; and miscellaneous. We choose variables without missing values between January 1964 and November 2007. In total, we identify 141 variables as potential predictors of future recession probability. All 141 variables are subject to three possible transformations following the standard practice: taking logarithms, taking the first difference, and screening for outliers. After these transformations, all series are standardized to have a sample mean of zero and a sample variance of one.

Owing to space limitations, the full list of all variables and the related transformations is available upon request. It is quite similar to the appendix in Stock and Watson (2002b).

3.2. Implementation strategy

We distinguish two modelling approaches in the literature: restricted and unrestricted. The restricted model includes only the current value of explanatory variables using monthly data (Filardo 1999; Marcellino 2006; Estrella and Mishkin 1998) or the three-month moving average of explanatory variables using quarterly data (Wright, 2006; Silvia, Bullard, and Lai 2008). The former approach restricts all of the coefficients related to the lagged explanatory variables to be zero, whereas the latter assumes that the current value of an explanatory variable has the same predictive power as its first and second lags and that its longer lags do not have any predictive power. The unrestricted model puts no restriction on the coefficients related to the current value or lags of an explanatory variable. In this case, we expect a better fit because we have greater freedom to choose parameters during optimization. For the five extant models examined, for comparison, we implement two restricted specifications (one-month restricted, three-month-average restricted) and one less restricted specification (denoted the unrestricted model). The unrestricted model allows the current month and up to five lags of the specified variables in the model specification, with the final specification determined by the BIC. These models are then compared with the Probit-DFM.

To evaluate the performance of each model, we consider both the in-sample and the out-of-sample criteria. Following Estrella (1998), we use the pseudo R^2

6 We find most variables in Stock and Watson (2002b) in the IHS Global Insight database. For the few variables not available from that database, we use closely related variables and do the transformation.

as a simple measure of goodness of fit for the non-linear Probit model, which is defined as

$$PseudoR^2 = 1 - \left(\frac{\log L}{\log L_0}\right)^{-\frac{2}{T}\log L_0}, \tag{8}$$

where T is the number of periods, log L is the value of the log likelihood function, and log L_0 is the value of the log likelihood function when all regression coefficients, except the intercept term, are zero. We also use other measures such as the McFadden R^2 or BIC to decide the specification.

Given the relatively short time series and limited number of recessions, overfitting is a big concern. That is, a model chosen by a large pseudo R^2 may poorly forecast a future recession. To address this problem, we consider an out-of-sample criterion: the root mean square error (RMSE). Assume that data are available between time $t = 1$ and $t = T$ for model building, where T represents the most recent month in which the values of all variables are available. For an integer $m \in (1, M)$, we choose data between time $t = 1$ and $t = T - m$ to build a Probit model and apply it to month $T - m + 1$ to forecast the probability of a recession occurring in any of the next h months. With M predictions, we can then calculate the RMSE. If $\hat{p}_t$is the fitted probability of a recession occurring between month $t + 1$ and month $t + h$, then the RMSE is calculated as

$$RMSE_h = \sqrt{\frac{1}{M}\sum_{t=1}^{M}(\hat{p}_t - y_{t+h|t})^2}. \tag{9}$$

The value of the RMSE is used to compare the out-of-sample performance of different model specifications.

We estimate each model specification using both the fixed sample and the varying sample approach. The fixed sample approach adopts the standard in-sample regression using the whole sample between January 1964 and November 2007. If a specification is one-month restricted or three-month-average restricted, then the specification is already known. If a specification is unrestricted or is the Probit-DFM, then the BIC is used to choose the exact final specification.

In the varying sample approach, multiple forecasts are generated for each sampled ending month and a simulated real-time RMSE is calculated by recursive parameter estimation and model selection. Except for the Probit-DFM, the first sample window is set to be between January 1964 and June 1979. The dependent variable,$y_{t+6|t}$, and the predictors of the model specification are estimated.[7] Then, the model is applied to the next month predictor values to generate a forecast probability. We recursively move the sample window one month forward to generate a series of forecast probabilities. The last recursive model is based on the

7 Similar to the fixed sample approach, the final specification for the unrestricted model is determined by the BIC with each sub-sample.

sample between January 1964 and October 2007, and the last forecast probability is evaluated using the November 2007 values of the predictors. These simulated real-time forecasts are graphed against the actual NBER recession dummy y_t for visual examination. A simulated real-time RMSE is calculated based on equation (9).

