Research Reports
Kansantaloustieteen tutkimuksia, No. 122:2010
Dissertationes Oeconomicae

HENRI NYBERG

STUDIES ON BINARY TIME SERIES MODELS WITH APPLICATIONS TO EMPIRICAL MACROECONOMICS AND FINANCE

ISBN: 978-952-10-5352-8 (nid.)
ISBN: 978-952-10-5353-5 (pdf)
ISSN: 0357-3257

Contents

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible.

It is difficult to overstate my gratitude to my thesis supervisors, Professor Markku Lanne and Professor Pentti Saikkonen, for excellent supervising and support throughout my postgraduate studies. They have supervised this thesis patiently and given important advice during this work. Their encouragement and guidance to the academic world have provided the best possible basis for this thesis.

I wish to thank the official examiners of my thesis, Professor Denise Osborn and Professor Markku Rahiala, for many insightful comments and interesting suggestions for future research. I am also grateful to Professor Heikki Kauppi and Dr. Pekka Pere for examining my Licenciate thesis in 2007–2008. Their comments have been valuable when writing this thesis. I have also greatly benefited numerous discussions and comments to my presentations in seminars and conferences organized in Helsinki, elsewhere in Finland and abroad.

I had a great privilege of writing this thesis while working as a full time doctoral student at the Department of Economics (2007–2009) and at the Department of Political and Economic Studies and its Economics discipline (from January 2010) in the University of Helsinki. Most of that time I have worked in the research project of "Econometrics of Macroeconomics and Finance, and the Interface between the Macroeconomy and Financial Markets" financed by the Academy of Finland. I am very grateful to all people who gave me these excellent research facilities. I also wish to thank the staff and my colleagues at the Economics discipline for a friendly and inspirational working environment.

The financial support from the Academy of Finland, the Research Foundation of the University of Helsinki, the Okobank Group Research Foundation and the Finnish Foundation for Advancement of Securities Markets is gratefully acknowledged. Their support have made it possible to concentrate on research and attend interesting international conferences, such as the Econometric Society meetings in Milan (2008) and Canberra (2009) as well as conferences in Luxembourg (2008), Manchester (2009) and Lund (2009), which have been useful when writing this thesis.

Finally, I would like to thank my family for support and encouragement in my studies ever since I started at the university.

Helsinki, March 2010

Henri Nyberg

Chapter 1

Introduction

1.1 Background

In recent decades, nonlinear time series models have attracted serious attention in the econometric literature. Various nonlinear models have been proposed to describe potential nonlinear characteristics in the underlying data generation process. Threshold (TAR) models (see, e.g., Tong, 1990), smooth transition (STAR) models (see, e.g., Granger and Teräsvirta, 1993), Markov switching models (Hamilton, 1989), as well as ARCH (Engle, 1982) and GARCH models (Bollerslev, 1986) are examples of nonlinear time series models that have been applied in numerous economic applications (see also, for example, the book by Franses and van Dijk, 2000).

In all of the above-mentioned models the dependent variable is "continuous" indicating that its observed value can, in principle, be any real number. However, the dependent variable can also be qualitative with only a limited number of possible outcomes. This is the case, for example, if the dependent variable is binary or a count variable. Models of this type are referred to as qualitative response models in the econometric literature. Qualitative time series models can also be seen as generalizations of the generalized linear models designed for cross-sectional data considered extensively in the previous literature in statistics.

The objective of this thesis is to consider time series models with a binary dependent variable. There are various potential empirical applications of these models. In this thesis, we are interested in applications to empirical macroeconomics and finance. For example, forecasting the recession periods of the economy or the signs of stock market returns are of interest for many economic decision makers. Policymakers, such as central banks and governments are interested in predicting the probability of a recession in the future or assessing the probability that the economy is already in recession. Correspondingly, financial investors can benefit from probability forecasts of the future developments in financial markets in making their investment decisions. For instance, binary time series models can be used to evaluate the probability that the stock market return is positive in the next period.

In this thesis, the main interest is on different dynamic extensions of the traditional "static" univariate probit model commonly used in the previous literature. The static model can be extended in various ways. Models with the so-called autoregressive model structure (see Kauppi and Saikkonen, 2008) are of particular interest throughout the thesis.

The rest of this introductory chapter consists of a survey of binary time series models and their economic applications. In addition, summaries of the contents of the four studies to be presented in Chapters 2–5 are provided. In Section 1.2, different binary time series models are discussed. Although the main interest is on different dynamic extensions to the static univariate probit model, a few closely related models, such as multinomial models and count data models, are also briefly reviewed. Section 1.3 provides an introduction to the empirical applications considered in Chapters 2–5. As in much of the previous literature, forecasting the recession periods of the economy is the main empirical application examined in the thesis, especially in Chapter 2. Finally, Section 1.4 gives a short summary of the main results and contributions of the thesis.

1.2 Binary and Qualitative Time Series Models

Many economic applications are concerned with variables whose range is discrete or limited. Time series of discrete directions of change constructed from observed continuous variables are also often studied for reasons of forecasting. For instance, in empirical finance, the directional predictability of financial asset return is suggested to be more closely related to the profitability than other statistical goodness-of-fit measures (see, e.g., Leitch and Tanner, 1991).

Qualitative response models, also known as discrete or categorical models, have been used in various cross-sectional and panel data applications. Amemiya (1980), Maddala (1983), Gourieroux (2000) and Wooldridge (2002, chapters 15 and 19) provide comprehensive introductions to these models from the viewpoint of econometric applications. In applied microeconometrics, the qualitative dependent models for cross-sectional and panel data are predominantly employed, as emphasized by Cameron and Trivedi (2005, chapter IV). In the context of panel data Honore and Kyriazidou (2000) and Honore and Lewbel (2002), among others, have proposed various dynamic models for binary dependent variables.

The models for cross-sectional and panel data mentioned above are closely related to the binary time series models considered in this thesis. In statistics, similar models have been studied under the heading of generalized linear models (see, e.g., McCullagh and Nelder, 1989). Li (1994), Shephard (1995) and Benjamin, Rigby and Stasinopoulos (2003), among others, have considered generalized linear models designed for time series. Overall, the literature on dynamic models for binary time series is scant, but a number of new models have recently been suggested. Some of these models are surveyed in more detail in the next section.

1.2.1 Static Univariate Probit Model

The simplest example of a qualitative time series model is a model where the dependent variable is binary. For simplicity, and without loss of generality, let 0 and 1 denote the values of the dependent variable. Typically, the value 1 indicates that some event occurs, and 0 that it does not occur. Now, let y_t, $t = 1, 2, ..., T$, be a univariate binary time series. We denote by Ω_{t-1} the information set containing, for example, lagged values of y_t and explanatory variables. Conditional on the information set Ω_{t-1}, y_t follows a Bernoulli distribution

$$y_t|\Omega_{t-1} \sim B(p_t), \tag{1.1}$$

where, according to the properties of Bernoulli distribution,

$$p_t = E_{t-1}(y_t) = P_{t-1}(y_t = 1). \tag{1.2}$$

In this expression, $E_{t-1}(\cdot)$ and $P_{t-1}(\cdot)$ signify the conditional expectation and the conditional probability given the information set Ω_{t-1}, respectively.

A probit model is obtained by specifying the conditional probability p_t in (1.2) as

$$p_t = \Phi(\pi_t), \tag{1.3}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and π_t is a linear function of variables included in the information set Ω_{t-1}. A logit model, where the function $\Phi(\cdot)$ is replaced by the logistic function, is an alternative specification.[1] In empirical applications, results obtained with logit and probit models usually yield very similar results (see, e.g., Maddala, 1983, 23; Davidson and MacKinnon, 1993, 516). In this thesis, we concentrate on probit models.

In a "static" probit model, the variable π_t is specified as

$$\pi_t = \omega + \boldsymbol{x}_{t-1}'\boldsymbol{\beta}, \tag{1.4}$$

where ω is a constant term and $\boldsymbol{x}_{t-1}'$ contains the explanatory variables. This specification has mainly been employed in the previous literature.[2] It is "static" in the sense that explanatory variables have an immediate effect on the conditional probability p_t which does not change unless values of the explanatory variables change.

[1] In this context, a classical linear model designed for continuous dependent variables is referred to the linear probability model (see, e.g., Maddala 1983, 15–16). However, it has various weaknesses (see Gourieroux 2000, 6–8). For example, in this binary case, the conditional probability p_t in (1.2) is not necessarily between zero and one.

[2] Note that in Chapter 3 we denote $\pi_t(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of parameters.

1.2.2 Extensions to the Static Probit Model

A limitation of the static model specification (1.4) is that it does not allow for the potential autocorrelation in y_t. Therefore, various dynamic extensions have been proposed, and we review some of them in this section.

An obvious dynamic extension of the static specification (1.4) is obtained by augmenting it with a lagged value of y_t. Cox (1981), Zeger and Qaqish (1988), Shephard (1995), and Dueker (1997), among others, have considered the "dynamic" model

$$\pi_t = \omega + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{1.5}$$

where a positive value of δ_1 implies that y_t tends to take the same value respectively. This kind of "clustering effect" appears quite plausible in many applications. In model (1.5), only one lagged value of y_t is included, but the model can be made more general by increasing the number of lags (see, e.g., Cox, 1981; Kaufmann, 1987).

In this thesis, the main interest is on probit models with an autoregressive structure. Following Kauppi and Saikkonen (2008), consider the model

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{1.6}$$

where the condition $|\alpha_1| < 1$ is assumed. The inclusion of the lagged value π_{t-1} on the right hand side yields a first-order autoregressive structure for π_t. Therefore, Kauppi and Saikkonen (2008) called model (1.6) the "dynamic autoregressive model".[3] On the other hand, Anatolyev (2009) refers to this model as the "generalized autoregressive model". This name is motivated by the analogy of the model to the GARCH model (Bollerslev, 1986) and the ACD model (Engle and Russell, 1998) used for continuous dependent variables.[4] For simplicity, only the first lag of π_t is included in (1.6), but it is also possible to have several lags of π_t in the model.

[3] Throughout this thesis we use the same terminology as in Kauppi and Saikkonen (2008).

[4] A model somewhat similar to (1.6) has been proposed by Russell and Engle (2005) for a multinomial dependent variable. However, these authors replace π_{t-1} by $p_{t-1} = \Phi(\pi_{t-1})$ on the right hand side of (1.6).

Using recursive substitution and the assumption $|\alpha_1| < 1$ in model (1.6) we obtain the equation

$$\pi_t \;=\; \sum_{i=1}^{\infty} \alpha_1^{i-1}\omega + \delta_1 \sum_{i=1}^{\infty} \alpha_1^{i-1} y_{t-i} + \sum_{i=1}^{\infty} \alpha_1^{i-1} \boldsymbol{x}_{t-i}^{'} \boldsymbol{\beta}, \tag{1.7}$$

which shows that π_t can be expressed in terms of the infinite history of y_t and the explanatory variables. This suggests that a model with an autoregressive structure (i.e. a model including π_{t-1}) may be a parsimonious alternative in cases where a large number of lagged values of y_t and the explanatory variables appear necessary.

A special case of model (1.6) is obtained by restricting the coefficient δ_1 to zero. This restriction yields the "autoregressive" model,

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \boldsymbol{x}_{t-1}^{'} \boldsymbol{\beta}, \tag{1.8}$$

where $|\alpha_1| < 1$ and the predictive power comes completely from the infinite history of the explanatory variables (cf. (1.7)).

In the previous literature, the so-called "generalized linear autoregressive" (GLAR) models closely related to model (1.6) have also been considered. Shephard (1995) formulated a class of GLAR models which also contains a moving average term. These models are referred to as GLARMA models.[5] Further, Rydberg and Shephard (2003) proposed the model

$$\pi_t = \omega + \boldsymbol{x}_{t-1}^{'} \boldsymbol{\beta} + g_t, \tag{1.9}$$

where

$$g_t = a_1 g_{t-1} + \delta_1 y_{t-1}.$$

It is also possible to augment the expression of g_t by including a moving average term based on the difference between y_{t-1} and $p_{t-1} = \Phi(\pi_{t-1})$ in the model. Note that Kauppi (2008) suggested the following alternative to (1.9)

$$\pi_t = \omega + \delta_1 y_{t-1} + g_t, \tag{1.10}$$

[5] Benjamin *et al.* (2003) derive very closely related "GARMA" (Generalized Autoregressive Moving Average) model. See also the multinomial models by Russell and Engle (2005) and Liesenfeld, Nolte and Pohlmeier (2006).

where

$$g_t = a_1 g_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}.$$

Compared with model (1.6), in models (1.9) and (1.10), the autoregression is related to the expression for g_t instead of π_t.

Model (1.6) and its special cases discussed above can be estimated by the method of maximum likelihood (ML).[6] Recently, de Jong and Woutersen (in press) have shown that, under appropriate regularity conditions, the conventional large sample theory of ML estimation applies to the dynamic probit model (1.5). Extending this theory to models with autoregressive structure, such as (1.6) and (1.8), still remains to be done.

An important advantage of model (1.6) and its special cases is that one-period and multiperiod forecasts can be computed by explicit formulae (see Kauppi and Saikkonen, 2008). This is in contrast to many other nonlinear time series models and alternative dynamic extensions of the static probit model (1.4) to be discussed in Section 1.2.3. This is a very useful property, especially in this thesis, because in our empirical applications forecasting binary time series is of interest (see Chapters 2, 4 and 5).

To illustrate the properties of the introduced probit models, we simulate realizations from the dynamic autoregressive model (1.6) and its special cases. In the general case, we use the process

$$\pi_t = 0.1 + 0.70\pi_{t-1} + 0.5y_{t-1} - 0.20x_{t-1}, \tag{1.11}$$

where the simulated explanatory variable x_t follows the AR(1) process

$$x_t = 0.1 + 0.95x_{t-1} + \eta_t, \quad \eta_t \sim \mathrm{NID}(0, 1). \tag{1.12}$$

The special cases of model (1.6) (i.e. models (1.4), (1.5) and (1.8)), are obtained by imposing zero restrictions on the coefficients in (1.11). For example, in the

[6] Note that alternative estimation methods have also been suggested for binary time series models (see, e.g., Manski 1975, 1985; Elliott and Lieli, 2007).

static model (1.4), π_t is generated from

$$\pi_t = 0.1 - 0.20x_{t-1}.$$

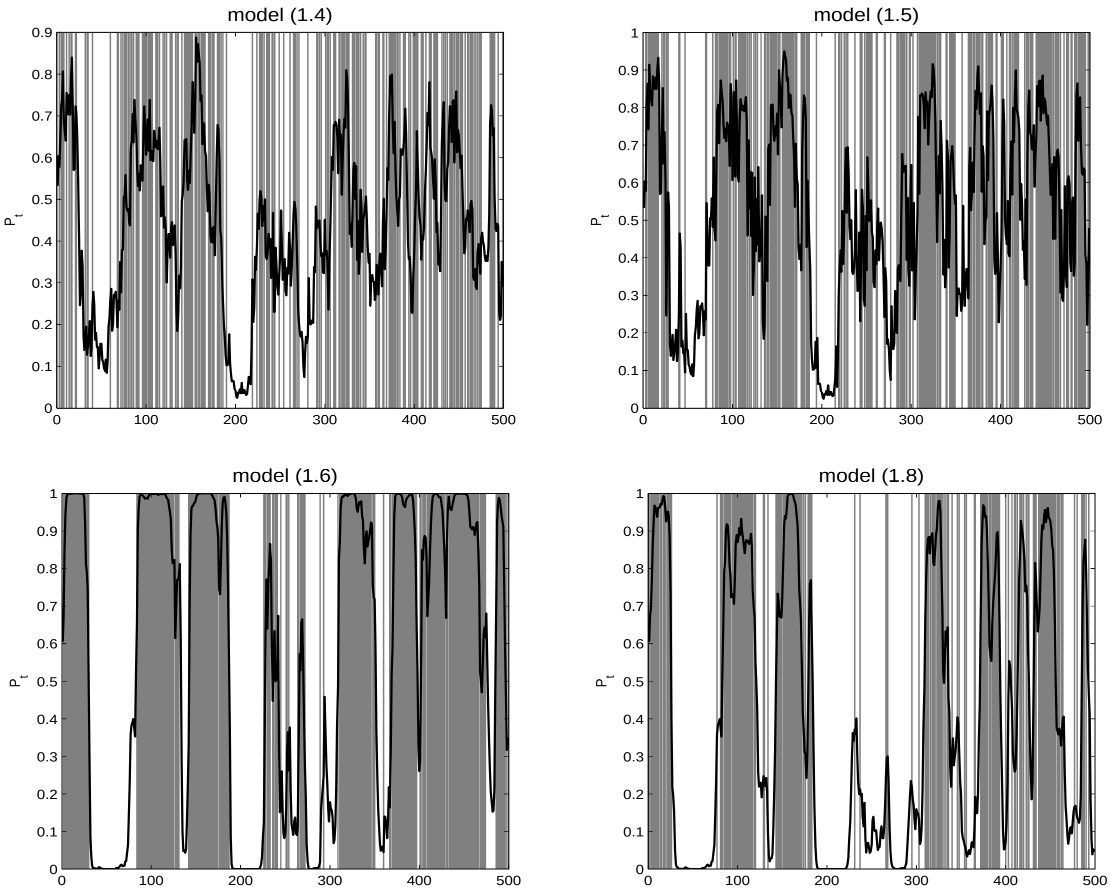

Figure 1.1: Simulated realizations of length 500 from different probit models. Conditional probability p_t (line) and the realized values of y_t (see (1.1)) with the shaded areas corresponding to $y_t = 1$ are depicted.

Figure 1.1 shows the fact that the conditional probability p_t is much more persistent in models with an autoregressive structure (models (1.6) and (1.8)). There seem to be periods characterized by a higher conditional probability with y_t typically taking the value 1, and similarly periods characterized by a lower conditional probability where $y_t = 0$. Thus, the realized values of the corresponding series are clustered in such a way that the outcome ($y_t = 1$ or $y_t = 0$) is typically followed by the same outcome also in the next period. Periods of economic recession and expansion, for example, follow a similar clustered pattern (see Chapters 2 and 5) which suggests that these models may be able to capture business cycle variation.

1.2.3 Latent Variable Approach

An alternative to the way the probit model was specified above is provided by latent variable approach (see, e.g., Davidson and MacKinnon, 1993, 514–515). In this approach, it is assumed that the outcome $y_t = 1$ is obtained when an unobserved variable y_t^* takes a positive value. In other words,

$$y_t = \begin{cases} 1, \text{if } y_t^* > 0, \\ 0, \text{if } y_t^* \leq 0. \end{cases} \tag{1.13}$$

The static probit model (1.4) can now be based on the equation

$$y_t^* = c + \boldsymbol{x}_{t-1}^{'}\boldsymbol{b} + u_t, \quad u_t \sim \text{NID}(0, 1), \tag{1.14}$$

where the normality assumption yields the probit model. As in Section 1.2.1, the conditional probability of the outcome $y_t = 1$ is (cf. (1.2) and (1.4))

$$P_{t-1}(y_t = 1) = P_{t-1}(y_t^* > 0) = \Phi(c + \boldsymbol{x}_{t-1}^{'}\boldsymbol{b}).$$

According to the classification of Cox (1981) (see also, e.g., Benjamin *et al.*, 2003), model (1.6) belongs to the class of "observation-driven" models, whereas the model based on the latent variable approach can be seen as a "parameter-driven" model.

Within the latent variable approach, various dynamic extensions of the static probit model have been suggested. Poirier and Ruud (1988) consider a model where the error term in (1.14) is assumed to follow a stationary first-order autoregressive process (see also Gourieroux, 2000, 35–36). Chauvet and Potter (2005) have proposed a dynamic model in which the latent variable y_t^* is generated by

$$y_t^* = c + \phi y_{t-1}^* + \boldsymbol{x}_{t-1}^{'}\boldsymbol{b} + \sigma_t u_t, \tag{1.15}$$

where the latent variable y_t^* follows a first-order autoregression and u_t is as in (1.14) but the variance of y_t^* may be time varying via the variable σ_t. Compared to the dynamic autoregressive model (1.6), model (1.15) is computationally more demanding. Unlike in model (1.6), the application of maximum likelihood methods requires high-dimensional multiple integrations over the unobserved lagged latent variable y^*_{t-1} and, therefore, Bayesian methods, such as the Gibbs sampler, have been used to evaluate the likelihood function. Note also that forecasts from model (1.6) can be computed by explicit formulae, which is not the case in model (1.15).

Kauppi (2008) has made some comparisons between the two alternative approaches to specify the probit model. He argues that although the latent variable y^*_t can have a meaningful economic interpretation, it is rather difficult to see what kind of dynamic properties of y_t are implied by a model such as (1.15). For instance, as the lagged value of y_t is not included in (1.15), it is not very easy to see how lagged values of y_t drive the conditional probability (1.2).

All in all, further research is needed to compare different dynamic probit models and their properties in these two approaches. As the applications of this thesis are based on the observation-driven models introduced in Section 1.2.2, such comparisons are left for future research.

1.2.4 Bivariate and Multivariate Models

So far, we have considered univariate binary time series models. In practice, we are often interested in joint dynamic behavior and interrelationships of several variables. Generalizing univariate binary time series models to the bivariate and multivariate cases is therefore of interest.

Bivariate and multivariate binary time series models have not much been examined in the previous literature. Ashford and Sowden (1970) have proposed the static bivariate probit model, where the joint conditional probabilities of the four possible outcomes of the random vector (y_{1t},y_{2t}) (cf. the univariate model in

(1.1)–(1.2)) are obtained as

$$
\begin{aligned}
P_{11,t} &= P_{t-1}(y_{1t}=1, y_{2t}=1) = \Phi_2(\pi_{1t}, \pi_{2t}, \rho), \\
P_{10,t} &= P_{t-1}(y_{1t}=1, y_{2t}=0) = \Phi_2(\pi_{1t}, -\pi_{2t}, -\rho), \\
P_{01,t} &= P_{t-1}(y_{1t}=0, y_{2t}=1) = \Phi_2(-\pi_{1t}, \pi_{2t}, -\rho), \\
P_{00,t} &= P_{t-1}(y_{1t}=0, y_{2t}=0) = \Phi_2(-\pi_{1t}, -\pi_{2t}, \rho).
\end{aligned} \tag{1.16}
$$

Here P_{ij} denotes the conditional probability of the outcome of vector $(y_{1t} = i, y_{2t} = j)$, $i, j = 0, 1$, and $\Phi_2(\cdot)$ is the bivariate standard normal cumulative distribution function. As in the univariate case, π_{1t} and π_{2t} are linear functions of variables in the information set Ω_{t-1}. Specifically, in the static bivariate model we have

$$
\begin{bmatrix} \pi_{1t} \\ \pi_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{x}'_{1,t-1} & 0 \\ 0 & \boldsymbol{x}'_{2,t-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}, \tag{1.17}
$$

where $\boldsymbol{x}'_{1,t-1}$ and $\boldsymbol{x}'_{2,t-1}$ are vectors of explanatory variables.

