The MIT Press

Predicting U.S. Recessions with Dynamic Binary Response Models

Author(s): Heikki Kauppi and Pentti Saikkonen

PREDICTING U.S. RECESSIONS WITH DYNAMIC BINARY RESPONSE MODELS

Heikki Kauppi and Pentti Saikkonen*

Abstract—We develop dynamic binary probit models and apply them for predicting U.S. recessions using the interest rate spread as the driving predictor. The new models use lags of the binary response (a recession dummy) to forecast its future values and allow for the potential forecast power of lags of the underlying conditional probability. We show how multiperiod-ahead forecasts are computed iteratively using the same one-period-ahead model. Iterated forecasts that apply specific lags supported by statistical model selection procedures turn out to be more accurate than previously used direct forecasts based on horizon-specific model specifications.

I. Introduction

WHAT is the probability of a recession a month ahead or a year ahead? This is an important question for central bankers, entrepreneurs, and consumers who make their current decisions on the basis of what they predict will happen to the economy in the future. There is a great body of recent empirical research indicating that financial variables such as the interest rate spread and stock prices provide useful information about whether the U.S. economy will be in a recession between one and eight quarters ahead (for example, Estrella & Mishkin, 1998; Dueker, 1997). Much of the earlier evidence is based on analyses where the probability of a recession in a given number of quarters ahead is modeled by a probit model in which current values of selected financial variables are used as regressors. In addition to this static modeling approach, where no information about the current or past state of the economy is exploited in predicting future recessions, some papers discuss and apply dynamic models where lagged values of the recession indicator are also included in the probit function (for example, Dueker, 1997; Moneta, 2003). Recently, Chauvet and Potter (2005) have applied even more general specifications of the probit model that allow for autocorrelated errors and multiple break points across business cycles (see also Dueker, 2005).

This paper extends the methodology of modeling and forecasting binary time series in a variety of ways. We develop a unified model framework that accommodates most previously analyzed dynamic binary time series models as special cases. Within this framework we then consider new model variants with richer forms of dynamics than previous ones. In one of the extensions the conditional probability of the binary response can also depend on its lagged values, not only on lagged values of the binary response itself. In another extension the impact of explanatory variables may depend on a lagged value of the binary response, that is, on a previous state of the economy. Parameters of all these models can be straightforwardly estimated by a maximum likelihood (ML) procedure described in the paper. Formulas for computing misspecification-robust standard errors are also provided.

We also develop alternative forecasting approaches for our new model variants. The central issue here is how to deal with multiperiod-ahead forecasts. In time series literature, two approaches have commonly been considered to obtain multiperiod-ahead forecasts of continuous variables. One is to compute "direct" forecasts in which multiperiod-ahead values of the response are regressed on their own current and past values and on other regressors. An analog of this procedure is applied in the aforementioned empirical papers on forecasting recessions by dynamic binary probit models. The other approach is to compute "iterated" multiperiod-ahead forecasts. Irrespective of the forecast horizon, a (typically) one-period-ahead model is always estimated in this approach and forecasts are computed iteratively for the desired number of periods. For this latter approach, which, to the best of our knowledge, has previously not been applied to binary time series models, formulas of the forecast functions are derived in the paper.

We apply our methodology to forecast recessions in the United States and compare different models and different forecasting procedures in terms of in-sample and out-of-sample performance. The results show that dynamic probit models outperform their static counterparts. Comparisons among different dynamic specifications indicate that models with lagged values of the binary response yield more accurate forecasts of U.S. recessions than models where dynamics enter only through a lagged probit probability. Like in recent studies on predicting continuous macroeconomic variables, we find that iterated multiperiod-ahead forecasts are better than direct forecasts based on horizon-specific model specifications. In all models, the interest rate spread is an important predictor, but as a new result we find that the choice of its lag order can have a substantial impact on forecasting accuracy. According to our results, it is better to apply specific lags supported by statistical model selection procedures rather than use lags that match the forecast horizon. As in previous studies, we also find that the 1990/91 and the 2001 recessions were hard to predict with any model.

The rest of the paper is organized as follows. The employed framework for dynamic binary time series models is introduced in section II and the related ML estimation and

forecasting procedures are described in sections III and IV, respectively. The empirical application of the paper is presented in section V. Section VI concludes.

II.   Model

Consider the stochastic processes y_t (scalar) and x_t ($k \times 1$) of which y_t is binary valued, that is, takes on only the values 0 and 1. We are interested in dynamic modeling of y_t using x_t as a vector of explanatory variables. Let $\mathcal{F}_t = \sigma\{(y_s, x_s), s \leq t\}$ be the information set available at time t and assume that, conditional on $\mathcal{F}_{t-1}$, y_t has a Bernoulli distribution with probability p_t or, in symbols,

$$y_t | \mathcal{F}_{t-1} \sim B(p_t). \tag{1}$$

Our aim is to model the conditional probability p_t. The standard way to do this is to use a monotonically increasing transformation function Φ and assume the relation $p_t = \Phi(\pi_t)$ where π_t is specified as a linear function of variables in $\mathcal{F}_{t-1}$.[1] The function Φ is chosen to be the cumulative distribution function of a continuous random variable, which guarantees that $\Phi(\pi_t)$ takes values in the unit interval $(0, 1)$. Thus, if $E_{t-1}(\cdot)$ and $P_{t-1}(\cdot)$ signify conditional expectation and conditional probability given $\mathcal{F}_{t-1}$, we have

$$E_{t-1}(y_t) = P_{t-1}(y_t = 1) = \Phi(\pi_t) = p_t. \tag{2}$$

In practice, one typically assumes that Φ is the cumulative distribution function of a standard normal distribution or a logistic distribution. The former assumption leads to the probit model and the latter to the logit model. In our empirical application the probit model is used but, in principle, any twice continuously differentiable cumulative distribution function could be employed.

Now consider specifying the generation mechanism of the conditional probability p_t or, equivalently, that of π_t. The most commonly used dynamic specification assumes that

$$\pi_t = \sum_{j=1}^{q} \delta_j y_{t-j} + x'_{t-1}\beta. \tag{3}$$

Of course, lagged values of the explanatory variables may be included in the vector x_{t-1} so that there is no need to make them explicit in the notation.

There are (at least) two conceivable ways to extend the previously used specification (3). First, one may add lagged values of π_t linearly to the right-hand side and thereby allow for richer dynamics in the process π_t and hence in the conditional probability p_t. If many lagged values of the

explanatory variables and the binary variable are needed in the standard dynamic specification (3), this may provide a more parsimonious and, therefore, preferable alternative. We shall return to this point in section IV when we discuss the use of our models for forecasting. The second extension is based on the idea that the impact of the explanatory variables may depend on a lagged value of the binary variable y_t. For instance, the impact of the interest rate spread on the recession indicator may be asymmetric and differ during recession and expansion periods, and this feature may be advantageous to take into account in forecasting.[2] A specification that accommodates both of these two extensions is

$$\pi_t = \sum_{j=1}^{p} \alpha_j \pi_{t-j} + \sum_{j=1}^{q} \delta_j y_{t-j} + x'_{t-1}\beta + y_{t-d} x'_{t-1}\gamma. \tag{4}$$

We note that, as part of their price movement model, Rydberg and Shephard (2003) consider a somewhat similar specification without the interaction term $y_{t-d} x'_{t-1}\gamma$. Their GLARMA binary model (with exogenous variables) assumes that $\pi_t = x'_{t-1}\beta + g_t$, where $g_t = \sum_{j=1}^{p} \alpha_j g_{t-j} + \sum_{j=1}^{q} \delta_j y_{t-j}$.

The statistical analysis of our dynamic binary response models is based on traditional ML procedures. It is worth noting, however, that one could also consider using a Bayesian approach, which has recently gained popularity when the model is used for real-time forecasting. We refer to Geweke and Whiteman (2006) for an overview of Bayesian forecasting and its potential advantages. Chauvet and Potter (2005) applied a Bayesian approach to a dynamic probit model in which an underlying latent variable is regressed on its lagged value and an exogenous regressor (interest rate spread). A practical difficulty with this model is that ML estimation is complicated, requiring multiple integration over the unobserved lagged variable. The authors used a Bayesian technique based on the Gibbs sampler. The needed computations can be quite extensive. For instance, obtaining estimates and forecasts with their most complicated model took around four hours (see their footnote 20). The dynamic formulation put forward in this paper avoids such complications. As will be seen in the following two sections, the computation of ML estimates requires only a straightforward application of standard numerical methods, and forecasts can be obtained from explicit formulas.

III.   Parameter Estimation

Suppose we have observed the processes y_t and x_t for $t = 1, \ldots, T$ and that q initial values $y_{-q+1}, \ldots, y_0$ are also

---

[1] This assumption means that the conditional probability p_t cannot depend on contemporaneous values of the components of the vector x_t. This, however, is only a matter of notation for one can always replace x_t by $x_t^* = x_{t+1}$.

[2] Such an asymmetric relation may be motivated by noticing that the interest rate spread can be regarded as a measure of the stance of monetary policy, while there is evidence that the impact of monetary policy on the real economy may be different during recessions and expansions (see Florio, 2004).

available. Define the parameter vector $\theta = [\alpha' \; \delta' \; \beta' \; \gamma']'$ where $\alpha = [\alpha_1 \cdots \alpha_p]'$ and $\delta = [\delta_1 \cdots \delta_q]'$. Then the log likelihood function (conditional on initial values) has the form

$$l(\theta) = \sum_{t=1}^{T} l_t(\theta)$$

$$= \sum_{t=1}^{T} [y_t \log \Phi(\pi_t(\theta)) + (1 - y_t) \log (1 - \Phi(\pi_t(\theta)))],$$

where $\pi_t(\theta)$ is given by the right-hand side of equation (4) or its restricted version (3). When the more general specification (4) is used, a choice for the initial values $\pi_{-p+1}, \ldots, \pi_0$ is needed. We choose these by using formulas that can be interpreted as estimates of the unconditional mean of π_t. For example, when $p = q = 1$, we set $\pi_0 = (\delta_1 \bar{y} + \bar{x}'\beta + \overline{yx}'\gamma)/(1 - \alpha_1)$, where a bar is used to signify the sample mean of the indicated variables.

