Helsinki Center of Economic Research

Discussion Papers

Yield-Curve Based Probit Models for Forecasting U.S. Recessions: Stability and Dynamics

Heikki Kauppi
University of Turku, RUESG and HECER

Discussion Paper No. 221
May 2008

ISSN 1795-0562

HECER – Helsinki Center of Economic Research, P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND, Tel +358-9-191-28780, Fax +358-9-191-28781,
E-mail info-hecer@helsinki.fi, Internet www.hecer.fi

HECER
Discussion Paper No. 221

Yield-Curve Based Probit Models for Forecasting U.S. Recessions: Stability and Dynamics*

Abstract

Recent research provides controversial evidence on the stability of yield-curve based binary probit models for forecasting U.S. recessions. This paper reviews so far applied specifications and presents new procedures for examining the stability of selected probit models. It finds that a yield-curve based probit model that treats the binary response (a recession dummy) as a nonhomogeneous Markov chain produces superior in-sample and out-of-sample probability forecasts for U.S. recessions and that this model specification is stable over time. Thus, the failure of yield-curve based forecasts to signal the 1990-1991 and 2001 recessions should not be attributed to parameter instability, instead the evidence suggests that these events were inherently uncertain.

JEL Classification: C22, C25, E32, E37

Keywords: Recession forecast, yield curve, dynamic probit models, parameter stability

Heikki Kauppi

Department of Economics
Assistentinkatu 7
FI-20014 University of Turku
FINLAND

e-mail: heikki.kauppi@utu.fi

* I thank Markku Lanne and Pentti Saikkonen for their useful comments and suggestions on an earlier draft. Financial support from the Research Unit of Economic Structures and Growth (RUESG), the Academy of Finland and the Okobank Group Research Foundation is gratefully acknowledged.

1 Introduction

Predicting recessions is an important task for business and policy makers that condition their decisions on their assessment of the future state of the economy. A number of papers indicate that a simple probit model using predictive information from the yield curve, the spread between short and long-term interest rates, provides superior probability forecasts for the NBER dated recessions in the U.S., at least for forecast horizons ranging from a month to a year ahead.[1] However, Chauvet and Potter (2005) note that the standard yield-curve based probit model has difficulty in signaling the 1990-1991 and 2001 recessions. They argue this negative result is due to parameter instability in the relationship between the yield curve and future economic activity. Against this assertion, Estrella, Rodrigues and Schich (2003) find no evidence of instability, while Chauvet and Potter (2005) argue their results may be misleading, because their breakpoint tests do not properly account for serial dependence in the errors of the probit model. Chauvet and Potter (2005) then apply a probit model formulated through an autoregressive latent variable with business cycle specific error variances and find that the predictive content of the yield curve for U.S. recessions is subject to structural breaks. Whether one assumes such breaks are present or not has marked implications for the way recession forecasts should be made.[2] Therefore, it is important to reassess the disparate evidence on the stability of the predictive relationship between the yield curve and U.S. recessions.

This paper conducts such an analysis by using an alternative statistical modeling approach to that of Chauvet and Potter (2005). The starting point of the analysis is to incorporate dynamics to the standard yield-curve based probit model by adding as a regressor a lagged value of the underlying binary recession indicator. Thus, effectively, the state of the economy is modeled by a nonhomogeneous Markov chain of order one, with transition probabilities changing with the value of the yield-curve. The performance of this simple dynamic probit model is assessed against a variety of models with richer forms of dynamics. Some of the more general models incorporate specific restrictions on

[1] E.g., Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), Estrella, Rodrigues and Schich (2003).

[2] In particular, if one assumes breaks, then future forecasts must rely heavily on the most recent data, while older data may be disregarded.

the duration of recession and expansion periods that result from the rules of the NBER dating of business cycle turning points. Of central importance are extensions that allow for parameter changes across business cycles, some with more general patterns of instability than those considered by Chauvet and Potter (2005). Finally, tests for breakpoints at known and unknown dates are conducted. By considering all of these extensions and stability tests we seek to obtain more robust inference about structural changes in the predictive content of the yield curve and the serial dependence of the U.S. business cycle phases.

Despite the fact that we consider a wide range of models that allow richer forms of parameter instability than previous models, we obtain no convincing evidence for breakpoints, especially when the serial dependence of the recession series is taken into account. It turns our that the above mentioned nonhomogeneous Markov chain of order one is sufficient for capturing the serial dependence of the recession series, richer forms of dynamics, like higher-order Markov chains, do not contribute to forecasting performance. Thus, it seems that simple first-order Markov dynamics provide a good approximation for the purpose of forecasting whether the recession series is truly a high-order process, or even an infinite-order process as in the model of Chauvet and Potter (2005). The bonus of the simple dynamic model is that it is straightforward to interpret, estimate and apply for making multiperiod ahead forecasts. As the paper illustrates, the same does not hold in the case of the model of Chauvet and Potter (2005).

As the final step, the paper compares the out-of-sample forecasting performance of the simple dynamic yield-curve based probit model with that of the standard static probit model. This exercise makes three points. The first point is to show how some practical puzzles, like the lack of real time recession observations, are resolved when dynamic models are applied. The second point is to show what type of probability forecasts of recessions are likely to be useful in practice. The third point is to demonstrate how the static model may yield misleading or implausible recession probability forecasts due to the fact that it neglects the apparent serial dependence of the business cycle phases of the economy. In particular, the static model tends to exaggerate the predictive content of the yield curve so as to produce false recession signals. By contrast, it is shown that the simple dynamic probit model produces forecasts that are more in line with the actual uncertainty that surround specific recessions. Overall, these considerations – together with the lack of evidence for structural instability – suggest that the failure of the yield curve to signal the most recent U.S. recessions may simply derive from the fact that these events were improbable ex-ante.

2 Recession Probability Forecasting

This section lays out the standard yield-curve based probit model for forecasting U.S. recessions and then describes how Chauvet and Potter (2005) extend the model so as take serial dependence and parameter instability into account. After pointing out puzzling features in the Chauvet and Potter (2005) approach, an alternative approach based on Markov type probit models is introduced.

2.1 Previous Approaches

The object of interest is the binary time series y_t that indicates the presence ($y_t = 1$) or absence ($y_t = 0$) of a recession in the U.S. at month t. The bulk of previous empirical analyses considers NBER dated recessions, and this line is followed in this paper as well. The main goal is to forecast the probability of a recession a year (i.e., 12 months) ahead. The key predictor is the yield curve, x_t, the spread between long- and short-term interest rates (see Section 3.1 for the applied interest rate data).

2.1.1 Standard Yield-Curve Probit Model and Its Critique

Consider forecasting the probability that a recession hits at month t (i.e., $y_t = 1$) given that observations until 12 months earlier (i.e., (y_s, x_s), $s \leq 12$) are available.[3] The standard yield-curve based recession forecasting model is a probit model of the form

$$P(y_t = 1) = \Phi(\beta_0 + \beta_1 x_{t-12}), \tag{1}$$

where $P(\cdot)$ denotes probability and $\Phi(\cdot)$ is the cumulative standard normal distribution function. A number of papers find the model (1) useful for forecasting U.S. recessions

[3] In practice, NBER business cycle turning points are announced with delay so that one is uncertain about whether the economy is currently in recession or not. This problem is discussed in Section 3.4.

a year (or less) ahead (the most cited papers are Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998)). Nevertheless, Chauvet and Potter (2005) argue that the predictive performance of the model (1) is not stable over time. In particular, they note that the model has difficulty in signalling the 1990 recession. They argue that this negative result might derive from structural changes in the predictive relationship between the yield curve and economic activity. They point out various potential reasons for such structural changes, like a shift in the volatility of the U.S. economy during 1980s, as documented by McConnell and Perez-Quiros (2000) and others.

The initial assertion of Chauvet and Potter (2005) is in contrast with evidence by Estrella, Rodrigues and Schich (2003). Estrella et al. (2003) examine the stability of (1) using classical tests for an unknown single breakpoint and find that the model is stable, even if there is evidence that corresponding yield-curve based forecasting models for continuous variables like the GDP growth rate are instable. As a response to this, Chauvet and Potter (2005) refer to their earlier paper Chauvet and Potter (2002) that finds strong evidence of structural instability when the probit model in (1) is estimated using the Gibbs sampler. Furthermore, although their applied models consider only a single break, they obtain evidence of the presence of multiple breakpoints. Chauvet and Potter (2005) argue that the difference in the results may derive from serial dependence in the errors of the probit model against which their Bayesian inference may be more robust than the classical tests of Estrella et al. (2003). They point out that the NBER recession indicator is necessarily serially dependent, because the NBER business cycle turning points are determined under the restriction that recessions and expansions are at least six months long, and that a complete cycle lasts at minimum 15 months. Thus, the model in (1) is misspecified in the sense it entails that any serial dependence in y_t derives entirely from that of the yield-curve.