For the Probit-DFM, our calculation of a simulated real-time RMSE is slightly different. Our first sample window is set from January 1964 to July 1979. PCA is performed to extract eight time series factors, $V_{1t}, V_{2t}, \ldots, V_{8t}$.[8] The BIC is applied to the dependent variable and the extracted factors between January 1964 and June 1979 to build a model. Then, the July 1979 values of the extracted factors are applied to this model to generate a Probit-DFM-based forecast probability. We recursively move this sample window one month forward to generate a series of forecast probabilities. A similar graph and the RMSE are presented.

4. Empirical results

4.1. Results of the Probit-DFM

In our Probit-DFM, several factors extracted from a large number of economic variables (X_t in equation (4)) are used in the Probit modelling framework. The starting point is that a few extracted factors account for a significant amount of data variation in the underlying variables. Following Ludvigson and Ng (2009),[9] we provide the summary statistics for eight estimated factors in table 1a, where R^2 in the ith row represents the percentage of variation in the data explained by factors 1 to i. These eight factors explain 56% of the data variation among the 141 variables included in X_t.

Table 1a also reports the first-order autoregressive coefficient for each factor. There is a significant heterogeneity among these factors, with AR(1) coefficients that range from −0.08 to 0.94. In terms of forecasting, however, we may not need to incorporate the lags of each factor into the Probit model. By construction, both the current month value and the relevant lags of an underlying dynamic factor are presented as extracted factors (Stock and Watson 2002a).

Table 1b and figure 1 present the estimation results for the Probit-DFM. The in-sample regression selects six out of the eight factors as predictors based on the BIC. The first, second, and seventh factors are significant, while the fifth and eighth factors are marginally significant in predicting the next six-month recession probability. The model has a pseudo R^2 as high as 0.62 and a simulated real-time RMSE as low as 0.24. Figure 1 shows that the model captures all four

8 In the next section, we discuss the determination of the number of extracted factors.

9 Ludvigson and Ng (2009) use the information criterion developed in Bai and Ng (2002) and determine r as eight. We attempt to determine r using the same criterion with our data, which results in an r bigger than eight. However, we decide that eight is an appropriate choice, since with an R^2 of 56%, these factors have already explained more data variation than those in Ludvigson and Ng (2009) (50%).

TABLE 1a
Summary statistics for factors

Factor number (i)	AR1	R2
1	0.82	0.18
2	0.94	0.28
3	−0.06	0.34
4	0.47	0.4
5	0.47	0.45
6	0.89	0.49
7	0.43	0.52
8	0.06	0.56

NOTES: The eight factors are estimated by PCA using a panel of 141 variables between January 1964 and November 2007. These variables are chosen from Stock and Watson (2002b) and transformed and standardized accordingly before estimation. The second column reports the first-order autocorrelation coefficient for each factor. R^2 is calculated as the portion of total data variation explained by factors 1 to *i*.

TABLE 1b
Probit estimation with extracted factors

Forecasting NBER Recessions within the next six months			
Coefficients	Estimate		t-statistic
Intercept	−1.65		−8.86
F1	−3.35		−8.44
F2	−3.41		−6.49
F5	1.12		1.89
F6	0.59		0.85
F7	−2.26		−2.83
F8	−1.13		−1.23
In-sample criteria			
Pseudo R^2		0.62	
McFadden R^2		0.63	
BIC		−118.3	
Out-of-Sample Criterion:			
RMSE		0.24	

NOTES: This table reports the in-sample estimation of a Probit model using factors extracted by PCA. One hundred and forty-one variables from a sample taken from between January 1964 and November 2007 are used, out of which eight factors are extracted (see Table 1a). The BIC is applied to these factors as candidate predictors. The final model has six factors.

recessions between 1980 and 2001 very well and generates only one false alert about one year before the 1991 recession.

The figure also shows that the Probit-DFM captures the 1980, 1981, and 1991 recessions exceptionally well. An interesting pattern is observed for the 2001 recession. In November 2001, the NBER's Business Cycle Dating Committee determined that a recession had started in March 2001. The forecast probabilities