To the best of our knowledge, only two bivariate and multivariate dynamic probit models have been proposed in the literature. Mosconi and Seri (2006) have suggested a dynamic extension to the static bivariate model defined in (1.16) and (1.17). Their model is based on the latent variable approach discussed in the univariate case in Section 1.2.3. Quite recently, Anatolyev (2009) has considered dynamic extension of the static multivariate models of Ekholm, Smith and McDonald (1995, 2000). Anatolyev's bivariate model is somewhat similar to the model proposed in Chapter 5, which can also be seen as an extension of the static bivariate model introduced above.

1.2.5 Qualitative Models

Although the main interest in this thesis is on binary time series models, it may also be useful to review other related qualitative time series models. A natural extension of binary models is a model, where the dependent variable is multinomial. In these models, there are at least three possible values that the dependent variable can take. Early work in this area includes Eichengreen, Watson and Grossman

(1985), Hausman, Lo and MacKinlay (1992) and Dueker (1999a). Recently, Russell and Engle (1998, 2005) and Kauppi (2007) have suggested new dynamic multinomial models with autoregressive dynamics of the same type as in the univariate model (1.6).

As discussed by Kauppi (2007), in economics multinomial models have typically been used to predict interest rate target changes made by central banks, especially the U.S. Federal funds rate target changes made by the Federal Reserve (see also the models proposed by Dueker (1999b), Hamilton and Jorda (2002), and Hu and Phillips (2004)). Predicting the direction of the financial asset returns is another example of potential application of multinomial models (see Liesenfeld *et al.*, 2006).

Models for count variables is another important class of qualitative time series models. In these models, the dependent variable is typically assumed to follow a conditional Poisson distribution leading to the Poisson regression model (see, e.g., Maddala, 1983, 51–54). The classification into observation-driven and parameter-driven models discussed in Section 1.2.3 is also applicable here. Parameter-driven models based on the latent variable have been considered by Zeger (1988), Chan and Ledolter (1995) and Davis, Dunsmuir and Wang (2000), whereas Zeger and Qaqish (1988) provide an early discussion on observation-driven models (see also Davis, Dunsmuir and Street, 2003). References to the count time series models can also be found in the literature of GLARMA and GARMA models discussed in Section 1.2.2 (see, e.g., Shephard, 1995; Benjamin *et al.* 2003). Furthermore, related to the observation-driven models, Heinen (2003) has proposed a univariate autoregressive conditional Poisson (ACP) model with an autoregressive structure similar to that in the dynamic autoregressive probit model (1.6). The ACP model can also be extended in various ways, such as to the multivariate case as in Heinen and Rengifo (2007).

In addition to the introduced multinomial and count data models there are also various other qualitative dependent models (see, e.g., the Qual VAR model of Dueker (2005)) examined in the econometric literature (see also, e.g., Maddala, 1983; Gourieroux, 2000).

1.3 Applications

In this section, we review three applications of binary time series models. In Section 1.3.1, we consider forecasting "classical" business cycle recession periods. In Section 1.3.2, the predictability of cycles in aggregate economic activity is discussed by considering growth rate cycles. The sign predictability of asset prices is a major issue in empirical finance. We concentrate on this application in Section 1.3.3 with emphasis on sign predictability of stock market returns.

1.3.1 Forecasting Business Cycle Recession Periods

Fluctuations in economic activity are a central topic in theoretical and empirical macroeconomic research. Predicting cycles in economic activity is one of the most challenging and also one of the most important applications in macroeconomic forecasting.[7] The importance of predictions of the macroeconomic state of the economy is based on the fact that business cycles, and especially recession periods, are costly. For example, central banks and governments may try to mitigate fluctuations by stabilization policy. However, if the current state of the the economy is unknown or forecasts for the future development are inaccurate, the timing of policy actions may not be optimal, and the policy actions may even amplify further business cycle fluctuations.

After the Great Depression in 1930s, the National Bureau of Economic Research (NBER) developed a methodology for the empirical analysis of cycles in the U.S. economic activity. NBER applies the definition of the business cycle, which was originally suggested by Burns and Mitchell (1946). They emphasized that business cycles are co-movements of several macroeconomic variables which determine the turning points, peaks and troughs, in aggregate economic activity. Recession starts just after the economy reaches a peak of a business cycle and ends at the trough,

[7] At the time of writing this thesis in 2008–2009, the financial crisis had taken place all around the world, and most of the industrialized countries have faced one of the most severe recessions since the Great Depression in 1930s. This has once again increased interest in forecasting recession periods.

and vice versa with expansionary period. Recessions and expansions are recurrent, but not periodic. The duration of a business cycle from peak to peak or trough to trough is at least one year, but it may be much longer, even more than ten years.

Forecasting the recession periods with binary time series models is based on a binary-valued recession indicator. Without loss of generality, the recession indicator takes the value 1 when the economy is in a recession, and 0 otherwise. As above, a recession begins when the aggregate economic activity has reached a peak and ends at the trough. Thus, to construct recession periods, it is first necessary to determine the peak and trough dates for business cycle phases.

Using the definition of Burns and Mitchell (1946), the Business Cycle Dating Committee of the NBER determines the U.S. business cycle turning points, that is, the peak and trough months. In contrast to the commonly used rule of two consecutive quarters of decline in real GDP, the NBER defines the recession as "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales." See Hall *et al.* (2003) for details.

The recession dating procedure applied by the NBER is informal in the sense that it is up to the judgment of the Business Cycle Dating Committee. However, due to the informational lags and revisions between the initial and final values of the macroeconomic variables given in the recession definition, the announcements of peak and trough months are available with a substantial delay. This, in turn, leads to a delay in the publication of the recession indicator. This delay is referred to as a "publication lag" in this thesis. The most recent announcements of peak and trough months presented in Table 1.1 (see also the left panel of Figure 1.2) show that the publication lag has been between six months and almost two years in the U.S.

In the literature on recession forecasting, the NBER recession periods remain the benchmark. At the time of writing (fall 2009), the latest business cycle turning point in the U.S. is the peak in December 2007, whereas the last trough month is November 2001 (see Hall *et al.*, 2008).

Table 1.1: NBER's U.S. business cycle peak and trough chronology from 1970s.

Peak	Trough	Recession time (months)	Publication lag (months) peak	trough
1973 M11	1975 M03	16	–	–
1980 M01	1980 M07	6	5	12
1981 M07	1982 M11	16	6	8
1990 M07	1991 M03	8	10	21
2001 M03	2001 M11	8	8	20
2007 M12			12	

Note: Recessions start at the peak of a business cycle and end at the trough. Publication lag is the time between the business cycle turning point month (peak or trough) and the month when the Business Cycle Dating Committee of the NBER has made an announcement of the turning point (see details in http://www.nber.org/cycles/cyclesmain.html [2 July 2009]). Note that publication lags for the first peak and trough months are not available.

In addition to the informal "NBER approach", there is also a large alternative literature on dissecting business cycles and determining recession periods. Bry and Boschan (1971) and Harding and Pagan (2002), among others, have developed formal mathematical methods to determine recession and expansion periods. These methods are typically based on the informational content of the same monthly macroeconomic variables that NBER uses in their approach.

Another way to locate the turning points is based on the use of detrending methods. The idea is to filter out the cyclical component of the output series (e.g. the real GDP), and then to apply a dating rule to locate the peaks and troughs. However, in this approach, as Canova (1998) has noted, the resulting turning points are very sensitive to the detrending method and dating rule employed. Therefore, depending on the exact methods, the obtained recession periods can be very different.

Overall, despite the large number of different methods available for finding the turning points, we take the NBER recession periods as given in this thesis. One reason for this is the fact that the Economic Cycle Research Institute (ECRI) uses the same type of recession definition as the NBER and in Chapter 2 we use the recession periods they provide for Germany along with the U.S. recession periods.

The literature on recession forecasting based on the binary recession indicator dates back to the beginning of the 1990s. To the best of our knowledge, the first authors in this area were Estrella and Hardouvelis (1991) who used the static probit model (1.4) to predict U.S. recession periods. The same model with different predictive variables was subsequently employed in various other studies (see, e.g., Estrella and Mishkin, 1998; Bernard and Gerlach, 1998; Estrella, Rodrigues and Schich, 2003, among others). In recent years the dynamic models discussed in Sections 1.2.2 and 1.2.3 have been employed (see, e.g., Chauvet and Potter, 2005; Dueker, 2005; Kauppi and Saikkonen, 2008; Startz, 2008).

Variables used to predict future economic activity are often referred to as "leading indicators". In fact, there is a branch in the literature in empirical macroeconomics focused on finding and constructing reliable leading indicators for business cycle fluctuations. Marcellino (2006) provides a comprehensive survey of this literature, especially from the viewpoint of constructing coincident and leading indices to economic activity (see also, e.g., Stock and Watson, 1989; Mariano and Murasawa, 2003; Aruoba, Diebold and Scotti, 2009).

Marcellino (2006) discusses properties that a useful leading indicator variable should have. First, a leading indicator should systematically anticipate recessions and expansions. Second, the potential predictive power of a leading indicator should be supported by economic theory. Third, values of a leading indicator should be regularly available without (major) subsequent revisions. The real time availability of financial variables supports the use of those variables as predictors in recession forecasting. Financial variables, such as interest rates and stock market returns, are available even at a very high frequency, and they are measured precisely without revisions.[8] This is not typically the case for most macroeconomic predictors.

Much of the previous research on the use of financial variables as predictors in recession forecasting lends support to the term spread between the long-term

[8] Present value models for the term structure of interest rates and the stock price provide theoretical basis for the potential predictive power of these variables.

and short-term interest rate being the main leading indicator (see, e.g., Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Bernard and Gerlach, 1998; Estrella, 2005). The usefulness of the term spread as a predictor in recession forecasting is illustrated in Figure 1.2, where the U.S. and German recession periods and term spreads are depicted. The term spread is constructed as the difference between the 10-year and three-month interest rates. It is seen that positive values close to zero or even negative values of the term spread have preceded recessions in both countries. This was also the case before the recent recession period that began in 2008. These preliminary findings suggest that the term spread is a reliable leading indicator of future recession periods.

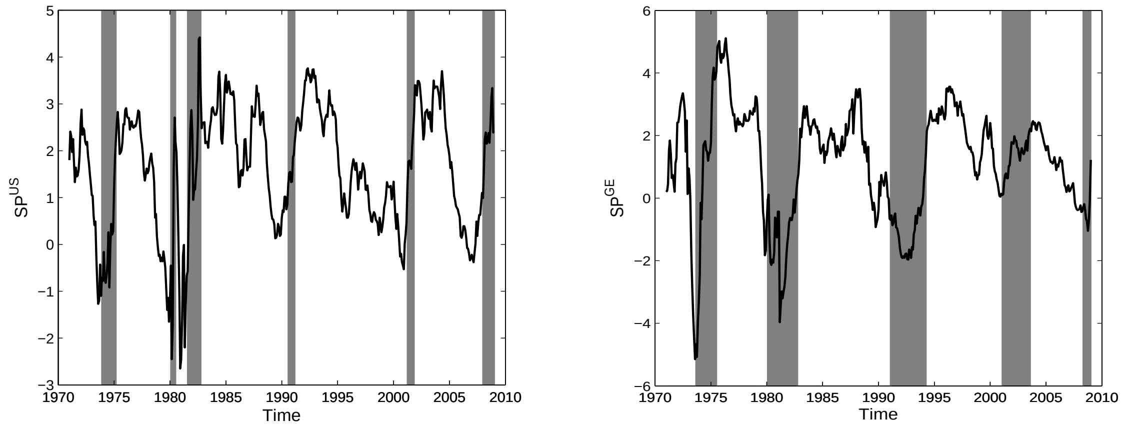

Figure 1.2: In the left panel, the U.S. recession periods (shaded areas) and the U.S. term spread (SP_t^{US}) from January 1971 to January 2009 are depicted. The German recession periods and the term spread (SP_t^{GE}) are shown in the right panel.

As an example of the use of univariate probit models in recession forecasting, we consider six-month-ahead recession forecasts obtained from the static (1.4) and the autoregressive model (1.8). In both models, the domestic term spread depicted in Figure 1.2 is employed as a predictor of the U.S. and German recession periods. Figure 1.3 shows the estimated recession probabilities. Intuitively, the recession probability should be as close as possible to one in recession and close to zero in expansion.

In Figure 1.3, we see that the autoregressive model (1.8) leads to more persistent recession forecast compared with the static model (1.4). Based on the pseudo-R^2 goodness-of-fit measure of Estrella (1998), which can be seen as a counterpart to the coefficient of determination in models for continuous variables, the autoregressive model outperforms the static model in both countries. Especially for the U.S. (German) recession periods, the autoregressive model yields a higher pseudo-R^2 value 0.391 (0.571) compared with the static model 0.181 (0.533). Overall, these results lead to the conclusion that the recession periods seem to be predictable with the predictive power provided by the term spread.

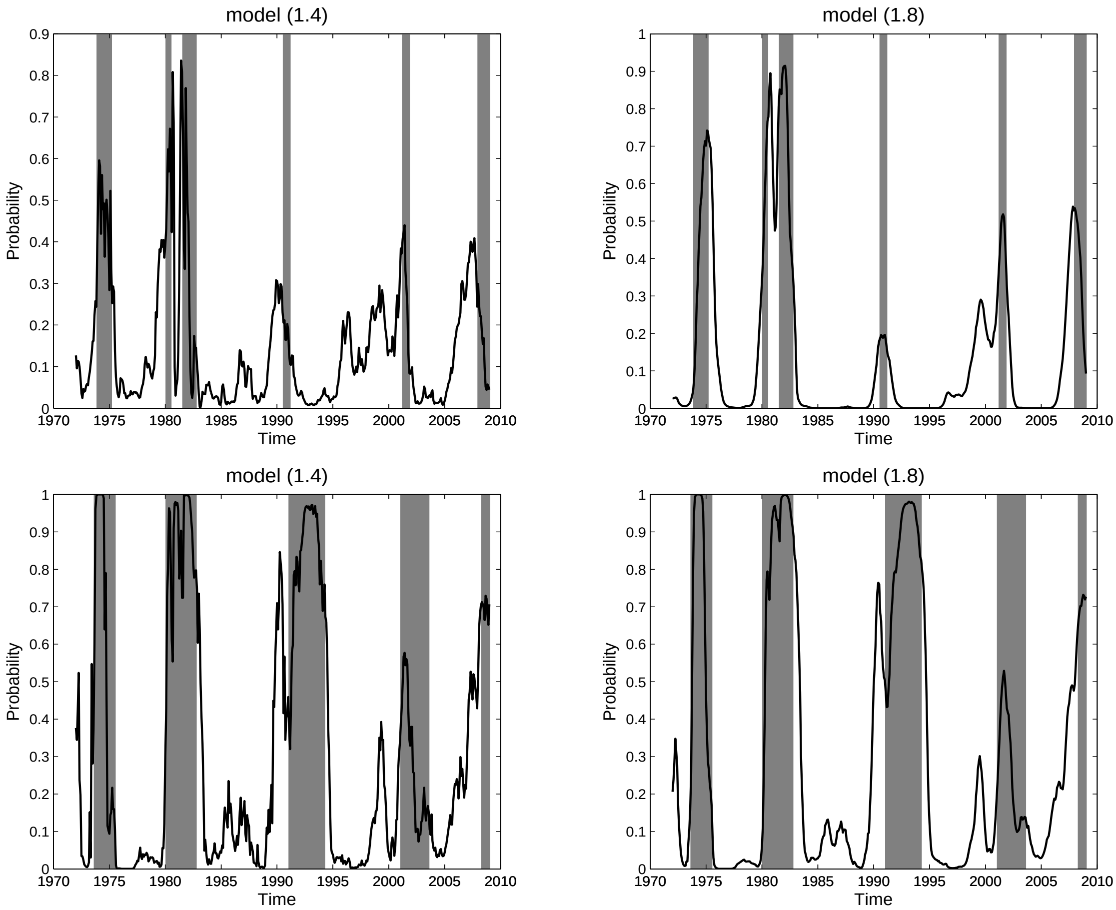

Figure 1.3: Estimated recession probabilities based on the static model (1.4) and the autoregressive model (1.8) where the term spread is employed as a predictor. In the upper panel, the U.S. recession periods are considered whereas in the lower panel the German recession periods are examined. The sample period is from January 1971 to January 2009 where the first 12 observations are used as initial values.

As emphasized by Marcellino (2006), although the term spread seems to a useful recession predictor, a model with a single predictor may yield unreliable forecasts for future recession periods. Experience has taught that each recession period has its own characteristics that a single explanatory variable may not be able to explain. Thus, in addition to the term spread, it is useful to consider also other financial variables as predictors in recession forecasting (see, e.g., the survey of Stock and Watson, 2003). This issue is one of the main objectives of Chapter 2.

1.3.2 Predicting Growth Rate Cycles

In recession forecasting, values of the recession indicator are related to the level of aggregate economic activity such that in a business cycle recession the level of economic activity is decreasing. However, at times the economy faces a slowdown that is not a decline in the level, but in the growth rate of economic activity. Such periods are called growth rate recessions, and this predictability can be studied by models similar to those discussed above.

Economic Cycle Research Institure (ECRI), among others, has determined a chronology of the growth rate cycle peak and trough months. The dating procedure is based on several macroeconomic variables and, in that sense, it is similar to the approach that the NBER uses to define business cycle turning points, as discussed in the previous section.[9] Layton and Moore (1989) and Banerji and Hiris (2001) have described in detail how the growth rate cycles can be determined. In summary, the methodology is based on the six-month smoothed growth rates of the macroeconomic variables included in the analysis. These smoothed growth rates together determine the cyclical turning points, i.e., peaks and troughs, in the growth rate cycles and hence also the growth rate "recession" and "expansion" periods.

Note that a distinction is made between growth cycles and growth rate cycles in the literature. According to Layton and Moore (1989) and Banerji and Hiris (2001) the difference between these two is that while growth cycles are deviations

[9] See details in http://www.businesscycle.com/resources/cycles [2 July 2009].

from the trend of economic activity, growth rate cycles are based on the suggested six-month or 12-month smoothed growth rate of economic activity. Thus, in the case of growth rate cycles, detrending is avoided, whereas it is inevitable in the case of growth cycles.[10]

In this thesis, we restrict ourselves to the ECRI growth rate cycles. To the best of our knowledge, Osborn, Sensier and van Dijk (2004) is the only study where growth rate cycle periods are predicted by using binary time series models. They used the static probit model (1.4) to predict ECRI growth rate cycle periods in a number of European countries. One of their main finding was that the growth rate cycles are closely related to the estimated output gap, which is the difference between the actual and potential output.

Although business cycle recessions and expansions have been studied much more than growth rate cycle periods, the latter are of importance for various reasons. As illustrated in Section 1.3.1, business cycle recessions are quite infrequent and, therefore, growth rate cycles give us information about the cyclical variation observed during business cycle expansions. In Chapter 5, we see that there is much more regime switching between recessions and expansions defined in terms of the growth rate cycle than those defined in terms of business cycle. This can be seen in Figure 1.4 where the values of both cycle indicators are illustrated with the business conditions index of Aruoba, Diebold and Scotti (2009) (hereafter ADS index).[11] This ADS index is an example of a coincident economic index discussed in Section 1.3.1 which reflects overall economic activity and is based on various macroeconomic variables measured at different data frequencies. The average value of the ADS index is set to zero. Thus, positive values indicate progressively better-than-average economic conditions, whereas negative values indicate progressively worse-than-average conditions.

In the left panel of Figure 1.4, we see that the ADS index is substantially

[10] Zarnowitz and Ozyildirim (2006) and Marcellino (2006), among others, have surveyed methods used to determine and characterize business cycles, growth rate cycles and growth cycles.

[11] See details at http://www.phil.frb.org/research-and-data/real-time-center/business-conditions-index [15 March 2010].

negative at business cycle recession periods whereas growth rate cycle recessions are closely related to the upswing and downswing periods in the growth rate of the index. A growth rate cycle recession is typically occurred at the same time with the decreasing values of the ADS index, and vice versa with expansion periods.

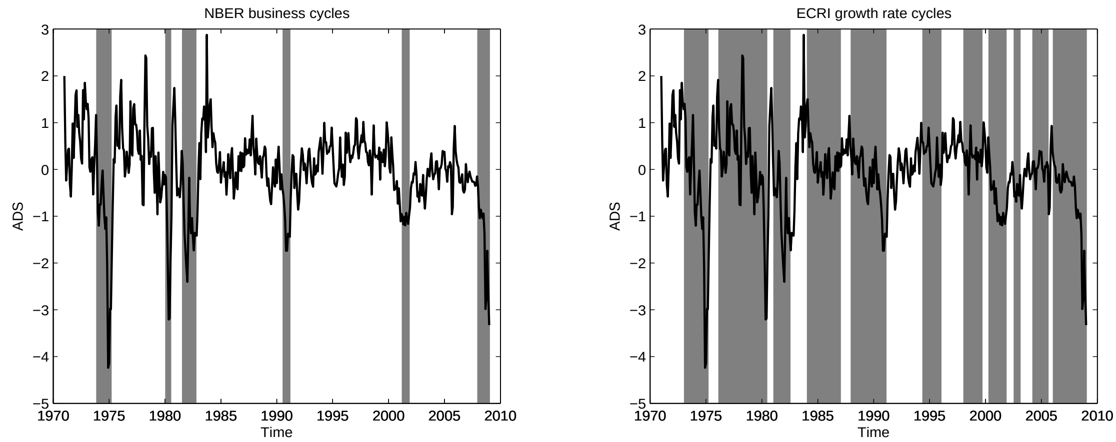

Figure 1.4: U.S. business cycle (left panel) and growth rate cycle (right panel) periods (shaded areas are the recession periods) with the real business conditions index of Aruoba, Diebold and Scotti (2009). The sample period is from January 1971 until January 2009.