2.1.2 Autoregressive Latent Variable Formulation with Time Varying Parameters

To obtain robust inference on the stability of yield-curve based recession forecasts, Chauvet and Potter (2005) extend the standard yield-curve probit model to account for: (i) time varying parameters due to the existence of multiple breakpoints across business cycles, and (ii) the presence of autocorrelated errors. According to their definition, a business cycle starts at the first month of an expansion period and lasts until the final month of the subsequent recession period. Let t_c indicate the last month of business cycle c and define

$$s_t = \sigma_c, \ 0 < \sigma_c < \infty, \ \text{for } t \in (t_{c-1}, t_c], \ c = 1, ..., n, \tag{2}$$

where n is the number of business cycles in the sample. Now, the extended probit model of Chauvet and Potter (2005) assumes

$$y_t = I(y_t^* > 0) \tag{3}$$

with

$$y_t^* = \alpha y_{t-1}^* + \beta_0 + \beta_1 x_{t-12} + s_t \varepsilon_t, \tag{4}$$

where ε_t is an i.i.d. standard normal variable. The variable y_t^* is regarded as a latent continuous stochastic process, and a recession hits whenever this unobserved process exceeds zero, otherwise there is an expansion. Notice that the model is equivalent to the standard probit model in (1) when $\alpha = 0$ and $s_t = 1$ in (4). The fact that y_t^* is in general autoregressive results in serial dependence in the binary series y_t. In addition, due to changes in s_t, the variance of the innovation process of the latent variable y_t^* is specific to the business cycle. These features are discussed in more detail below.

2.1.3 Issues with Latent Autoregressive Variable Probit Model

The latent variable in (4) may be interpreted as the state of the economy that depends on various macroeconomic variables, like the real GDP and the rate of unemployment. Given that the state of the economy exhibits persistence, it is natural to model it by an autoregressive model as in (4). While this framework has intuitive appeal it is rather difficult too see what the dynamics of the (observed) binary series are, and how the predictor, here the yield curve, affects the probability of a recession in the future. One may wish, for example, to understand how past recession observations drive the probability of a recession.

To obtain some insights into the model of Chauvet and Potter (2005), consider a simple case, where the latent variable is a mean zero Gaussian AR(1) process (i.e., $\beta_0 = \beta_1 = 0$ and $s_t = 1$ in (4)). The first point to note is that while in this case y_t^* is a Markov chain, the 'clipped' series y_t is not, because the underlying transformation is not one-to-one (see Kedem 1980). Thus, it turns out that the probability distribution of y_t depends on the whole history $y_{t-1}, y_{t-2}, \ldots$. To understand the dynamics of the binary series and to forecast its future values, one wishes to know the joint probability of the binary series. For two- and three dimensional probabilities, closed form expressions are known, while for four and higher dimensions no closed form expressions are available (see Kedem (1980)). Thus, figuring out conditional probabilities of future values of the binary series given past observations is challenging even if α was known.

It is possible to connect moments of the latent Gaussian AR(1) process with moments of the clipped process as follows

$$\rho_k = \frac{2}{\pi} \sin^{-1}(\rho_k^*), \tag{5}$$

where ρ_k and ρ_k^*, respectively, is the autocorrelation function of y_t and y_t^* (Kedem (1980, p. 34)). From equation (5) one sees the inequality $|\rho_k| \leq |\rho_k^*|$, which shows that pairwise dependence in the clipped series is weaker than that in the latent series. Note that because y_t is stationary with $E(y_t = 1) = P(y_t = 1) = \frac{1}{2}$ and $\rho_1 = \alpha$, we have the relationship $P(y_t = 1 | y_{t-1} = 1) = \frac{1}{2} + \frac{1}{\pi} \sin^{-1}(\alpha)$. Thus, the parameter α can be estimated by estimating the transition probability $P(y_t = 1 | y_{t-1} = 1)$ using observations on the clipped series. Kedem (1976) shows that a consistent estimator for $P(y_t = 1 | y_{t-1} = 1)$ is obtained by treating the binary series as a first-order Markov chain. In line with this results, Keenan (1982) finds that for predicting a future value of the binary process, it suffices to treat the binary series as a Markov chain. These points suggest that for forecasting as well as for understanding the dynamics of the binary process obtained by clipping a Gaussian AR(1) process, it suffices to treat the binary series as a Markov chain.

When the regressor is included in the latent variable model or when the variance of the underlying innovation series is allowed to change over time, it is even more difficult to analyze the dynamic properties of the binary process than in the above case, where the latent variable is a mean zero Gaussian AR(1). Results like (5) are not available and there is no way to obtain as simple estimation procedures as the one above. To obtain a feasible approach, Chauvet and Potter (2005) apply Bayesian numerical methods. Their estimation and forecasting procedures entail multiple integration over the unobserved state, and therefore, are computationally demanding. Related complications are likely to reduce one's scope for alternative model specifications and forecasts. Thus, there is a greater risk that an inferior forecast model is chosen. In view of these points, the model of Chauvet and Potter (2005) is somewhat troublesome to apply for modeling and forecasting binary time series. On the other hand, the above points on the simple latent variable model suggest that treating the underlying binary series as Markovian may be a reasonable approach for capturing the dynamics of the binary series and for forecasting its future values, even if the true process was of infinite order. Of course, the true process is unknown and therefore one should nevertheless consider alternative forecasting models. These points in part motivate the approach discussed in the subsequent section.

2.2 Markov Chain Approach

2.2.1 Basic Models and an Extension

The starting point is a simple dynamic generalization of the conventional static probit model in (1). Let $\mathcal{I}_t = \{y_t, y_{t-1}..., x_t, x_{t-1}, ...\}$ be the information set available at time t. Then, consider the one-period ahead probit model

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi(\beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-12}). \tag{6}$$

The model in (6) is analogous to one applied by Kauppi and Saikkonen (2007) for forecasting U.S. recessions at the quarterly frequency. As the models in the previous section, it is designed for making recession forecasts 12 months ahead, while the subsequent section shows how such forecasts are obtained from the model in an iterative manner. It is easy to see that under (6) the binary series y_t is governed by a first-order Markov chain, with transition probabilities varying as a function of the regressor x_{t-12}, the lagged yield curve. Given that the model (6) turned out to have superior predictive performance in the analysis of Kauppi and Saikkonen (2007), it is regarded as the baseline dynamic probit model in what follows. Nevertheless, the results of Kauppi and Saikkonen (2007) do not prove that (6) is superior in the case of monthly data, and therefore, we will assess the model against various alternative dynamic specifications.

First, to capture more complicated dynamic dependencies, one can extend (6) so as to attain higher-order Markov chains. For example, the model

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi(\beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-1} y_{t-2} + \beta_4 x_{t-12})$$

results in a nonhomogeneous Markov chain of order two, with time varying transition probabilities. By adding further lags of y_t and all of their interaction terms one may specify a Markov chain of any desired order (e.g., Kaufmann 1987). One problem with this route is that the number of parameters grows exponentially with the order of the Markov chain. Thus, in order to estimate the parameters of a high-order Markov chain one must have a large number of observations. In the application of the present paper, one is faced with the fact that various interaction terms between lagged y_t's, which are needed for higher-order Markov chains, tend to be linearly dependent so that it is impossible to estimate higher-order Markov chains without considerable (zero) restrictions on the coefficients of the interaction terms, or equivalently, on the underlying transition probabilities.

An alternative strategy for increasing the order of the process is to employ autoregressive formulations for the modeling of the dependence of the conditional probability $P(y_t = 1|\mathcal{I}_{t-1})$ on lagged y_t's. Such extensions are considered by Kauppi and Saikkonen (2007) and Rydberg and Shephard (2003). Although these models can break the Markov property, they are straightforward to estimate using standard techniques and to apply for computing multiperiod ahead forecasts (see Kauppi and Saikkonen (2007)). However, in the present application it turns out that such extensions do not yield superior forecasting performance compared with finite order models. Thus, we do not consider such extensions in this paper.

At this point, it is useful to note that the Chauvet and Potter (2005) model formulates the impact of the regressor x_{t-12} in the fashion of an autoregressive distributed lag model. Such a formulation may offer a parsimonious and preferable alternative, if many lags of the regressor are needed for predicting the binary response. Even if this was the case, one may want to keep with the Markov property of the underlying binary series. This wish is met by the following new extension to the baseline model (6)

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi(\beta_0 + \beta_1 y_{t-1} + \upsilon_t), \tag{7}$$

where

$$\upsilon_t = \alpha_1 \upsilon_{t-1} + ... + \alpha_p \upsilon_{t-p} + \beta_2 x_{t-12}. \tag{8}$$

To ensure stationarity, one assumes that $\alpha_1, ..., \alpha_p$ in (8) are such that the roots of the characteristic equation $1 - \alpha_1 z - ... - \alpha_p z^p$ lie outside the unit circle. Unlike the autoregressive models considered by Kauppi and Saikkonen (2007) and Rydberg and Shephard (2003), the one defined by (7) and (8) does not break the Markov property of the underlying binary series.