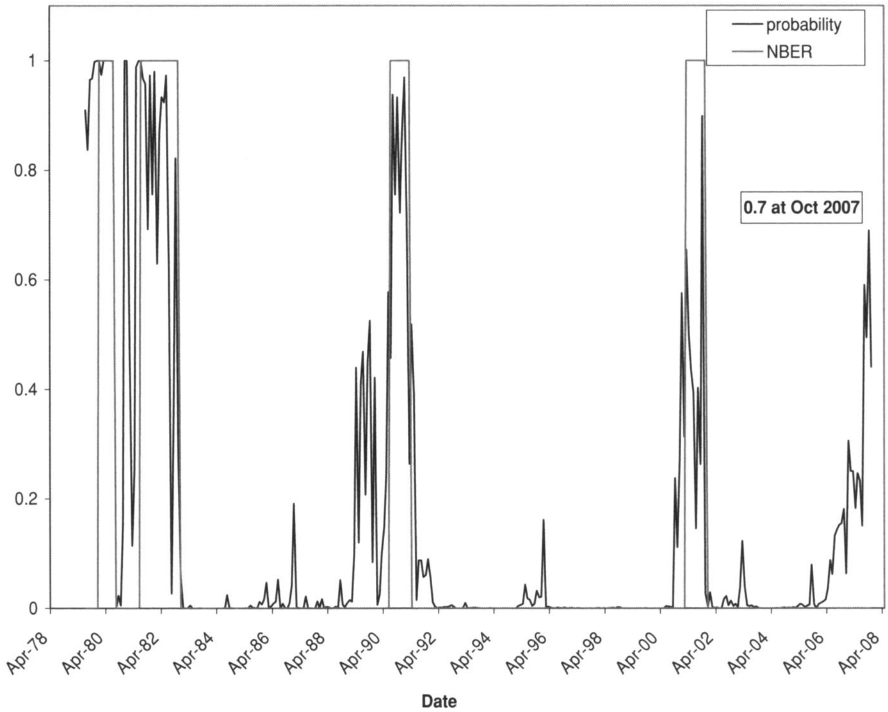

FIGURE 1 Recession probability of the dynamic factor model (RMSE = 0.24)

generated by our model for several months around that time are somewhat elevated, with those of January, March, April, and October of 2001 exceeding 50%. In retrospect, economic activity during this period was mixed: GDP declined in the second quarter but rose in the third quarter and declined again in the fourth quarter, largely because of the impact of the 11 September attacks. The committee suggested that it was possible that the decline in economic activity would have been too mild to qualify as a recession without the unexpected impact of the 11 September attacks (see Hall et al. 2001). Indeed, no professional forecasters forecast that recession correctly (Stock and Watson 2003). With this background in mind, the higher than normal probabilities generated by the Probit-DFM, although not as high as those for the previous three recessions, would have stood out as strong signals of the weak economy and potential recession that followed.

Our Probit-DFM also forecasts the most recent economic recession extremely well. The next six-month recession probability forecast based on the model picks up slightly after July 2006, but does not reach a significantly high threshold until August 2007, the month that the credit crunch started. The probabilities based on data up to each month between July 2007 and November 2007 are p(2007M7) = 0.15, p (2007M8) = 0.59, p(2007M9) = 0.50, p(2007M10) = 0.7, and p(2007M11) = 0.45. These elevated forward-looking probabilities are consistent with the subsequent onset of this recession in December 2007. After one full year, the committee made the formal declaration of recession. During that period, economists had been debating the existence and timing of a potential recession. The elevated forward-looking probabilities from our model, however, perfectly predict the onset of the recent recession on a real-time basis with the data available up to November 2007.

4.2. Comparison of the Probit-DFM with other recession forecast models

We compare the in-sample and simulated out-of-sample regression results of our Probit-DFM and five existing recession forecast models. Those five models and the explanatory variables included in each of them are as follows:

1) Yield spread model: includes the yield spread between three-month treasury bills and a 10-year government bond.
2) Estrella and Mishkin (1998) model: includes the yield spread in (1) plus the month-over-month percentage change in the S&P 500 index.
3) Wright (2006) model: includes the yield spread in (1) plus the fed funds rate.
4) LEI model: includes the month-over-month change in the Conference Board's Leading Economic Index (LEI).
5) Silvia, Bullard, and Lai (2008) model: includes the month-over-month change in the LEI and Chicago PMI employment index, and the month-over-month percentage change in the S&P 500 index.

Table 2 presents the results of these five non-Probit-DFM models. Only the results of the unrestricted specifications are reported, because the unrestricted specifications always perform at least as well as the restricted ones.[10] The in-sample regressions are based on a sample between 1964M1 and 2007M11. Three in-sample statistics are presented: the pseudo R^2, MacFadden R^2, and BIC.[11] Each model's out-of-sample statistic, the simulated real-time RMSE, is presented in the last row.[12] Corresponding to each RMSE is a graph of $\hat{p}_t$versus y_t. A small RMSE reflects a good fit between $\hat{p}_t$ and $y_{t+6|t}$, which reflects a higher $\hat{p}_t$one to six months before a $y_t$value changes from zero to one. This relationship can be observed in the graph. Figures 2(a)–2(e) correspond respectively to the models in columns 1–5 in table 2.