As argued by Osborn *et al.* (2004), the relationship between growth rate recessions and the estimated output gap makes growth rate cycles of interest because the output gap is an important variable to policymakers, especially those concerned with monetary policy. Evidence of this kind can be found in Figure 1.5, where we depict the estimated U.S. output gap and growth rate cycle recession and expansion periods. The monthly output gap is approximated by the difference between the U.S. industrial production, which is used to measure economic activity in a monthly basis, and its trend component extracted by using the Hodrick-Prescott filter. It appears that the growth rate cycle periods have typically been related to the turning points of the estimated output gap. A growth rate recession is often began after the output gap has reached its local maximum and vice versa with expansion periods.

It should be pointed out that despite the importance of the output gap, it is not available in real time because of the delays and revisions in real GDP or industrial production typically employed to construct the output gap. Therefore, if the growth rate cycle phases, i.e. the turning points in the estimated output gap, are predictable, the predictions may yield important information for monetary policy decisions.

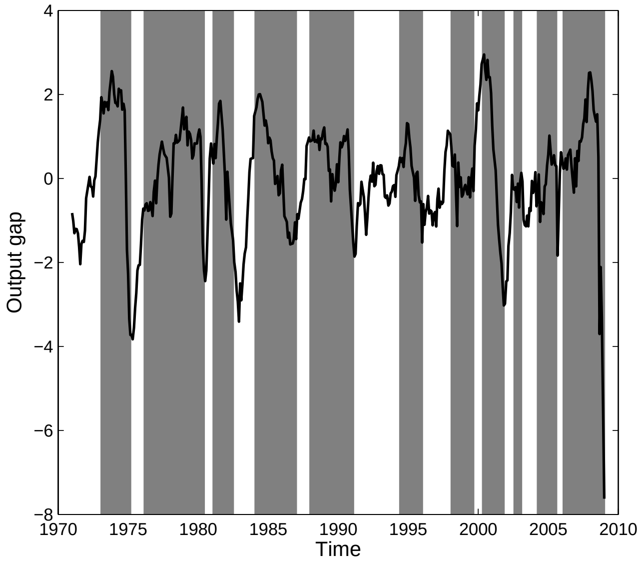

Figure 1.5: Estimated monthly U.S. output gap and the U.S. growth rate cycle recession (shaded areas) and expansion periods.

1.3.3 Directional Predictability in Stock Market Returns

According to recent empirical evidence, stock market returns or, more precisely excess stock market returns over the risk-free rate, are predictable, although the predictability is rather weak out of sample. Macroeconomic variables included in traditional linear regression models have been capable of predicting U.S. stock returns (see, e.g., Chen, Roll and Ross, 1986; Fama and French, 1989; Chen, 1991; Pesaran and Timmermann, 1995), and the recent work of Rapach, Wohar and Rangvid (2005), Ang and Bekaert (2007), among others, suggests that interest rates are the best predictors.

Christoffersen and Diebold (2006) and Christoffersen *et al.* (2007) argue that the sign of stock market returns may be predictable even if the returns themselves are not predictable. The arguments used to support this view have been based on models with continuous dependent variables. For instance, Christoffersen *et al.* (2006, 2007) have shown that the predictability of stock return volatility implies the sign predictability of the returns.

Despite the potential sign predictability of stock returns, surprisingly few studies have considered binary time series models in this context. However, they have been incorporated into models where the stock return is decomposed into different components. The simplest decomposition is

$$r_t = \mathrm{sign}(r_t)|r_t|, \tag{1.18}$$

where a return (r_t) consists of a sign component ($\mathrm{sign}(r_t)$) and an absolute value component ($|r_t|$). This decomposition has been employed by Anatolyev and Gospodinov (2010) who specified a probit model for the sign component.

Rydberg and Shephard (2003) suggest a different decomposition where the return is decomposed into three distinct processes when analyzing the dynamics of (high frequency) trade-by-trade price movements in asset prices. This decomposition can be presented as follows,

$$r_t = A_t D_t S_t, \tag{1.19}$$

where a binary variable A_t takes value one when the price moves and zero otherwise. Conditional on $A_t = 1$, the second binary variable D_t indicates the direction-of-change ($D_t = 1$ if the price moves upwards). Finally, conditional on the price movement and the direction-of-change, the variable S_t gives the size of the movement. The dynamics of the two binary variables can be modeled by binary time series models. Similarly, Liesenfeld *et al.* (2006) incorporate two binary models into their trinomial autoregressive conditional model (see also the model of Russell and Engle, 1998, 2005).

Instead of using binary models, it is also possible to extract sign forecasts from models built for the level of a stock return. Sign forecasts are obtained by the decision rule that, say, a positive forecast is a signal for a positive stock return. To the best of our knowledge, Leung, Daouk and Chen (2000) is the only study to compare the predictive performance of continuous dependent models and different discrete models, including probit models. They find that discrete models and various classification-based methods outperform traditional continuous models in terms of the number of correct sign forecasts and subsequent investment returns. In Chapter 4, we compare sign forecasts from binary and continuous predictive models for U.S. stock returns.

Finally, binary time series models may also be used to predict "bull" and "bear" market phases in stock markets which can be determined in the same fashion as recession and expansion periods of the macroeconomy (see Section 1.3.1). Maheu and McCurdy (2000), Pagan and Sossounov (2003), and Candelon, Piplack, and Straetmans (2008) have suggested methods to determine the bull and bear markets. Recently, Chen (2009) considered a Markov switching model and the static probit model (1.4) to forecast bear market phases. He concluded that the probit model is capable of predicting the stock market phases and that the term spread and inflation rate are the most useful predictors.

1.4 Contributions of the Thesis

In this thesis, we concentrate on different probit models based on the dynamic autoregressive probit model (1.6) and its special cases. We also suggest new model variants, whose most important feature is the inclusion of an autoregressive structure to the model. In this section, we introduce and summarize the main contributions to be presented in Chapters 2–5.

1.4.1 Recession Forecasts for the U.S. and Germany

In Chapter 2, we explore various probit models and their ability to predict the U.S. and German recession periods. The monthly data from January 1972 to March 2007 consists of observations of dependent recession indicators and financial predictive variables, such as interest rates and stock market returns.

The empirical findings show that models with an autoregressive structure outperform their competitors, especially the static model (1.4) used in many previous studies. We also propose a new "autoregressive interaction" model

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \boldsymbol{x}_{t-1}^{'} \boldsymbol{\beta} + y_{t-1} \boldsymbol{z}_{t-1}^{'} \boldsymbol{\gamma}, \tag{1.20}$$

where the autoregressive model (1.8) is augmented by the interaction term $y_{t-1}\boldsymbol{z}_{t-1}^{'}$ in which $\boldsymbol{z}_{t-1}^{'}$ includes the predictors. Compared with model (1.6), y_{t-1} is now excluded from the model, but the predictive variables included in the vector $\boldsymbol{z}_{t-1}^{'}$ are allowed to have an asymmetric effect depending on the lagged state of the economy (y_{t-1}).

As suggested in the previous literature, the term spread between the long-term and short-term interest rates turns out to be the main predictor, but various other financial variables are also found to have useful predictive power. As already suggested by Estrella and Mishkin (1998) in the case of the static probit model (1.4), stock market returns have predictive power for both countries. In addition to the term spread and stock market returns, the foreign term spread is a statistically significant predictor for both countries. We also find that the short-term interest rate differential between the U.S. and Germany is a statistically significant predictor for recession periods for Germany. Our findings show that these additional predictive variables have also useful out-of-sample predictive power over and above the term spread for both countries.

In an out-of-sample forecasting exercise covering the period from January 1995 to March 2007, the best probit models give good forecasts for the state of the business cycle at least six months ahead. The best out-of-sample forecasts are obtained from the interaction model (1.20), where the domestic or the foreign term spread is included in the interaction term $y_{t-1}\boldsymbol{z}_{t-1}^{'}$.

In Chapter 2, we also report out-of-sample forecasts for the years 2006–2008. This period is of interest given that the latest recession period in the U.S. began in December 2007 and in April 2008 in Germany. Our results show that the best dynamic probit models, such as model (1.20), were able to predict the beginning of the recent recession for both countries.

1.4.2 LM Test for the Autoregressive Structure

The findings of this thesis lend support to the usefulness of models with an autoregressive structure. Therefore, it is of interest to test the statistical significance of the autoregressive structure when specifying the model. From the practical point of view, this is useful because parameter estimation in model (1.6) is more complicated than, say, in model (1.5) where estimation can be carried out by the procedures available in standard econometric software packages.

In Chapter 3, we propose a LM test for testing the adequacy of the restricted model (1.5) with no autoregressive structure. Thus, the null hypothesis imposed by model (1.5) is

$$H_0 : \alpha_1 = 0. \tag{1.21}$$

The rejection of the null hypothesis would provide evidence in favor of the unrestricted model (1.6). As a matter of fact we propose two LM test statistics for the null hypothesis (1.21). Asymptotically these test statistics are equivalent, but their small-sample properties seem to be different. Our simulation results show that both LM tests can be severely oversized in small samples. As a remedy, we propose a parametric bootstrap method which makes the finite sample size of tests acceptable. Both tests have reasonable size-adjusted power, even in sample as small as 150 observations.

As an illustration, we test for significance of the autoregressive model structure in various binary models for U.S. recession periods. The data set consists of the monthly observations of the U.S. recession indicator and explanatory variables from January 1953 until December 2006. Thus, the sample size of 634 observations is somewhat larger than in Chapter 2. The main findings are, however, the same as in Chapter 2. The autoregressive structure is found to be a statistically significant addition to the traditional static recession prediction model, possibly augmented with the forecast horizon specific predictor y_{t-15}, used in many previous studies. However, when the first lag of the recession indicator (y_{t-1}) is employed in the dynamic autoregressive model (1.6), it is the main predictor and the autoregressive structure is in fact statistically insignificant indicating that model (1.6) reduces to the dynamic model (1.5). It is worth noting that, in both cases, the proposed LM tests yield the same conclusion as the Wald and likelihood ratio tests.

1.4.3 Sign Predictability of the U.S. Stock Returns

As discussed in Section 1.3.3, some predictability in excess stock returns over the risk-free interest rate has been found in the previous literature. The out-of-sample predictability has mostly been rather weak leading to the suggestion that it may only be the sign of an excess return that is predictable.

In Chapter 4, we study the in-sample and out-of-sample sign predictability of U.S. stock returns using different probit specifications. The data set consists of monthly U.S. data from January 1968 to December 2006. The monthly excess stock return is constructed as the difference between the U.S. nominal return on the S&P500 index and the risk-free interest rate. Several financial variables included in the data set are considered as predictors. The models are estimated with the data up to December 1988 and the remaining sample period is left for out-of-sample forecasts. Thus, the first out-of-sample forecasts are made for January 1989 and last ones for December 2006.

We propose a new framework where the recession forecast (see Chapter 2) is used as a predictor in a probit model to predict the sign of the return. Second, alternative dynamic probit specifications are considered in sign forecasting. To the best of our knowledge, this type of model has not been considered previously. This model can also be interpreted as a "structural" model where a recursive structure is imposed on two binary variables. First, we construct forecasts for the recession periods, and thereafter compare the forecasts for the sign of the return using the generated recession forecast as a predictor.

It turns out that the six-month recession forecast is the main predictor of the sign of the return. This indicates that forecasts of the future state of the economy are informative for the sign of the next month's stock return. In addition to recession forecasts, the differenced short-term and long-term interest rates have statistically significant predictive power out of sample.

Among the alternative probit models considered, our new "error correction" model yields the best out-of-sample sign forecasts. This is obtained by imposing the restriction $\delta_1 = 1 - \alpha_1$ on model (1.6),

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + (1 - \alpha_1) y_{t-1} + \boldsymbol{x}'_{t-1}\boldsymbol{\beta}. \tag{1.22}$$

We refer this model as the error correction model because it can be written as error correction form

$$\Delta \pi_t = \omega + \boldsymbol{x}'_{t-1}\boldsymbol{\beta} + (1 - \alpha_1)(y_{t-1} - \pi_{t-1}),$$

where the difference $(y_{t-1} - \pi_{t-1})$ is interpreted as a "disequilibrium error".

In addition to probit models, in terms of out-of-sample sign forecasting, model (1.22) also outperforms a number of other models, such as ARMAX models and predictive models based on volatility forecasts (cf. Christoffersen *et al.*, 2006, 2007). These findings are in line with those of Leung *et al.* (2000).

1.4.4 Nowcasting Cycles in the U.S. Economic Activity with Bivariate Autoregressive Probit Model

Chapter 5 extends the univariate analysis of binary time series considered in Chapters 2–4 to the bivariate case by introducing a new bivariate autoregressive probit model. This model is an extension of the static bivariate model of Ashford and Sowden (1970) described briefly in Section 1.2.4. Our model is given by

$$\begin{bmatrix} \pi_{1t} \\ \pi_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \pi_{1,t-1} \\ \pi_{2,t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{x}'_{1,t-k} & 0 \\ 0 & \boldsymbol{x}'_{2,t-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}, \tag{1.23}$$

where π_{1t} and π_{2t}, with the correlation coefficient ρ that appears in bivariate cumulative normal distribution function define the joint conditional probabilities given in (1.16) for the four possible outcomes of the vector (y_{1t}, y_{2t}). The main difference between the static bivariate probit model (1.17) and the bivariate autoregressive model (1.23) is the absence of an autoregressive structure in the former. Model (1.23) can also be interpreted as a bivariate extension to the univariate autoregressive model (1.8). Like its univariate counterpart (1.7), model (1.23) has an infinite-order representation implying that π_{1t} and π_{2t} depend on the infinite history of the predictive variables included in $\boldsymbol{x}_{1,t-k}$ and $\boldsymbol{x}_{2,t-k}$.

Before applying the bivariate autoregressive probit model it may be of interest to test whether the correlation coefficient ρ in the bivariate cumulative normal distribution (1.16) defining the model is zero ($\rho = 0$). If this is the case, the bivariate probit model reduces to two univariate models, which simplifies parameter estimation. To this end, we propose a Lagrange Multiplier test for testing the hypothesis of $\rho = 0$.

We apply the bivariate model to nowcasting the current state of the U.S. economy defined in terms of business cycle and growth rate cycle recession periods. By nowcasting we mean that the forecast horizon is one month ($h = 1$) indicating that we are interest in forecasts of the state of the economy at the current time. The data set covers the U.S. data from January 1972 to December 2005. This type of bivariate analysis of the two binary cycle indicators has not been considered in the previous literature. It extends the univariate recession forecasting studies of the previous literature and Chapter 2.

Empirical results show that additional predictive power can be gained by nowcasting business cycle and growth rate cycle periods jointly within a bivariate model. The bivariate autoregressive probit model (1.23) outperforms the univariate models yielding the best in-sample and out-of-sample predictions for the state of the U.S. economy.

As in the previous literature and also in Chapter 2, the U.S. term spread is the main predictive variable for business cycle recession periods. For growth rate cycles, however, the lagged differenced Federal funds rate and stock market returns are the best predictors. The U.S. term spread also has some predictive ability for growth rate cycles, but it is less powerful than the lagged differenced Federal funds rate and stock market returns. Furthermore, the best out-of-sample nowcasts for growth rate cycle recessions are obtained from the bivariate autoregressive probit model in which the effect of "Great Moderation" time period is also taken into account.

References

Amemiya T. 1980. Qualitative response models: A survey. *Journal of Economic Literature* **19**: 1483–1536.

Anatolyev S. 2009. Multi-market direction-of-change modeling using dependence ratios. *Studies in Nonlinear Dynamics and Econometrics* **13**, article 5.

Anatolyev S, Gospodinov N. 2010. Modeling financial return dynamics via decomposition. *Journal of Business and Economic Statistics* **28**: 232–245.

Ang A, Bekaert G. 2007. Stock return predictability: Is it there? *Review of Financial Studies* **20**: 651–707.

Aruoba SB, Diebold FX, Scotti R. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* **27**: 417–427.

Ashford JR, Sowden RR. 1970. Multivariate probit analysis. *Biometrics* **26**: 535–546.

Banerji A, Hiris L. 2001. A framework for measuring business cycles. *International Journal of Forecasting* **17**: 333–348.

Benjamin MA, Rigby RA, Stasinopoulos DM. 2003. Generalized autoregressive moving average models. *Journal of American Statistical Association* **98**: 214–223.

Bernard H, Gerlach S. 1998. Does the term structure predict recessions? The international evidence. *International Journal of Finance and Economics* **3**: 195–215.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **52**: 5–59.

Bry G, Boschan C. 1971. *Cyclical analysis of time series: Selected procedures and computer programs.* National Bureau of Economic Research, Columbia University Press.

Burns AF, Mitchell WC. 1946. *Measuring Business Cycles.* National Bureau of Economic Research, New York.

Cameron CA, Trivedi PK. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press, New York.

Candelon B, Piplack J, Straetmans S. 2008. On measuring synchronization of bulls and bears: The case of East Asia. *Journal of Banking and Finance* **32**: 1022–1035.

Canova F. 1998. Does detrending matter for the determination of the reference cycle and the selection of turning points? *Economic Journal* **109**: 126–150.

Chan KS, Ledolter J. 1995. Monte Carlo EM estimation for time series models involving counts. *Journal of American Statistical Association* **90**: 242–251.

Chauvet M, Potter S. 2005. Forecasting recession using the yield curve. *Journal of Forecasting* **24**: 77–103.

Chen N. 1991. Financial investment opportunities and the macroeconomy. *Journal of Finance* **46**: 529–554.

Chen N, Roll R, Ross SA. 1986. Economic forces and the stock market. *Journal of Business* **56**: 383–403.

Chen SS. 2009. Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking and Finance* **33**: 211–223.

Christoffersen PF, Diebold FX. 2006. Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science* **52**: 1273–1287.

Christoffersen PF, Diebold FX, Mariano RS, Tay AS, Tse YK. 2007. Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: International evidence. *Journal of Financial Forecasting* **1**: 3–24.

Cox DR. 1981. Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* **8**: 93–115.

Davidson R, MacKinnon JG. 1993. *Estimation and Inference in Econometrics.* Oxford University Press, New York.

Davis RA, Dunsmuir WTM, Wang Y. 2000. On autocorrelation in a Poisson regression model. *Biometrika* **87**: 491–505.

Davis RA, Dunsmuir WTM, Streett SB. 2003. Observation-driven models for Poisson counts. *Biometrika* **90**: 777–790.

de Jong RM, Woutersen TM. In press. Dynamic time series binary choice. *Econometric Theory*, forthcoming.

Dueker MJ. 1997. Strengthening the case for the yield curve as a predictor of U.S. recessions. *Federal Reserve Bank of St.Louis Review* **79**: 41–51.

Dueker M. 1999a. Conditional heteroscedasticity in qualitative response models of time series: A Gibbs-Sampling approach to the Bank Prime Rate. *Journal of Business and Economic Statistics* **17**: 466–472.

Dueker M. 1999b. Measuring monetary policy inertia in target fed funds rate changes. *Federal Reserve Bank of St. Louis Review* **81**: 3–10.

Dueker MJ. 2005. Dynamic forecasts of qualitative variables: A Qual VAR model of U.S. recessions. *Journal of Business and Economic Statistics* **23**: 96–104.

Eichengreen B, Watson MW, Grossmann RS. 1985. Bank rate policy under the interwar gold standard. *Economic Journal* **95**: 725–745.

Ekholm A, Smith PWJ, McDonald JW. 1995. Marginal regression analysis of a multivariate binary response. *Biometrika* **82**: 847–854.

Ekholm A, Smith PWJ, McDonald JW. 2000. Association models for a multivariate binary response. *Biometrics* **56**: 712–718.

Elliott G, Lieli RP. 2007. Predicting binary outcomes. Unpublished manuscript, University of Texas.

Engle RF. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**: 987–1008.

Engle RF, Russell J. 1998. Autoregressive conditional duration: A new model for irregular spaced transaction data. *Econometrica* **66**: 1127–1162.

Estrella A. 1998. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics* **16**: 198–205.

Estrella A. 2005. The yield curve as a leading indicator: Frequently asked questions. Federal Reserve Bank of New York. Available at http:/www.newyorkfed.org/research/capital_markets/ycfaq.pdf. [1 October 2009].

Estrella A, Hardouvelis GA. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* **46**: 555–576.

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45–61.

Estrella A, Rodrigues AP, Schich S. 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.

Fama EF, French KR. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* **25**: 23–49.

Franses PH, van Dijk D. 2000. *Non-linear Time Series Models in Empirical Finance.* Cambridge University Press, Cambridge.

Gourieroux C. 2000. *Econometrics of Qualitative Dependent Variables.* Cambridge University Press, Cambridge.

Granger CWJ, Teräsvirta T. 1993. *Modelling Nonlinear Economic Relationships.* Oxford University Press, New York.

Hall RE, Feldstein M, Frankel J, Gordon R, Romer C, Romer D, Zarnowitz V. 2003. The NBER's recession dating procedure. Business Cycle Dating Committee, National Bureau of Economic Research. Available at http://www.nber.org/cycles/recessions.pdf. [2 July 2009].

Hall RE, Feldstein M, Frankel J, Gordon R, Poterba J, Romer D, Zarnowitz V. 2008. Determination of the December 2007 peak in economic activity. Business Cycle Dating Committee, National Bureau of Economic Research. Available at http://www.nber.org/dec2008.pdf. [2 July 2009].

Hamilton JD. 1989. A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica* **57**: 357–384.

Hamilton JD, Jorda O. 2002. A Model of the Federal funds rate target. *Journal of Political Economy* **110** : 1135-1167.

Harding D, Pagan A. 2002. Dissecting the cycle: a methodological investigation. *Journal of Monetary Economics* **49**: 365–381.

Hausman J, Lo A, MacKinlay C. 1992. An ordered probit analysis of transaction stock prices. *Journal of Financial Economics* **31**: 319–379.

Heinen A. 2003. Modelling time series count data: An autoregressive conditional Poisson model. Core discussion paper 62.

Heinen A, Rengifo E. 2007. Multivariate autoregressive modeling of time series count data using copulas. *Journal of Empirical Finance* **14**: 564–583.

Honore BE, Kyriazidou E. 2000. Panel data discrete choice models with lagged dependent variables. *Eonometrica* **68**: 839–874.

Honore BE, Lewbel A. 2002. Semiparametric binary choice panel data models without strictly exogenous regressors. *Eonometrica* **70**: 2053–2063.

Hu L, Phillips PCB. 2004. Dynamics of the Federal Funds target rate: A nonstationary discrete choice approach. *Journal of Applied Econometrics* **19**: 851–867.