2.2.2 Minimum Duration Restrictions

There is one more interesting extension to the baseline model (6). This derives from the fact, pointed out by Chauvet and Potter (2005), that the NBER business cycle turning points are determined under the restriction that recessions and expansions are at least six months long, and that a complete business cycle lasts at least 15 months. Interestingly, while Chauvet and Potter (2005) regard these duration restrictions as a reason for serial dependence in the recession series, their proposed dynamic model does not impose these restrictions. In fact, the minimum duration restrictions on the binary series are likely to entail that the latent series model in (4) is augmented with complicated additional nonlinearities. Such a model is likely to be even more difficult to handle than the one applied by Chauvet and Potter (2005). By contrast, it is straightforward to incorporate duration restrictions into the Markov type models considered here. For example, to account for the minimum duration of expansion and recession periods, one may apply the model

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi\left[(\beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-12})\left(1 - I_t^0 - I_t^1\right) + \beta_3 I_t^1 + \beta_4 I_t^0\right], \tag{9}$$

where I_t^1 (I_t^0) is an indicator function for whether the economy has been in a recession (expansion) at least one and at most five most recent months prior to month t. To ensure that a recession and an expansion lasts at least six months, one sets $\beta_3 = \infty$ and $\beta_4 = -\infty$, respectively, so that $P(y_t = 1|\mathcal{I}_{t-1}) = j$, if $I_t^j = 1$, $j = 1, 0$. Notice that the model (9) implies that the yield curve has no impact on the probability of a recession when $I_t^1 = 1$ or $I_t^0 = 1$. The idea of the model is to avoid conditioning the recession state on the regressors y_{t-1} and x_{t-12} in situations where y_t is 'predetermined' by the NBER business cycle dating rules. Intuitively, the model may allow one to obtain more accurate estimates of the degree of serial dependence of the recession series and the predictive content of the yield curve.

2.2.3 Structural Changes Across Business Cycles

This section introduces model extensions for capturing various forms of structural changes. In line with the model of Chauvet and Potter (2005), one may replace the baseline model (6) with

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi(\frac{\beta_0}{s_t} + \frac{\beta_1}{s_t} y_{t-1} + \frac{\beta_2}{s_t} x_{t-12}), \tag{10}$$

where s_t is given in (2).[4] Clearly, changes in s_t translate into changes in the scale of the parameters of the regressors in (10). Without loss of generality let the first business cycle be the reference period and normalize $\sigma_1 = 1$. Then, if $\sigma_c, (c > 1)$, is larger (smaller) than one, the intercept and the coefficients of both of the regressors y_{t-1} and x_{t-12} are smaller (larger) during the business cycle c than they are during the first business cycle of the sample.

The above formulation of instability across business cycles is restrictive in the sense that the scale of all of the regression coefficients is governed by one parameter, σ_c. To allow for more flexibility, one may apply the following model

$$P(y_t = 1|\mathcal{I}_{t-1}) = \Phi(\sum_{j\in C_0} \beta_{0j} c_{jt} + \sum_{j\in C_1} \beta_{1j} c_{jt} y_{t-1} + \sum_{j\in C_2} \beta_{2j} c_{jt} x_{t-12}), \tag{11}$$

where c_{jt} are business cycle specific indicator functions such that $c_{jt} = 1$ for the months of the jth business cycle and $c_{jt} = 0$ otherwise and the sets C_k contain the applied selections of business cycle indices in each case. By the model (11) the intercept and the coefficients

[4] Notice that (10) is equivalent to assuming $P(y_t = 1|\mathcal{I}_{t-1}) = \Phi_t(\beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-12})$, where $\Phi_t(\cdot)$ is the cumulative distribution function of $N(0, s_t^2)$.

of the regressors y_{t-1} and x_{t-12} may change across business cycles in a variety of ways. In practice, one must specify the index sets C_k in a parsimonious manner and so that the resulting regressors are not linearly dependent.

2.2.4 Multiperiod Ahead Forecasts

This section shows how the models considered in the previous section are applied for making probability forecasts for an h periods ahead observation of a binary series given information available at the time of forecasting. First, notice that, in the mean square sense, an optimal h periods ahead forecast of y_t based on information at time $t-h$ is $E(y_t|\mathcal{I}_{t-h}) = P(y_t = 1|\mathcal{I}_{t-h})$. By this relation and the law of iterated conditional expectations, we have

$$E(y_t|\mathcal{I}_{t-h}) = E(P(y_t = 1|\mathcal{I}_{t-1})|\mathcal{I}_{t-h}) = E(\Phi\left(z_t\right)|\mathcal{I}_{t-h}), \tag{12}$$

where z_t is determined by the considered model (e.g., $z_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-12}$). Clearly, any of the specifications in the previous section for z_t give readily the optimal one-step ahead prediction.

Multiperiod ahead forecast with $h \geq 2$ are computed iteratively. To illustrate this, consider the new model given by equations (7) and (8). In this case, z_t in (12) can be written as

$$z_t = \beta_0 + \beta_1 y_{t-1} + \sum_{s=1}^{t} \rho_s \beta_2 x_{t-12-s} \tag{13}$$

where $\rho_j = \alpha_1 \rho_{j-1} + ... + \alpha_p \rho_{j-p}$, for $j > 1$, $\rho_1 = 1$, and $\rho_j = 0$ for $j < 1$. Provided that $h \leq 12$, the variables x_{t-12-s} in (12) are available at the time of forecasting, while the variable y_{t-1} is not observed at date $t-h$. Hence, to evaluate the conditional expectation in (12) one must compute the probabilities of all possible 'paths' or realizations of $y_{t-h+1}, y_{t-h+2}, ..., y_{t-1}$ that lead to $y_t = 1$. Define the vector notation

$$y_{t-k}^{t} = (y_{t-k}, y_{t-k+1}, ..., y_t) \text{ for } k = 0, 1, 2, ...$$

and the Cartesian product $B_k = \{1, 0\}^k$ for $k = 1, 2, ...$ In other words, the set B_k contains all possible k-vectors with components either zero or one ($k = 1, 2, ...$). Then notice that

$$P(y_t|\mathcal{I}_{t-h}) = \sum_{y_{t-h+1}^{t-1} \in B_{h-1}} P(y_{t-h+1}|\mathcal{I}_{t-h}) \prod_{j=1}^{h-1} P(y_{t-h+1+j}|\mathcal{I}_{t-h}, y_{t-h+1}^{t-h+j}), \text{ for } h \geq 2, \tag{14}$$

where $P(y_{t-h+1+j}|\mathcal{I}_{t-h}, y_{t-h+1}^{t-h+j})$ is the conditional probability of $y_{t-h+1+j}$ given $\mathcal{I}_{t-h}$ and the event y_{t-h+1}^{t-h+j}. Each of the conditional probabilities on the right hand side of (14) can be computed straightforwardly using the underlying specification of z_t, while the forecast is $E(y_t|\mathcal{I}_{t-h}) = P(y_t = 1|\mathcal{I}_{t-h})$.

In addition to forecasting whether a particular month is a recession month one is often interested in forecasting the probability that an expansion continues until a specific month (See Section 3.4). Such probabilities are straightforward to compute by removing probabilities of specific realizations of $(y_{t-k}, y_{t-k+1}, ..., y_t)$ from the sum in (14) (cf. Kauppi and Saikkonen (2007)).

3 Empirical Analysis

This section applies the above presented Markov chain approach to the real data, which are described in Section 3.1. Section 3.2 discusses estimation results for baseline models and their extensions, while Section 3.3 examines the stability of leading model variants. Finally, Section 3.4 illustrates how certain practical issues in forecasting are resolved and then conducts an analysis of out-of-sample forecasting performance.

3.1 The Data

As noted above, we analyze a monthly binary time series for U.S. recessions that is obtained from the NBER business cycle turning points. A recession period starts from an NBER 'trough' month and lasts until the month preceding the subsequent NBER 'peak' month.[5] All those months that are not included in a recession period are classified as expansion months.

As for the yield curve, we apply the difference between the ten year Treasury bond rate (constant maturity) and the three month Treasury bill rate (secondary market).[6] Estrella and Trubin (2006) find that this definition of the yield curve is superior in comparison with various alternative long- and short-term interest rates.

For the most part, the analysis is conducted using recession observations on the period

[5] For the dates of the peaks and troughs see http://www.nber.org/cycles/.

[6] The raw data are available at http://www.federalreserve.gov/releases/h15/data.htm.

from January 1955 through November 2001. This period matches closely with the one considered by Chauvet and Potter (2005) and covers eight U.S. recessions. Given this data period, the first business cycle covers the expansion period from January 1955 through July 1957 and the subsequent recession period from August 1957 through April 1958. The second business cycle starts in May 1958, lasts through the corresponding expansion period and the subsequent recession period, and so on for the remaining business cycles. The final complete business cycle starts in April 1991, right after the 1990-1991 recession, and lasts until November 2001, the last month of the 2001 recession. At the time this analysis is completed (April 2008), it is commonly believed that the U.S. economy has not been in a recession from December 2001 through December 2007, while many observers suspect it may turn so during 2008. Nevertheless, the analysis here focuses on the sample that ends in November 2001.