The maximization of the likelihood function is clearly a highly nonlinear problem but can be straightforwardly carried out by standard numerical methods. In the case of the specification (3), de Jong and Woutersen (forthcoming) showed that, under appropriate regularity conditions, the conventional large sample theory applies to the ML estimator of the parameter vector θ, denoted by $\hat{\theta}$. The assumed regularity conditions include stationarity of the explanatory variables and correctness of the assumed probit specification. The obtained result reads as

$$T^{1/2}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{I}(\theta)^{-1}), \tag{5}$$

where $\mathcal{I}(\theta) = p \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} (\partial l_t(\theta)/\partial \theta)(\partial l_t(\theta)/\partial \theta')$. We use this result also in the case of the more general specification (4), although no formal proof of its validity is presently available. The same remark applies to the asymptotic results discussed below. As in de Jong and Woutersen (forthcoming), the development of a rigorous asymptotic estimation theory for our general dynamic model would first require obtaining appropriate mixing results or equivalent, a task beyond the scope of this paper.

De Jong and Woutersen (forthcoming) note that their results on ML estimation can be extended to the case where the employed probit specification does not hold but, for instance, a logit model would be correct. There is also another kind of misspecification that may result when the model is used for forecasting, as is done in this paper. Specifically, suppose that at time t the purpose is to forecast the value of $y_{t+h}(h \geq 1)$. This requires forecasting the explanatory variables. However, one may be unwilling to build a forecasting model for the explanatory variables but, as an alternative, one modifies the specification (3) or (4) by replacing x_{t-1} by x_{t-h} (cf. Estrella & Mishkin, 1998). A similar approach may be applied to lagged values of y_t.

Thus, if the data-generation process is given by equation (3) or (4) such modified models are misspecified when $h > 1$.

Motivated by the preceding discussion we also consider the standard extension of equation (5) given by

$$T^{1/2}(\hat{\theta} - \theta_*) \xrightarrow{d} N(0, \mathcal{I}(\theta_*)^{-1}\mathcal{I}(\theta_*)\mathcal{I}(\theta_*)^{-1}), \tag{6}$$

where $\mathcal{I}(\theta) = -p \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} \partial^2 l_t(\theta)/\partial \theta \partial \theta'$ 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 we have $\mathcal{I}(\theta) = \mathcal{I}(\theta)$, and consistent estimators of this matrix are given by both $T^{-1} \sum_{t=1}^{T} (\partial l_t(\hat{\theta})/\partial \theta)(\partial l_t(\hat{\theta})/\partial \theta')$ and

$$\hat{\mathcal{I}}(\hat{\theta}) = T^{-1} \sum_{t=1}^{T} \partial^2 l_t(\hat{\theta})/\partial \theta \partial \theta'.$$

In the case of a misspecified model the estimator $\hat{\mathcal{I}}(\hat{\theta})$ still estimates the matrix $\mathcal{I}(\theta_*)$ consistently, but consistent estimation of the matrix $\mathcal{I}(\theta)$ is more complicated. For simplicity, denote $\partial l_t(\hat{\theta})/\partial \theta = \hat{d}_t$. Then a general estimator is given by

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

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 our empirical application we use the Parzen kernel function (see Davidson, 2000, p. 227) and, following the suggestion of Newey and West (1987), we select m_T according to the rule $m_T = floor(4(T/100)^{2/9})$, where the function $floor(x)$ rounds x to the nearest integer less than or equal to x (cf. Eviews, 2004, p. 457).

Using the estimators $\hat{\mathcal{I}}(\hat{\theta})$ and $\hat{\mathcal{I}}(\hat{\theta})$ in conjunction with the asymptotic results (5) and (6) 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 $\hat{\mathcal{I}}(\hat{\theta})^{-1}\hat{\mathcal{I}}(\hat{\theta})\hat{\mathcal{I}}(\hat{\theta})^{-1}$ or, if a correct specification is assumed, from the diagonal elements of the matrix $\hat{\mathcal{I}}(\hat{\theta})^{-1}$.

IV. Forecasting Procedures

Our intention is to forecast U.S. recessions, a binary variable that equals 1 if there is a recession, and 0 otherwise. We examine and compare the predictive performance of alternative model specifications, all embedded in the general model given by equations (1), (2), and (4). In this section, we introduce the models employed in our empirical

application and the corresponding forecast functions, one at the time.

Previous literature has examined the predictive content of several financial variables for future recessions. In general, the interest rate spread has proved to be the most useful variable for predicting U.S. recessions (for example, Estrella & Mishkin, 1998; Dueker, 1997). Estrella and Mishkin (1998) point out that, even if many financial variables result in a relatively good in-sample fit, in most cases, a parsimonious probit model with the interest rate spread as the sole predictor tends to produce the most accurate out-of-sample predictions of U.S. recessions. Due to its dominant predictive power we only use the interest rate spread (and a constant term) in our forecasting models. This also enables us to focus on examining the predictive performance of different models, not of different variables. Henceforth, let y_t denote the recession indicator and x_t the value of the interest rate spread (see the subsequent section for precise definitions and data sources).[3]

In the mean square sense, an optimal h quarters ahead forecast of y_t based on information at time $t - h$ is the conditional expectation $E_{t-h}(y_t) = P_{t-h}(y_t = 1)$. By the law of iterated conditional expectations and the relations in equation (2), we have

$$E_{t-h}(y_t) = E_{t-h}(P_{t-1}(y_t = 1)) = E_{t-h}(\Phi(\pi_t)), \qquad (7)$$

where, and also henceforth, Φ is the cumulative distribution function of the normal distribution and π_t is determined by the considered specification. This relation will be used in the subsequent discussion.

A. Static Forecasting

The benchmark forecasts are obtained from the static binary probit model, which assumes that the conditional probability in equation (2) is given by

$$P_{t-1}(y_t = 1) = \Phi(\omega + \beta x_{t-k}), \qquad (8)$$

where k is the employed lag order of the interest rate spread. In practice, one may wish to choose $k \geq h$, where h is the forecast horizon, because then the employed value of the spread is known at the time of forecasting, that is, at time $t - h$. Indeed, from equation (7) it is seen that the corresponding h-quarters-ahead forecast is given directly by the right-hand side of equation (8). In previous studies on forecasting U.S. recessions it has been common to choose the most recent value of the interest rate spread by setting $k = h$. This means that the employed lag of x_t is tailored to match the forecast horizon. Obviously, this approach only provides a correct specification and optimal forecasts for one value of h. Our empirical application demonstrates that

such an artificial lag order selection procedure may lead to very poor forecasts, and that better results can be obtained by using lags supported by statistical model selection procedures.

The limitation of specification (8) is its static nature, which means that the same h-quarters-ahead forecast is obtained independently of whether the economy is in recession or not, that is, whether $y_{t-h} = 1$ or $y_{t-h} = 0$. The dynamic model framework of section II offers several possible specifications for predicting future recessions that use information about the current and past states of the economy.

B. Dynamic Forecasting

First consider a specification in which the static probit model is augmented by using a lagged value of the recession indicator as an additional regressor. This model, referred to as dynamic probit, assumes that the conditional probability in equation (2) equals

$$P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-1} + \beta x_{t-k}), \qquad (9)$$

where, by equation (7), the right-hand side gives the one-quarter-ahead forecast based on information at time $t - 1$ (provided that $k \geq 1$). As in section II, more than a single lag of the recession indicator could be employed. We shall not consider that extension to simplify exposition and because no evidence of its need was found in our empirical application.

Now suppose one wants to forecast several quarters ahead. There are two basic ways to do that. One can use a direct forecast based on the assumption that, instead of equation (9), the conditional probability in equation (2) is given by

$$P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-l} + \beta x_{t-k}). \qquad (10)$$

Under the condition $l, k \geq h$, one again sees from equation (7) that the right-hand side gives "directly" the h-step-ahead forecast made at time $t - h$, so the forecast depends on only the values of the recession indicator and the interest rate spread that are known at the time of forecasting.[4] Dueker (1997) applied model (10) by choosing $l = k = h$. In this approach, the employed model is again horizon specific, and a correct specification and optimal forecasts can only be obtained for one value of h.

The other possibility is to use an iterative approach in which the lag order of the recession indicator need not be

---

[3] The forecast formulas presented in this section and the appendix can be used in an obvious way when more than a single explanatory variable is used.

[4] When recessions are predicted in practice, one must account for the fact that the value of the recession indicator becomes known with a delay. Thus, the present forecasting setup should be thought of as one of predicting unknown values, whether past, current, or future, using the most recent observations of the recession indicator. This means that, in spite of the delay, our forecasting formulas can be used as such if only the date of forecasting is chosen to be the date in the recent past where the value of the recession indicator can be assumed known. See section V for an application.

related to the forecast horizon. For example, the h-quarters-ahead forecast can be based on the assumption that the conditional probability in equation (2) is given by the right-hand side of equation (10) with y_{t-l} replaced by y_{t-1} (that is, model (9)). This makes the computation of the forecasts more difficult than in the direct approach. Forecasts made at time $t - h$ require evaluating

$$E_{t-h}(y_t) = E_{t-h}(\Phi(\omega + \delta y_{t-1} + \beta x_{t-k})) \qquad (11)$$

(see equation (7)). Thus, the problem is similar to that in forecasting several steps ahead in conventional nonlinear autoregressive models (see Tong, 1990, p. 346). However, unlike in many nonlinear autoregressions the binary nature of the response variable makes it possible to compute forecasts explicitly. For instance, consider the case $h = 2$. From the preceding equation one obtains

$$E_{t-2}(y_t) = \sum_{y_{t-1} \in \{0,1\}} P_{t-2}(y_{t-1})\Phi(\omega + \delta y_{t-1} + \beta x_{t-k}), \qquad (12)$$

where

$$P_{t-2}(y_{t-1}) = \Phi(\omega + \delta y_{t-2} + \beta x_{t-k-1})^{y_{t-1}}$$
$$\times [1 - \Phi(\omega + \delta y_{t-2} + \beta x_{t-k-1})]^{(1-y_{t-1})}$$

gives the conditional probabilities of the two possible outcomes of y_{t-1} given information known at the forecast period. Thus, basically, the forecast takes account of the two possible paths through which we can enter a recession in two periods' time: either we go through a recession in the quarter between or we don't. For $h > 2$ the number of possible paths is larger and the situation gets more complicated; this is discussed in the appendix.