Kaufmann H. 1987. Regression models for nonstationary categorical time series: Asymptotic estimation theory. *Annals of Statistics* **18**: 79–98.

Kauppi H. (2007): Predicting the Fed's target rate decisions, HECER Discussion Paper, 182. Helsinki Center of Economic Research.

Kauppi H. (2008): Yield-curve based probit models for forecasting U.S. recessions: Stability and dynamics. HECER Discussion Paper, 221. Helsinki Center of Economic Research.

Kauppi H, Saikkonen P. 2008. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics* **90**: 777–791.

Layton AP, Moore GH. 1989. Leading indicators for the service sector. *Journal of Business and Economic Statistics* **7**: 379–386.

Leitch G, Tanner JE. 1991. Economic forecast evaluation: Profit versus the conventional error measures. *American Economic Review* **81**: 580–590.

Leung MT, Daouk H, Chen AS. 2000. Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting* **16**: 173–190.

Li WK. 1994. Time series models based on generalized linear models: Some further results. *Biometrics* **50**: 506–511.

Liesenfeld R, Nolte I, Pohlmeier W. 2006. Modelling financial transaction price movements: A dynamic integer count data model. *Empirical Economics* **30**: 795–825.

Maddala GS. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, New York.

Maheu JM, McCurdy TH. 2000. Identifying bull and bear markets in stock returns. *Journal of Business and Economic Statistics* **18**: 100–112.

Manski CF. 1975. Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics* **3**: 205–228.

Manski CF. 1985. Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics* **27**: 313–333.

Marcellino M. 2006. Leading Indicators in *Handbook of Economic Forecasting* eds. Elliott G, Granger CWJ, and Timmermann A. Volume 1, Elsevier.

Mariano, RS, Murasawa, Y. 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* **18**: 427–443.

McCullagh P, Nelder JA. 1989. *Generalized Linear Models*. Second Edition, Chapman and Hall. London.

Mosconi R, Seri R. 2006. Non-causality in bivariate binary time series. *Journal of Econometrics* **132**: 379–407.

Osborn DR, Sensier M, van Dijk D. 2004: Predicting growth regimes for European countries, in Lucrezia Reichlin (eds.), *The Euro Area Business Cycle: Stylized Facts and Measurement Issues*. Centre for Economic Policy Research.

Pagan AR, Sossounov KA. 2003. A simple framework for analyzing bull and bear markets. *Journal of Applied Econometrics* **18**: 23–46.

Pesaran HM, Timmermann A. 1995. Predictability of stock returns: Robustness and economic significance. *Journal of Finance* **50**: 1201–1228.

Poirier DJ, Ruud PA. 1988. Probit with dependent observations. *Review of Economic Studies* **55**: 593–614.

Rapach DE, Wohar ME, Rangvid J. 2005. Macro variables and international stock return predictability. *International Journal of Forecasting* **21**: 137–166.

Russell JR., Engle RF. 1998. Econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new multinomial autoregressive conditional multinomial model. Economics Working Paper Series, 10, University of California, San Diego.

Russell JR, Engle RF. 2005. A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial-autoregressive conditional duration model. *Journal of Business and Economic Statistics* **23**: 166–180.

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

Shephard N. 1995. Generalized linear autoregressions. Unpublished manuscript, Nuffield College, Oxford.

Startz R. 2008. Binomial autoregressive moving average models with an application to U.S. Recessions. *Journal of Business and Economic Statistics* **26**: 1–8.

Stock JH, Watson MW. 1989. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual* **4**: 351–409.

Stock JH, Watson MW. 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* **41**: 788–829.

Tong H. 1990. *Non-linear time series: A Dynamic System Approach*, Oxford University Press, London.

Wooldridge JM. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT press, Cambridge Massachusetts.

Zarnowitz V, Ozyildirim A. 2006. Time series decomposition and measurement of business cycles, trends and growth cycles. *Journal of Monetary Economics* **53**: 1717–1739.

Zeger SL. 1988. A regression model for time series of counts. *Biometrika* **75**: 621–629.

Zeger SL, Qaqish B. 1988. Markov regression models for time series: A quasi-likelihood approach. *Biometrics* **44**: 1019–1031.

Chapter 2

Dynamic Probit Models and Financial Variables in Recession Forecasting

Abstract[1]

In this chapter, various financial variables are examined as predictors of the probability of a recession in the U.S. and Germany. We propose a new dynamic probit model that outperforms the standard static model, giving accurate out-of-sample forecasts in both countries for the recession period that began in 2001, as well as the beginning of the recession in 2008. In accordance with previous findings, the domestic term spread proves to be an important predictive variable, but stock market returns and the foreign term spread also have predictive power in both countries. In the case of Germany, the interest rate differential between the U.S. and Germany is also a useful additional predictor.

[1] A paper "Dynamic Probit Models and Financial Variables in Recession Forecasting" based on this chapter has been published in the *Journal of Forecasting*, 29, 215–230, 2010, Wiley-Blackwell, ©[2009], John Wiley & Sons, Ltd.

2.1 Introduction

A substantial amount of research has considered the predictive ability of various financial variables to predict the economic growth and recession periods in different countries. Much of the previous analysis is focused on time series models where the dependent variable is "continuous", such as the growth rate of real GDP (see, for example, the survey by Stock and Watson, 2003). However, in the recent econometric literature, forecasting a binary recession indicator with probit or logit models has attracted attention and, consequently, new time series models for binary dependent variables have been introduced. Dueker (2002, 2005), Chauvet and Potter (2005) and Startz (2008), among others, have proposed dynamic extensions to the standard static probit model used by Estrella and Hardouvelis (1991), Bernard and Gerlach (1998), and Estrella and Mishkin (1998), among others, to predict recession periods. The main objective of this study is to apply the dynamic models suggested by Kauppi and Saikkonen (2008) to predict monthly recession periods in the U.S. and Germany.

Among various financial explanatory variables considered, the term spread, which is the difference between the long-term and short-term interest rate, has proved to be a useful predictor of future economic growth and recession periods (see, for example, Estrella and Mishkin, 1998; Estrella, 2005a). However, other financial predictors have also been suggested. For instance, if the domestic spread is a useful predictor, then the foreign spread may also have predictive power (Bernard and Gerlach, 1998). A potentially useful alternative is the interest rate differential between the considered countries, which to the best of our knowledge has not been used in recession forecasting prior to this study. Furthermore, as a forward-looking variable, the stock market return should also have additional predictive power in addition to interest rate-based predictive variables (Estrella and Mishkin, 1998).

Our findings extend the earlier literature in several ways. We confirm that the domestic term spread is the primary predictive variable, but we also find stock returns to have statistically significant predictive power for both countries. Furthermore, in the case of German recessions the interest rate differential between the U.S. and Germany is also a useful predictor, whereas the German term spread helps predict the U.S. recessions. The U.S. term spread is also a statistically significant explanatory variable in all predictive models fitted for German recession periods, but its out-of-sample predictive performance seems to be poor. Out-of-sample forecasts also lend some support to an asymmetric impact of the term spread on the recession probability dependent on the state of the economy. Overall, dynamic probit models outperform the standard static recession prediction models in terms of both in-sample and out-of-sample predictions. The best models also provide accurate out-of-sample forecasts and recession signals for the beginning of the recession in 2008.

The rest of this chapter is organized as follows. Section 2.2 presents the probit models to be used in forecasting, and provides a brief discussion on multiperiod forecasts of the recession indicator. In Section 2.3 the results of the in-sample and out-of-sample predictions of recession periods in the U.S. and Germany are provided. Finally, Section 2.4 concludes.

2.2 Dynamic Probit Models

2.2.1 Models

In binary time series analysis, the dependent variable y_t, $t = 1, 2, ..., T$, is a realization of a stochastic process that only takes on values one and zero. In recession forecasting, the value of an observable binary recession indicator depends on the state of the economy in the following way

$$y_t = \begin{cases} 1, & \text{if the economy is in a recessionary state at time } t, \\ 0, & \text{if the economy is in an expansionary state at time } t. \end{cases} \tag{2.1}$$

In other words, conditional on the information set Ω_{t-1}, y_t has a Bernoulli distribution

$$y_t|\Omega_{t-1} \sim B(p_t). \tag{2.2}$$

Let $E_{t-1}(\cdot)$ and $P_{t-1}(\cdot)$ denote the conditional expectation and conditional probability given the information set Ω_{t-1}, respectively. In the probit model the conditional probability that y_t takes the value 1 can be written as

$$p_t = E_{t-1}(y_t) = P_{t-1}(y_t = 1) = \Phi(\pi_t), \tag{2.3}$$

where π_t is a linear function of variables included in the information set Ω_{t-1} and $\Phi(\cdot)$ is a standard normal cumulative distribution function.

In the previous recession forecasting research, the standard ‘static” model has been the most commonly used model (see, for example, Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Bernard and Gerlach, 1998). That is,

$$\pi_t = \omega + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}, \tag{2.4}$$

where $\boldsymbol{x}_{t-k}$ is the vector of explanatory variables. In $\boldsymbol{x}_{t-k}$, the ith ($i = 1, ..., n$) explanatory variable, $x_{i,t-k}$, should satisfy the condition $k \geq h$, where h is the forecast horizon and the employed lag k may be different in different predictive variables.

A major shortcoming of the static model (2.4) is that it does not take the autocorrelation structure of the binary time series into account (Dueker, 1997). In recession forecasting, this means that the previous states of the economy are not included in the model. Therefore, a natural dynamic extension to the static model (2.4) is obtained by adding a lagged value of the dependent time series, y_{t-l}, to the right hand side of (2.4). This yields the “dynamic” probit model[2]

$$\pi_t = \omega + \delta_1 y_{t-l} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}, \tag{2.5}$$

where $l \geq 1$. Kauppi and Saikkonen (2008) extend this model by adding a lagged value of π_t. The resulting “dynamic autoregressive” model is given by

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \delta_1 y_{t-l} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}. \tag{2.6}$$

One can also consider an “autoregressive” model obtained by restricting the coefficient δ_1 in (2.6) to zero, that is,

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}. \tag{2.7}$$

[2] We use the terminology of Kauppi and Saikkonen (2008). All extensions of the static model (2.4) are called dynamic models although model (2.5) is referred to as the “dynamic” probit model.

When $|\alpha_1| < 1$, it can be seen that by recursive substitution

$$\pi_t = \sum_{i=1}^{\infty} \alpha_1^{i-1}\omega + \delta_1 \sum_{i=1}^{\infty} \alpha_1^{i-1} y_{t-l-i+1} + \sum_{i=1}^{\infty} \alpha_1^{i-1} \boldsymbol{x}'_{t-k-i+1}\boldsymbol{\beta},$$

so that the dynamic autoregressive model (2.6) is an "infinite"-order extension of the dynamic model (2.5). This presentation indicates that the autoregressive models, (2.6) and (2.7), may be useful and parsimonious specifications if a large number of explanatory variables are helpful in forecasting. Rydberg and Shephard (2003) proposed a model somewhat similar to (2.6), but their model does not imply dependence on the infinite history of the explanatory variables. Throughout this chapter, only one lagged value of π_t and of the recession indicator y_t are assumed, but including several lags is of course possible.

An interesting extension of model (2.7) is obtained by including an interaction term

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta} + y_{t-a}\boldsymbol{z}'_{t-k}\boldsymbol{\gamma}, \tag{2.8}$$

where $a \geq 1$. Note that the explanatory variables included in $\boldsymbol{z}_{t-k}$ may be different from those in $\boldsymbol{x}_{t-k}$. If $\boldsymbol{z}_{t-k} = \boldsymbol{x}_{t-k}$, the impact of the explanatory variables in $\boldsymbol{x}_{t-k}$ is allowed to depend on the state of the economy (cf. Kauppi and Saikkonen, 2008). Of course, it is also possible to augment model (2.8) by the lagged value y_{t-l}, $l \geq 1$.

The parameters of models (2.4)–(2.8) can be estimated by the method of maximum likelihood (ML).[3] Unfortunately, there is no formal proof of the asymptotic properties of the maximum likelihood estimator in models (2.6)–(2.8) with an autoregressive structure. However, the results of Estrella and Rodrigues (1998) and de Jong and Woutersen (in press) indicate that under reasonable regularity conditions, such as the stationarity of the specification of π_t and explanatory variables, the ML estimator is consistent and asymptotically normal. Robust standard errors allowing for autocorrelation can be obtained as in Kauppi and Saikkonen (2008).

[3] In models with the autoregressive structure ((2.6), (2.7) and (2.8)) a choice for the initial value π_0 is needed. As suggested by Kauppi and Saikkonen (2008), the initial value in model (2.6), for example, is set to $\pi_0 = (\omega + \delta_1\bar{y} + \bar{\boldsymbol{x}}'_{t-k}\boldsymbol{\beta})/(1-\alpha_1)$, where a bar is used to denote the sample mean of the considered variables.

2.2.2 Forecasts for the Recession Indicator

Kauppi and Saikkonen (2008) show how one period and multiperiod forecasts in models (2.4)–(2.8) can be constructed by explicit formulae. A practical problem with recession forecasting is that realized values of the recession indicator y_t defined in (2.1) are known after a considerable delay. The initial announcements of many major indicators of economic activity are preliminary and often subject to substantial revision. Thus it is difficult to identify the turning points in real time. For instance, the most recent announcements of business cycle peak and trough months in the U.S. have taken place from five up to twenty months after the business cycle turning point occurred.[4]

In this study, the "publication lag" in the recession indicator is assumed to be nine months. Owing to this assumed delay, the forecast horizon h consists of two periods. The first nine months $h = 1, 2, ..., 9$, are related to predictions of the most recent past values and the current value of the recession indicator. The longer-horizon forecasts ($h \geq 10$) are presumably the most interesting ones. Later in this study, this "ahead" forecast horizon is denoted by h^f, and defined as $h^f = h - 9$ for $h \geq 10$, where the number 9 is the assumed publication lag.

Kauppi and Saikkonen (2008) propose two methods of computing multiperiod recession forecasts, termed "direct" and "iterative" (cf. forecasts in continuous dependent time series models in, for example, Marcellino, Stock and Watson, 2006). A "direct" forecast is obtained by employing lagged values of the dependent variable y_{t-l} and explanatory variables $\boldsymbol{x}_{i,t-k}$ when $k \geq h^f$, provided that $k \geq 1$ and $l \geq h$. This forecast is direct in the sense that the right-hand side of model (2.6), for example, gives the h-step forecast "directly". An "iterative" forecast at time $t-h$ is obtained by accounting for all possible paths and their probabilities between y_{t-h} and y_t using the same one period model iteratively. Typically, y_{t-1} is used in the one-period model instead of forecast horizon-specific predictor y_{t-l} employed in direct forecasts.

[4] See details at http://www.nber.org/cycles/cyclesmain.html [20 March 2009].

2.3 Empirical Analysis of Recession Periods in the U.S. and Germany

2.3.1 Data and Predictive Variables

We consider the domestic and foreign term spreads (SP_t), stock market returns (r_t) and the interest rate differential between the U.S. and Germany (IS_t) as predictive variables in probit models. The term spread, defined as the difference between the long-term and the short-term interest rates, has been the most commonly used predictor in recession forecasting. The study of Estrella and Hardouvelis (1991) was among the first to find the term spread a useful predictor of economic growth and recession periods in the U.S. Bernard and Gerlach (1998), for instance, present similar evidence for Germany. Estrella (2005a, 2005b) and the references therein provide an extensive literature review and the main theoretical basis for the predictive power of the term spread.

Using the static probit model (2.4), Bernard and Gerlach (1998) show that, in addition to the domestic term spread, foreign term spreads are also useful predictors in some considered countries. Estrella and Mishkin (1998) find that the stock return is the only variable that has out-of-sample predictive power beyond the domestic term spread to predict U.S. recession periods in model (2.4). These variables have not been considered as predictors in dynamic probit models previously. Davis and Fagan (1997) include the interest rate differentials between EU countries to predict the output growth, but the evidence in favor of its predictive ability is quite weak. To the best of our knowledge, interest rate differentials between different countries have not been considered previously in recession prediction models.

The data set includes values of the recession indicator y_t and the considered explanatory variables $\boldsymbol{x}_t$ in the U.S. and Germany covering the period from January 1972 to December 2007. We adopt the recession periods defined by the National Bureau of Economic Research (NBER) for the U.S. and the Economic Cycle Research Institute (ECRI) for Germany. The data set of predictive variables is collected from various sources.[5]

2.3.2 In-Sample Results and Model Selection

In the in-sample analysis, the sample period from January 1972 to December 1994 is used to examine the performance of different probit models with various combinations of explanatory variables. In model evaluation, the main goodness-of-fit measure is the pseudo-R^2 measure suggested by Estrella (1998). Values of some other statistical measures are also presented in Table 2.1. We experiment with different lag orders k and l of the explanatory variables $\boldsymbol{x}_{t-k}$ and the lagged dependent variable y_{t-l}, respectively, with k and l varying between one and 12. In practice, it has been common to set k and l equal to the forecast horizon h. On the other hand, Estrella and Mishkin (1998) and Kauppi and Saikkonen (2008) have emphasized that the latest values of the predictive variables included in the information set at the time the forecast is made are not necessarily the best ones in terms of predictive power. This indicates that better results may be obtained by employing lags supported by model selection.

Tables 2.1 and 2.2 show the estimation results of the best in-sample models for both countries.[6] The sixth lags of the domestic and foreign term spreads performed consistently better, on the average, than the alternative lag orders in different probit models for both countries. Based on the model selection, the best lag orders for the stock returns and the interest rate differential are also used in the estimation results presented in Tables 2.1 and 2.2.

[5] Recession periods for the U.S. are obtained from http://www.nber.org/cycles/cyclesmain.html and German recession periods from http://www.businesscycle.com/resources/cycles. Interest rates are taken from http://www.federalreserve.gov/releases/h15/data.htm (10-year Treasury Bond and three-month Treasury Bill rate) and http://www.bundesbank.de/statistik/statistik (10-year interest rate and three-month money market rate). Stock returns are log-differences from the $S\&P500$ index (http://www.finance.yahoo.com) and German MSCI index (http://www.mscibarra.com) [20 March 2009].

[6] Details on all model selection results are available upon request. All estimations have been executed with Matlab 7.4.0 and its BFGS optimization routine in the Optimization Toolbox.

Table 2.1: In-sample results from recession prediction models for the U.S.

model	static(2.4)	static(2.4)	dynamic(2.5)	dyn.auto(2.6)	auto(2.7)	auto.int(2.8)
constant	-0.26	-0.54	-0.17	-2.37	-0.11	-0.06
	(0.24)	(0.28)	(0.03)	(0.46)	(0.09)	(0.08)
SP_{t-6}^{US}	-0.61	-0.42	-0.41	-0.51	-0.15	-0.32
	(0.13)	(0.13)	(0.14)	(0.20)	(0.04)	(0.09)
SP_{t-6}^{GE}		-0.38	-0.31	-0.31	-0.11	-0.07
		(0.09)	(0.09)	(0.11)	(0.04)	(0.05)
r_{t-2}^{US}		-0.05	-0.12	-0.12	-0.13	-0.16
		(0.02)	(0.02)	(0.03)	(0.03)	(0.04)
r_{t-4}^{US}		-0.10	-0.17	-0.17	-0.08	-0.08
		(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
r_{t-6}^{US}		-0.08	-0.07	-0.07	-0.05	-0.02
		(0.03)	(0.04)	(0.06)	(0.03)	(0.03)
π_{t-1}				-0.13	0.79	0.80
				(0.21)	(0.03)	(0.03)
y_{t-1}^{US}			3.88	4.55		
			(0.53)	(0.90)		
$y_{t-1}^{US} SP_{t-6}^{US}$						0.47
						(0.12)
log-L	-84.64	-58.93	-16.99	-16.79	-28.88	-22.34
psR^2	0.29	0.49	0.83	0.84	0.73	0.79
adj - psR^2	0.29	0.48	0.83	0.83	0.72	0.78
AIC	86.64	64.93	23.99	24.79	35.88	30.34
BIC	90.26	75.79	36.66	39.27	48.55	44.83
QPS	0.19	0.13	0.04	0.04	0.07	0.05
$CR^{50\%}$	0.89	0.90	0.98	0.97	0.95	0.96
$CR^{25\%}$	0.81	0.89	0.97	0.97	0.94	0.95

Notes: The models are estimated using monthly observations of the recession indicator (2.1) and explanatory variables from 1972 M1 to 1994 M12, $T = 276$. Robust standard errors suggested by Kauppi and Saikkonen (2008) are reported in parentheses. In the table, psR^2 reflects the pseudo-R^2 and adj - psR^2 the adjusted pseudo-R^2 (Estrella, 1998), which is calculated as $1 - (1 - psR^2)\frac{T-1}{T-K-1}$, where K is the number of estimated parameters. In addition, AIC and BIC are the values of the Akaike (1974) and Schwarz (1978) information criteria, QPS is the quadratic probability score (Diebold and Rudebusch, 1989). Furthermore, $CR^{50\%}$ and $CR^{25\%}$ indicate the ratio of correct predictions with 50 and 25 percent threshold values in the classification of recession probabilities.

Table 2.2: In-sample results from recession prediction models for Germany.

model	static(2.4)	static(2.4)	dynamic(2.5)	dyn.auto(2.6)	auto(2.7)	auto.int(2.8)
constant	0.13	0.66	-0.17	0.09	0.69	1.72
	(0.22)	(0.38)	(0.07)	(0.34)	(0.22)	(0.93)
SP_{t-6}^{US}		-0.55	-2.67	-1.10	-0.56	-2.53
		(0.20)	(0.78)	(0.68)	(0.19)	(0.68)
SP_{t-6}^{GE}	-0.95	-0.94	-0.67	-0.46	-0.47	-0.70
	(0.16)	(0.21)	(0.28)	(0.14)	(0.13)	(0.45)
r_{t-3}^{GE}		0.01	0.04	-0.06	-0.08	-0.15
		(0.02)	(0.04)	(0.05)	(0.03)	(0.06)
r_{t-6}^{GE}		-0.04	-0.17	-0.18	-0.10	-0.45
		(0.03)	(0.07)	(0.06)	(0.05)	(0.22)
r_{t-9}^{GE}		-0.06	-0.23	-0.35	-0.25	-0.73
		(0.03)	(0.07)	(0.15)	(0.06)	(0.23)
IS_{t-6}		-0.27	-0.79	-0.44	-0.26	-0.80
		(0.10)	(0.23)	(0.22)	(0.08)	(0.21)
π_{t-1}				0.64	0.80	0.77
				(0.10)	(0.01)	(0.03)
y_{t-1}^{GE}			7.48	2.38		
			(1.35)	(1.76)		
$y_{t-1}^{GE} SP_{t-6}^{US}$						1.65
						(0.48)
log-L	-72.14	-54.26	-11.98	-7.65	-12.56	-1.37
psR^2	0.69	0.78	0.97	0.98	0.97	0.99
adj - psR^2	0.68	0.77	0.97	0.98	0.97	0.99
AIC	74.14	61.26	19.98	16.65	20.56	10.37
BIC	77.76	73.93	34.46	32.94	35.04	26.66
QPS	0.16	0.13	0.03	0.02	0.03	0.01
$CR^{50\%}$	0.71	0.90	0.98	0.99	0.98	1.00
$CR^{25\%}$	0.70	0.89	0.97	0.99	0.97	1.00

Note: See notes to Table 2.1.