3.2 Baseline Estimation Results

3.2.1 Benchmark Models

Estimation results for the probit model in (6) are given in Table 1. The results here and below are obtained by using the maximum likelihood estimation procedures described in the appendix. The estimates of column (1) of Table 1 are for the static model that assumes (6) with the restriction $\beta_2 = 0$, while the results of column (2) are for the dynamic model without such restriction. In both models, the parameter estimates are significantly different from zero at standard confidence levels. A decrease in the yield curve at month $t-12$ increases the likelihood of a recession at month t. The estimation results of the dynamic probit model indicate positive serial dependence in the recession series: the likelihood of a recession at month t is much larger when the economy was in a recession at the previous month than it it is otherwise. The pseudo R^2 reported in the table is a measure of the over-all fit of the model.[7] As the R^2 in an OLS regression, it lies between 0 and 1 and corresponds roughly to the hypothesis that all coefficients except for

[7] Denote by L_u the unconstrained maximum value of the likelihood function L and by L_c the corresponding maximum value under the constraint that all coefficients are zero except for the constant. The pseudo R^2 measure is defined as pseudo $R^2 = 1 - (\log(L_u)/\log(L_c))^{-2\log(L_c)/T}$, where T denotes the sample size (Estrella 1998).

the constant term are zero. According to the pseudo R^2, the dynamic probit yields more accurate in-sample predictions than the static model.

Figure 1 plots the estimated probabilities that the economy is in a recession state in a particular month from January 1955 to November 2001, for the models in columns 1 and 2 of Table 1. These are probabilities of recessions at t conditional on the value of the yield curve at $t - 12$ and whether the economy is in recession or not at $t - 1$ (dynamic probit). Clearly, the dynamic probit model captures the recession series more accurately than the static probit model. The fit of the static probit is indistinguishable from the one of the corresponding static probit model of Chauvet and Potter (2005, panel (a) of Figure 2). Interestingly, the fit of the dynamic model is very similar to those of the latent autoregressive probit models of Chauvet and Potter (2005, panel (c) and (d) of Figure 2). A close inspection indicates that the present specification provides a slightly better fit to the recession data than the latent autoregressive probit models. This suggests that the present model does not fail to capture any patterns in the recession series that are captured by the latent variable autoregressive probit models. Finally, it must be noted that Figure 1 does not yet illustrate how forecasts based on the dynamic probit perform out-of-sample. In particular, multiperiod ahead forecasts cannot condition on the lagged recession state (i.e., y_{t-1}) and thus the iterative forecast formulae of Section 2.3 must be applied. The performance of out-of-sample forecasts is analyzed in Section 3.4.

3.2.2 Alternative Dynamic Models

The analysis above demonstrates that the simple dynamic model in (6) provides much better in-sample performance than the standard static probit model. It is reasonable to ask whether alternative and more general dynamic specifications might yield even better in-sample performance than the simple models. Table 2 presents estimation results for models where the impact of the yield curve is formulated in an autoregressive manner (see equations (7) and (8)). Column (1) of Table 2 reports estimation results for a model that assumes (7) with $\beta_1 = 0$ and (8) with $p = 1$. While the autoregressive parameter is positive and statistically significant, the Schwarz (1978) Bayesian information criterion (BIC) indicates that the model is inferior to the baseline static model of column (1) of Table 1.[8] When the lagged recession series is allowed in (7), the estimate of the autoregressive coefficient in (8) is negative. This result may be difficult to interpret, while again the BIC suggests that the simple dynamic formulation without the autoregressive formulation is better. Column (3) of Table 2 shows that adding another autoregressive lag in (8) does not improve the overall performance of the model. Similar considerations with other dynamic specifications, like models with higher-order Markov chains, lead to the same conclusion – the simple dynamic model ranks the best especially when the in-sample performance is evaluated by BIC.

Finally, to explore whether restrictions on the duration of business cycle phases make a difference to the in-sample performance, the specification in (9) is estimated under the restrictions $\beta_3 = \infty$ and $\beta_4 = -\infty$ so that $P(y_t = 1|\mathcal{I}_{t-1}) = j$, if $I_t^j = 1$, $j = 1, 0$.[9] The estimates (and robust standard errors) of the unrestricted parameters in (9) are -1.63 (0.20) for β_0, 2.75 (.24) for β_1, and $-.405$ (.14) for β_2. Comparing these estimates with those in column (2) of Table 1, one sees that the coefficient of the yield-curve is slightly larger (in absolute value) and the one of the lagged recession is a bit smaller in the presence than in the absence of the duration restrictions. While these differences make sense, they are very small. The parameter estimates do not make a noticeable difference to measures of in-sample predictive accuracy. Hence, we conclude that the minimum duration restrictions on expansions and recessions are likely to be of little importance for forecasting in practice.

3.3 Stability Analysis

This section investigates whether the baseline forecasting models considered in the previous section are stable over time. As in Chauvet and Potter (2005), we first examine the possibility that the predictive content of the yield curve and the serial dependence of the recession series change across business cycles.

[8] See Inoue and Kilian (2006) for motivation to using BIC as a criterion for selecting a forecasting model.

[9] The corresponding (log) likelihood function is obtained from (8) in the appendix by removing "predetermined" observations with $I_t^j = 1$, $j = 0, 1$.

3.3.1 Business Cycle Specific Parameters

As the first cut, we apply the model formulation in (10) that allows for the scale of the regression coefficients to change with the business cycle. Table 3 reports corresponding estimation results for the static and dynamic probit specifications. In the case of the static model (column 1), the scale parameters, σ_c, are allowed to vary across each of the business cycles of the sample with the exception that the 1980 and 1982 business cycles are combined together. This additional restriction is imposed, because otherwise the estimated value of the scale parameter of the 1980 business cycle turns out to be excessively large.[10] The fact that the estimated scale coefficients in column (1) differ rather much from unity suggests that the impact of the yield curve changes across business cycles. Also, a robust Wald test for the hypothesis that all of the scale coefficients are jointly equal to one rejects the null hypothesis at the 1% significance level.[11] These results are in line with the Bayesian evidence of Chauvet and Potter (2002).

The picture changes quite a bit when the static specification is replaced by the simple dynamic specification (see column (2) of Table 3). Indeed, in the presence of y_{t-1}, the estimates of the scale coefficients are fairly close to one. The largest deviation from unity is obtained for the 1980-82 business cycle. Nevertheless, the corresponding 95% confidence interval (the estimate plus-minus two times its standard error) as well as those of the other scale coefficients cover unity. Also, a robust Wald test for the hypothesis that all of the scale coefficients are equal to one is no longer rejected. These findings indicate that the impact of the regressors do not depend on the business cycle provided that the serial dependence of the recession series is taken into account. Thus, the evidence here gives no support to the form of business cycle specific breaks considered by Chauvet and Potter (2005).

In addition to the model (10), in which only the scale of the parameters changes across

[10] This result may indicate that the relationship between the yield curve and the economy is completely ambiguous during the early 1980s. On the other hand, the result may derive from the fact that this time period contains relatively few observations, which may result in additional uncertainty to the estimation. Nevertheless, it makes sense to combine the 1980 and 1982 business cycles into one and thereby avoid estimation uncertainty coming from too few observations.

[11] The Wald test was constructed using a robust estimator of the variance covariance matrix of the maximum likelihood estimates (see the appendix).

business cycles, several versions of the more flexible specification in (11) were experimented in an initial analysis. Each of the three types of interaction terms in (11) was allowed at the time and various alternative choices of business cycles (sets C_k) were considered. None of the estimated models indicate statistically significant changes in the coefficients of the model with the exception of cases related to early 1980s. For example, according to the estimation results of Table 4, the predictive impact of the spread variable is different during the business cycle starting in 1980 compared with its estimated effect during other business cycle periods in the sample. This observation is consistent with what is observed in the context of the estimation results in column (2) of Table 3. Overall, the evidence here suggests that the predictive relationship between U.S. recessions and the yield curve experiences a transitory break in the beginning of the 1980s.

3.3.2 Breakpoint Tests

The above analysis focuses on searching for structural breaks that are tied to business cycle periods. Alternatively, there may be a structural change that is not related to a specific business cycle and it is worth investigating whether the models studied above are subject to instabilities at any date during the sample period. Estrella, Rodrigues and Schich (2003) conduct such tests for the static formulation of the yield-curve based probit model. The following conducts similar tests for the static and the dynamic model formulation in the present sample.

Tests are conducted for the presence of known as well as unknown breakpoints. Following Andrews and Fair (1988), the applied tests derive from the Lagrange multiplier (LM) statistic

$$LM = \frac{1}{T\omega_1(1-\omega_2)} \mathcal{S}_1(\hat{\theta})' \widehat{\mathcal{J}}(\hat{\theta})^{-1} \mathcal{S}_1(\hat{\theta})$$

where ω_i indicates the proportion of the data before ($i = 1$) or after ($i = 2$) the breakpoint, $\omega_1 + \omega_2 = 1$. The vector $\mathcal{S}_1(\hat{\theta})$ is obtained from the first derivative of the log likelihood function (the score function) given in (16) in the appendix, where the sum is taken over the first portion (ω_1) of the full sample and the parameter vector θ is replaced by its full sample maximum likelihood estimate $\hat{\theta}$. The matrix $\widehat{\mathcal{J}}(\hat{\theta})$ is a misspecification robust estimator of the covariance matrix of the score function (see (19) in the appendix). Andrews and

Fair (1998) show that under regularity conditions with potential breakpoints known a priori, LM has asymptotic chi-squared distribution with degrees of freedom equal to the number of parameters, assuming all can change across subsamples. To test for a break when the breakpoint is unknown one can apply the sup of LM, where the sup is taken over an interior portion of the full sample that excludes observations (a fraction ω_0 of the total observations) at each end. Andrews (1993) shows that the LM statistic converges in distribution to the square of a standardized tied-down Bessel process under general conditions. For a fixed breakpoint, this process has a chi-squared distribution. Tables of critical values corresponding to the distribution of the sup of this process are tabulated in Andrews (1993, 2003) and Estrella (2003).