Thus, either a direct or an iterative procedure can be used when the forecast horizon is longer than a period, and it is not obvious which one of these two procedures should be employed in practice. A similar problem arises when one forecasts continuous time series with autoregressive models. In the presence of estimation uncertainty, the ranking between the direct and iterative procedures largely depends upon whether the model used in iterative forecasting is close to or far from the true data-generating process. If it is close, iterated forecasts tend to be more efficient than direct forecasts, and vice versa (see Bhansali, 1999). Because the ranking in efficiency between direct and iterative forecasts is theoretically ambiguous, it has proved useful to study their relative forecasting merits empirically (cf. Marcellino, Stock, & Watson, 2006). In our application, iterative forecasts turn out to be clearly superior to direct forecasts.

Next consider forecasts based on a formulation in which the conditional probability in equation (2) is assumed to be given by

$$P_{t-1}(y_t = 1) = \Phi(\omega + \alpha \pi_{t-1} + \beta x_{t-k}), \qquad (13)$$

where $k \geq h$ is assumed. We label this model specification as autoregressive probit. To obtain h-quarters-ahead forecasts based on information at time $t - h$, we need to evaluate

$$E_{t-h}(y_t) = E_{t-h}(\Phi(\omega + \alpha \pi_{t-1} + \beta x_{t-k}))$$

(again, see equation [7]). In this case the required computations are simple for any value of h. Using the equation $\pi_t = \omega + \alpha \pi_{t-1} + \beta x_{t-k}$ and repetitive substitution yields

$$E_{t-h}(y_t) = E_{t-h}\Phi\left(\alpha^h \pi_{t-h} + \sum_{j=1}^{h} \alpha^{j-1}(\omega + \beta x_{t-k+1-j})\right)$$

$$= \Phi\left(\alpha^h \pi_{t-h} + \sum_{j=1}^{h} \alpha^{j-1}(\omega + \beta x_{t-k+1-j})\right),$$

where the latter equation is based on the fact that π_{t-h} is a function of past values of x_t and the initial value π_0. Thus, in contrast to the basic static specification (8), forecasts now use information of a series of past values of the interest rate spread. Of course, one could also use several lagged values of the interest rate spread in the basic static specification but, if many lags are really useful for forecasting future recessions, the autoregressive specification may be a better alternative. The reason is that, because of the involved parameter restrictions, the resulting forecasts are based on a more parsimonious model, which may be useful in practice where estimates of parameters have to be used. Indirect support for this is provided by Estrella and Mishkin (1998) who find that, in general, parsimonious models produce the most accurate out-of-sample forecasts of U.S. recessions. However, the autoregressive specification does not directly incorporate information from the lagged values of the dependent variable. For that, lagged values of y_t have to be used too.

Including a lagged recession indicator in the autoregressive specification (13) gives the dynamic autoregressive specification

$$P_{t-1}(y_t = 1) = \Phi(\omega + \alpha \pi_{t-1} + \delta y_{t-1} + \beta x_{t-k}), \qquad (14)$$

where $k \geq h$ is again assumed. At time $t - 1$ forecasts one quarter ahead are obtained directly from this equation. For longer forecast horizons, direct forecasts can be computed in the same way as in the case of the autoregressive specification (13). One only needs to change y_{t-1} in equation (14) to y_{t-k}. For iterative forecasting the needed formulas are somewhat more complicated than previously, although the principle is still quite straightforward. To see the basic idea, consider forecasting two quarters ahead, assuming $k = 2$. Because now $\pi_t = \omega + \alpha \pi_{t-1} + \delta y_{t-1} + \beta x_{t-2}$, it follows from equation (7) that we have to compute

$$E_{t-2}(y_t) = E_{t-2}\Phi(\alpha^2\pi_{t-2} + \delta y_{t-1} + \omega + \beta x_{t-2}$$

$$+ \alpha(\delta y_{t-2} + \omega + \beta x_{t-3})) \overset{def}{=} E_{t-2}\tilde{\Phi}(y_{t-1}),$$

where

$$\tilde{\Phi}(y_{t-1}) = \begin{cases} \Phi(\alpha^2\pi_{t-2} + \omega + \delta + \beta x_{t-2} \\ \quad + \alpha(\omega + \delta y_{t-2} + \beta x_{t-3})), & \text{if } y_{t-1} = 1 \\ \Phi(\alpha^2\pi_{t-2} + \omega + \beta x_{t-2} \\ \quad + \alpha(\omega + \delta y_{t-2} + \beta x_{t-3})), & \text{if } y_{t-1} = 0, \end{cases}$$

and use has been made of the fact that $y_{t-1}|\mathscr{F}_{t-2} \sim B(p_{t-1})$ with $p_{t-1} = \Phi(\omega + \alpha\pi_{t-2} + \delta y_{t-2} + \beta x_{t-3})$. Note that $\tilde{\Phi}(y_{t-1})$ is a binary variable with two possible values, $\tilde{\Phi}(1)$ and $\tilde{\Phi}(0)$, that realize with probabilities p_{t-1} and $1 - p_{t-1}$. Thus,

$$E_{t-2}(y_t) = p_{t-1}\tilde{\Phi}(1) + (1 - p_{t-1})\tilde{\Phi}(0).$$

When the forecast horizon increases the number of possible paths within the forecast period increases. Consequently, the forecast function is more complicated. For details see the appendix.

It is straightforward to formulate the above forecast procedures to allow for the interaction term in equation (4) (i.e., $\gamma \neq 0$). Suppose that, in addition to p and q, also $d = 1$, and first assume that $\alpha_1 = 0$, that is, the autoregressive term π_{t-1} is not present in equation (4). In direct forecasting one replaces y_{t-1} and $y_{t-1}x_{t-1}$ by y_{t-h} and $y_{t-h}x_{t-h}$, respectively, where h is again a given forecast horizon. The underlying conditional probability is then similar to equation (10) in that the h-step-ahead forecast made at time $t - h$ depends on only the known values of the recession indicator, the interest rate spread, and their interaction term at the time of forecasting. In iterative forecasting one always uses y_{t-1} so that, when $\alpha_1 = 0$ is assumed, forecasts made at time $t - h$ require evaluating

$$E_{t-h}(y_t) = E_{t-h}(\Phi(\omega + \delta y_{t-1} + \beta x_{t-h} + \gamma y_{t-1}x_{t-h}))$$

$$= E_{t-h}(\Phi(\omega + (\delta + \gamma x_{t-h})y_{t-1} + \beta x_{t-h})).$$

Notice that this conditional expectation differs from equation (11) only in that the value of the coefficient of y_{t-1} depends on the value of x_{t-h}. Modifying the previous iterative h-step-ahead forecast to the present case is therefore simple. For example, when $h = 2$, one uses equation (12), but with δ replaced by the time variant coefficient $\delta + \gamma x_{t-h}$. When the autoregressive term is present in equation (4), this replacement is made in equation (14), after which the formulas presented in the previous paragraph and in the appendix apply without any further change.[5]

---

[5] Note that the same formulas apply whether $\delta = 0$ or $\delta \neq 0$.

C. Hitting Times

The above forecasts are based on the conditional probability that the economy is in a recession in period t, given information in period $t - h$. In this approach no attention is paid to the particular sequence of values (path) of the recession indicator during periods between $t - h$ and t. However, in some cases particular sequences of the recession indicator and their probabilities may also be of interest. For example, Chauvet and Potter (2005) propose forecasting future recessions using what they call hitting probabilities. Basically, a hitting probability is defined as the probability that the economy is hit by a recession no earlier (and no later) than in a given period in the future. More formally, at time $t - h$ the first hitting time to a recession is defined by

$$H(t - h) = \min \{\tau \geq 1 : y_{t-h+\tau} = 1\}.$$

Thus, if $H(t - h) = k$ we observe the sequence $y_{t-h+k} = 1, y_{t-h+k-1} = \cdots = y_{t-h+1} = 0$ and the associated hitting probability is

$$\rho(k, t - h) = P_{t-h}(H(t - h) = k)$$

$$= P_{t-h}(y_{t-h+k} = 1|y_{t-h+k-1} = 0, \ldots$$

$$\ldots, y_{t-h+1} = 0)(1 - \rho(k - 1, t - h)),$$

where $\rho(0, t - h) = 0$, $k = 1, 2, \ldots$

In practice it is usually not relevant to compute probabilities for very remote hitting times. Chauvet and Potter (2005) assume a fixed forecast horizon, say h, and compute $\rho(k, t - h)$ for $k = 1, 2, \ldots, h$. An advantage of our model is that such probabilities are easy to compute by modifying the forecast formulas presented above and in the appendix. For example, if $h = 2$ and the conditional probability is specified as

$$P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-1} + \beta x_{t-2}),$$

we obtain

$$\rho(2, t - 2) = P_{t-2}(y_{t-1} = 0)\Phi(\omega + \beta x_{t-2}),$$

where

$$P_{t-2}(y_{t-1} = 0) = 1 - \Phi(\omega + \delta y_{t-2} + \beta x_{t-3}).$$

Notice that the same value of $\rho(2, t - 2)$ can be obtained from the "standard" forecast formula (12) by setting $P_{t-2}(y_{t-1} = 1) = 0$, that is, by ignoring the probability of the path $\{y_{t-1} = 1, y_t = 1\}$. The same principle of ignoring irrelevant paths applies to computing hitting probabilities for any other forecast horizon and model specification presented above. All computations can be performed with explicit formulas and there is no need for (multiple) integration and simulation needed in the autoregressive latent variable probit model of Chauvet and Potter (2005).