The main findings are very much the same for both countries. According to the model selection criteria, the first lag of the dependent variable y_{t-1} is superior for both countries, and it is a highly statistically significant predictor. This is in line with the findings of Kauppi and Saikkonen (2008), and it gives tentative evidence that the iterative multiperiod forecasts could be superior to horizon-specific direct forecasts in out-of-sample forecasting.

The domestic term spread is the primary financial explanatory variable, but the foreign term spread and most stock return lags are also statistically significant predictors. The signs of the estimated coefficients are negative as expected, indicating an increased probability of recession when the values of the term spreads are relatively low. Negative stock returns also increase the probability of recession, and it appears that the predictive power is distributed among several preceding stock market returns. The interest rate differential is a statistically significant predictor in the case of Germany. Its negative coefficient means that the recession probability increases when the short-term interest rate is higher in Germany than in the U.S. However, in the U.S. the interest rate differential turned out to be a statistically insignificant predictor and is, therefore, not included in the reported models in Table 2.1.

Overall, based on the in-sample evidence for both countries, it is clear that the foreign term spreads, several lagged stock returns and the interest rate differential in Germany add significantly to the predictive power of a model that contains only the domestic term spread as a single explanatory variable. The dynamic models (2.5)–(2.8) outperform the static model (2.4) in terms of in-sample performance. However, the static model augmented with the above-mentioned additional explanatory variables also outperforms the traditional static model where the domestic term spread is the only predictor (the first and the second models in Tables 2.1 and 2.2).

The best in-sample fit for the U.S. is obtained from the dynamic model (2.5) with the y_{t-1} predictor. On the other hand, the "pure" autoregressive model (2.7) also yields good in-sample fit with a relatively large and highly statistically significant estimate of the autoregressive coefficient α_1, indicating that the statistical improvement compared with the static model (2.4) is clear.

In the autoregressive interaction model (2.8) it seemed reasonable to use an interaction term of the form $y_{t-1}\boldsymbol{z}_{t-k}$. The first lag of the dependent variable y_{t-1} as such was excluded because its inclusion rendered the interaction term statistically insignificant, reducing the model to the dynamic model (2.5). In the models presented in Tables 2.1 and 2.2, the sixth lag of the term spread is used in both $\boldsymbol{z}_{t-k}$ and $\boldsymbol{x}_{t-k}$. For both countries the estimate of the interaction term coefficient is statistically significant, suggesting that the U.S. term spread has an asymmetric effect on recession probability, with the asymmetry depending on the state of the economy. As a matter of fact, this model yields the best in-sample fit for Germany. Interestingly, the U.S. term spread has a stronger asymmetric effect than the domestic term spread on German recession periods. The evidence of the asymmetric effect of the term spread is in accordance with monetary policy having a similar asymmetric effect on the real economy (see, e.g., Morgan, 1993; Florio, 2004).

An important issue in specifying a model for recession forecasting is the stability of the relationship between the explanatory variables and the recession indicator (2.1). We examined this by using the LM test proposed by Andrews and Fair (1988) and applied by Kauppi (2008) in the context of probit model when testing potential structural break dates suggested by Estrella, Rodrigues and Schich (2003). We found no evidence of structural breaks in the models presented in Tables 2.1 and 2.2 at conventional significance levels.[7] The evidence is in line with the findings of Estrella *et al.* (2003), Wright (2006) and Kauppi (2008).

Figure 2.1 depicts the in-sample recession probabilities of the static model (2.4) and the autoregressive model (2.7) (the first and the fifth models in Tables 2.1 and 2.2). These models are used as such also in out-of-sample forecasting in the following section.[8] It can be seen that in the autoregressive model the recession

[7] Further details are available upon request.

[8] This is not the case in the models employing y_{t-1} as a predictive variable because those one-period models are used iteratively in out-of-sample forecasting to obtain the iterative recession forecast.

probability matches better with the realized values of the recession indicator than in the static model. In recession periods, the recession probabilities are also higher in the autoregressive model. When the economy is in an expansionary state, the recession probability is constantly higher in the static model, whereas in the autoregressive model, it is very close to zero, as it should be.

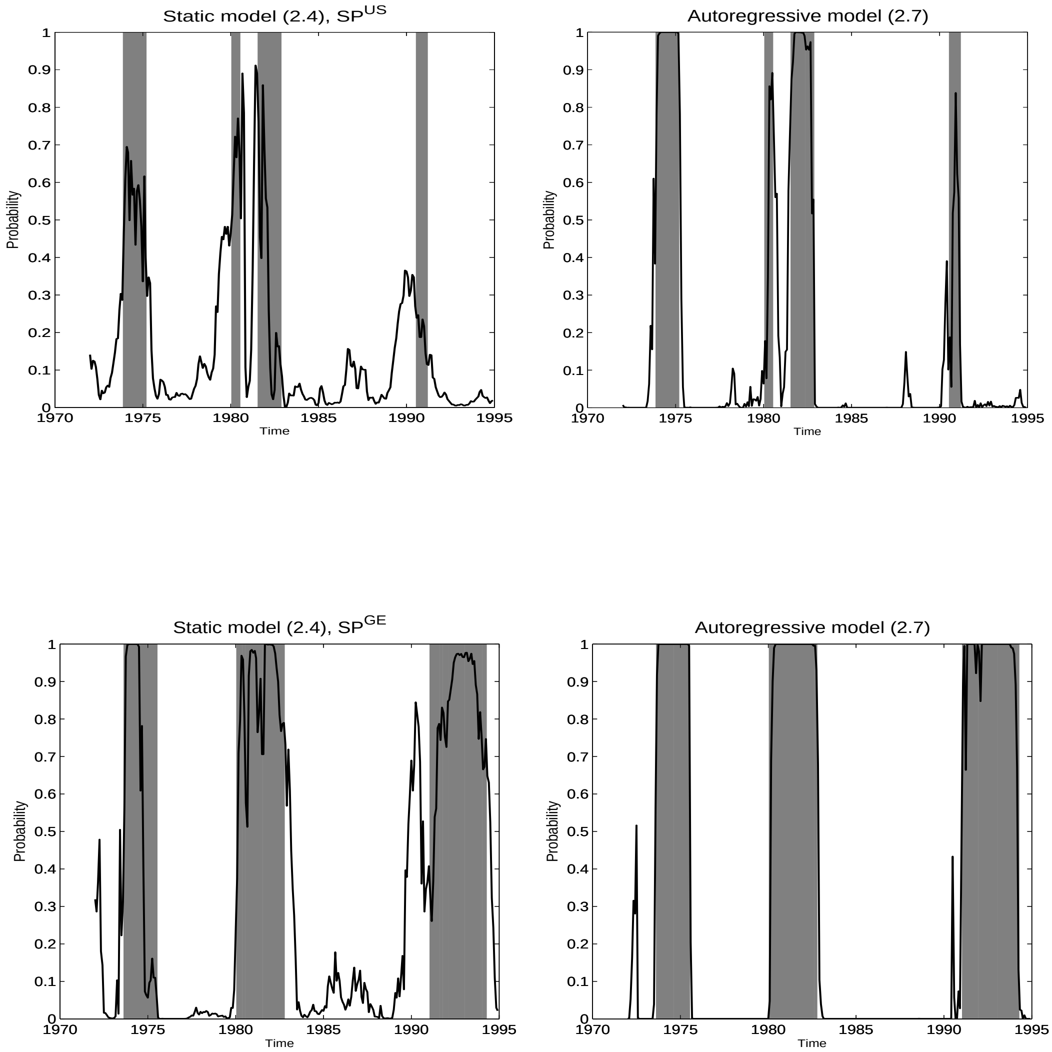

Figure 2.1: In-sample recession probabilities implied by the static model (2.4), with the domestic term spread as the only predictor and the autoregressive model (2.7) with additional explanatory variables for the U.S. (upper panel) and Germany (lower panel) as given in Tables 2.1 and 2.2.

2.3.3 Out-of-Sample Forecasting Results

The in-sample evidence shows a great deal of predictability for recession periods in the U.S and Germany. However, in-sample predictability does not necessarily mean out-of-sample predictability. For instance, some of the static probit models considered by Estrella and Mishkin (1998) for the U.S. recession periods provide the best in-sample fitted values, but perform quite poorly out of sample.

In this chapter, the first out-of-sample predictions are made for January 1995 and last ones for March 2007. The forecast period thus contains the recession period that began in both countries in 2001. In other months, both economies are in an expansionary state. The parameters are estimated recursively. In other words, after adding one month to the previous estimation period and re-estimating the parameters, forecasts for the next month are computed. This procedure is repeated recursively until the end of the forecast period.

We examine the out-of-sample predictive ability of the models that turned out to be the best ones according to the in-sample results presented in Tables 2.1 and 2.2. Therefore, we employ the following explanatory variables,

$$\boldsymbol{x}_{t-k}^{US} = \left(SP_{t-6}^{US}, \quad SP_{t-6}^{GE}, \quad r_{t-2}^{US}, \quad r_{t-4}^{US}, \quad r_{t-6}^{US}\right)' \tag{2.9}$$

and

$$\boldsymbol{x}_{t-k}^{GE} = \left(SP_{t-6}^{GE}, \quad SP_{t-6}^{US}, \quad r_{t-3}^{GE}, \quad r_{t-6}^{GE}, \quad r_{t-9}^{GE}, \quad IS_{t-6}\right)'. \tag{2.10}$$

Models where the foreign term spread is excluded are also examined. In these cases, the vectors of the explanatory variables are denoted by $\boldsymbol{x}_{t-k}^{US*}$ and $\boldsymbol{x}_{t-k}^{GE*}$. When the domestic term spread is the only predictor, the corresponding vectors are denoted by $\boldsymbol{v}_{t-k}^{US}$ and $\boldsymbol{v}_{t-k}^{GE}$.[9]

With the forecast horizon h, the lags in the explanatory variables should be tailored so that only the information included in the information set Ω_{t-h} at the forecast time $t-h$ is used. For example, if the forecast horizon is 16 months ($h = 16$), and because the publication lag is assumed to be nine months, it means

[9] The lag orders used in the explanatory variables in these two cases are the same as in the vectors $\boldsymbol{x}_{t-k}^{US}$ and $\boldsymbol{x}_{t-k}^{GE}$ presented in (2.9) and (2.10).

that we are interested in forecasting the value of the U.S. recession indicator seven months ($h^f = 7$) ahead, and the vector $\boldsymbol{x}_{t-k}^{US}$ is given by

$$\boldsymbol{x}_{t-k}^{US} \;=\; \left(SP_{t-7}^{US}, \quad SP_{t-7}^{GE}, \quad r_{t-7}^{US} \right)'.$$

Forecast accuracy is evaluated by using the same statistical goodness-of-fit measures as in the in-sample analysis. In addition, 50% and 25% threshold values are used to classify recession probabilities and to construct "strong", "weak" and "no" recession signals. For example, if the recession probability is between 25 and 50 percent, the model gives a "weak" recession signal. Related to these signal forecasts an asymmetric "forecasting point" scheme is applied. The idea, illustrated in Table 2.4, is to put greater emphasis on correct forecasts (cf. Dueker, 2002). It also favors a false recession alarm compared with a missed recession month. One rationale behind this is that firms or policymakers, for example, are possibly willing to take a "recession insurance" and accept a possible false alarm rather than be caught by an unexpected recession.

As discussed earlier in the context of multiperiod forecasting, the most interesting forecasts are typically the future values of the recession indicator. Thus we concentrate on forecasts with forecast horizon $h \geq 10$ ($h^f \geq 1$). The results concerning shorter horizons ($h \leq 9$) are available upon request. It is worth noting that, in practice, iterative forecasts are computationally very demanding when the forecast horizon is as long as 21 months in Tables 2.3 and 2.4.[10] This is a difficulty of the iterative forecasting approach employed in the dynamic models (2.5) and (2.8) with y_{t-1} as a predictor. Therefore, only forecasts based on the static model (2.4) and the autoregressive model (2.7) are considered when the forecast horizon is so long that iterative forecasting becomes computationally infeasible.

Based on the adjusted-pseudo-R^2 the best predictive models yield good out-of-sample forecasts for the state of the U.S. economy. In Table 2.3, the highest values of the adjusted-pseudo-R^2 in different probit models are obtained with the forecast horizon of 15 months ($h^f = 6$).[11] At this forecast horizon, the autoregressive

[10] In iterative forecasts with $h = 21$, 2^{21} different paths need to be computed and the computational burden increases rapidly if longer publication lags are considered.

[11] The values of the adjusted-pseudo-R^2 are adjusted to the number of parameters estimated.

Table 2.3: Adjusted pseudo-R^2 measures of out-of-sample predictive performance for different models in the U.S.

model	h / h^f	10 / 1	11 / 2	12 / 3	13 / 4	14 / 5	15 / 6	16 / 7	21 / 12
static (2.4); $\boldsymbol{x}_{t-k}^{US}$		0.31	0.31	0.26	0.29	0.26	0.26	0.26	0.28
dyn.iter (2.5); $\boldsymbol{x}_{t-k}^{US}$		0.50	0.52	0.49	0.49	0.37	0.37	0.37	–
auto (2.7); $\boldsymbol{x}_{t-k}^{US}$		0.46	0.45	0.46	0.46	0.45	0.45	0.44	0.26
auto.int (2.8); $\boldsymbol{x}_{t-k}^{US}, SP_{t-6}^{US}$		0.33	0.30	0.51	0.51	0.53	0.53	0.53	–
auto.int (2.8); $\boldsymbol{x}_{t-k}^{US}, SP_{t-8}^{US}$		0.10	0.05	0.45	0.44	0.53	0.53	0.54	–
static (2.4); $\boldsymbol{x}_{t-k}^{US*}$		0.18	0.18	0.16	0.16	0.10	0.10	0.15	0.17
dyn.iter (2.5); $\boldsymbol{x}_{t-k}^{US*}$		0.43	0.45	0.41	0.40	0.26	0.26	0.25	–
auto (2.7); $\boldsymbol{x}_{t-k}^{US*}$		0.42	0.42	0.39	0.39	0.38	0.38	0.34	0.14
auto.int (2.8); $\boldsymbol{x}_{t-k}^{US*}, SP_{t-6}^{US}$		0.21	0.18	0.40	0.38	0.43	0.43	0.36	–
static (2.4); $\boldsymbol{v}_{t-k}^{US}$		0.08	0.08	0.08	0.07	0.07	0.08	0.13	0.18
dyn.iter (2.5); $\boldsymbol{v}_{t-k}^{US}$		0.24	0.24	0.24	0.24	0.24	0.24	0.23	–
auto (2.7); $\boldsymbol{v}_{t-k}^{US}$		0.26	0.26	0.26	0.26	0.26	0.26	0.25	0.16
auto.int (2.8); $\boldsymbol{v}_{t-k}^{US}, SP_{t-6}^{US}$		0.31	0.30	0.29	0.29	0.28	0.28	0.23	–

Notes: The table presents the adjusted pseudo-R^2 values (see Table 2.1 and Estrella, 1998) of different models in out-of-sample predictions. The probit model is denoted at the left with the explanatory variables included in the model. As in (2.9), and in the subsequent discussion, $\boldsymbol{x}_{t-k}^{US}$ include all explanatory variables, in $\boldsymbol{x}_{t-k}^{US*}$ the German term spread is excluded and in $\boldsymbol{v}_{t-k}^{US}$ only the U.S. term spread is employed in the model. In the autoregressive interaction model (2.8), the term spread that is used in the interaction term $\boldsymbol{z}_{t-k}$ is also mentioned. In dynamic model (2.5) and in autoregressive interaction model (2.8) the first lagged value of the recession indicator y_{t-1} is used in the model.

model (2.7) and the autoregressive model with the interaction term (2.8) outperform the dynamic model (2.5), yielding the best out-of-sample forecasts given the explanatory variables included in the vector $\boldsymbol{x}_{t-k}^{US}$ in (2.9). In fact, the autoregressive interaction model (2.8) is the best model when the forecast horizon is between

The evidence from different goodness-of-fit measures, such as forecasting points, information criteria or QPS, is the same. Further, and also in the case of Germany in Table 2.4, results from the dynamic model (2.5) with y_{t-h} and the dynamic autoregressive model (2.6) are excluded because forecasts from the restricted static model (2.4) and the dynamic model (2.5) with y_{t-1} yield almost the same or even better predictions than these more general models.

12 to 16 months, providing evidence that the asymmetric predictive power of the U.S. term spread found in the in-sample analysis also shows up in out-of-sample predictions.

Overall, the models with the U.S. stock return ($\boldsymbol{x}_{t-k}^{US*}$) and the models that also include the German term spread ($\boldsymbol{x}_{t-k}^{US}$) outperform the models with the U.S. term spread ($\boldsymbol{v}_{t-k}^{US}$) as the only predictor across all probit model specifications and forecast horizons. This suggests that these additional financial variables have not only in-sample but also out-of-sample predictive content for the U.S. recessions.

In Germany, the recession in the out-of-sample period lasted considerably longer than in the U.S, but the essential conclusions between different predictive models are parallel to those for the U.S. However, because the term spreads soared immediately after the recession began, the recession probability decreased amidst the recession period. Consequently, the negative values of the adjusted-pseudo-R^2 were obtained for some models, making comparisons difficult. Therefore, the forecasting points presented in Table 2.4 are the main model evaluation measure for Germany.

The results of Table 2.4 confirm the previous in-sample findings. Even out of sample, the interest rate differential between the U.S. and Germany and the German stock return clearly have additional predictive power beyond the German term spread in all probit models. As in the case of the U.S., when the forecast horizon increases towards 15 or 16 months the models with an autoregressive structure, (2.7) and (2.8), seem to outperform their competitors. Interestingly, the U.S. term spread, which is a statistically significant predictor in sample, seems to be a rather poor predictor out of sample, because the forecasting results are much better without it. The statistical significance of the interest rate differential, however, suggests that the U.S. monetary policy has an impact on the probability of recession in Germany via the U.S. short-term interest rate.

Paap, Segers and van Dijk (2009) propose a model that allows for asymmetries such that, for example, the term spread has a different lead time in recession and expansion. This issue can be examined with model (2.8) by selecting different lag orders of explanatory variables included in $\boldsymbol{z}_{t-k}$ and in $\boldsymbol{x}_{t-k}$. In Tables 2.3 and

Table 2.4: Out-of-sample forecasting points of employed predictive models for Germany.

model	h / h^f	10 / 1	11 / 2	12 / 3	13 / 4	14 / 5	15 / 6	16 / 7	21 / 12
static (2.4); $\boldsymbol{x}_{t-k}^{GE}$		0.53	0.49	0.49	0.50	0.48	0.50	0.48	0.50
dyn.iter (2.5); $\boldsymbol{x}_{t-k}^{GE}$		0.66	0.64	0.63	0.61	0.60	0.60	0.55	–
auto (2.7); $\boldsymbol{x}_{t-k}^{GE}$		0.74	0.74	0.74	0.73	0.73	0.70	0.71	0.52
auto.int (2.8); $\boldsymbol{x}_{t-k}^{GE}, SP_{t-6}^{US}$		0.65	0.63	0.60	0.60	0.60	0.60	0.49	–
auto.int (2.8); $\boldsymbol{x}_{t-k}^{GE}, SP_{t-6}^{GE}$		0.60	0.52	0.52	0.50	0.50	0.48	0.45	–
static (2.4); $\boldsymbol{x}_{t-k}^{GE*}$		0.60	0.58	0.58	0.58	0.58	0.58	0.58	0.53
dyn.iter (2.5); $\boldsymbol{x}_{t-k}^{GE*}$		0.72	0.71	0.67	0.69	0.67	0.67	0.61	–
auto (2.7); $\boldsymbol{x}_{t-k}^{GE*}$		0.81	0.81	0.81	0.71	0.71	0.71	0.70	0.53
auto.int (2.8); $\boldsymbol{x}_{t-k}^{GE*}, SP_{t-6}^{GE}$		0.66	0.67	0.62	0.75	0.75	0.75	0.70	–
auto.int (2.8); $\boldsymbol{x}_{t-k}^{GE*}, SP_{t-7}^{GE}$		0.85	0.84	0.85	0.73	0.73	0.73	0.70	–
static (2.4); $\boldsymbol{v}_{t-k}^{GE}$		0.29	0.29	0.29	0.30	0.29	0.30	0.41	0.30
dyn.iter (2.5); $\boldsymbol{v}_{t-k}^{GE}$		0.56	0.55	0.52	0.48	0.49	0.48	0.43	–
auto (2.7); $\boldsymbol{v}_{t-k}^{GE}$		0.44	0.44	0.44	0.44	0.42	0.42	0.41	0.42
auto.int (2.8); $\boldsymbol{v}_{t-k}^{GE}, SP_{t-6}^{GE}$		0.52	0.52	0.52	0.52	0.52	0.52	0.50	–

Notes: Forecasting points are obtained from the point scheme presented below by dividing the sum of individual points by the number of maximum points obtained when predicting the state of the economy correctly in every month. Forecast point scheme is as follows:

signal		recession ($y_t = 1$)	expansion ($y_t = 0$)
"strong" recession signal	$p_t \geq 0.50$	1	-1
"weak" recession signal	$0.25 \leq p_t < 0.50$	1/2	0
"no" recession signal	$p_t < 0.25$	-1	1/2,

where p_t is the recession probability obtained from (2.3). As in (2.10), and in the subsequent discussion, $\boldsymbol{x}_{t-k}^{GE}$ include all explanatory variables, in $\boldsymbol{x}_{t-k}^{GE*}$ the U.S. term spread is excluded and in $\boldsymbol{v}_{t-k}^{GE}$ only the German term spread is employed in the model. See also notes to Table 2.3.