10 The restricted specifications are available upon request.
11 For each non-Probit-DFM, two restricted specifications are fixed, but the unrestricted specification is determined by the Schwarz-Bayesian criterion (SBC). Notice also that the yield spread and LEI models include transformations of one variable. In contrast, the Estrella and Mishkin (1998), Wright (2006), and Silvia et al. (2008) models include transformations of two or three different variables. We use two or three sub-columns to present these models for convenience.
12 For each model, we present the in-sample and the out-of-sample results in the same table for convenience.

TABLE 2
Performance of the five recession forecast models

	(1)	(2)		(3)		(4)	(5)		
	Yield Spread	Yield Spread	S&P500	Yield Spread	Fed Funds	LEI	LEI	Chicago PMI	S&P500
Intercept	−0.15	−0.06		−2.7		−0.85		5.89	
	(−0.85)	(−0.35)		(−5.07)		(−4.61)		(4.19)	
Current Month						−1.23	−2.28		
						(−4.58)	(−7.20)		
One Month Lag			−10.69			−1.56	−2.69		
			(−3.73)			(−5.49)	(−7.0)		
Two Month Lag			−9.83		0.35	−1.17	−2.12		
			(−4.24)		(4.78)	(−5.46)	(−6.51)		
Three Month Lag			−10.55			−0.93	−1.3	−0.14	
			(−4.49)			(−4.39)	(−3.66)	(−4.75)	
Four Month Lag			−9.3			−1.06	−1.16		
			(−4.21)			(−4.58)	(−3.36)		
Five Month Lag	0.78	0.82	−10.47	0.74		−0.95			
	(5.48)	(5.7)	(−3.57)	(4.31)		(−3.65)			
In-Sample Statistics									
Pseudo R^2	0.3	0.48		0.53		0.57		0.72	
MacFadden R^2	0.3	0.49		0.54		0.58		0.73	
BIC	−186.19	−154.54		−127.71		−130.29		−90.95	
Out-of-Sample Statistic									
RMSE	0.33	0.31		0.28		0.29		0.25	

NOTES: Column (1) is the yield spread model, which includes the yield spread between three-month Treasury bills and a 10-year government bond; Column (2) is the Estrella and Mishkin (1998) model, which includes the yield spread in (1) plus the month-over-month percentage change in the S&P 500 index; Column (3) is the Wright (2006) model, which includes the yield spread in (1) plus the fed funds rate; Column (4) is the LEI model, which includes the month-over-month change in the Conference Board's Leading Economic Index (LEI); Column (5) is the Silvia et al. (2008) model, which includes the month-over-month change in the LEI and Chicago PMI employment index, and month-over-month percentage change in the S&P 500 index. The in-sample regression results are based on the sample taken from between January 1964 and November 2007. The out-of-sample statistic, the RMSE, in the last row, is calculated using the forecast errors from recursive model estimation and forecasting by varying sample windows. The t-statistics are in parenthesis below the parameters.

Column 1 in table 2 reports the results of the yield spread model. The fifth lag of the yield spread is chosen as a predictor. This is consistent with the notion that a time lag exists between a yield curve inversion and a recession (see Labonte 2008). Regarding the out-of-sample performance, the simulated real-time RMSE is 0.33. Figure 2(a) shows that the yield spread model successfully forecasts the two recessions between 1980 and 1981 – a period of high inflation and high interest rates paradigm. However, it performs poorly in forecasting the 1991 and 2001 recessions.

Columns 2 and 3 in table 2 and figures 2b and 2c present respectively the results of the two extended yield spread models: the Estrella and Mishkin (1998) model using the yield spread plus the S&P 500 index and the Wright (2006) model using the yield spread plus the fed funds rate. The five different lags of the S&P 500 index contribute significantly to the predictive power of the Estrella and Mishkin (1998) model: the pseudo R^2 increases from 0.3 in column 1 to 0.48 in column 2. Figure 2b shows that this model performs reasonably well, forecasting well the two recessions in the early 1980s and the 2001 recession.[13] In contrast, the Wright (2006) model performs relatively poorly in forecasting the two most recent recessions.

That the yield spread forecasts well the high-interest and high-inflation regime of the recessions in the early 1980s but fails to forecast two more recent recessions indicates that the structural change of economic variables can erode the forecast ability of the yield spread model.

The Leading Economic Index (LEI) is widely used as a natural starting point for recession forecasting (Stock and Watson 1989; Filardo 1999; Marcellino 2006). Using quarterly frequency data and stepwise regression, Silvia, Bullard, and Lai (2008) identified the LEI as having the greatest predictive power among single explanatory variable models in forecasting recession in the next two quarters.[14] Column 4 in table 2 presents a specification closely related to the best performing model identified by Silvia et al. Using monthly data, our specification selects the current month value and five lags of the LEI variable as predictors. It has a pseudo R^2 as high as 0.57 and outperforms all three yield spread-related models. Our results further confirm that the LEI does indeed serve as a good leading indicator of overall economic activity, as it was designed to do.