Estrella et al. (2003) argue that October 1979 and October 1982, both associated with specific shifts in the Federal Reserve's monetary policy practices, are plausible candidates for breakpoints in a yield-curve based forecasting model for U.S. recessions. The corresponding LM test statistics for the dynamic (static) model are 1.20 (1.92) for the former date and 7.30 (1.14) for the latter date. These test statistic values appear to be insignificant when evaluated against the null distribution of no structural change, the chi-squared distribution with four (three) degrees of freedom for dynamic (static) model. In the case of the static model, the sup LM statistic (assuming $\omega_0 = .25$) is equal to 3.41 with the implied breakpoint date being December 1969, while the test does not reject the null of no structural change even at the 10% significance level (the critical value being 9.23, Estrella (2003, p. 1136)). The corresponding sup LM statistic for the dynamic model is equal to 8.82 with the implied breakpoint date in November 1982. While the estimated breakpoint in November 1982 is in line with common expectations, again, the test does not reject the null of no structural change at the 10% significance level (with the critical value being 11.47, Estrella (2003, p. 1136)). The main conclusion from these test results is that they support the view that the above favored simple dynamic probit model using the yield-curve does not experience a structural change during the sample period. At most, given the above analysis of business cycle specific effects, there is weak evidence for a temporary break in the beginning of the 1980s, while this break should not have substantial implications for the predictive relationship in the long term.

3.4 Out-of-Sample Forecasts

3.4.1 Initial Arguments for Dynamic Models

The above analysis suggests that the predictive content of the yield curve for U.S. recessions is stable over time especially when the applied model allows for serial dependence in the recession series. Nevertheless, recent research has mostly applied static probit models for predicting recessions. Some authors (e.g., Estrella et al. (2003)) motivate this approach by arguing that dynamic models are unrealistic models for forecasting recessions, because they assume that the forecaster knows whether the previous or very recent months were recessions. The beginning of a recession can usually be identified only some time after the recession has started. Moreover, recession dating from NBER is typically available with a lag of six months or more. Despite possible delays in recession dating, there are at least two important points that motivate predicting recessions using a dynamic rather than a static model.

The first point is that, if one does not know whether the economy is currently in recession or not, it is natural to modify the forecast horizon so as to start from the most recent observation available. That is, in the presence of a publication lag in recession dating, one is interested in predicting past, current as well as future states of the economy conditional on all available information. In many cases, however, forecasts for future recessions are made in a situation, where one assumes (even if this information is uncertain) that the economy is in an expansion at the time of forecasting.[12] Whatever the case, nothing prevents from making the prediction based on a dynamic model conditional on available information or conditional on alternative scenarios. Also, it should be pointed out that recent research offers various alternatively procedures for dating business cycle turning points that work well in real time even if they cannot forecast future turning points (see Chauvet and Piger (2008)). Thus, the publication lag of the NBER dating of business cycles is not necessarily an insuperable obstacle for making forecasts based on real time recession data.

[12] For example, in February 5, 2008, in a discussion at *Econbrowser*, Michael Dueker (from Federal Reserve) says that one can be reasonably certain that the NBER will not classify the fourth quarter of 2007 as a recessionary period and thus one can condition out-of-sample forecasts accordingly (see http://www.econbrowser.com/archives/2008/02/predicting_rece.html).

The second, even more important, point for favoring dynamic forecasting models is the fact that static models may yield misleading probability predictions in the presence of serial dependence of the recession and expansion states. To illustrate this point, it is useful to consider actual out-of-sample forecasts. Suppose one wishes to predict whether the economy is turning into a recession at any month from $t-d+1$ to $t-d+h$ $(d \geq 0, h > 0)$ conditional on yield data through month t and knowing the state of the economy through month $t-d$. Here t may be regarded as the month where the forecast is made and d as the information lag in recession dating. Chauvet and Potter (2005) consider cases where $d = 3$ and $h = 15$. For example, they argue that in March 2000 the public was certain that the economy was in an expansion through December 1999, while there was lots of uncertainty about the state of the economy from January 2000 on. In practice, d may vary over time depending on circumstances in the economy. For simplicity, the following forecasts assume $d = 3$ and $h = 15$ as in Chauvet and Potter (2005). Given that the yield curve is known through month t, one can then generate the desired probability forecasts for $y_{t-2}, y_{t-1}, y_t, y_{t+1}, ..., y_{t+12}$ using one of the considered model specifications.

3.4.2 Month-by-Month Probability Forecasts

Recession forecasts for two 15 month periods are given in Figure 2, using the static and dynamic baseline specifications. Column (1) (panels (a) and (c)) of the figure plots predicted recession probabilities for each month from January 2000 to March 2001 using yield data up to March 2000 and recession dates through December 1999, while column (2) (panels (b) and (d)) plots corresponding probabilities for each month from January 2001 to March 2002 using yield data up to March 2001 and recession dates through December 2000. The first period covers all the 15 months immediately preceding the latest known recession that started in April 2001 and finished in November 2001, while the second forecast period covers the 2001 recession altogether. Independent of the applied model, it is rather difficult to interpret the month-by-month predictions shown in Figure 2. All of the predicted recession probabilities in Figure 2 are below 0.5. Some of the recession probabilities for actual recession months are smaller than those for some expansion months. The figure illustrates the fact that none of the models is very good at distinguishing whether individual months are in a recession or not. This is not surprising given that the yield curve evolves smoothly rather than in a discrete manner. One would expect that the yield curve carries predictive power for the overall risk that the economy is turning into a recession, while nothing suggests it could pinpoint the precise date at which this happens. To see this point, alternative forecasting approaches must be applied.

3.4.3 Probabilities of Continued Expansion

One possibility is to consider forecasting the probability that an expansion continues, say, 15 months, as in Chauvet and Potter (2005). Figure 3 plots such probabilities over a period of nine years in advance to the 2001 recession. That is, at each month t in the figure, the filled circle at the top of the stem indicates the probability that the economy stays in an expansion from month $t-2$ to month $t+12$ conditional on being in an expansion at $t-3$. Panel (a) of Figure 3 shows that according to the static probit model the probability of continued expansion next 15 months is well below 0.5 over a long period before the economy really turns into a recession. For example, forecasts made during 1996 predict that a recession hits in 15 months with less than 10% probability, while during 1998-1999 the probability of continued expansion is as small as 2-3%. Clearly, predictions based on the static probit tend to alarm recessions too promptly. Given these false recession signals long before the economy turns into a recession, it is not a big gain that the static probit forecasts the 2001 recession right before it actually happens. In fact, the certainty at which the static probit predicts the 2001 recession a year in advance is not in line with conventional wisdom. Indeed, various authors argue that the 2001 recession was very difficult to anticipate well in advance.

The predictions of the dynamic probit model in panel (b) of Figure 3 are more consistent with the reality. First, the predicted probabilities of continued expansion next 15 months remain relatively high during 1996 and 1989-1999. Thus, the dynamic probit model seems to avoid making false recession signals. On the other hand, the probability of continued expansion decreases in advance to the actual recession and thereby provides a reasonable warning of an upcoming recession. The fact that the probability of continued expansion remains above 0.5 is consistent with the common view that it was uncertain whether the economy is turning into recession or not in 2000.

Figures 4 and 5 present similar illustrations for the 1990-1991 and 1980 recessions. The former recession is also commonly regarded as difficult to forecast early in advance. Still, predictions based on the static probit produce very sharp recession calls already in 1986. Also, the sudden decline of the probability of continued expansion in late 1988 seems to call for a recession many months earlier than it actually happens. By contrast, again, the dynamic probit does not produce too early signals of recessions, while it gives a reasonable warning a year in advance to the 1990-1991 recession. As in the case of the 2001 recession, the weakness of the signal is consistent with the common view that the 1990-1991 recession was difficult to predict. Figure 5 yields a similar conclusion on the performance of the two forecasting models for the 1980 recession; the static model seems to give strong recession signals too early. On the other hand, notice that the prediction of the dynamic model is now sharper in advance to the 1980 recession than in the above cases; this is consistent with the fact that the 1980 recession is commonly regarded as easier to forecast than the 1990-1991 and 2001 recessions.