V. Empirical Analysis

Our empirical application to forecasting U.S. recessions uses a recession indicator, y_t, obtained from the business cycle turning points determined by the National Bureau of Economic Research (NBER).[6] It is important to note that the NBER business cycle turning points are announced with a delay. For example, the most recent turning point (known in September 2006) was the end of recession in November 2001, which was announced as late as in July 2003.[7] Thus, in principle, we do not know for sure whether the next turning point is a past quarter, the ongoing quarter (2006: Q3), or some quarter in the future. Even if there is some uncertainty regarding the current and most recent quarters, we are (and the public is) confident that the U.S. economy was still in expansion in the last quarter of 2005; many observers think it was still so in the second quarter of 2006. To make a safe choice, we regard it as known that the U.S. economy has been in an expansion from November 2001 until the end of 2005. The key predictor, that is, the interest rate spread, x_t, is constructed as the difference between the ten-year Treasury bond rate (constant maturity) and the three-month Treasury bill rate (secondary market).[8] Given that these financial variables are available from the second quarter of 1953 and that some observations are needed to initialize their lagged values, our full sample below amounts to cover the period from 1955:Q4 to 2005:Q4. We start by analyzing the in-sample performance of various models.

A. In-Sample Results

Even though we are mainly interested in out-of-sample forecasts it is useful to first compare the in-sample performance of the employed models over the entire sample period. Such comparisons are also performed in a true forecasting situation where the first step is to select a forecasting model out of different alternatives.

We mainly evaluate and compare the in-sample performance of the employed models using the pseudo R^2 measure of fit developed by Estrella (1998) (cf. Estrella & Mishkin, 1998).[9] This measure takes on values between 0 and 1 and it can be interpreted in the same fashion as the

TABLE 1.—PSEUDO R^2 MEASURES OF IN-SAMPLE FIT FOR DIFFERENT FORECASTING MODELS

		k					
		1	2	3	4	5	6
	—	.165	.125	.189	.233	.211	.121
	1	.376	.465	.406	.407	.366	.328
	2	.105	.245	.256	.252	.219	.137
l	3	.022	.133	.202	.233	.231	.132
	4	.019	.127	.192	.233	.237	.176
	5	.029	.134	.203	.250	.220	.146
	6	.019	.125	.189	.236	.213	.123

Notes. The table entries report the pseudo R^2 values for the model $P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-l} + \beta x_{t-k})$ with different choices of l and k. The results of the first row (denoted by —) are obtained under the restriction $\delta = 0$. All results are computed from the sample 1955:Q4–2005:Q4.

coefficient of determination in a linear regression. We also made similar comparisons using the Akaike (1973, 1974) information criterion, the Schwarz (1978) Bayesian information criterion, and a variety of specialized probability forecast accuracy measures, like the quadratic probability score and the log probability score that have been used in earlier studies (see Diebold & Rudebusch, 1989). All these measures lead to the same conclusion about the ranking of the employed models, and the same was true when (robust) standard errors (and corresponding t-ratios) of the estimated parameters were considered.

As a first cut, table 1 presents the pseudo R^2 values for models with two regressors, y_{t-l} and x_{t-k}, with selected choices of the lag orders l and k. The first row of the table presents results for the static model with the employed lag order of x_t varying between one and six, while the remaining rows present results for the dynamic specification with the chosen lag order of y_t changing between one and six. Among the static models, the one with the fourth lag of x_t has the best in-sample fit. The remaining rows show that the in-sample fit of the static model can be improved clearly by adding the first lag of y_t on the right-hand side, while not that much is gained by using a longer lag of y_t. Overall, the results of table 1 indicate that the best in-sample fit is obtained by applying the first lag of y_t and the second lag of x_t as regressors. On the other hand, almost an equal in-sample fit is obtained by replacing the second lag of x_t by its third or fourth lag.

Next we analyze the additional predictive power obtained by including a lag of π_t in the regressor set of the models considered in table 1. The pseudo R^2 values obtained for these models are displayed in table 2, where the rows again vary according to the employed lag order of y_t (the results of the first row assume the restriction $\delta = 0$) while the columns vary according to the chosen lag order of x_t. As in the preceding case, the results indicate that the first lag of y_t is what counts in forecasting its future values and the second, third, and fourth lags of x_t tend to produce the best in-sample fit. We found no evidence that longer lags of π_t would improve the in-sample performance of the model.

As a whole, the in-sample results indicate that for each regressor there is just one or a couple of lag order choices

---

[6] We classify a given quarter as the first quarter of a recession period if its first month or the preceding quarter's second or third month is classified as the NBER business cycle *peak*, while we classify a given quarter as the last quarter of a recession period if its second or third month or the subsequent quarter's first month is classified as the NBER business cycle *trough*. All those quarters that are not included in a recession period are classified as expansion quarters. For the dates of the peaks and troughs see http://www.nber.org/cycles/.

[7] The average announcement delay is about twelve months. See http://www.nber.org/cycles/ for the actual announcement dates.

[8] Source: http://www.federalreserve.gov/releases/h15/data.htm.

[9] 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 0 except for the constant. The latter model is thought of as a benchmark and 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).

TABLE 2.—PSEUDO R^2 MEASURES OF IN-SAMPLE FIT FOR DIFFERENT
FORECASTING MODELS

				k			
		1	2	3	4	5	6
	—	.223	.286	.285	.264	.212	.124
	1	.378	.466	.407	.422	.392	.346
	2	.223	.288	.286	.264	.230	.142
l	3	.227	.289	.291	.278	.243	.162
	4	.230	.290	.300	.278	.243	.192
	5	.224	.286	.289	.278	.221	.156
	6	.224	.294	.286	.266	.214	.128

Notes. The table entries report the pseudo R^2 values for the model $P_{t-1}(y_t = 1) = \Phi(\omega + \alpha\pi_{t-1} + \delta y_{t-l} + \beta x_{t-k})$ with different choices of l and k. The results of the first row (denoted by —) are obtained under the restriction $\delta = 0$. All results are computed from the sample 1955:Q4–2005:Q4.

TABLE 3.—IN-SAMPLE ESTIMATION RESULTS FOR SELECTED MODELS

Predictor	(1)	(2)	(3)	(4)
Spread, x_{t-4}	−0.77	−0.55	−0.57	−0.64
	(0.21)	(0.19)	(0.20)	(0.20)
Recession, y_{t-1}		1.79		2.40
		(0.25)		(0.27)
Probability, π_{t-1}			0.44	−0.31
			(0.20)	(0.14)
Log likelihood	−65.28	−40.86	−53.74	−39.56
Pseudo R^2	0.23	0.41	0.26	0.42
AIC	58.68	43.86	56.74	43.56
SBC	61.98	48.81	62.00	50.20

Notes. The models are estimated using quarterly recession indicator values from 1955:Q4 to 2005:Q4. Robust standard errors are given in parentheses and are computed with the procedures of section III using the Parzen kernel function.
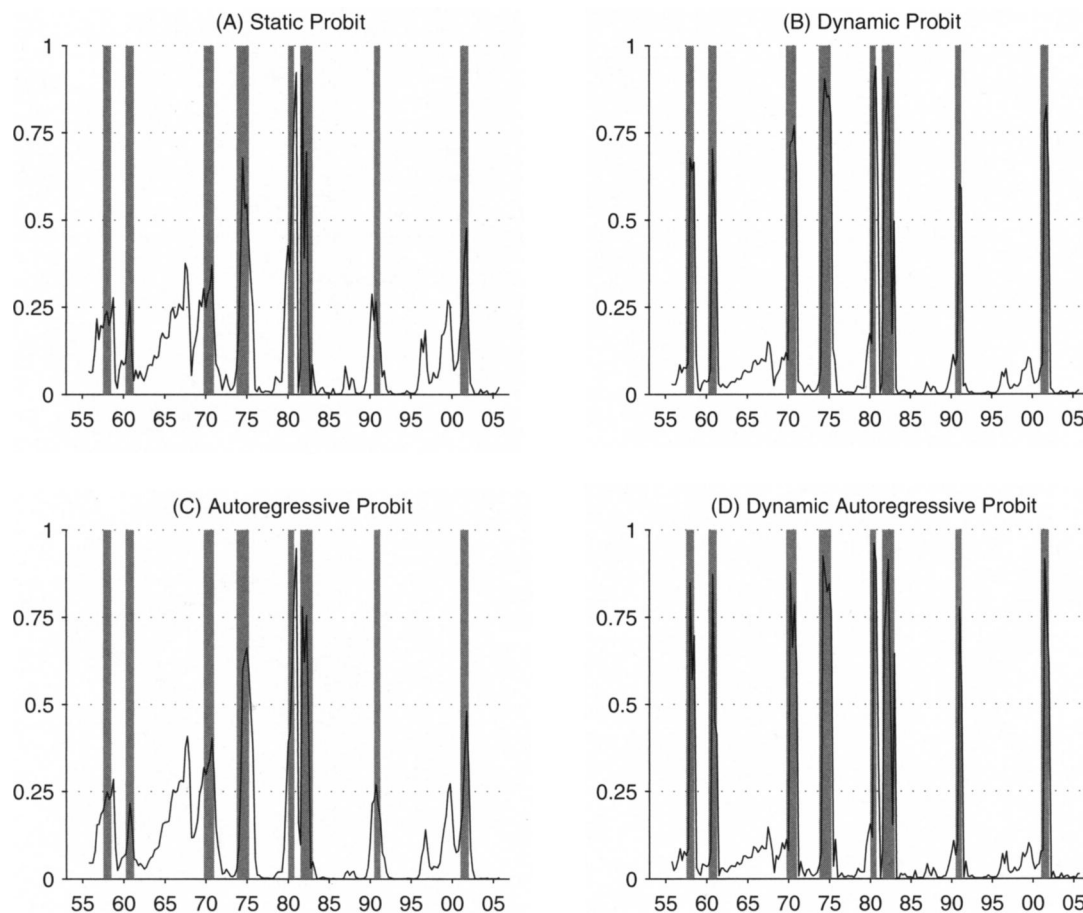
that produce the most accurate in-sample predictions independent of what lag orders are chosen for the other regressors. This particularly holds for the binary indicator y_t itself, for which the first lag is clearly superior independent of the employed lags of x_t and π_t. We also found this to be the case in out-of-sample forecasts over different forecast horizons. Recall that in out-of-sample forecasting, the use of the first lag of y_t entails computing multiperiod-ahead forecasts iteratively, while longer lags of y_t enable computing the corresponding forecasts directly. Thus, our results clearly indicate that "direct" forecasts based on a horizon-specific lag of y_t are inferior to "iterative" forecasts based on the first lag of y_t (numerical examples will be given in the next section in footnote 11).