2.4 two models are considered where the predictive lag of the term spread in $\boldsymbol{z}_{t-k}$ is selected based on the in-sample model selection. For instance, the selection $\boldsymbol{z}_{t-k} = SP_{t-8}^{US}$ seems to be the best one for the U.S. However, according to the out-of-sample results of both countries, differences between the presented two versions of model (2.8) using different lag orders in $\boldsymbol{z}_{t-k}$ are minor.

For both countries, the best autoregressive interaction model (2.8) seems to generate somewhat better out-of-sample forecasts than the dynamic model (2.5). This is in contrast with the findings of Kauppi and Saikkonen (2008) and may be due to the additional financial explanatory variables used in this study. However, it should be pointed out that the autoregressive model (2.7) produces almost as good out-of-sample predictions as model (2.8) for both countries. Moreover, forecasts obtained with model (2.7) do not require computationally intensive iterative methods. When the forecast horizon is 21 months, which is the longest horizon considered, the static model (2.4) without any dynamics turns out to be an adequate model. However, also with this forecast horizon, the additional explanatory variables have useful predictive power.

Figure 2.2 illustrates the out-of-sample performance of the static model (2.4) with only the domestic term spread and the autoregressive interaction model (2.8) with additional explanatory variables. The forecast horizon is 15 months ($h = 15$). The performance of the static model considered in many previous studies is inferior to the autoregressive interaction model for both countries. The latter model has predictive power, especially at the beginning and at the end of the latest recession period for both countries.

In recession forecasting the probability of continued expansion (a time period where the economy is in an expansionary state every month) is of particular interest (see Chauvet and Potter, 2005). Continued expansion probabilities give a similar impression of expansionary and recessionary periods as the month-to-month predictions discussed above.[12] The predictive ability of the different models appears to depend on the state of the economy. During expansion the static model seems to overpredict the recession probability, whereas the dynamic models perform better in this respect (see, for example, Figure 2.2). On the other hand, in recession periods the dynamic model (2.5) constantly gives the highest probabilities of continued expansion. Thus also according to the continued expansion probabilities, the autoregressive probit models (2.7) and (2.8), including the domestic term spread and

[12] The results are available upon request.

other explanatory variables, seem to yield the most reliable predictions.

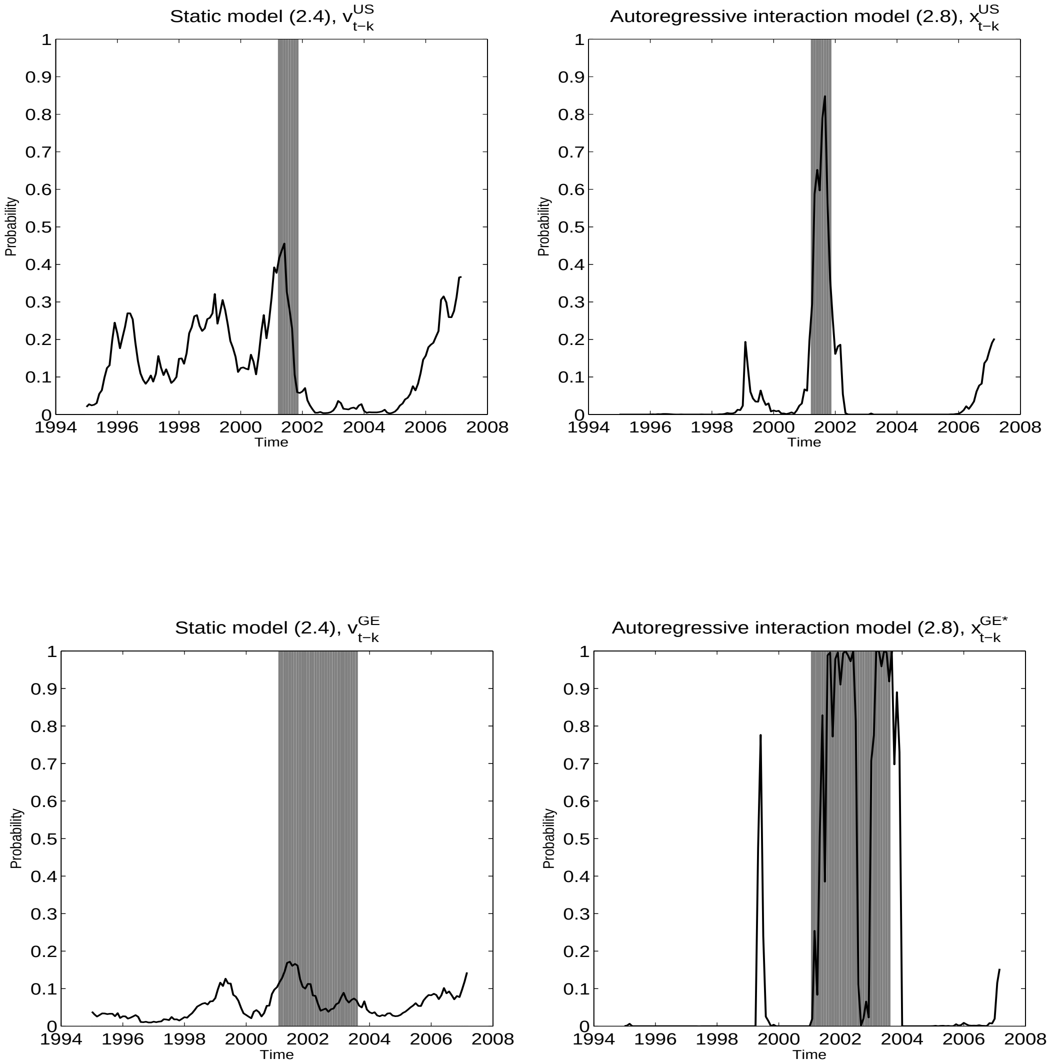

Figure 2.2: Out-of-sample recession forecasts ($h = 15, h^f = 6$) for the U.S. (upper panel) and Germany (lower panel) based on the standard static model (2.4) with the domestic term spread as the only predictor (left) and from the autoregressive interaction model (2.8) (right), where the employed explanatory variables are given in (2.9) and (2.10), and in the subsequent discussion.

2.3.4 Recession Probabilities in 2006–2008

In this section, we consider out-of-sample recession forecasts up to June 2008. Figure 2.3 depicts the recession probabilities from the beginning of year 2006 for both countries. The forecast horizon is again 15 months ($h = 15, h^f = 6$), indicating that the latest forecasts for June 2008 are based on predictive information from December 2007. In December 2008 the NBER announced that a peak in U.S. eco-

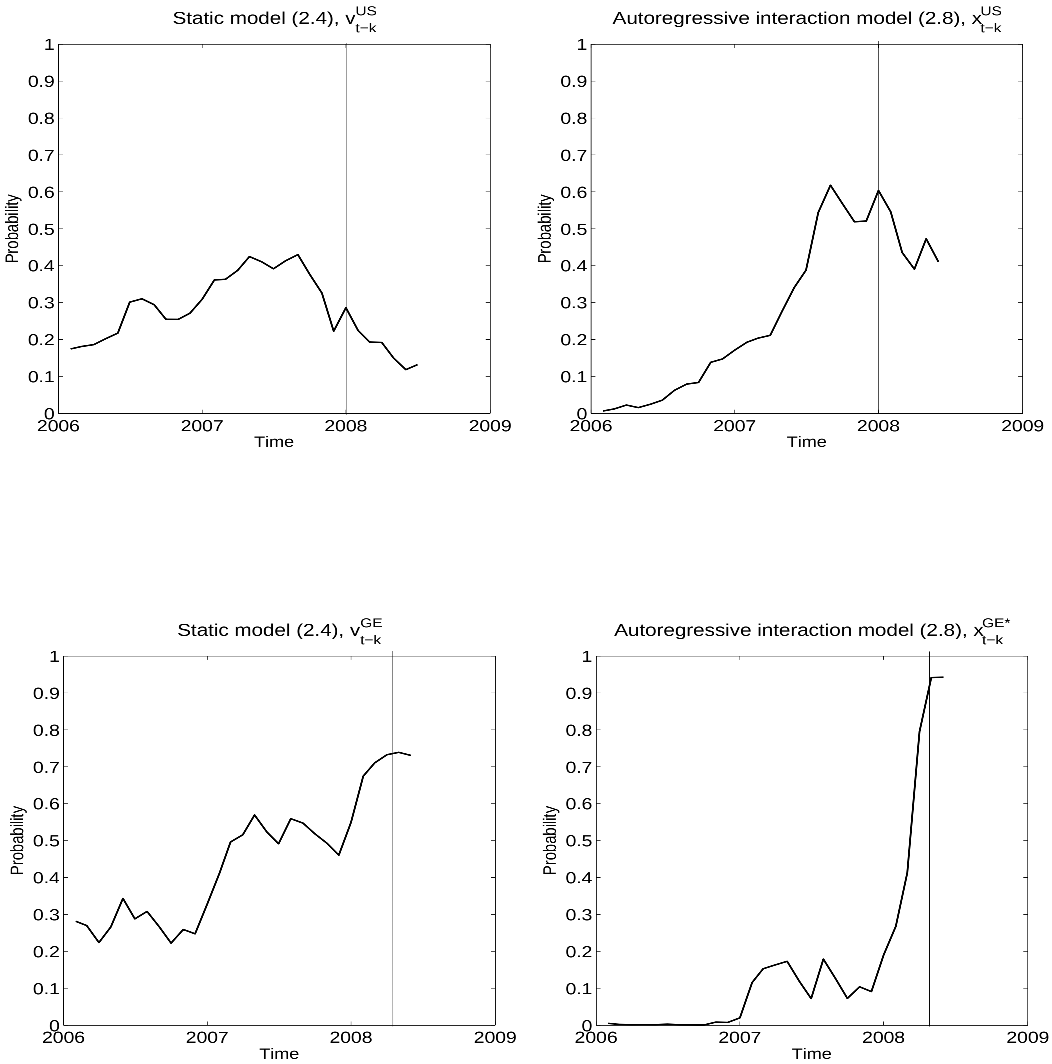

Figure 2.3: Recession forecasts ($h = 15, h^f = 6$) for the U.S. (upper panel) and Germany (lower panel) in 2006 M2–2008 M6. Forecasts from the static model (2.4) (left) with only the domestic term spread employed and the autoregressive interaction model (2.8) (right), where the employed explanatory variables are given in (2.9) and (2.10), and in the subsequent discussion. The beginning of a recession is indicated by a vertical line.

nomic activity had occurred in December 2007. Similarly, ECRI made an announcement that for Germany a peak had occurred in April 2008. After these peak months, both countries have been in a recession.

As seen in Figure 2.3, the autoregressive interaction model (2.8) predicts the beginning of the recession well for both countries, especially for Germany. For the U.S., the recession forecast exceeded the 50% threshold value in August 2007 and thereafter the recession probability has been relatively high. The static model (2.4) with the domestic term spread as the only explanatory variable, which has been the standard recession prediction model, does not give as precise recession and expansion signals as the autoregressive interaction model. As in the previous section, the performance of this standard model appears disappointing in comparison with different dynamic models, such as model (2.8).

2.4 Conclusions

We examine the performance of recession prediction models that include a number of financial explanatory variables. The results indicate that, compared with the standard static recession prediction model used in many previous studies, statistically significant additional predictive power is obtained by allowing for dynamic structures in the model. In particular, models with an autoregressive structure outperformed the static model and they were also somewhat better than other dynamic models considered in terms of out-of-sample performance. The best model for the U.S. and Germany turned out to be an autoregressive interaction model in which the term spread between the long-term and short-term interest rate has an asymmetric effect on recession probability, with the asymmetry depending on the state of the economy.

In accordance with previous studies, the term spread is found a useful predictor for both the U.S. and German recession periods, but for both countries, additional predictive power is provided by stock returns. For Germany, the short-term interest rate differential between the U.S and Germany also has substantial predictive power in both in-sample and out-of-sample prediction. The same holds for the

German term spread when forecasting the U.S. recessions. Furthermore, the U.S. term spread is a statistically significant predictor in the case of Germany, but its out-of-sample predictive power appears poor.

References

Andrews DWK, Fair RC. 1988. Inference in non-linear econometric models with structural change. *Review of Economic Studies* **55**: 615–640.

Akaike H. 1974. A new look at statistical model identification. *IEEE Transactions on Automatic Control* **19**: 713–723.

Bernard H, Gerlach S. 1998. Does the term structure predict recessions? The international evidence. *International Journal of Finance and Economics* **3**: 195–215.

Chauvet M, Potter S. 2005. Forecasting recession using the yield curve. *Journal of Forecasting* **24**: 77–103.

Davis EP, Fagan G. 1997. Are financial spreads useful indicators of future inflation and output growth in EU countries? *Journal of Applied Econometrics* **12**: 701–714.

de Jong RM, Woutersen TM. In press. Dynamic time series binary choice. *Econometric Theory*, forthcoming.

Diebold FX, Rudebusch GD. 1989. Scoring the leading indicators. *Journal of Business* **62**: 369–391.

Dueker MJ. 1997. Strengthening the case for the yield curve as a predictor of U.S. recessions. *Federal Reserve Bank of St.Louis Review* **79**: 41–51.

Dueker MJ. 2002. Regime-dependent recession forecasts and the 2001 recession. *Federal Reserve Bank of St.Louis Review* **84**: 29–36.

Dueker MJ. 2005. Dynamic forecasts of qualitative variables: A qual VAR model of U.S. recessions. *Journal of Business and Economic Statistics* **23**: 96–104.

Estrella A. 1998. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics* **16**: 198–205.

Estrella A. 2005a. The yield curve as a leading indicator: Frequently asked questions. Federal Reserve Bank of New York. http:/www.newyorkfed.org/research/capital_markets/ycfaq.pdf. [30 October 2008].

Estrella A. 2005b. Why does the yield curve predict output and inflation? *Economic Journal* **115**: 722–744.

Estrella A, Hardouvelis GA. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* **46**: 555–576.

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45–61.

Estrella A, Rodrigues AP. 1998. Consistent covariance matrix estimation in probit models with autocorrelated disturbances. *Federal Reserve Bank of New York Staff Reports No.* 39.

Estrella A, Rodrigues AP, Schich S. 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.

Florio A. 2004. The asymmetric effects of monetary policy. *Journal of Economic Surveys* **18**: 409–426.

Kauppi H. 2008. Yield-Curve Based Probit Models for Forecasting U.S. Recessions: Stability and Dynamics. HECER Discussion Paper, 221. Helsinki Center of Economic Research.

Kauppi H, Saikkonen P. 2008. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics* **90**: 777–791.

Marcellino M, Stock JH, Watson MW. 2006. A comparison of direct and iterated AR methods for forecasting macroeconomic time series. *Journal of Econometrics* **135**: 499–526.

Morgan DP. 1993. Asymmetric effects of monetary policy. *Federal Reserve Bank of Kansas City Economic Review* **78**: 21–33.

Paap R, Segers R, van Dijk D. 2009. Do Leading Indicators Lead Peaks More Than Troughs? *Journal of Business and Economic Statistics* **27**: 528–543.

Rydberg T, Shephard N. 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.

Startz R. 2008. Binomial autoregressive moving average models with an application to U.S. Recessions. *Journal of Business and Economic Statistics* **26**: 1–8.

Stock JH, Watson MW. 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* **41**: 788–829.

Wright JH. 2006. The yield curve and predicting recessions. *Finance and Economics Discussion Series No.* 7, Board of Governors of the Federal Reserve System.

Chapter 3

Testing an Autoregressive Structure in Binary Time Series Models

Abstract[1]

This chapter introduces a Lagrange Multiplier (LM) test for testing an autoregressive structure in a binary time series model proposed by Kauppi and Saikkonen (2008). Simulation results indicate that the two versions of the proposed LM test have reasonable size and power properties when the sample size is large. A parametric bootstrap method is suggested to obtain approximately correct sizes also in small samples. The use of the test is illustrated by an application to recession forecasting models using monthly U.S. data.

[1] An earlier version of this chapter has been published in HECER Discussion Papers, No. 243, 2008.

3.1 Introduction

Recently, Rydberg and Shephard (2003), Chauvet and Potter (2005) and Startz (2008), among others, have introduced new time series models for binary dependent variables. In this chapter, the "dynamic autoregressive" probit model suggested by Kauppi and Saikkonen (2008) is considered. We develop Lagrange Multiplier (LM) test which can be used to test the adequacy of a restricted model in which the autoregressive structure is excluded.

The proposed LM test is attractive because it only requires estimates from the restricted models, which can be obtained by using standard econometric software packages. According to our simulations, the two versions of the LM test considered have reasonable size and high power, especially in large samples. In small samples, a parametric bootstrap method is proposed to obtain critical values which are more reliable than the asymptotic ones. In an empirical application, the LM tests are used to assess recession forecasting models for the U.S.

The plan of this chapter is as follows. The probit model is introduced in Section 3.2 and the LM tests are developed in Section 3.3. Results of the simulation and bootstrap experiments are provided in Section 3.4 and the empirical example is presented in Section 3.5. Finally, Section 3.6 concludes.

3.2 Model

Consider the binary valued stochastic process y_t, $t = 1, 2, ..., T$, and let $E_{t-1}(\cdot)$ and $P_{t-1}(\cdot)$, respectively, signify the conditional expectation and the conditional probability given the information set Ω_{t-1}. Conditional on Ω_{t-1}, y_t has a Bernoulli distribution, that is,

$$y_t|\Omega_{t-1} \sim B(p_t). \tag{3.1}$$

In the probit model

$$p_t = E_{t-1}(y_t) = P_{t-1}(y_t = 1) = \Phi(\pi_t(\boldsymbol{\theta})), \tag{3.2}$$

where $\Phi(\cdot)$ is a standard normal cumulative distribution function. The model $\pi_t(\boldsymbol{\theta})$ is a linear function of variables in the information set Ω_{t-1} and the parameter vector $\boldsymbol{\theta}$.

In the previous literature, the "static" model

$$\pi_t(\boldsymbol{\theta}) = \omega + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{3.3}$$

has been the most commonly used specification. It has been employed in various applications, such as forecasting the recession periods of an economy (see, e.g., Estrella and Mishkin, 1998). A natural extension of the static model (3.3) is the dynamic specification

$$\pi_t(\boldsymbol{\theta}) = \omega + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{3.4}$$

where the lagged value of the dependent variable is also assumed to belong to the information set (see, e.g., Cox, 1981).

Kauppi and Saikkonen (2008) generalize the dynamic model (3.4) by adding a lagged value of $\pi_t(\boldsymbol{\theta})$, giving

$$\pi_t(\boldsymbol{\theta}) = \omega + \alpha_1 \pi_{t-1}(\boldsymbol{\theta}) + \delta_1 y_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{3.5}$$

where $|\alpha_1| < 1$.[2] This induces a first-order autoregressive structure to the model equation. It is worth noting that alternative, but very similar, models have been proposed by Rydberg and Shephard (2003) and Kauppi (2008). The LM tests developed in the next section can straightforwardly be extended to these models as well.

The parameters of models (3.3)–(3.5) can conveniently be estimated by the method of maximum likelihood (ML). Conditional on initial values, the log-likelihood function is

$$l(\boldsymbol{\theta}) = \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \Big(y_t \log(\Phi(\pi_t(\boldsymbol{\theta}))) + (1-y_t)\log(1-\Phi(\pi_t(\boldsymbol{\theta}))) \Big), \tag{3.6}$$

where $l_t(\boldsymbol{\theta})$ is the log-likelihood for t:th observation. The score function is

$$s(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^{T} s_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \Big(\frac{y_t - \Phi(\pi_t(\boldsymbol{\theta}))}{\Phi(\pi_t(\boldsymbol{\theta}))(1-\Phi(\pi_t(\boldsymbol{\theta})))} \phi(\pi_t(\boldsymbol{\theta})) \frac{\partial \pi_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big), \tag{3.7}$$

[2] For simplicity, only the first lags of the dependent variable, y_{t-1}, and $\pi_t(\boldsymbol{\theta})$, $\pi_{t-1}(\boldsymbol{\theta})$, are employed.

where $\phi(\cdot)$ signifies the probability density function of the standard normal distribution and an explicit expression of the derivative term $\partial\pi_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ will be given in the next section. The ML estimator $\hat{\boldsymbol{\theta}}$, which solves the first order condition $s(\hat{\boldsymbol{\theta}}) = 0$, is found by maximizing the log-likelihood function (3.6) with numerical methods.

3.3 LM Tests

In applications, model (3.5) may be a superior to its restricted version (3.4) but, on the other hand, its ML estimation is more complicated and no estimation procedures are readily available in standard econometric software packages. Thus, it is of interest to start with the simpler model (3.4) and check for its adequacy by testing whether the autoregressive coefficient α_1 in (3.5) is zero. The null hypothesis of interest is therefore

$$H_0 : \alpha_1 = 0. \tag{3.8}$$

In this context, the LM test is attractive because it only requires the estimation of the parameters of model (3.4). The general LM test statistic (see, e.g., Engle, 1984) for the null hypothesis (3.8) can be written as

$$LM = s(\tilde{\boldsymbol{\theta}})'\mathcal{I}(\tilde{\boldsymbol{\theta}})^{-1}s(\tilde{\boldsymbol{\theta}}), \tag{3.9}$$

where $\tilde{\boldsymbol{\theta}}$ is the restricted ML estimate of $\boldsymbol{\theta}$ restricted by (3.8), $s(\tilde{\boldsymbol{\theta}})$ is the score vector (3.7) evaluated at $\tilde{\boldsymbol{\theta}}$, and $\mathcal{I}(\tilde{\boldsymbol{\theta}})$ is a consistent estimate of the information matrix $\mathcal{I}(\boldsymbol{\theta})$. Under H_0, the test statistic (3.9) has an asymptotic χ^2_1 distribution.