4 Conclusion

Recent research provides disparate evidence on the stability and dynamics of yield-curve based probit models for forecasting U.S. recessions. This paper reviewed this evidence and underlying modeling approaches. In particular, it was illustrated that dynamic probit models obtained by clipping a latent autoregressive process have problems in their interpretation, practical implementation and flexibility. As an alternative approach, we considered probit models with Markovian type dynamics and showed how such models can be extended to capturing various forms of structural changes. We applied the new approach for examining whether the predictive content of the yield-curve for U.S. recessions is stable over time. According to the empirical results, there is no evidence for parameter instability provided that the apparent serial dependence of the recession indicator is taken into account. It turned out that for forecasting purposes it is sufficient to apply a probit model that treats the recession indicator as a nonhomogeneous first-order Markov chain where transition probabilities change with the yield curve.

Finally, the paper illustrated the out-of-sample predictive performance of the simple dynamic probit specification for U.S. recessions and compared it with that of the static probit model applied in a number of previous empirical studies. The analysis showed how the static probit model tends to exaggerate the predictive content of the yield curve so as to produce false or too prompt recession signals and that the dynamic probit model produces probability predictions that are more in line with the actual uncertainty that surround specific recessions. In particular, the results are consistent with the assessment that the 1990-1991 and 2001 recessions were inherently uncertain and thus difficult to predict in advance.

Appendix: Estimation Procedures

This section shows how the parameters of the Markov type models considered in section 2.2 are estimated by maximum likelihood (ML) and how corresponding robust standard errors are obtained. While most of the procedures are readily available in the literature (e.g., Kauppi and Saikkonen (2007)), this section shows what modifications are needed when the regressor is specified in an autoregressive form.

Consider the specification given by equations (7) and (8). One observes the series y_t and x_t for $t = 1, ..., T$ and the initial values $y_0, x_0, ..., x_{-p+1}$. Let $\theta = (\beta_0, \beta_1, \beta_2, \alpha_1, ..., \alpha_p,)'$. Then the log-likelihood function (conditional on the initial values) is

$$l(\theta) = \sum_{t=1}^{T} l_t(\theta) = \sum_{t=1}^{T} \left[y_t \log \Phi\left(z_t(\theta)\right) + (1 - y_t) \log\left(1 - \Phi\left(z_t(\theta)\right)\right)\right], \tag{15}$$

where z_t is given in (13).[13] The first derivative of the log-likelihood, or the score vector, is given by

$$\mathcal{S}_T(\theta) = \frac{\partial l(\theta)}{\partial \theta} = \sum_{t=1}^{T} \frac{[y_t - \Phi(z_t)]\,\phi(z_t)}{\Phi(z_t)[1 - \Phi(z_t)]} \frac{\partial z_t}{\partial \theta}, \tag{16}$$

[13] The likelihood function in (15) is sometimes called the partial likelihood to reflect the fact that it does not require the complete knowledge of the joint distribution of the covariate, x_t. Basically, partial likelihood takes into account only what is known to the observer up to the time of actual observation. (see Fokianos and Kedem 1998).

where $\phi(\cdot)$ is the density function of the standard normal and

$$\frac{\partial z_t}{\partial \theta} = \begin{bmatrix} \partial z_t/\partial \beta_0 \\ \partial z_t/\partial \beta_1 \\ \partial z_t/\partial \beta_2 \\ \partial z_t/\partial \alpha \end{bmatrix} = \begin{bmatrix} 1 \\ y_{t-1} \\ \sum_{s=1}^{t} \rho_s x_{t-s} \\ \sum_{s=1}^{t} \frac{\partial \rho_s}{\partial \alpha} \beta_2 x_{t-12-s} \end{bmatrix}$$

Here $\partial \rho_s/\partial \alpha$ is the vector of derivatives, $(\partial \rho_s/\partial \alpha_1, ..., \partial \rho_s/\partial \alpha_p)$, with

$$\frac{\partial \rho_s}{\partial \alpha_i} = \alpha_1 \frac{\partial \rho_{s-1}}{\partial \alpha_i} + ... + \alpha_p \frac{\partial \rho_{s-p}}{\partial \alpha_i} + \rho_{s-i}, \quad \frac{\partial \rho_j}{\partial \alpha_i} = 0, \ j \leq 1.$$

The ML estimator $\hat{\theta}$ of θ is obtained by maximizing the log-likelihood function in (15), or equivalently, by solving the first order conditions $\mathcal{S}_T(\theta) = 0$, e.g., by applying the BHHH algorithm. To enforce v_t in (8) obeys stationarity conditions, one can reparametrize $\alpha_1, ..., \alpha_p$ in terms of partial correlations and then restrict these to lie within the interval $[-1, 1]$ (see Barndorff-Nielsen and Schou (1973) and Monahan (1984)).

Asymptotic theory for $\hat{\theta}$ is studied by Fokianos and Kedem (1998). They prove existence, consistency and asymptotic normality of $\hat{\theta}$ under regularity conditions. When the model is correctly specified, we have the result

$$T^{1/2}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{J}(\theta)^{-1}), \tag{17}$$

where $\mathcal{J}(\theta) = \text{plim}_{T\to\infty} T^{-1} \sum_{t=1}^{T} (\partial l_t(\theta)/\partial \theta)(\partial l_t(\theta)/\partial \theta')$. In practice, the applied forecasting model may be misspecified. Thus, it is useful to consider the standard extension of (17) given by

$$T^{1/2}(\hat{\theta} - \theta_*) \xrightarrow{d} N(0, \mathcal{H}(\theta_*)^{-1} \mathcal{J}(\theta_*) \mathcal{H}(\theta_*)^{-1}), \tag{18}$$

where $\mathcal{H}(\theta) = -\text{plim}_{T\to\infty} T^{-1} \sum_{t=1}^{T} \partial^2 l_t(\theta)/\partial\theta\partial\theta'$ and θ_* is a value in the parameter space of θ assumed to maximize the probability limit of $T^{-1} l(\theta)$ (for details, see Section 9.3 of Davidson (2000)). In the case of a correctly specified model $\mathcal{J}(\theta) = \mathcal{H}(\theta)$ and consistent estimators of this matrix are given by both $T^{-1} \sum_{t=1}^{T} \partial^2 l_t(\hat{\theta})/\partial\theta\partial\theta'$ and

$$\widehat{\mathcal{H}}(\hat{\theta}) = T^{-1} \sum_{t=1}^{T} (\partial l_t(\hat{\theta})/\partial \theta)(\partial l_t(\hat{\theta})/\partial \theta')$$

In the case of a misspecified model, the estimator $\widehat{\mathcal{H}}(\hat{\theta})$ still estimates the matrix $\mathcal{H}(\theta_*)$ consistently but consistent estimation of the matrix $\mathcal{J}(\theta)$ must account for potential serial dependence of the first order conditions. For simplicity, denote $\partial l_t(\hat{\theta})/\partial\theta = \hat{d}_t$. Then a general estimator is given by

$$\widehat{\mathcal{J}}(\hat{\theta}) = T^{-1}\left(\sum_{t=1}^{T}\hat{d}_t\hat{d}_t' + \sum_{j=1}^{T-1} w_{Tj}\sum_{t=j+1}^{T}\left(\hat{d}_t\hat{d}_{t-j}' + \hat{d}_{t-j}\hat{d}_t'\right)\right), \tag{19}$$

where $w_{Tj} = k(j/m_T)$ for an appropriate function $k(x)$ referred to as a kernel function. The quantity m_T is the so-called bandwidth which for consistency is assumed to tend to infinity with T but at a slower rate. In the empirical application, the Parzen kernel function (see Davidson (2000, p. 227)) is applied and, following the suggestion of Newey and West (1994), m_T is selected according to the rule $m_T = int(4(T/100)^{2/9})$, where $int(x)$ returns the integer part of x.

Using the estimators $\widehat{\mathcal{H}}(\hat{\theta})$ and $\widehat{\mathcal{J}}(\hat{\theta})$ in conjunction with the asymptotic results (17) and (18) one can construct standard Wald tests for hypotheses on the parameter vector θ. In particular, approximate standard errors for the components of the ML estimator $\hat{\theta}$ can be obtained in the usual way from the diagonal elements of the matrix $\widehat{\mathcal{H}}(\hat{\theta})^{-1}\widehat{\mathcal{J}}(\hat{\theta})\widehat{\mathcal{H}}(\hat{\theta})^{-1}$ or, if a correct specification is assumed, from the diagonal elements of the matrix $\widehat{\mathcal{H}}(\hat{\theta})^{-1}$. Section 3.3.2 investigates parameter instability by applying an LM type test statistic that uses the above results.

References

Andrews, D. W. K. (1993) "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61, 821–856.

Andrews, D. W. K. (2003) "Tests for parameter instability and structural change with unknown change point: Corrigendum," *Econometrica*, 71, 395–397.

Andrews, D. W. K., and R. C. Fair (1988) "Inference in non-linear econometric models with structural change," *Review of Economic Studies*, 55, 615–640.

Barndorff-Nielsen, O., and G. Schou (1973) "On the parametrization of autoregressive models by partial autocorrelations," *Journal of Multivariate Analysis*, 3, 408–419.

Chauvet, M., and J. Piger (2008) "A comparison of the real-time performance of business cycle dating methods," *Journal of Business Economics and Statistics*, 26, 42–49.

Chauvet, M., and S. Potter (2002) "Predicting a recession: evidence from the yield curve in the presence of structural breaks," *Economics Letters*, 77, 245–253.