The top model in Silvia, Bullard, and Lai (2008) also includes two variables related to the Chicago PMI employment index and the S&P 500 index, respectively. Column 5 in table 2 shows that using these variables as indicators leads to significant improvement in the performance of this model over that of the corresponding single variable (LEI) model – the pseudo R^2 increases to 0.72.[15]

13 Despite a clear false alert in 1987 due to the stock market crash.

14 The quarter-to-quarter change in the LEI was used as the predictor, where the three-month moving average of the LEI was treated as the quarterly LEI.

15 Note that the lags only of the LEI and Chicago PMI employment index (not the S&P 500 index) are included in the model using the pre-specified sample range between January 1964 and

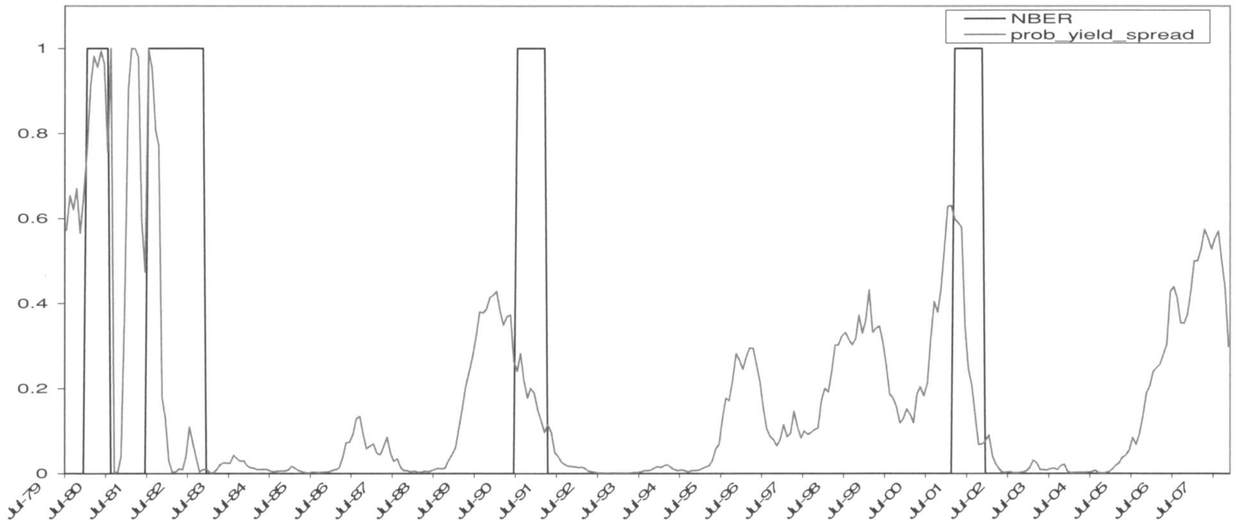

FIGURE 2(a) Recession probability of the yield spread model (RMSE = 0.33)

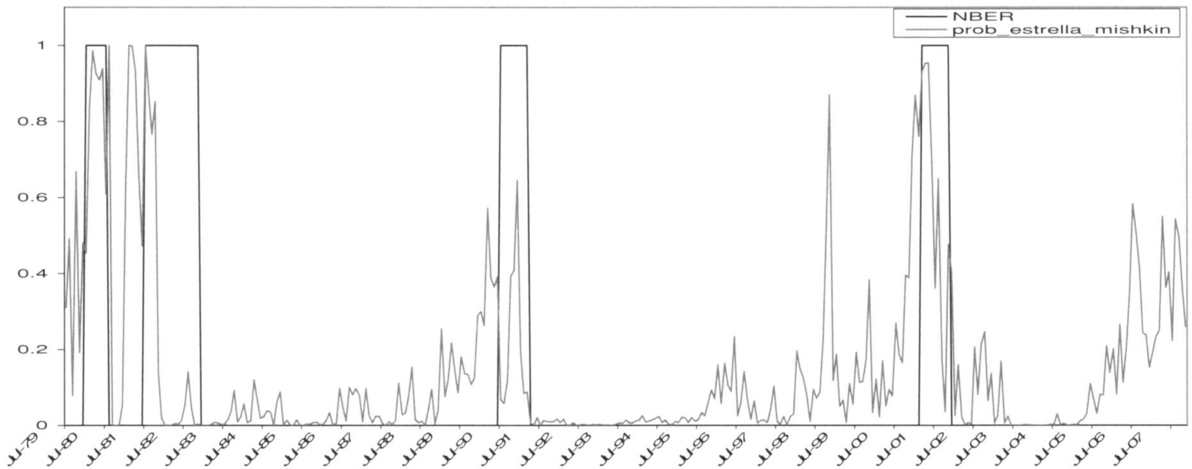

FIGURE 2(b) Recession probability of the Estrella and Mishkin (1998) model (RMSE = 0.31)