In the case of the interest rate spread, x_t, we find that, in most cases, the second and fourth lags stand out producing the best or the second best in-sample performance. Thus if one wants to choose a single lag order, one is faced with the question of which one of the two one should use in making out-of-sample forecasts. In practice, the fourth lag of x_t has the advantage that it allows computing longer forecasts than the second lag, unless one is willing to forecast x_t itself. We leave the latter possibility for future research. Thus, our subsequent out-of-sample analysis compares x_{t-2} with x_{t-4} only in situations where the value of x_{t-2} is available. We regard the fourth lag of x_t as the benchmark case, because it allows making longer forecasts than the second lag, and because it is the best choice for the static model.

In what follows we take a closer look at the performance of the four baseline model specifications; their estimation results are summarized in table 3. Column 1 shows the results of the static probit model using the fourth lag of x_t as the sole regressor. Columns 2–4 present results for different dynamic specifications that each employ the same fourth lag of x_t as the driving predictor. The model of column 2 induces dynamics through the first lag of y_t, while the one of column 3 is based on the autoregressive formulation and uses the first lag of π_t, instead. Column 4 combines the two and includes both y_{t-1} and π_{t-1} as regressors in addition to x_{t-4}.

Some interesting findings emerge from table 3. First, in each case a negative and statistically significant estimate is

obtained for the coefficient of the interest rate spread.[10] This is theoretically expected and implies that the interest rate spread has predictive power that does not vanish even if dynamic features are included in the model. Yet, the results indicate that significant additional predictive power may be obtained by allowing for dynamic effects. In particular, the coefficient estimates of the lagged recession indicator are clearly statistically significant in the relevant models (columns 2 and 4). In both models, the likelihood of a recession in the current quarter would be much higher if the economy was already in a recession in the previous quarter than in the case where the economy was in an expansion. The coefficient estimate of π_{t-1} is also significant in the relevant models (columns 3 and 4). Interestingly, it is positive in the pure autoregressive probit model (column 3) but becomes negative when y_{t-1} is added to the regressor set. The positive estimate seems more natural, because it indicates that the probability of a recession in this quarter is higher, the higher the probability of a recession is in the previous quarter. The estimated negative coefficient in column 4 may indicate that the two variables y_{t-1} and π_{t-1} interact in a complicated way.

According to the values of the pseudo R^2 and the employed information criteria, the two models that use y_{t-1} as a regressor have the best in-sample fit, while the static model has the poorest, and the pure autoregressive model locates somewhere in between. Comparing the two best-fitting models, the Akaike information criterion is in favor of the less restricted specification, while the Schwarz Bayesian information criterion favors the more parsimonious specification without the autoregressive term.

To further illustrate the in-sample performance of different models, figure 1 shows estimated conditional probabilities that the economy is in a recession in a particular quarter given the relevant values of the regressors. These estimates are obtained from the models of table 3. They run from the beginning to the end of the sample and track the actual recession periods (shaded areas) with quite different pat-

[10] The robust standard errors are not sensitive to the choice of the bandwidth for the Parzen kernel function.

FIGURE 1.—PROBABILITY OF RECESSION, IN-SAMPLE PREDICTION

(A) Static Probit

(B) Dynamic Probit

(C) Autoregressive Probit

(D) Dynamic Autoregressive Probit

terns across models. First, notice that the fit of the static model matches rather poorly with the actual recessions. The estimated recession probabilities are more or less above 50% only for the 1974–1975, the 1980, and the 1981–1982 recessions. The pure autoregressive model seems to produce a somewhat similar fit. Compared with these two models, the dynamic models that employ the lagged recession indicator perform much better in that the estimated recession probabilities are much closer to the realized (two) values of the recession indicator.

In summary, the in-sample results indicate that the use of a lagged recession indicator can significantly improve on the in-sample performance of the static model and the pure autoregressive model.

An additional issue is whether the impact of the interest rate spread depends on the current state of the economy or the value of the recession indicator. To study this, we estimated the previous models with the additional regressor $y_{t-1}x_{t-4}$. The coefficient estimate of this regressor turned out to be statistically significant when added in the static specification, but it played no role in models that included the lagged value of the recession variable. These findings indicate that the lagged value of the recession indicator is what counts in forecasting future recessions.

B. Out-of-Sample Results

In this section, we examine the out-of-sample forecasting performance of the model specifications discussed in the previous section. As noted earlier, we must account for the fact that data on recessions become known with a delay. One approach to analyze real forecasting performance is to consider specific forecast situations and make reasonable assumptions about what is known at the time of forecasting. Such assessments are made in a recent paper by Chauvet and Potter (2005) and we do our analysis along these lines too.

As in Estrella and Mishkin (1998), we first consider a series of out-of-sample forecasts obtained from separate samples running from 1955:Q4–1977:Q4 to 1955:Q4–2003: Q4. Using each of the samples at a time, we estimate four models, and make forecasts one to eight quarters ahead from the last observation in the estimation period. For example, if the estimation period is 1955:Q4–1977:Q4, forecasts are made for eight subsequent quarters 1978:Q1–1979:Q4. Consistent with the average announcement delay of the NBER business cycle turning points, we assume that recession dates are known with a lag of one year (four quarters), while the value of the interest rate spread is known in real

TABLE 4.—PSEUDO R^2 MEASURES OF OUT-OF-SAMPLE FIT FOR DIFFERENT FORECASTING MODELS

Model specification	Forecast Horizon, Quarters							
	1	2	3	4	5	6	7	8
Static probit	.117	.055	.077	.083	.062	.059	.063	.022
Dynamic probit (x_{t-4})	.320	.126	.190	.258	.258	.261	.270	.269
Dynamic probit (x_{t-2})	.410	.196	.152	.175	.187	.189	—	—
Autoregressive probit	.164	.094	.109	.106	.086	.075	.077	.030
Dyn. autoregressive probit	.325	.110	.164	.221	.213	.206	.206	.201

Notes. The out-of-sample forecasts are based on estimation samples running from 1955:Q4–1977:Q4 to 1955:Q4–2003:Q4. Forecast horizon refers to the number of quarters from the last observation in the estimation sample. The value of the interest rate spread is assumed to be known four quarters in advance to the recession data (that is, recession data are announced with a delay of one year). The rows refer to the model $P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-1} + \alpha \pi_{t-1} + \beta z_t)$ under the following restrictions and choices: static probit $\delta = \alpha = 0$, $z_t = x_{t-4}$, dynamic probit (x_{t-4}) $\alpha = 0$, $z_t = x_{t-4}$, dynamic probit (x_{t-2}) $\alpha = 0$, $z_t = x_{t-2}$, autoregressive probit $\delta = 0$, $z_t = x_{t-4}$, and dyn. autoregressive probit $z_t = x_{t-4}$, respectively. Forecasts based on the dynamic models (the last four rows) are computed by the iterative formulas described in section IV and the appendix.

time. For example, if the estimation period is 1955:Q4–1977:Q4, the interest rate spread is known up to the quarter 1978:Q4. Therefore, all eight forecasts can be based on the fourth lag of the spread, which is our baseline choice (see previous section).

We use the maximum value of the likelihood function as the main criteria for evaluating out-of-sample forecasts and, to facilitate interpretation, we rescale it into a pseudo R^2 in the same way as in the case of the in-sample results in the previous section. Without a firm theoretical justification, Estrella and Mishkin (1998) also use the pseudo R^2 measure to evaluate out-of-sample forecasts, and we refer to their paper for a further discussion about the motivation of this approach. We also computed other summary measures of forecasting accuracy such as the quadratic probability score and the log probability score, which have been used in earlier studies (see Diebold & Rudebusch, 1989). However, the conclusions obtained with these alternative measures were the same as with the pseudo R^2 and, therefore, are not reported (they are available upon request).

Table 4 summarizes the out-of-sample forecast performance of the employed models. Different models are now listed on different rows, while the columns show results over different forecast horizons. If we compare the baseline specifications (rows 1,2 and 4,5) based on the fourth lag of x_t, the simple dynamic model outperforms its competitors across all forecast horizons, with the minor exception of one-quarter-ahead forecasts for which the more complicated model performs slightly better. However, the simple dynamic model is superior over longer forecast horizons. The pure autoregressive model produces more accurate out-of-sample forecasts than the static model, but the improvement in accuracy remains smaller than that obtained when the lagged recession indicator is added to the static model. Interestingly, the forecast accuracy of the simple dynamic model remains quite good even over longer forecast horizons, while especially the static model and the pure autoregressive model seem to lose their forecast accuracy rather quickly as the forecast horizon increases.[11]

In the previous section, we pointed out that a slightly better in-sample fit can be obtained by combining the first lag of the recession indicator with the second lag of the interest rate spread rather than with its fourth lag. To examine the difference between these two choices, the third row of table 4 reports values of the pseudo R^2 for the former specification as well. Notice that with the second lag of the interest rate spread we can forecast up to six quarters ahead. Comparing the figures on the second row of table 4 with those on the third row (forecast horizons 1–6) reveals that forecasts based on the second lag of the interest rate spread can be more accurate than obtained with the fourth lag. However, the advantage seems to be limited to one-quarter-ahead and two-quarters-ahead forecasts, while in longer-term forecasts the fourth lag of the interest rate spread performs better. This observation suggests that the fourth lag of the interest rate spread is the most robust single choice.[12]

The results of table 4 strongly indicate that the apparent serial dependence of the recession indicator can be successfully used to improve recession predictions at least as far as eight quarters ahead. Recall that the underlying dynamic forecasts are based on an iterative procedure, where multi-period-ahead forecasts depend on forecasts of the preceding states. If the model asserts a high probability for a recession in the next quarter, then it does so also for the quarter that follows. This gives some intuition as to why dynamic forecasts can improve static forecasts that do not take previous states (and their serial dependence) into account.