Chauvet, M., and S. Potter (2005) "Forecasting recessions using the yield curve," *Journal of Forecasting*, 24, 77–103.

Davidson, J. (2000) "Econometric theory," Oxford, UK: Blackwell Publishers Ltd.

Estrella, A. (1998) "A new measure of fit for equations with dichotomous dependent variables," *Journal of Business & Economic Statistics*, 16, 198–205.

Estrella, A. (2003) "Critical values and p values of Bessell process distributions: computation and application to structural break tests," *Econometric Theory*, 19, 1128–1143.

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

Estrella, A., and F. S., Mishkin (1998) "Predicting U.S. recessions: financial variables as leading indicators," *Review of Economics and Statistics*, 80, 45–61.

Estrella, A., A. P. Rodrigues, and S. Schich (2003) "How stable is the predictive power of the yield curve? Evidence from Germany and the United States," *Review of Economics and Statistics*, 85, 629–644.

Estrella, A., and M. R. Trubin (2006) "The yield curve as a leading indicator: some practical issues," Federal Reserve Bank of New York, *Current Issues in Economics and Finance*, 12, 1–7.

Fahrmeir, L., and H. Kaufmann (1987) "Regression models for non-stationary categorical time series," *Journal of Time Series Analysis*, 8, 147–160.

Fokianos, K., and B. Kedem (1998) "Prediction and classification of non-stationary categorical time series," *Journal of Multivariate Analysis*, 67, 277–296.

Harding, D., and A. Pagan (2002) "A comparison of two business cycle dating methods," *Journal of Economic Dynamics & Control*, 27, 1681–1690.

Inoue, A., and L. Kilian (2006) "On the selection of forecasting models," *Journal of Econometrics*, 130, 273–306.

Kaufmann, H. (1987) "Regression models for nonstationary categorical time series: Asymptotic estimation theory," *Annals of Statistics*, 15, 79–98.

Kedem, B. (1980), "*Binary time series*," Dekker, New York.

Keenan, D. M. (1982), "A time series analysis of binary data," *Journal of American Statistical Association*, 77, 816–821.

McConnell, M. M., and G. Perez-Quiros (2000) "Output fluctuations in the United States: what has changed since the early 1980's," *American Economic Review*, 90, 1464–1476.

Monahan, J. F. (1984) "A note on enforcing stationarity in autoregressive -moving average models," *Biometrica*, 71, 403–404.

Newey, W. K., and K. D. West (1994) "Automatic lag selection in covariance matrix estimation," *Review of Economic Studies*, 61, 631–653.

Rydberg, T. H., and N. Shephard (2003) "Dynamics of trade-by-trade price movements: decomposition and models," *Journal of Financial Econometrics*, 1, 2–25.

Schwarz, G. (1978) "Estimating the dimension of a model," *Annals of Statistics*, 6, 461–464.

Table 1. Estimation Results for Baseline Probit Models

	(1)		(2)	
	Static		Dynamic	
Predictor	coeff.	s.e.	coeff.	s.e.
Constant	$-.39$	.15	-1.75	.19
Yield curve, x_{t-12}	$-.82$	.12	$-.33$	.15
Recession, y_{t-1}	—		3.2	.22
Pseudo R^2	.23		.68	
Log-likelihood	-171.6		-60.8	
BIC	177.9		70.3	

Notes: The models are estimated using monthly data from January 1955 through November 2001 (563 observations). The reported standard errors (s.e.'s) are robust to misspecification and are computed with procedures described in the appendix.

Table 2. Estimation Results for Probit Models with Autoregressive Effects

	(1)		(2)		(3)	
Predictor	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
Constant	−.35	.17	−1.77	.21	−1.76	.24
Recession, y_{t-1}	—		3.2	.23	3.2	.26
Yield curve, x_{t-12}	−.52	.15	−.55	.27	−.51	.53
Autoreg. lag 1, υ_{t-1}	.40	.21	−.76	.36	−.66	.86
Autoreg. lag 2, υ_{t-2}	—		—		.11	1.03
Pseudo R^2	.24		.68		.68	
Log-likelihood	−170.2		−60.6		−60.6	
BIC	179.7		73.3		76.4	

Notes: The models are given by equations (7) and (8), and are estimated using monthly data from January 1955 through November 2001 (563 observations). The reported standard errors (s.e.'s) are robust to misspecification and are computed with procedures described in the appendix.

Table 3. Estimation Results for Probit Models with Business Cycle Specific Parameters

	(1)		(2)	
Predictor	coeff.	s.e.	coeff.	s.e.
Constant	$-.10$	.13	-1.67	.38
Recession, y_{t-1}	—		3.3	.76
Yield curve, x_{t-12}	-1.18	.84	$-.41$	.27
Scale coeff. (σ_c):				
Jan 55 - Apr 58	1		1	
May 58 - Feb 61	2.31	2.52	1.19	.34
Mar 61 - Nov 70	.34	.27	.83	.23
Dec 70 - Mar 75	1.40	1.16	1.17	.33
Apr 75 - Jul 80	.93	.84	1.11	.33
Aug 80 - Mar 91	2.10	1.56	—	
Aug 80 - Nov 82	—		1.41	.43
Dec 82 - Mar 91	—		.95	.30
Apr 91 - Nov 01	.56	.45	.94	.25
Pseudo R^2	.23		.72	
Log-likelihood	-185.1		-59.0	
BIC	191.4		71.53	

Notes: The models are estimated using monthly data from January 1955 through November 2001 (563 observations). The reported standard errors (s.e.'s) are robust to misspecification and are computed with procedures described in the appendix.

Table 4. Estimation Results for Probit Models with a Transient Structural Break

	(1)		(2)	
Parameter	coeff.	s.e.	coeff.	s.e.
Constant	-1.5	.20	-1.5	.2
Recession, y_{t-1}	3.1	.25	3.1	.26
Interaction term $y_{t-1} \cdot c_{80}$			.20	.59
Yield curve, x_{t-12}	$-.63$	.21	$-.62$	.21
Interaction term $x_{t-12} \cdot c_{80}$	.87	.29	.89	.29
Pseudo R^2	.70		.70	
Log-likelihood	-56.4		-56.3	
BIC	69.0		72.1	

Notes: The models are estimated using monthly data from January 1955 through November 2001 (563 observations). '$c80$' is an indicator variable that equals 1 for August 1980 through November 1982 and 0 otherwise. The reported standard errors (s.e.'s) are robust to misspecification and are computed with procedures described in the appendix.

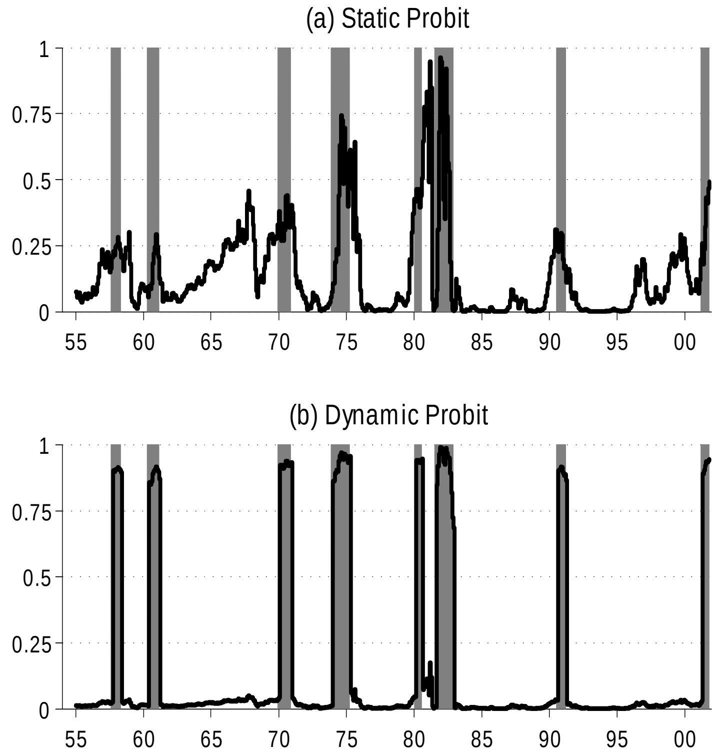

Figure 1: Probability of Recession, In-sample Prediction (the shaded area indicate NBER-dated recessions)

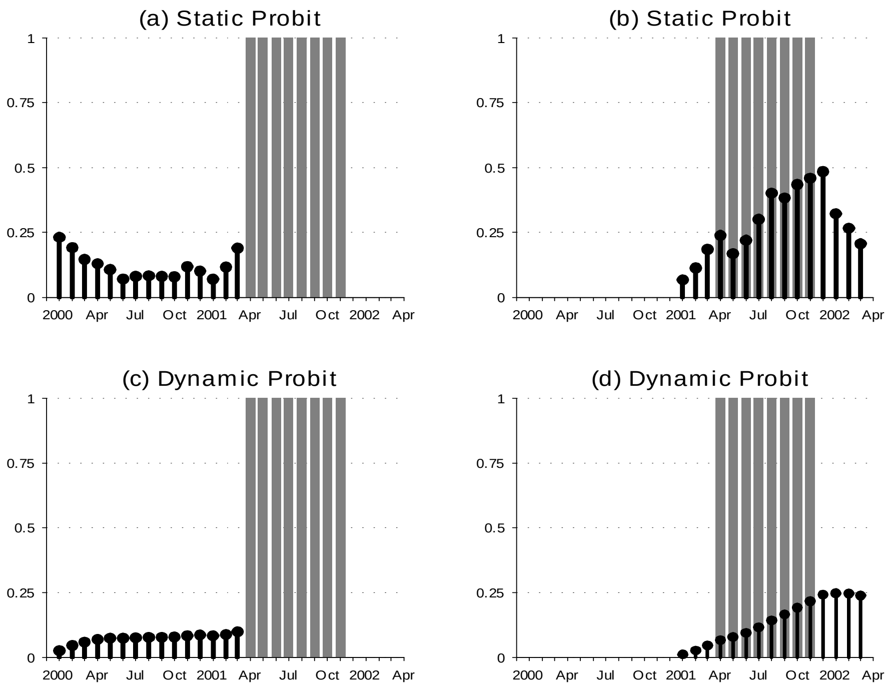

Figure 2: Probability of Recession, Out-of-sample Predictions for January 2000 through March 2001 (panels a and c) and January 2001 through March 2002 (panels b and d). The shaded bars indicate NBER-dated recession months.