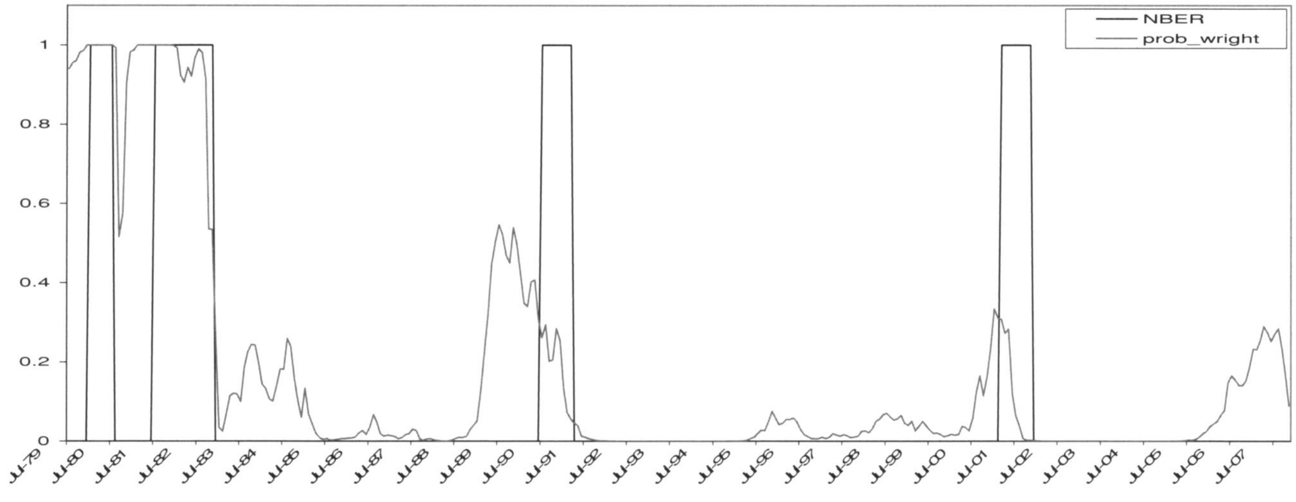

FIGURE 2(c) Recession probability of the Wright (2006) model (RMSE = 0.28)

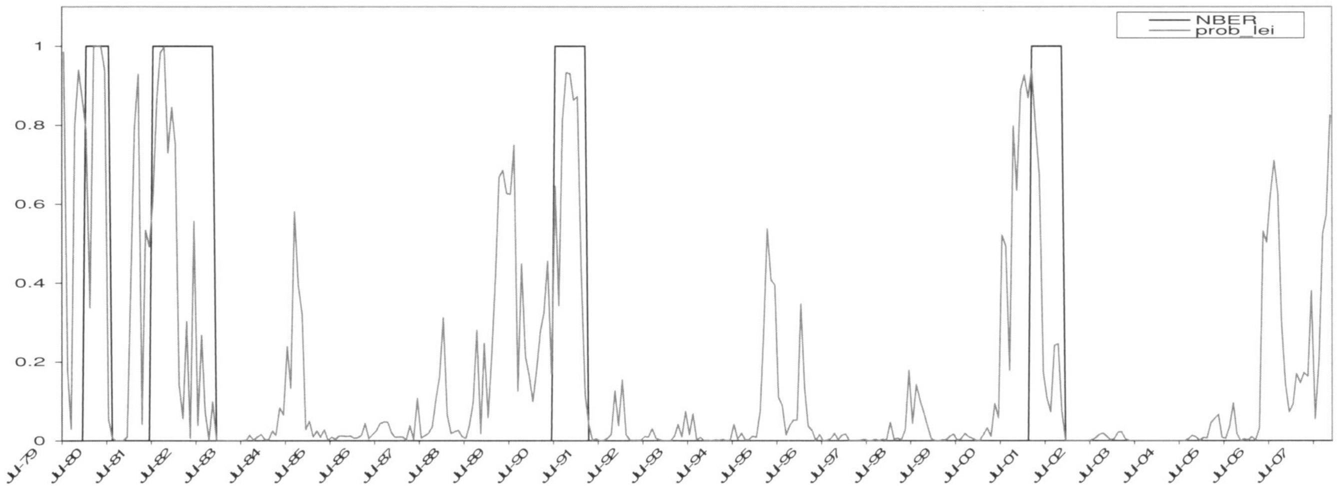

FIGURE 2(d) Recession probability of the LEI model (RMSE = 0.29)