Figure 2 illustrates the out-of-sample predictive performance of different models when the forecast horizon is two quarters ahead. The predictions of figure 2 derive from the model specifications in figure 1, but now each forecast is based on data available at the time the forecast is made.

---

[11] Above we pointed out that according to in-sample results iterative forecasts should be superior to direct forecasts that apply horizon-specific lag orders for the predictors. This conclusion is indeed warranted. For example, if we make the same pseudo out-of-sample analysis as in table

4 by applying direct forecasts based on the specification $P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-h} + \beta x_{t-4})$, where h is the forecast horizon, the corresponding pseudo R^2 measures of out-of-sample fit would be 0.046 for $h = 2$, 0.0025 for $h = 3$, and a negative number for any longer forecast horizons. Thus, direct forecasts tend to have the worst out-of-sample performance compared with any of the considered alternative forecasting procedures.

[12] We note that, in all of our forecasting exercises, the second, third, and fourth lag of x_t beat all other lag choices independent of the forecast horizon.

FIGURE 2.—PROBABILITY OF RECESSION TWO QUARTERS AHEAD, OUT-OF-SAMPLE PREDICTION

(A) Static Probit

(B) Dynamic Probit

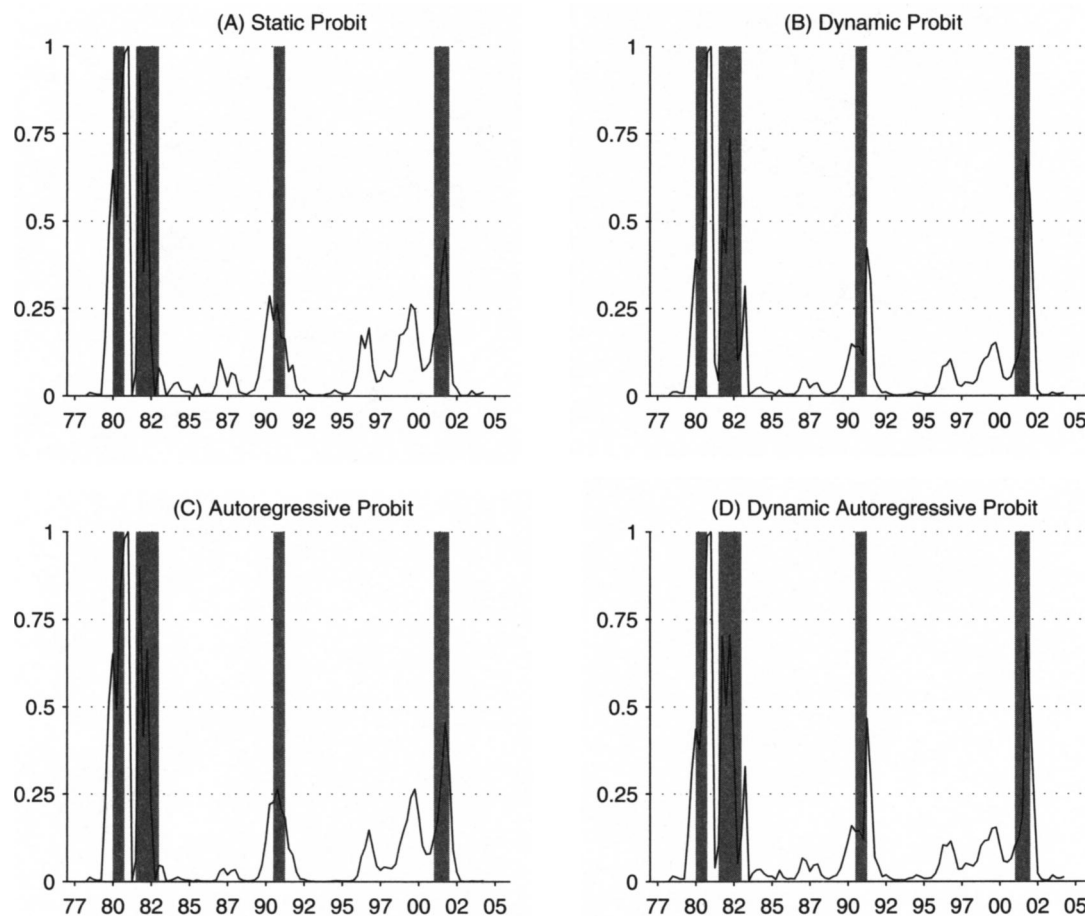(C) Autoregressive Probit
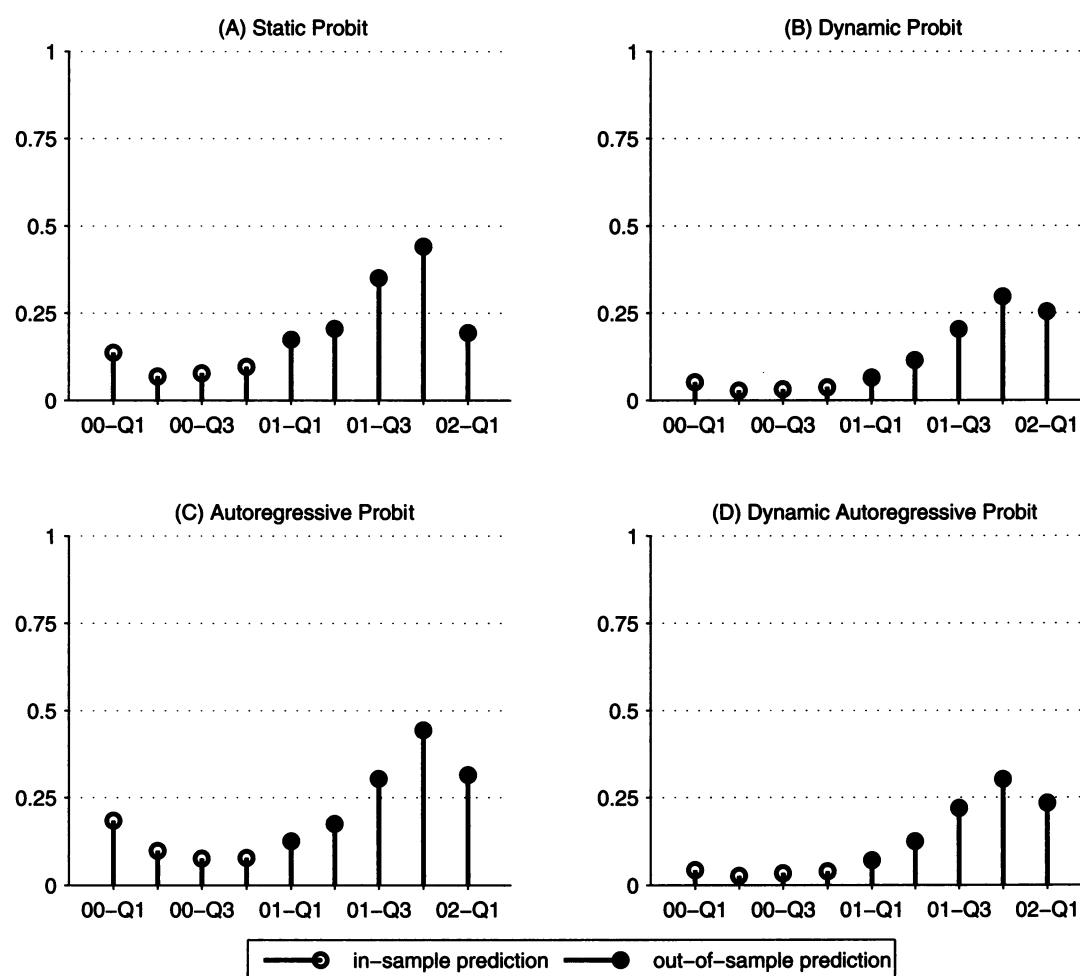
(D) Dynamic Autoregressive Probit

Figure 2 indicates that all four models have predictive content for the two recessions in the beginning of the 1980s, while none produces clear early warnings of the last two recessions. This has been a common finding in previous studies, too. Note, however, that in general the estimated probabilities based on the two dynamic models that use the lagged recession indicator as a predictor match better with the realized recession indicators than the estimated probabilities based on the static model and the pure autoregressive model.

Next we consider more detailed forecasts around the 2001 recession, the latest recession known by September 2006. The situation we analyze is similar to the baseline case of Chauvet and Potter (2005, p. 94). They assumed that in March 2001 market participants were aware that the U.S. economy had not been in a recession by the end of 2000, while there was considerable uncertainty whether the expansion was to continue from 2001:Q1 on. In the end, the NBER business cycle committee declared that the economy was in a recession from 2001:Q2 until 2002:Q1. The task here is to check to what extent our benchmark models would have been able to forecast these recession quarters in real time. To this end, forecasts are first made for the period 2001:Q1–2002:Q1 using information that was known in 2001:Q1 (March 2001).