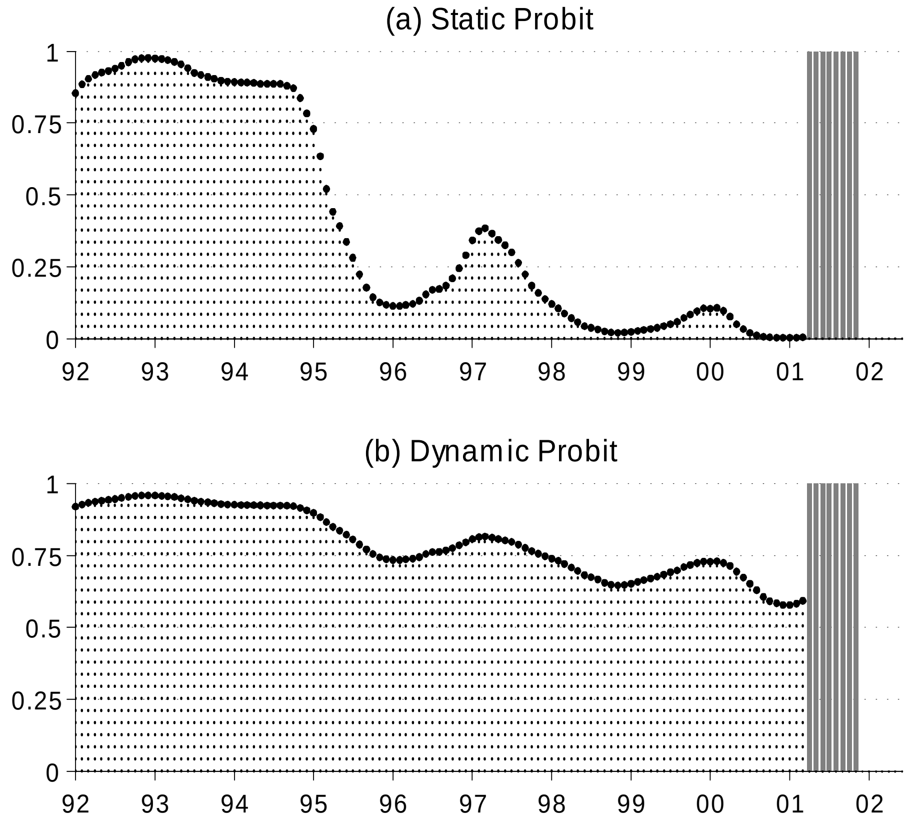

Figure 3: Probability of Continuing Expansion Next 15 Months, Rolling Out-of-sample Prediction (the shaded bars indicate NBER-dated recession months)

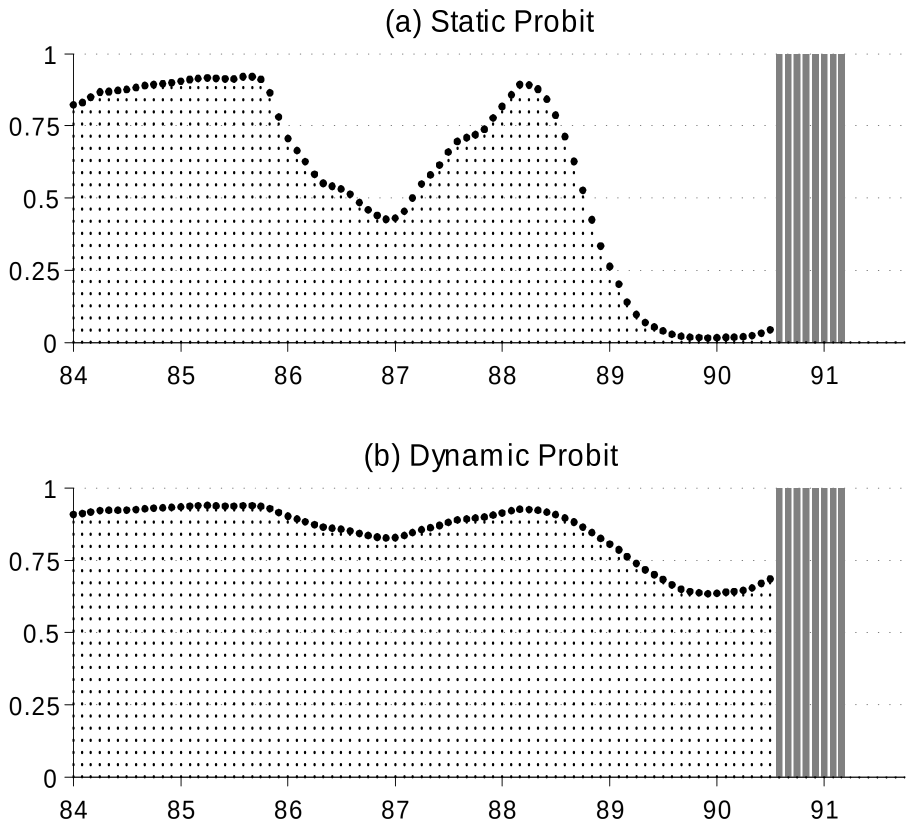

Figure 4: Probability of Continuing Expansion Next 15 Months, Rolling Out-of-sample Prediction (the shaded bars indicate NBER-dated recession months)

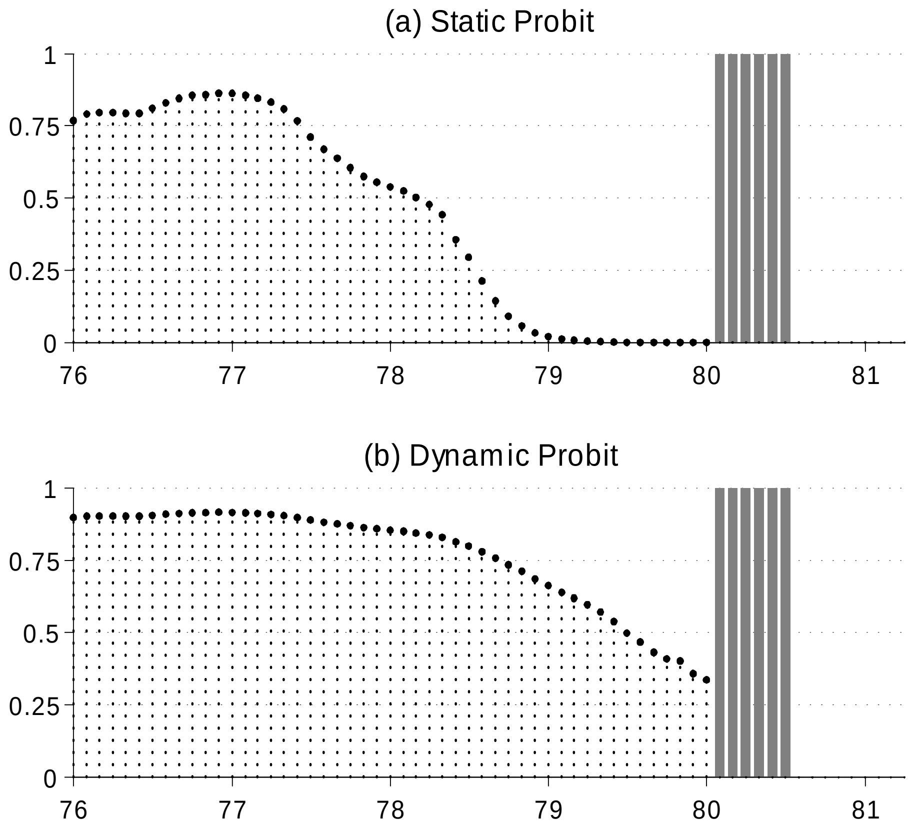

Figure 5: Probability of Continuing Expansion Next 15 Months, Rolling Out-of-sample Prediction (the shaded bars indicate NBER-dated recession months)