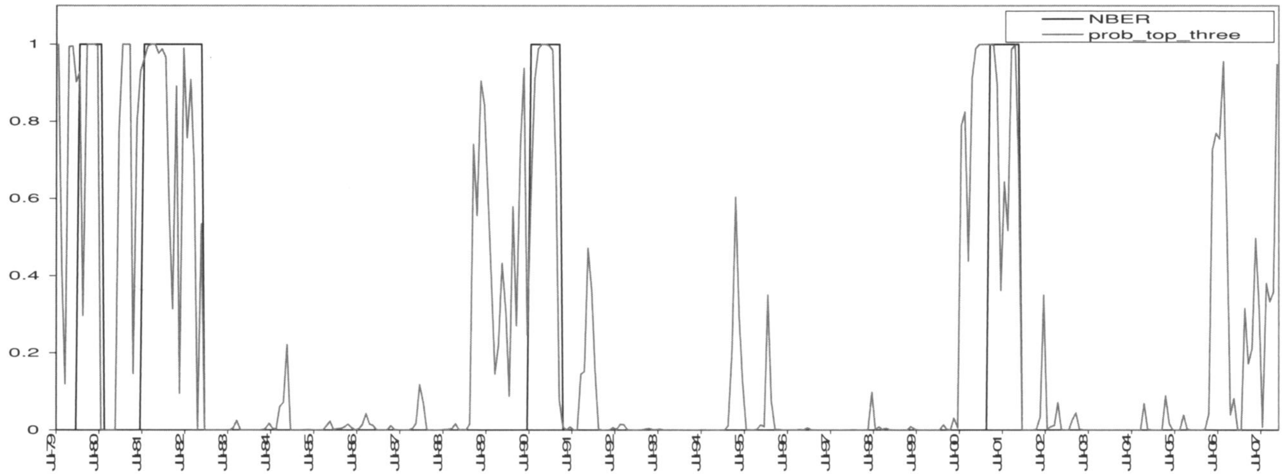

FIGURE 2(e) Recession probability of the Silvia, Bullard, and Lai (2008) model (RMSE = 0.25)

The LEI and Silvia et al. models also demonstrate good out-of-sample performance. Figures 2d and 2e show that they fit the data much better than the three yield spread-related models. The simulated real-time RMSE for the top model is 0.25, the lowest among all models we have examined so far.

Despite its good overall performance, the top model related to the specification of Silvia, Bullard, and Lai (2008) generates several false alerts.[16] A caveat is that the model was chosen using a backward-looking data mining process. We cannot rule out the possibility that data peculiarities, not economic driving forces, may underlie the coincidence between the onset of historical recessions and the model's forecasts. In addition, even if these variable combinations do explain past recessions well, potential structural change in any of these variables could cause poor forward-looking forecast ability.

Now, we compare the performance of our Probit-DFM with that of the foregoing models. The three in-sample statistics of the Probit-DFM perform better than four of the five models presented, the exception being the Silvia, Bullard, and Lai (2008) model. The pseudo R^2 of the probit-DFM is 0.62 versus 0.72 for the unrestricted specification of Silvia et al.

However, the Probit-DFM stands out as the best performing model based on the out-of-sample criterion. Its simulated real-time RMSE is 0.24, the lowest among all models examined. Figure 1 shows that the model captures all four recessions between 1980 and 2001 very well. Except for a false alert about one year before the 1991 recession, it does not generate other false alerts as does the Silvia, Bullard, and Lai (2008) specification. This indicates that the better in-sample statistics of the Silvia et al. specification could be due to overfitting. Our model is also more robust in the recursive modelling process. That is, if both models were applied to forecast the next six-month recession probabilities from late 1979, then the Probit-DFM would perform better.

5. Caveat

Despite the significant improvements of our Probit-DFM over existing models, we offer one important caveat. In evaluating model performance with in-sample and out-of-sample statistical criteria, we use final, revised data. In real-time forecasting, however, the unrevised real-time values of the model predictors are usually used as input.

Using a combined DFM and vector autoregressive model framework, Bernanke and Boivin (2003) found that the impact of data revision on forecasting is insignificant. This is encouraging. In addition, the inclusion of a large number of variables may lessen the impact of data revision. Although in a new

November 2007. With a different sample range in the recursive modelling, however, the lags related to all three variables have been included.