Figure 3 depicts the estimated recession probabilities for individual quarters in 2001:Q1–2002:Q1. The four stems with unfilled circles denote in-sample figures (quarters 2000:Q1–2000:Q4), and the following five stems with filled circles denote out-of-sample figures (quarters 2001:Q1– 2002:Q1). The simple static model and the pure autoregressive model produce rather similar results. The estimated recession probabilities obtained with these models are higher than those obtained with the two specifications that use a lagged recession indicator. Thus, in this situation, the former models give stronger signals of a forthcoming recession than the latter models. On the other hand, the dynamic models with a lagged recession indicator give much smaller recession probabilities in the estimation period than the static model and the pure autoregressive model (see quarters 2000:Q1–2000:Q4). This indicates that a potential weakness of the static model and the pure autoregressive model is that they tend to produce relatively strong recession signals even when the expansion continues or only slows down without turning into a recession (for a similar point, see Chauvet & Potter, 2005). Looking at this from another angle, the dynamic models with a lagged recession indicator have the advantage that there is usually a sharper difference between recession signals for

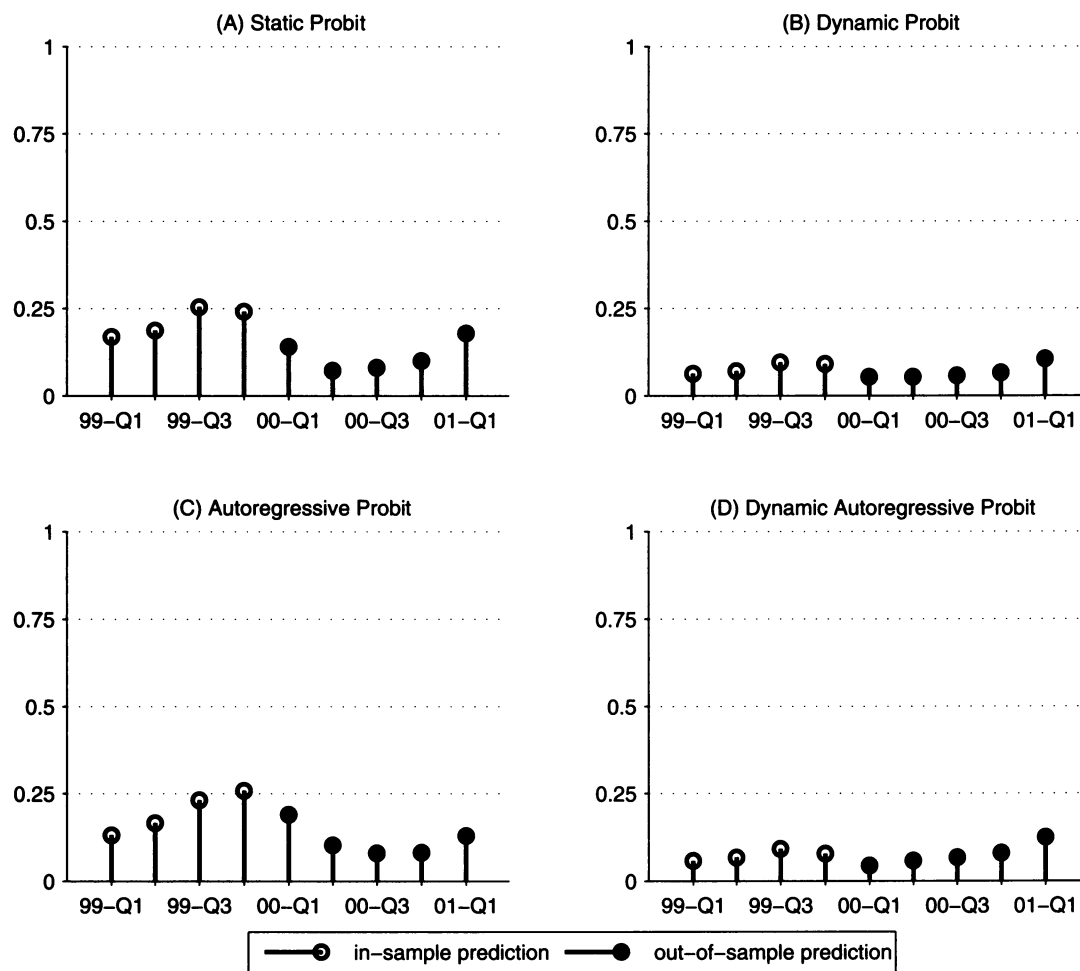FIGURE 3.—PROBABILITIES OF RECESSION BASED ON FOUR MODEL SPECIFICATIONS

expansions and actual recessions than in the case of the static model and the pure autoregressive model.

As a further illustration, we do a similar analysis for the period a year before. Thus, we make forecasts for the period 2000:Q1–2001:Q1 using observations of the recession indicator until 1999:Q4 and observations of the interest rate spread until 2000:Q1. This setup is again consistent with one of Chauvet and Potter (2005, p. 95). Figure 4 plots the corresponding estimates for the in-sample and out-of-sample recession probabilities. Clearly, the static model and the pure autoregressive model give much stronger recession signals for specific quarters than the two dynamic models with a lagged recession indicator. Given that no recession took place during these quarters, the former models would have produced rather fuzzy signals. It is difficult to see what one would conclude from such fluctuating predictions in a real forecasting situation. By contrast, the dynamic models with a lagged recession indicator would have produced a fairly stable pattern of weak recession signals, the overall conclusion being much clearer. This suggests that in a forecasting situation it is useful to check how the out-of-sample prediction compares with previous periods' in-sample predictions.

An alternative way to summarize information about the likelihood of future recessions is to apply the hitting probabilities defined in section IV. Table 5 reports real-time forecasts for the probability of continued expansion for five different periods. Two of the periods are the same as above, while the remaining three are similar to those considered by Chauvet and Potter (2005, p. 95). The underlying probabilities of continued expansion are defined as the complement of the cumulative distribution of the hitting probabilities of the individual quarters in the selected period. The general observation from these measures is roughly the same as above. According to the static model and the pure autoregressive model there was about a one in five chance that the economy remains in expansion from 2001:Q1 until 2002: Q1, while the dynamic models with a lagged recession indicator would have predicted a chance of two in five. Thus, again, the static model and the pure autoregressive model would have asserted a higher probability of a forthcoming recession. A somewhat similar picture is obtained from the period 1990:Q1–1991:Q1 including the 1990–1991 recession. On the other hand, for the remaining periods 1989:Q1–1990:Q1, 1999:Q1–2000:Q1, and 2000:Q1–2001:Q1 that do not include recession quarters, the dynamic

FIGURE 4.—PROBABILITIES OF RECESSION BASED ON FOUR MODEL SPECIFICATIONS

models with a lagged recession indicator provide higher probabilities of continued expansion than the static model and the pure autoregressive model. Thus the former models seem to distinguish true recessions from slowdowns better than the latter models.

Finally, we make genuine real-time recession forecasts up to the second quarter of 2007. The employed models and parameter estimates, based on the full sample, are given in table 3. The estimated recession probabilities are plotted in figure 5 and, consistent with the recent decline of the interest rate spread, they all steadily increase from the beginning of the year. Interestingly, the recession signals are rather similar across different models, whether we look at in-sample or out-of-sample figures. The estimated probability of continued expansion until 2007:Q2 is 0.34 for the static model, 0.44 for the pure autoregressive model, and 0.67 for both of the remaining models. In light of these figures, there is a marked risk that the current expansion period is about to end soon.
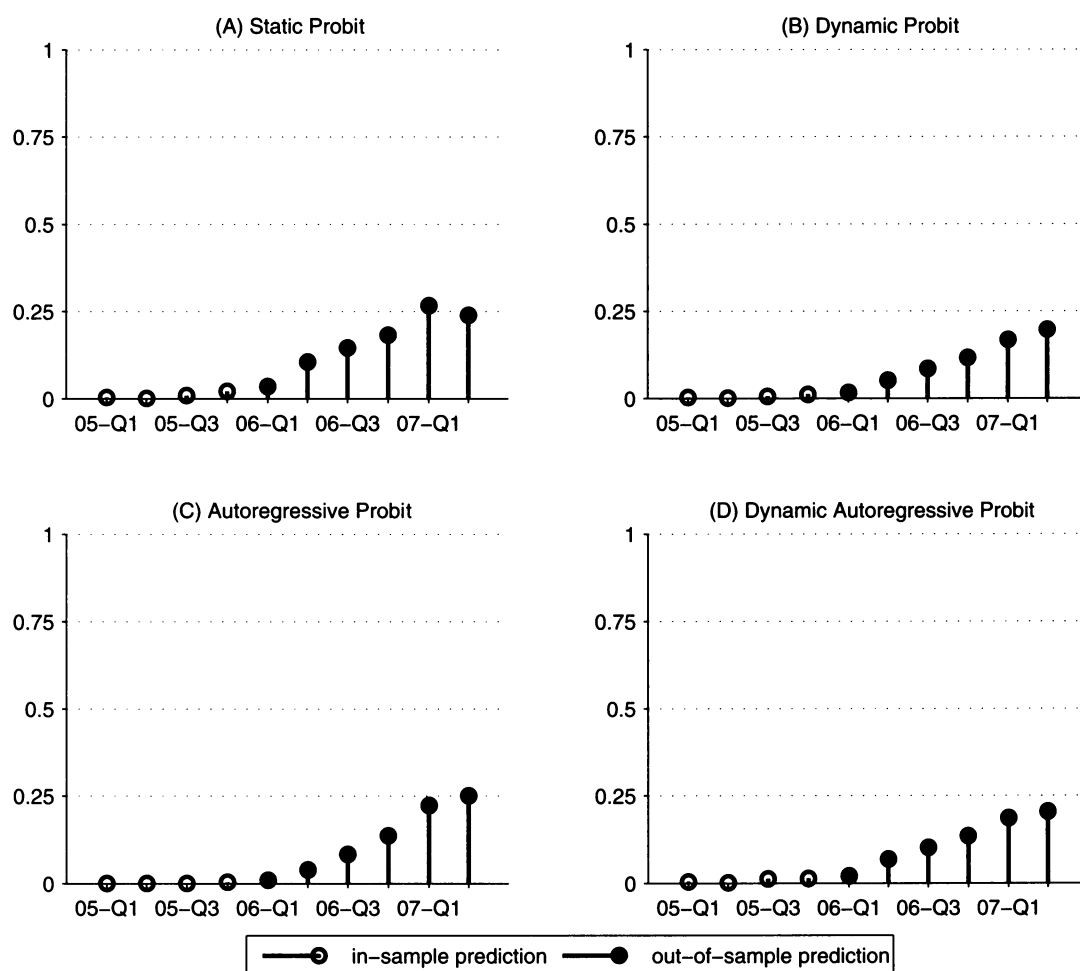
VI.   Conclusions

This paper introduces new dynamic models for binary time series and applies them to forecasting a binary

TABLE 5.—PROBABILITIES OF CONTINUED EXPANSION FOR SELECTED PERIODS

Model specification	Period				
	89:Q1–90:Q1	90:Q1–91:Q1	99:Q1–00:Q1	00:Q1–01:Q1	01:Q1–02:Q1
Static probit	.65	.27	.31	.55	.19
Dynamic probit	.85	.65	.67	.80	.42
Autoregressive probit	.82	.30	.32	.56	.19
Dyn. autoregressive probit	.83	.66	.70	.80	.61

Notes. The reported probabilities of continued expansion are obtained using spread data up to the first quarter of the indicated forecast period, while the underlying model estimates are based on recession data up to the preceding quarter (1988:Q4 for the first forecast period, and so on). The model specifications refer to the equation $P_{t-1}(y_t = 1) = \Phi(\omega + \delta y_{t-1} + \alpha \pi_{t-1} + \beta x_{t-4})$ with restriction $\delta = \alpha = 0$ for static probit, $\alpha = 0$ for dynamic probit, $\delta = 0$ for autoregressive probit, and no restriction for dyn. autoregressive probit, respectively. For the dynamic models (the last three rows) the underlying forecasts are computed by the iterative formulas described in section IV and the appendix.