Following Davidson and MacKinnon (1984) we can construct two LM test statistics for the null hypothesis (3.8). The first one is

$$LM_1 = \boldsymbol{\iota}'S(\tilde{\boldsymbol{\theta}})\Big(S(\tilde{\boldsymbol{\theta}})'S(\tilde{\boldsymbol{\theta}})\Big)^{-1}S(\tilde{\boldsymbol{\theta}})'\boldsymbol{\iota}, \tag{3.10}$$

where $\boldsymbol{\iota}$ is a vector of ones and the matrix $S(\tilde{\boldsymbol{\theta}})$ is given by

$$S(\tilde{\boldsymbol{\theta}}) = \Big(s_1(\tilde{\boldsymbol{\theta}}) \quad s_2(\tilde{\boldsymbol{\theta}}) \ldots s_T(\tilde{\boldsymbol{\theta}})\Big)'.$$

Expression (3.10) can also be seen as the regression sum of squares from the artificial linear regression

$$\boldsymbol{\iota} = S(\tilde{\boldsymbol{\theta}})a + error.$$

Using the symbols $\tilde{\Phi}_t = \Phi(\pi_t(\tilde{\boldsymbol{\theta}}))$ and $\tilde{\phi}_t = \phi(\pi_t(\tilde{\boldsymbol{\theta}}))$, a second LM test statistic can be based on the artificial regression

$$r(\tilde{\boldsymbol{\theta}}) = R(\tilde{\boldsymbol{\theta}})b + error, \tag{3.11}$$

where

$$R(\tilde{\boldsymbol{\theta}}) = \Big(R_1(\tilde{\boldsymbol{\theta}})' \quad R_2(\tilde{\boldsymbol{\theta}})' \ldots R_T(\tilde{\boldsymbol{\theta}})'\Big)'$$

with

$$R_t(\tilde{\boldsymbol{\theta}}) = \Big(\tilde{\Phi}_t(1-\tilde{\Phi}_t)\Big)^{-1/2}\tilde{\phi}_t\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}$$

and

$$r(\tilde{\boldsymbol{\theta}}) = \Big(r_1(\tilde{\boldsymbol{\theta}}) \quad r_2(\tilde{\boldsymbol{\theta}}) \ldots r_T(\tilde{\boldsymbol{\theta}})\Big)'$$

with

$$\begin{aligned} r_t(\tilde{\boldsymbol{\theta}}) &= y_t\Big(\frac{1-\tilde{\Phi}_t}{\tilde{\Phi}_t}\Big)^{1/2} + (y_t-1)\Big(\frac{\tilde{\Phi}_t}{1-\tilde{\Phi}_t}\Big)^{1/2} \\ &= \Big((1-\tilde{\Phi}_t)\tilde{\Phi}_t\Big)^{-1/2}\Big(y_t - \tilde{\Phi}_t\Big). \end{aligned}$$

Running the artificial regression (3.11) and computing the regression sum of squares yields the test statistic

$$LM_2 = r(\tilde{\boldsymbol{\theta}})'R(\tilde{\boldsymbol{\theta}})\Big(R(\tilde{\boldsymbol{\theta}})'R(\tilde{\boldsymbol{\theta}})\Big)^{-1}R(\tilde{\boldsymbol{\theta}})'r(\tilde{\boldsymbol{\theta}}). \tag{3.12}$$

Because $R(\tilde{\boldsymbol{\theta}})'r(\tilde{\boldsymbol{\theta}}) = s(\tilde{\boldsymbol{\theta}}) = S(\tilde{\boldsymbol{\theta}})'\boldsymbol{\iota}$, it can be seen that the test statistics LM_1 and LM_2 only differ in the way the information matrix estimate $\mathcal{I}(\tilde{\boldsymbol{\theta}})$ is constructed.

Note that the test statistics LM_1 and LM_2 can also be expressed as

$$LM_1 = \sum_{t=1}^{T}\tilde{d}_t\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)'\Bigg(\sum_{t=1}^{T}\tilde{d}_t^2\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)'\Bigg)^{-1}\sum_{t=1}^{T}\tilde{d}_t\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big),$$

and

$$LM_2 = \sum_{t=1}^{T}\tilde{d}_t\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)'\Bigg(\sum_{t=1}^{T}\frac{\tilde{\phi}_t^2}{\tilde{\Phi}_t(1-\tilde{\Phi}_t)}\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big)'\Bigg)^{-1}\sum_{t=1}^{T}\tilde{d}_t\Big(\frac{\partial \pi_t(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\Big),$$

where

$$\tilde{d}_t = \frac{y_t - \tilde{\Phi}_t}{\tilde{\Phi}_t(1-\tilde{\Phi}_t)}\tilde{\phi}_t.$$

This shows that the derivative term $\partial\pi_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ evaluated at $\tilde{\boldsymbol{\theta}}$ is central for the test statistics. From (3.5), the derivative is defined as

$$\frac{\partial\pi_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial\pi_t(\boldsymbol{\theta})}{\partial\omega} \\ \frac{\partial\pi_t(\boldsymbol{\theta})}{\partial\alpha_1} \\ \frac{\partial\pi_t(\boldsymbol{\theta})}{\partial\delta_1} \\ \frac{\partial\pi_t(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} 1+\alpha_1\frac{\partial\pi_{t-1}(\boldsymbol{\theta})}{\partial\omega} \\ \pi_{t-1}(\boldsymbol{\theta})+\alpha_1\frac{\partial\pi_{t-1}(\boldsymbol{\theta})}{\partial\alpha_1} \\ y_{t-1}+\alpha_1\frac{\partial\pi_{t-1}(\boldsymbol{\theta})}{\partial\delta_1} \\ \boldsymbol{x}_{t-1}+\alpha_1\frac{\partial\pi_{t-1}(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}} \end{pmatrix}$$

and under H_0, it is

$$\frac{\partial\pi_t(\tilde{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial\pi_t(\tilde{\boldsymbol{\theta}})}{\partial\omega} \\ \frac{\partial\pi_t(\tilde{\boldsymbol{\theta}})}{\partial\alpha_1} \\ \frac{\partial\pi_t(\tilde{\boldsymbol{\theta}})}{\partial\delta_1} \\ \frac{\partial\pi_t(\tilde{\boldsymbol{\theta}})}{\partial\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} 1 \\ \pi_{t-1}(\tilde{\boldsymbol{\theta}}) \\ y_{t-1} \\ \boldsymbol{x}_{t-1} \end{pmatrix}.$$

3.4 Simulation Results

The two LM tests described in the previous section are asymptotically equivalent. In this section, their finite small-sample properties are studied by simulation.[3] We simulated realizations from the Bernoulli distribution (3.1) using two different models[4]

$$\pi_t(\boldsymbol{\theta}) = -0.30 + \alpha_1\pi_{t-1}(\boldsymbol{\theta}) + 0.50\,y_{t-1}, \tag{3.13}$$

[3] Matlab version 7.5.0 and the BFGS algorithm in the Optimization Toolbox is used in simulation and estimation. Eviews code for computing LM tests (3.10) and (3.12) is also available upon request.

[4] The initial value $\pi_0(\boldsymbol{\theta})$ in (3.5) is set to $\pi_0(\boldsymbol{\theta}) = (\omega + \delta_1\bar{y} + \bar{\boldsymbol{x}}_{t-k}\boldsymbol{\beta})/(1-\alpha_1)$ with the parameter values used in (3.13) and (3.14) (see Kauppi and Saikkonen, 2008). A bar is used to denote the sample mean of the considered variables.

and

$$\pi_t(\boldsymbol{\theta}) = -0.30 + \alpha_1 \pi_{t-1}(\boldsymbol{\theta}) + 1.00\, y_{t-1} - 0.20\, x_{t-1}. \tag{3.14}$$

Since many macroeconomic and financial time series exhibit rather strong persistence we assume the following AR(1) process for the explanatory variable x_t,

$$x_t = 0.1 + 0.90x_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, 1).$$

Positive coefficients for the lagged y_{t-1} in (3.13) and (3.14) indicate that the realized values of y_t, i.e. zeros and ones, tend to cluster in the similar way as, for example, recession periods of the economy (see Section 3.5, and Chapters 2 and 5 in this thesis).

We provide simulation evidence for sample sizes 150, 300, 500, 1000 and 2000. For all generated series, 200 extra observations were simulated and discarded from the beginning of every sample to avoid initialization effects. We report empirical sizes of the models at 10%, 5% and 1% significance levels. All results are based on 2000 replications. However, in some cases, a little more than 2000 replications (about 20–30 extra replications) are needed because of numerical difficulties in the optimization of the log-likelihood function (3.6).

Empirical sizes of the LM tests for selected parameter values in (3.13) and (3.14) are presented in Tables 3.1 and 3.2. The empirical sizes range between sample sizes. Both tests seem to be rather severely oversized in small samples ($T = 150$, $T = 300$ and $T = 500$), but for larger samples, the empirical sizes are rather close to the nominal levels.

Table 3.1: Empirical size of the LM_1 and LM_2 tests in the model (3.13).

T	LM_1			LM_2		
	10%	5%	1%	10%	5%	1%
150	28.5	15.3	3.1	28.9	14.7	3.2
300	19.6	9.0	1.5	19.3	8.9	1.4
500	17.0	8.5	1.4	16.8	8.4	1.3
1000	14.3	6.6	1.1	14.3	6.6	1.1
2000	10.3	5.3	1.2	10.3	5.4	1.2

Notes: In size simulations, $\alpha_1 = 0$. The results are based on the 2000 replications.

Table 3.2: Empirical size of the LM_1 and LM_2 tests in the model (3.14).

T	LM_1			LM_2		
	10%	5%	1%	10%	5%	1%
150	42.8	26.3	7.1	41.6	23.0	5.0
300	30.1	15.2	3.3	28.4	14.5	2.3
500	22.0	10.8	2.2	21.0	10.3	2.0
1000	14.0	7.6	1.5	13.7	7.3	1.3
2000	11.4	5.7	0.9	11.4	5.3	0.9

Note: See notes to Table 3.1.

Rejection rates presented in Tables 3.1 and 3.2 are based on the critical values from the asymptotic χ^2_1 distribution. However, one can use a parametric bootstrap method to obtain alternative, potentially more accurate, critical values than the asymptotic ones. The employed procedure is the following. ML estimates $\tilde{\boldsymbol{\theta}} = (\tilde{\omega} \quad \tilde{\delta} \quad \tilde{\boldsymbol{\beta}})'$ and LM test statistics are computed under the null hypothesis $\alpha_1 = 0$. Bootstrap samples y^b_τ and the values of test statistics LM^b_1 and LM^b_2, $b = 1, 2, ..., B$, are then generated from the data-generating process

$$y^b_\tau \sim B(\Phi(\pi^b_\tau(\tilde{\boldsymbol{\theta}}))), \tag{3.15}$$

where $\tau = 1, 2, ..., T$, and

$$\pi^b_\tau(\tilde{\boldsymbol{\theta}}) = \tilde{\omega} + y^b_{\tau-1}\tilde{\delta} + \boldsymbol{x}'_{\tau-1}\tilde{\boldsymbol{\beta}}.$$

Finally, bootstrap critical values at different significance levels are obtained from the empirical distribution of the test statistics LM^b_1 and LM^b_2. The number of bootstrap replications B is set to 500 and the simulation is carried out for 500 replications.

As an illustration for the usefulness of the proposed bootstrap method, Table 3.3 presents the rejection rates based on the bootstrap critical values instead of the asymptotic ones. Compared with the results shown in Table 3.2, the empirical sizes of the LM tests are now much closer to the nominal values.

Table 3.3: Empirical size of the LM_1 and LM_2 tests using the model (3.14) and bootstrap critical values.

	LM_1			LM_2		
T	10%	5%	1%	10%	5%	1%
150	9.0	4.2	1.0	9.6	6.2	1.0
300	9.6	5.4	1.6	9.0	6.2	1.0
500	12.2	5.0	1.2	11.2	6.4	0.4

Size-adjusted empirical power functions with different sample sizes T at the 5% level are depicted in Figures 3.1 and 3.2 using (3.13) and (3.14) with different values of α_1. It is expected that in many applications the parameter α_1 is non-negative and, therefore, we concentrate on values from $\alpha_1 = 0.00$ up to $\alpha_1 = 0.80$.[5] The power seems to increase rather quickly when the value of α_1 increases, in particular when the explanatory variable x_t is employed in the model. It appears that the power of LM_2 is typically slightly higher that of LM_1 in both cases when $\alpha_1 > 0$. However, the differences are not very large.

Even for smaller sample sizes reasonable power is obtained although the power of tests is slightly decreasing at very high values of α_1. As Kauppi and Saikkonen (2008) note, a potential reason for this finding is that $\pi_{t-1}(\boldsymbol{\theta})$ and y_{t-1} may interact in a complicated way which could affect the statistical significance of $\pi_{t-1}(\boldsymbol{\theta})$, especially in small samples.

[5] The evidence appears to be rather the same with the negative values of α_1, especially in large samples.

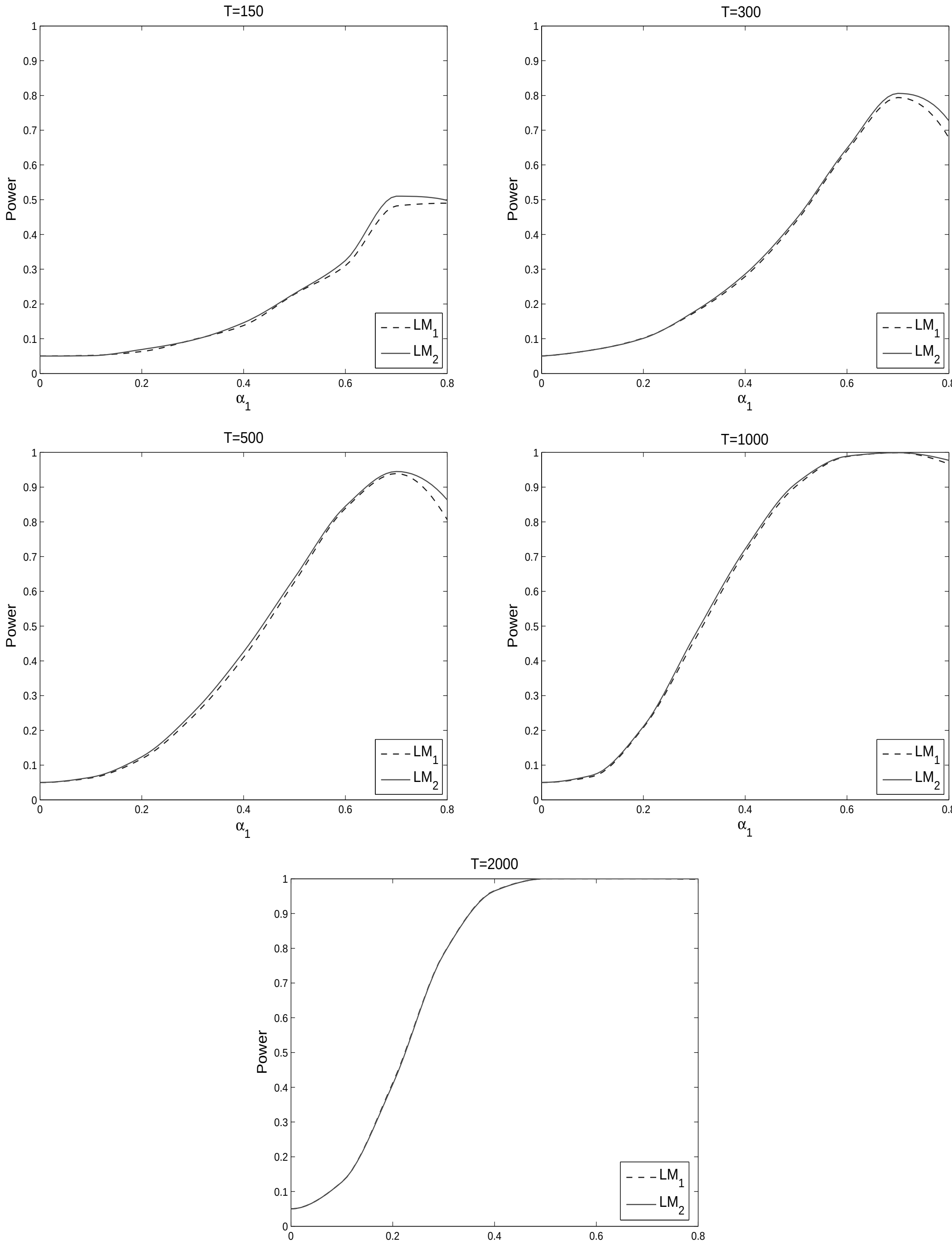

Figure 3.1: Empirical power in the case of model (3.13).

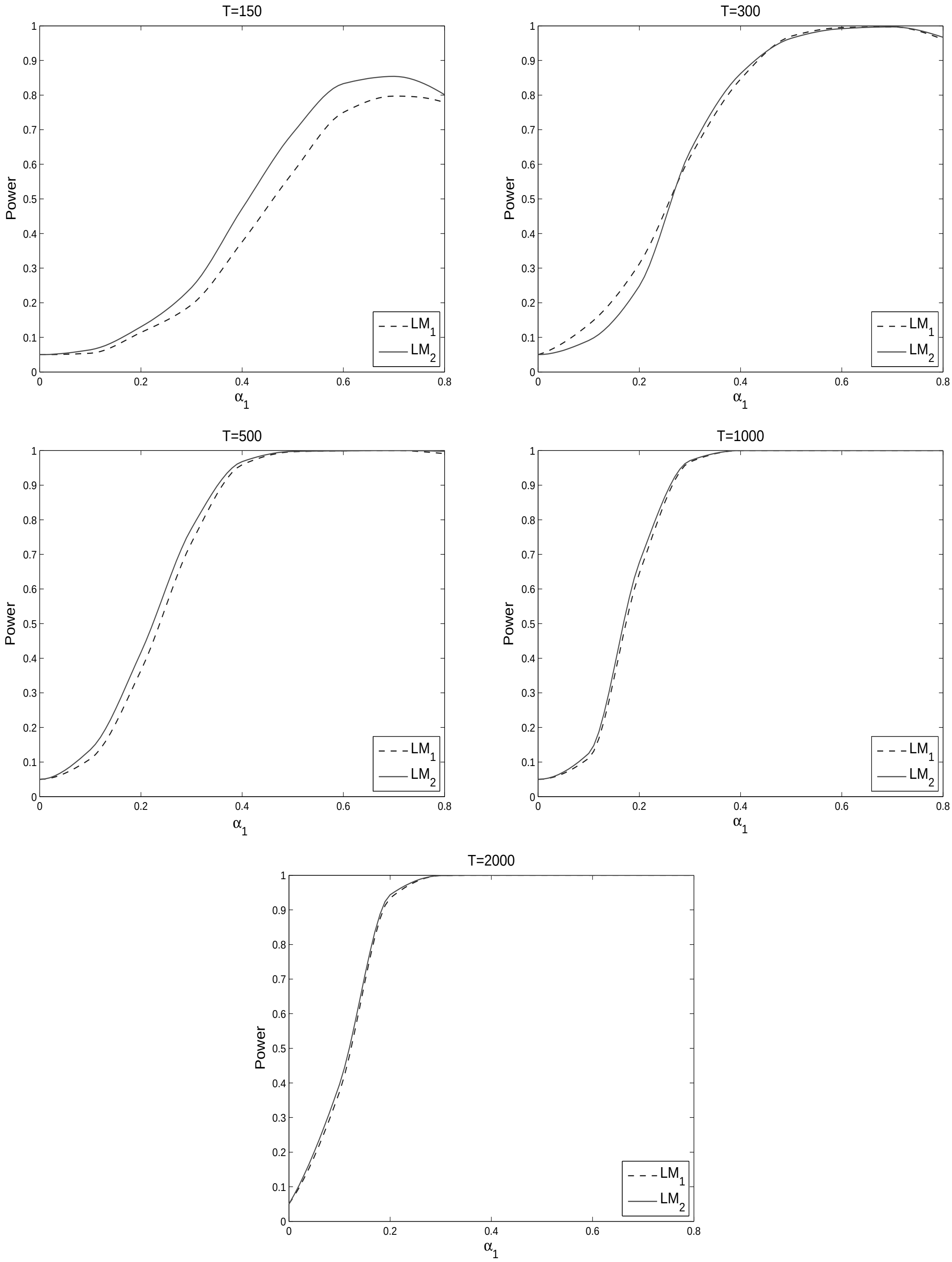

Figure 3.2: Empirical power in the case of model (3.14).

3.5 Application: U.S. Recession Forecasting Models

Forecasting recession periods has been one of the most common empirical applications of binary time series models. Predicting the direction-of-change in stock market returns is an example of another potential application (see, e.g., Leung, Daouk and Chen, 2000; Rydberg and Shephard, 2003; Nyberg, in press).

In recession forecasting, the dependent variable is the recession indicator

$$y_t = \begin{cases} 1, & \text{if the economy is in a recession state at time } t, \\ 0, & \text{otherwise.} \end{cases} \tag{3.16}$$

Although in this chapter we are not interested in out-of-sample forecasting, we consider forecasting models behind the "direct" (using horizon-specific predictor y_{t-15}) and "iterative" (y_{t-1}) multi-step forecasts for the binary response (3.16) (for details, see Kauppi and Saikkonen, 2008). The difference between "direct" and "iterative" forecasts is similar to that in time series models for traditional continuous variables (see, e.g., Marcellino, Stock, and Watson, 2006).

In this application, the recession periods identified by the National Bureau of Economic Research (NBER) for the U.S. between January 1954 and December 2006 are employed. The term spread (SP_t) between the long-term and short-term interest rate and the stock market return (r_t) are found to be useful predictive variables (see, e.g., Estrella and Mishkin, 1998; Nyberg, 2010). Explanatory variables are described in more detail in Table 3.4. Table 3.5 presents the estimated predictive models when the forecast horizon h is assumed to be six months ($h = 6$). The fact the NBER business cycle turning points are announced with a delay is taken into account in "direct" forecasting models.[6]

[6] We assume that this "publication lag" is nine months. Thus, in direct forecasting models the lag y_{t-15} is used given the six-month forecast horizon. For further details see, for example, Kauppi and Saikkonen (2008), Kauppi (2008), and Nyberg (2010).

Table 3.4: Explanatory variables.

R_t	10-year Treasury bond yield rate, constant maturity
i_t	Three-month Treasury bill rate, secondary market
SP_t	Term spread, $R_t - i_t$
r_t	Monthly stock market return, log-difference of the S&P500 index

Notes: Interest rates are from http://www.federalreserve.gov/releases/h15/data.htm. S&P500 stock index is taken from http://finance.yahoo.com and http://www.econstats.com. [2 July 2009].

In a direct forecasting (Model 1) shown in the first column of Table 3.5, the p-values of the two LM tests based on asymptotic critical values are zero indicating that the inclusion of an autoregressive structure gives a better model. The same conclusion is drawn by using bootstrap critical values. Further, when $\pi_{t-1}(\boldsymbol{\theta})$ is included in the model, it is a statistically significant predictor (Model 2) according to the Wald-type of test comparing the estimated coefficient of α_1 and its robust standard error or using likelihood ratio test between the competitive models. The values of the estimated log-likelihood function and the pseudo-R^2 measure (Estrella, 1998) are also substantially higher in the latter model.