16 False alerts are generated in March 1989, November 1991, April 1995, and May 2006.

monthly release, a specific variable may be revised up or down in the following months, it is unlikely that all variables will be revised in the same direction, as the 141 variables included in X_t are categorized and from different sources. That is, the noise related to each of the revisions of these variables would tend to be cancelled out. As a result, an extracted factor may be more robust to data revision than an individual variable.

We acknowledge that the impact of data revision remains an issue for our Probit-DFM, but perhaps less so compared with existing models with a small number of predictors.

6. Conclusion

In this paper, we model the widely accepted NBER chronology of business cycle by viewing the task of the NBER Dating Committee as a classification problem. We apply a Probit-dynamic factor model to forecast the probability of an economic recession. To the best of our knowledge, we are the first to do so to forecast U.S. recessions. In a data-rich environment, this model can capture the underlying dynamics of an economy and resolve the problems of previous models: too few explanatory variables and too restricted specifications. Our model outperforms popular recession forecast models based on both in-sample and out-of-sample criteria.

Simulated real-time analysis captures all of the recessions and generates only one false alert over the study period. Our model also estimates a high six-month recession probability based on data up to November 2007, which is highly consistent with the onset of the most recent recession, which started in December 2007. This finding is noteworthy, as economists debated the possibility and timing of the recession for one whole year after it had already started.

References

Bai, Jushan, and Serena Ng (2002) 'Determining the number of factors in approximate factor models,' *Econometrica* 70, 191–221

— (2006) 'Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions,' *Econometrica* 74, 1133–50

— (2008) 'Extremum estimation when the predictors are estimated from large panels,' *Annals of Economics and Finance* 9, 201–22

Bernanke, Ben S., and Jean Boivin (2003) 'Monetary policy in a data-rich environment,' *Journal of Monetary Economics* 50, 525–46

Birchenhall, Chris R., Hans Jessen, Denise R. Osborn, and Paul Simpson (1999) 'Predicting U.S. business cycle regimes,' *Journal of Business and Economic Statistics* 17, 313–23

Chauvet, Marcelle (1999) 'An econometric characterization of business cycle dynamics with factor structure and regime switching,' *International Economic Review* 39, 969–96

Chauvet, Marcelle, and Simon Potter (2005) 'Forecasting recessions using the yield curve,' *Journal of Forecasting* 24, 77–103
Estrella, Arturo (1998) 'A new measure of fit for equations with dichotomous dependent variables,' *Journal of Business and Economic Statistics* 16, 198–205
Estrella, Arturo, and Frederic S. Mishkin (1998) 'Predicting U.S. recessions: financial variables as leading indicators,' *Review of Economics and Statistics* 80, 45–61
Filardo, Andrew J. (1999) 'How reliable are recession prediction models?' *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter
Hall, Robert, Martin Feldstein, Ben Bernanke, Jeffrey Frankel, Robert Gordon, and Victor Zarnowitz (2001) 'The business-cycle peak of March 2001,' Business Cycle Dating Committee, National Bureau of Economic Research
Hamilton, James (1989) 'A new approach to the economic analysis of nonstationary time series and the business cycle,' *Econometrica* 57, 357–81
Labonte, Marc (2008) 'Evaluating the potential for a recession in 2008,' CRS Report for Congress RL34484
Ludvigson, Sydney C., and Serena Ng (2009) 'Macro factors in bond risk premia,' *Review of Financial Studies* 22, 5027–67
Marcellino, Massimiliano (2006) 'Leading Indicators,' in *Handbook of Economic Forecasting*, Vol.1, ed. G. Elliott, C. Granger, and A. Timmermann (Amsterdam: North-Holland)
Silvia, John, Sam Bullard and Huiwen Lai (2008) 'Forecasting U.S. Recessions with Probit Stepwise Regression Models,' Business Economics, 43, 7–18
Stock, James H., and Mark W. Watson (1989) 'New indexes of coincident and leading economic indicators,' in *NBER Macroeconomics Annual*, ed. J. Blanchard and S. Fischer (Cambridge, MA: MIT Press)
— (1999) 'Forecasting inflation,' *Journal of Monetary Economics* 44, 293–335
— (2002a) 'Forecasting using principal components from a large number of predictors,' *Journal of the American Statistical Association* 97, 1167–79
— (2002b) 'Macroeconomic forecasting using diffusion indexes,' *Journal of Business and Economic Statistics* 20, 147–62
— (2003) 'How did leading indicator forecasts do during the 2001 recession?' *Economic Quarterly – Federal Reserve Bank of Richmond* 89, 71–90
Wright, Jonathan H. (2006) 'The yield curve and predicting recessions,' Finance and Economics Discussion Series, Federal Reserve Board, February