FIGURE 5.—PROBABILITIES OF RECESSION IN THE NEAR FUTURE

indicator of U.S. recessions. The parameters of the new models can be straightforwardly estimated by ML, and misspecification-robust standard errors can be readily computed. The paper also shows how multistep-ahead forecasts can be computed iteratively for all horizons. This approach offers an alternative to the previously used direct approach in which the lag order of the binary response is always tailored to match the forecast horizon.

In our application to forecasting U.S. recessions we compare different probit models and different forecasting procedures in terms of in-sample and out-of-sample performance by using the interest rate spread as a sole exogenous predictor. According to our results, dynamic probit models outperform the static model in terms of both in-sample and out-of-sample predictions, and dynamic models with lagged values of the binary response variable outperform models in which dynamics enter only through a lagged probit probability. Iterated multiperiod-ahead forecasts based on a one-period-ahead model are found to be superior to previously used direct forecasts. Unlike in previous studies, we also find that the lag order chosen for the interest rate spread can have a substantial impact on forecasting accuracy. Our experience is that experimenting with different lags and

using statistical model selection procedures is preferable to the commonly used practice to choose lags to match the forecast horizon. Finally, our results support the previous finding that, irrespective of the employed model, the very recent U.S. recessions were harder to predict than the earlier ones.

The current research can be extended in several directions. In this paper we have focused on comparing different dynamic models in which the interest rate spread is the only exogenous predictor. It might be worthwhile to also consider other previously used exogenous predictors such as stock returns. Another issue of interest concerns the treatment of the exogenous predictor. It would be interesting to examine whether more accurate long-term forecasts could be obtained by using a forecast of the exogenous predictor rather than its realized value known at the time of forecasting. Finally, there are different ways to generalize the process π_t. These topics are left for future research.

REFERENCES

Akaike, Hirotugu, "Information Theory and an Extension of the Maximum Likelihood Principle" (pp. 267–281), in Boris N. Petrov and

Frigyes Csáki (Eds.), *2nd International Symposium on Information Theory* (Budapest, Hungary: Akademiai Kiado, 1973).

——— "A New Look at the Statistical Identification Model," *IEEE: Trans. Auto Control* AC-19 (December 1974), 713–723.

Bhansali, Rajendra J., "Parameter Estimation and Model Selection for Multistep Prediction of a Time Series: A Review" (pp. 201–225), in Subir Ghosh (Ed.), *Asymptotics, Nonparametrics, and Time Series* (New York: Marcel Dekker, 1999).

Chauvet, Marcelle, and Simon Potter, "Forecasting Recessions Using the Yield Curve," *Journal of Forecasting* 24 (March 2005), 77–103.

Davidson, James, *Econometric Theory* (Oxford, UK: Blackwell Publishers, 2000).

de Jong, Robert M., and Tiemen M. Woutersen, "Dynamic Time Series Binary Choice," *Econometric Theory* (forthcoming).

Diebold, Francis X., and Glenn D. Rudebusch, "Scoring the Leading Indicators," *Journal of Business* 62 (July 1989), 369–391.

Dueker, Michael, "Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions," *Federal Reserve Bank of St. Louis Economic Review* 79 (March/April 1997), 41–51.

——— "Dynamic Forecasts of Qualitative Variables: A Qual VAR Model of U.S. Recessions," *Journal of Business & Economic Statistics* 23 (January 2005), 96–104.

Estrella, Arturo, "A New Measure of Fit for Equations with Dichotomous Dependent Variables," *Journal of Business & Economic Statistics* 16 (April 1998), 198–205.

Estrella, Arturo, and Frederic S. Mishkin, "Predicting U.S. Recessions: Financial Variables as Leading Indicators," this REVIEW 80 (February 1998), 45–61.

Eviews 5 User's Guide (Irvine, CA: Quantitative Micro Software, 2004).

Filardo, Andrew J., "How Reliable Are Recession Prediction Models?" *Federal Reserve Bank of Kansas Economic Review* 84 (second quarter 1999), 35–55.

Florio, Anna, "The Asymmetric Effects of Monetary Policy," *Journal of Economic Surveys* 18 (July 2004), 409–426.

Geweke, John, and Charles Whiteman, "Bayesian Forecasting" (pp. 3–80), in Graham Elliott, Clive W. J. Granger, and Allan Timmerman (Eds.), *Handbook of Economic Forecasting* (Amsterdam: Elsevier, 2006).

Marcellino, Massimiliano, James H. Stock, and Mark W. Watson, "A Comparison of Direct and Iterated AR Methods for Forecasting Macroeconomic Time Series," *Journal of Econometrics* 135 (November–December 2006), 499–526.

Moneta, Fabio, "Does the Yield Spread Predict Recessions in the Euro Area?" ECB working paper no. 294 (December 2003).

Newey, Whitney K., and Kenneth D. West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55 (May 1987), 703–708.

Rydberg, Tina H., and Neil Shephard, "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models," *Journal of Financial Econometrics* 1 (spring 2003), 2–25.

Schwarz, Gideon, "Estimating the Dimension of a Model," *Annals of Statistics* 6 (March 1978), 461–464.

Tong, Howell, *Non-Linear Time Series Analysis: A Dynamic System Approach* (London: Oxford University Press, 1990).

APPENDIX

Details on Iterated Multiperiod Forecasts

This appendix shows how iterated multiperiod-ahead forecasts are computed for probit models in which dynamics enter through the first lag of the binary response alone or together with the first lag of the probit probability. Iterated multiperiod forecasts based on models with more lags are formed using the same principle, but involve somewhat more complex formulas and notation.

Consider making an h-period-ahead forecast using the dynamic probit model with the first lag of the binary response. Proceeding as in section IV in the case $h = 2$, we obtain

$$P_{t-h}(y_t = 1) = E_{t-h}(\Phi(\omega + \delta y_{t-1} + \beta x_{t-h}))$$
$$= \sum_{y_{t-1} \in \{0,1\}} P_{t-h}(y_{t-1})\Phi(\omega + \delta y_{t-1} + \beta x_{t-h}),$$

where $P_{t-h}(y_{t-1})$ defines the probabilities for $y_{t-1} = 1$ and $y_{t-1} = 0$ conditional on information known in the forecast period $t - h$. To compute $P_{t-h}(y_{t-1})$ we must account for all possible paths of realizations of $y_{t-h+1}, y_{t-h+2}, \ldots, y_{t-2}$ that lead to a given value of y_{t-1}. Define the vector notation

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

and the Cartesian product $B_k = \{0, 1\}^k$ for $k = 1, 2, \ldots$. In other words, the set B_k contains all possible $k \times 1$ vectors with components either 0 or 1 ($k = 1, 2, \ldots$). Notice that $y_{t-h+k}|\mathscr{F}_{t-h-1+k} \sim B(p_{t-h+k})$ with

$$p_{t-h+k} = \Phi(\omega + \delta y_{t-h-1+k} + \beta x_{t-2h+k}).$$

Then, we have

$$P_{t-h}(y_{t-1}) = \sum_{y_{t-h+1}^{t-1} \in B_{h-1}} \prod_{j=1}^{h-1} (p_{t-h+j})^{y_{t-h+j}}(1 - p_{t-h+j})^{(1-y_{t-h+j})}.$$

Next, consider making an h-period-ahead forecast using the autoregressive formulation (14). In the same way as in section IV, a repetitive substitution of equation $\pi_t = \omega + \alpha \pi_{t-1} + \delta y_{t-1} + \beta x_{t-h}$ leads to the expression

$$E_{t-h}(y_t) = E_{t-h}\Phi\left(\alpha^h \pi_{t-h} + \sum_{j=1}^{h} \alpha^{j-1}(\omega + \delta y_{t-j} + \beta x_{t-h+1-j})\right)$$
$$= \sum_{y_{t-h+1}^{t-1} \in B_{h-1}} P_{t-h}(y_{t-h+1}^{t-1})\Phi\left(\alpha^h \pi_{t-h} + \sum_{j=1}^{h} \alpha^{j-1}\right.$$
$$\left. \times (\omega + \delta y_{t-j} + \beta x_{t-h+1-j})\right),$$

where π_{t-h} is a function of past values of x_t and the initial value π_0. Note that $P_{t-h}(y_{t-h+1}^{t-1})$ defines probabilities for all possible realizations of $y_{t-h+1}, y_{t-h+2}, \ldots, y_{t-1}$ conditional on the information set available in period $t - h$. Because now $y_{t-h+k}|\mathscr{F}_{t-h-1+k} \sim B(p_{t-h+k})$ with

$$p_{t-h+k} = \Phi\left(\alpha^k \pi_{t-h} + \sum_{j=1}^{k} \alpha^{j-1}(\omega + \delta y_{t-h-1+j} + \beta x_{t-2h+j})\right)$$

we have

$$P_{t-h}(y_{t-h+1}^{t-1}) = \prod_{j=1}^{h-1} (p_{t-h+j})^{y_{t-h+j}}(1 - p_{t-h+j})^{(1-y_{t-h+j})}.$$