When the "iterative" predictive model is considered (Model 3), the values of the test statistics LM_1 and LM_2 are statistically insignificant at 5% significance level compared with both asymptotic and bootstrap critical values. When the autoregressive structure is imposed on the model the coefficient for $\pi_{t-1}(\boldsymbol{\theta})$ is indeed statistically insignificant in the extended Model 4.

In conclusion, in these two examples, outcomes of the two LM tests are in accordance with the Wald test when testing the statistical significance of the autoregressive structure. The recommendation is that an autoregressive model structure is worth considering as an alternative to the static recession prediction model (see, e.g., Estrella and Mishkin, 1998), possibly augmented by the forecast horizon-specific lagged value of y_t as presented in the first two models of Table 3.5. However, the lagged state of the economy, y_{t-1}, seems to be the main dynamic part in the iterative forecasting model (Models 3 and 4).

Table 3.5: Estimation results for the recession prediction models.

Model		1	2	3	4
constant		-0.50	-0.02	-1.71	-2.00
		(0.16)	(0.04)	(0.17)	(0.24)
SP_{t-6}		-0.61	-0.21	-0.58	-0.67
		(0.13)	(0.05)	(0.14)	(0.16)
r_{t-6}		-0.05	-0.08	-0.01	-0.01
		(0.02)	(0.02)	(0.03)	(0.03)
$\pi_{t-1}(\boldsymbol{\theta})$			0.81		-0.17
			(0.03)		(0.09)
y_{t-1}				3.38	3.95
				(0.25)	(0.35)
y_{t-15}		-0.41	-0.07		
		(0.36)	(0.16)		
log-L		-185.94	-136.60	-55.63	-55.29
pseudo $-R^2$		0.192	0.367	0.689	0.691
LM_1		25.25		2.75	
p-value		0.000		0.097	
LM_2		36.18		0.65	
p-value		0.000		0.420	
Bootstrap critical values					
LM_1	10%	3.06		5.10	
	5%	3.93		6.88	
	1%	5.90		9.61	
LM_2	10%	2.71		2.30	
	5%	3.75		3.07	
	1%	5.37		6.45	

Notes: Models are estimated using U.S. data from 1954 M01 to 2006 M12 ($T = 636$). First 21 months are used as initial values. Robust standard errors (see Kauppi and Saikkonen, 2008) are reported in parentheses. The estimated value of the log-likelihood function (3.6) and the pseudo-R^2 measure (Estrella, 1998) are also provided as well as the values of the LM_1 and LM_2 test statistics, their p-values based on the asymptotic χ^2_1 distribution, and critical values obtained by bootstrap.

3.6 Conclusions

We have proposed LM tests for testing an autoregressive model structure in binary time series models. Based on a limited simulation study, the tests appear to have reasonable empirical size, especially in large samples, and high power. For small samples, the proposed bootstrap simulation method provides improved empirical sizes. An empirical example of recession forecasting models for the U.S. illustrates that the inclusion of an autoregressive model structure may be a useful addition to the recession prediction model.

References

Chauvet M, Potter S. 2005. Forecasting recession using the yield curve. *Journal of Forecasting* **24**: 77–103.

Cox DR. 1981. Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* **8**: 93–115.

Davidson R, MacKinnon JG. 1984. Convenient specification tests for logit and probit models. *Journal of Econometrics* **25**: 241–262.

Engle RF. 1984. Wald, likelihood ratio and Lagrange Multiplier tests in econometrics, in Griliches, Z. and Intriligator MD. (eds) in *Handbook of Econometrics*, Vol. II. Amsterdam, North-Holland.

Estrella A. 1998. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics* **16**: 198–205.

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45–61.

Kauppi H. 2008. Yield-curve based probit models for forecasting U.S. recessions: Stability and dynamics. HECER Discussion Paper, 221. Helsinki Center of Economic Research.

Kauppi H, Saikkonen P. 2008. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics* **90**: 777–791.

Leung MT, Daouk H, Chen AS. 2000. Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting* **16**: 173–190.

Marcellino M, Stock JH, Watson MW. 2006. A comparison of direct and iterated AR methods for forecasting macroeconomic time series. *Journal of Econometrics* **135**: 499–526.

Nyberg H. 2010. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting* **29**: 215–230.

Nyberg, H. In press. Forecasting the direction of the U.S. stock market with dynamic binary probit models. *International Journal of Forecasting*, forthcoming.

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

Startz R. 2008. Binomial autoregressive moving average models with an application to U.S. Recessions. *Journal of Business and Economic Statistics* **26**: 1–8.

Chapter 4

Forecasting the Direction of the U.S. Stock Market with Dynamic Binary Probit Models

Abstract[1]

Several empirical studies have documented that the signs of excess stock returns are, to some extent, predictable. In this chapter, we consider the predictive ability of the binary dependent dynamic probit model in predicting the direction of monthly excess stock returns. The recession forecast obtained from the model for a binary recession indicator appears to be the most useful predictive variable, and once it is employed, the sign of the excess return is predictable in sample. A new dynamic "error correction" probit model proposed in the chapter yields better out-of-sample sign forecasts with the resulting average trading returns higher than those of the buy-and-hold strategy or trading rules based on ARMAX models.

[1] A paper based on this chapter has been accepted for publication in the *International Journal of Forecasting*, Elsevier (forthcoming).

4.1 Introduction

In the financial econometric literature there is considerable evidence that excess stock market returns are, to some extent, predictable. The main objective has been to predict the overall level, the conditional mean, of excess stock returns. It is emphasized that even though the predictability is statistically weak, it can be economically meaningful.

However, many studies have documented that only the direction of excess stock returns or other asset returns are predictable (see, among others, Breen, Glosten, and Jagannathan, 1989; Hong and Chung, 2003; Christoffersen and Diebold, 2006). A possible explanation for this is that the noise in the observed returns is too high for the accurate forecasting of the overall return. Leitch and Tanner (1991) find that the direction of the change is the best criterion for predictability because traditional statistical summary statistics may not be closely related to the profits that investors are seeking in the financial market. Directional predictability is also important for market timing, which is crucial for asset allocation decisions between stock and risk-free interest rate investments.

The previous findings of directional predictability are mainly based on time series models for the excess stock return. For instance, Christoffersen and Diebold (2006) and Christoffersen *et al.* (2007) have considered the theoretical connection between asset return volatility and asset return sign predictability and verified that volatility and higher-order conditional moments of returns have statistically significant explanatory power in sign prediction. Even though there is not much previous research, binary dependent time series models provide an another way to forecast the direction of excess stock returns. Various classification-based qualitative models, such as traditional static logit and probit models, were considered by Leung, Daouk, and Chen (2000), whereas Hong and Chung (2003), Rydberg and Shephard (2003), and Anatolyev and Gospodinov (2010) used the so-called autologistic model to predict the return direction. In the last two papers the return is decomposed into a sign component and an absolute value component, which are modeled separately before the joint forecast is constructed.

We consider various commonly used financial variables as explanatory variables for forecasting the signs of the one-month U.S. excess stock returns from the S&P500 index and size-sorted small and large firms stock indices in probit models. This chapter introduces a model in which the recession forecast constructed for a binary recession indicator is used as an explanatory variable in the predictive model. To the best of our knowledge, this kind of approach has not been applied earlier to forecast the stock return sign. As a motivation for this model, for example, Fama and French (1989) and Chen (1991) propose that business conditions are important determinants of expected stock returns and, therefore, the recession forecast may be a useful predictive variable in our model. Further, Chauvet and Potter (2000) have stressed that the stock market "cycle" leads the business cycle. This argument is based on the fact that the expectations about changes in future economic activity could have important predictive power to predict excess stock returns. If there are expectations of a coming recession, excess stock returns are low, and after a recession period stock returns should be positive.

In this chapter, new dynamic probit models suggested by Kauppi and Saikkonen (2008) are employed and further extended. Since there is not much earlier evidence on suitable explanatory variables in sign prediction with probit models, various explanatory variables and their in-sample predictive performance are first evaluated. After that, the out-of-sample directional predictability for the excess stock return sign is considered. It is not evident, however, how much the in-sample evidence should be emphasized in assessing overall return predictability because it does not guarantee out-of-sample predictability, as emphasized in many previous studies (see discussion, for example, Goyal and Welch, 2008; Campbell and Thompson, 2008).

The results show that the probit models have statistically significant in-sample predictive power for the signs of excess stock returns. A proposed new "error correction" model outperforms the other probit and alternative predictive models, such as "continuous" ARMAX models, out of sample. The received excess investment returns over the buy-and-hold trading strategy are economically significant. Comparisons between different probit models indicate that the forecasting framework based on the constructed recession forecasts yields more accurate sign predictions than the models where only financial explanatory variables are employed. Especially in the case of small and large firms, the excess stock return signs seem to be predictable also out of sample.

This chapter proceeds as follows. The employed forecasting model with recession forecasts, suggested dynamic probit models, and in particular, the new error correction model is presented in Section 4.2. In Section 4.3, the goodness-of-fit evaluation of sign forecasts and statistical tests for sign predictability are introduced. The empirical evidence on the directional predictability of the U.S. excess stock returns is reported in Section 4.4. Section 4.5 concludes.

4.2 Forecasting Model

4.2.1 Dynamic Probit Models in Directional Forecasting

Let r_t be the excess stock return over the risk-free interest rate. In many studies, the directional predictability of excess stock returns is examined by using models for continuous dependent variables. For example, Christoffersen *et al.* (2007) proposed a method of forecasting the direction of excess stock return, where they first model the conditional variance σ_t^2 and the conditional mean μ_t. Assuming that the data generating process of r_t is

$$r_t = \mu_t + \sigma_t \varepsilon_t,$$

where $\varepsilon_t \sim \text{IID}(0,1)$, the conditional probability of a positive return given the information set Ω_{t-1} is

$$\begin{aligned} P_{t-1}(r_t > 0) &= 1 - P_{t-1}(r_t \leq 0) \\ &= 1 - P_{t-1}\Big(\varepsilon_t \leq \frac{-\mu_t}{\sigma_t}\Big) \\ &= 1 - F_\varepsilon\Big(\frac{-\mu_t}{\sigma_t}\Big), \end{aligned} \tag{4.1}$$

where $F_\varepsilon(\cdot)$ is the cumulative distribution function of the error term ε_t. If the conditional probability of positive excess return (4.1) varies with the information set Ω_{t-1}, then the sign of the return should be predictable.

In this chapter, the main interest is to study the directional predictability using probit models, where the dependent variable is the binary sign return indicator

$$I_t = \begin{cases} 1, & \text{if } r_t > 0, \\ 0, & \text{if } r_t \leq 0, \end{cases} \tag{4.2}$$

which takes the value one when the excess stock return is positive and zero otherwise. Thus, I_t is a binary-valued stochastic process. Conditional on the information set Ω_{t-1}, which includes the predictive variables and lagged values of the stock indicator (4.2), it has a Bernoulli distribution with probability p_t^I, that is

$$I_t|\Omega_{t-1} \sim B(p_t^I).$$

Let E_{t-1} denote the conditional expectation given the information set Ω_{t-1}. In the probit model the conditional probability of positive excess stock return ($I_t = 1$) satisfies

$$p_t^I = E_{t-1}(I_t) = P_{t-1}(I_t = 1) = P_{t-1}(r_t > 0) = \Phi(\pi_t^I), \tag{4.3}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. The conditional probability is modeled by specifying a model for π_t^I which is supposed to be a function of variables in the information set.[2]

Previous literature indicates that there is not much autocorrelation between the two successive values of excess stock returns. Thus, the benchmark forecasting model is the static model,

$$\pi_t^I = \omega + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}, \tag{4.4}$$

where the employed explanatory variables are collected in the vector $\boldsymbol{x}_{t-1}$. Because of the expected lack of correlation between I_{t-1} and I_t, this static model, without any dynamic structure, might be adequate. In order to investigate this, the value of the lagged return indicator I_{t-1} can be included in the model. This yields the dynamic probit model

$$\pi_t^I = \omega + \delta_1 I_{t-1} + \boldsymbol{x}_{t-1}^{'}\boldsymbol{\beta}. \tag{4.5}$$

If the coefficient δ_1 is statistically significant, then the lagged direction of the stock return is a useful predictor of the future direction of excess stock returns.

[2] The superscript "I" in π_t^I refers to excess stock return sign forecasting.

In the last few years, new binary time series models have been introduced. We concentrate on the model variants suggested by Kauppi and Saikkonen (2008). They add the lagged value π_{t-1}^I, referred to as autoregressive structure, to the model equation. Thus the static model (4.4) and the dynamic model (4.5) are extended to the autoregressive model[3]

$$\pi_t^I = \omega + \alpha_1 \pi_{t-1}^I + \boldsymbol{x}_{t-1}' \boldsymbol{\beta} \tag{4.6}$$

and to the dynamic autoregressive model

$$\pi_t^I = \omega + \alpha_1 \pi_{t-1}^I + \delta_1 I_{t-1} + \boldsymbol{x}_{t-1}' \boldsymbol{\beta}, \tag{4.7}$$

respectively. By recursive substitution, and assuming $|\alpha_1| < 1$, the latter model can be rewritten as

$$\pi_t^I = \sum_{i=1}^{\infty} \alpha_1^{i-1} \omega + \delta_1 \sum_{i=1}^{\infty} \alpha_1^{i-1} I_{t-i} + \sum_{i=1}^{\infty} \alpha_1^{i-1} \boldsymbol{x}_{t-i}' \boldsymbol{\beta}. \tag{4.8}$$

Therefore, if several lagged values of the stock indicator (4.2) or explanatory variables $\boldsymbol{x}_t$ are useful in forecasting, the autoregressive specifications (4.6) and (4.7) could be useful as parsimonious forecasting models.

Parameters of the probit models (4.4)–(4.7), as well as the case of a new model presented in the next section, can be estimated by the method of maximum likelihood as described in de Jong and Woutersen (in press) and Kauppi and Saikkonen (2008). We assume needed regularity conditions, such as the stationarity of π_t^I, so that the usual results on maximum likelihood estimation holds.

4.2.2 An Error Correction Model

Based on, for example, the principles of efficient market theory, the lagged values of the stock indicator (4.2) should not have predictive power to predict future market directions. This indicates that the estimated coefficient of the lagged return indicator δ_1 may be zero or close to zero. Therefore, in the dynamic autoregressive model (4.7), if $\delta_1 = 0$ and there are no explanatory variables in the model, that is

[3] In this chapter, the same model attributes as Kauppi and Saikkonen (2008) are used.

$\boldsymbol{\beta} = \mathbf{0}$, then the autoregressive parameter α_1 is not identified as seen from equation (4.8) by imposing the above-mentioned restrictions.[4] Even if the coefficient δ_1 is just close to zero, there is a potential identification problem that can affect parameter estimation and have implications on the forecasting accuracy of excess return sign predictions.

Imposing the restriction $\delta_1 = 1 - \alpha_1$ in the unrestricted dynamic autoregressive model (4.7) and assuming $|\alpha_1| < 1$, a new "restricted" dynamic autoregressive model can be formulated as

$$\pi_t^I = \omega + \alpha_1 \pi_{t-1}^I + (1 - \alpha_1) I_{t-1} + \boldsymbol{x}_{t-1}' \boldsymbol{\beta}. \tag{4.9}$$

Because of the assumption $|\alpha_1| < 1$, the coefficient for the lagged return indicator I_{t-1}, $1 - \alpha_1$, is always positive. In model (4.9), α_1 can also be interpreted as a "weight" between π_{t-1}^I and I_{t-1}. It is expected that in our application α_1 will be positive and quite high (since $\delta_1 \approx 0$). This leads to the fact that the predictive power is distributed over the longer history of I_t and explanatory variables $\boldsymbol{x}_t$. On the other hand, if α_1 is "small", then the first lag I_{t-1} is more useful than in the case of a higher value of α_1.

For simplicity, we refer to model (4.9) as an "error correction" model (ecm) in this study. The reason is that by adding $-\pi_{t-1}^I$ to both sides of equation (4.9), we obtain the error correction form

$$\Delta \pi_t^I = \omega + (1 - \alpha_1)(I_{t-1} - \pi_{t-1}^I) + \boldsymbol{x}_{t-1}' \boldsymbol{\beta}, \tag{4.10}$$

where $\Delta \pi_t^I = \pi_t^I - \pi_{t-1}^I$. Thus, the difference between I_{t-1} and π_{t-1}^I measures the long-run relationship between the value of the stock return indicator and the transformed probability $\pi_{t-1}^I = \Phi^{-1}(p_{t-1}^I)$ in the probit model. Rewriting model (4.10) as

$$\pi_t^I = \omega + \pi_{t-1}^I + (1 - \alpha_1)(I_{t-1} - \pi_{t-1}^I) + \boldsymbol{x}_{t-1}' \boldsymbol{\beta},$$

it can be seen that, for α_1 close to one, the model can be expected to exhibit "near unit root" behavior, implying rather strong persistence for the variable π_t^I. This

[4] Note that when explanatory variables $\boldsymbol{x}_{t-1}$ are included in the model and $\boldsymbol{\beta} \neq 0$, then there is no such identification problem, even though $\delta_1 = 0$.

means that the conditional probability of positive excess stock return does not change much between successive time periods.

As seen from equation (4.9), the parameter $1-\alpha_1$ is always positive, and it can be interpreted as the proportion of the disequilibrium between I_{t-1} and π^I_{t-1} in period $t-1$. A positive value of the error correction term $(I_{t-1}-\pi^I_{t-1})$ increases the probability of positive excess stock return in the following period and, of course, vice versa if the the error correction term is negative. Later on, we will see that the conditional probability of positive excess stock return p^I_t is typically close to 0.50 in most models, which means that π^I_t is close to zero.

It is worth noting that the error correction model (4.9) is somewhat the same as the autoregressive conditional multinomial model (ACM) suggested by Russell and Engle (2005). In their model, the term $(I_{t-1}-\pi^I_{t-1})$ is replaced by $(I_{t-1}-\Phi(\pi^I_{t-1}))$. The model (4.9) without the term $\boldsymbol{x}'_{t-1}\boldsymbol{\beta}$ is also similar to the IGARCH model, suggested by Engle and Bollerslev (1986), for conditional heteroskedasticity in models for continuous variables.

4.2.3 Recession Forecast as an Explanatory Variable

A novel idea of this study is to consider whether recession forecasts have explanatory power to forecast the direction of excess stock returns. In the empirical finance literature, it is shown that as forward-looking variables, lagged stock returns should provide information about the future evolution of economic activity and potential recession periods (see, e.g., Pesaran and Timmermann, 1995; Estrella and Mishkin, 1998; Nyberg, 2010). Therefore, if the expectations of future economic activity are correct, the movements of the stock market should lead movements in economic activity (see, e.g., Fama 1990). Theoretically, this relation can be justified by present value or discounted-cash-flow models, where the price of a stock is equal to expected future dividends which are assumed to be related to the future economic activity and the profitability of firms.

Our main goal is to forecast recession periods and use the potential explanatory power of the obtained recession forecasts to make better forecasts for the sign of the excess stock returns. This is done by using the binary recession indicator

$$y_t = \begin{cases} 1, & \text{if the economy is in a recession at month } t, \\ 0, & \text{if the economy is in an expansion at month } t. \end{cases} \tag{4.11}$$

In this chapter, recession dates defined by the NBER are used. As Chauvet and Potter (2000) argue, one feature, but also a potential problem, with the NBER recession dates is that they do not reflect short-lived contraction periods in the economy, which could have notable explanatory power for predicting excess stock returns. Further, Chauvet and Potter (2000) construct the transition probabilities of the "bear" and "bull" state of the stock market using Markov chain methods. They find that bear markets generally start a couple of months before an economic slowdown or recession period and end some months before a recession period ends. Thus, it seems evident that movements in the stock market should lead the business cycle. Evidence of this kind can be seen in Figure 4.1.[5] The U.S excess stock returns are often negative before a recession period begins. On the other hand, it seems that the returns are positive in the last few recession months, indicating expectations about recovery in economic activity.

According to this idea, for example, in the general dynamic autoregressive model (4.7), the estimated recession forecast p^y_{t+5}, constructed using the model (4.12), may be included in the vector $\boldsymbol{x}_{t-1} = (p^y_{t+5} \quad \tilde{\boldsymbol{x}}_{t-1})'$, where $\tilde{\boldsymbol{x}}_{t-1}$ contains other financial explanatory variables. Therefore, a predictive probit model contains the fitted values of the binary explanatory recession indicator (4.11) (cf. Maddala 1983, 122–123). Parameter estimation and forecasting is carried out with a two-step procedure where the recession and the stock return sign prediction models are estimated separately. It is worth noting, however, that in this kind of model, the usual asymptotic distribution of the maximum likelihood estimate may not apply because the recession probability forecast included in the model is based on the estimated model (cf. Pagan, 1984).

The forecast horizon in recession forecasting is assumed to be six months. A six-month recession forecast for the value of the recession indicator (4.11) at time

[5] Details on dataset are given in Table 4.1.

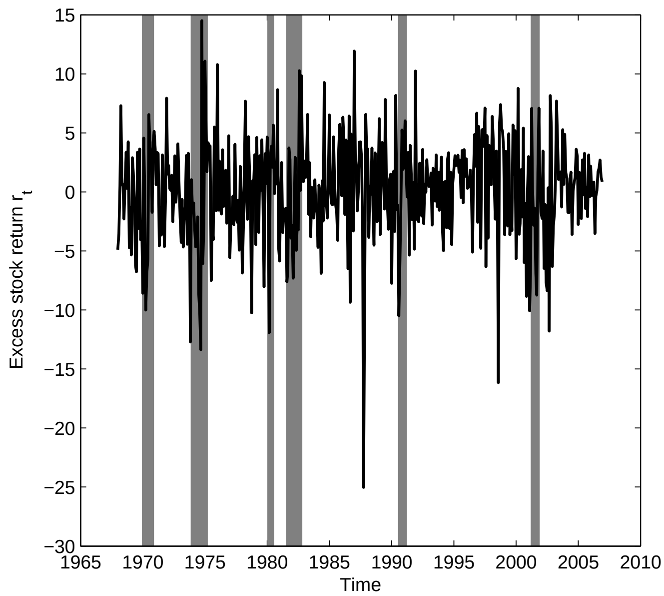

Figure 4.1: U.S. excess stock returns r_t and the NBER recession periods y_t (shaded areas) for a sample period of 1968 M1–2006 M12.

$t+5$, based on the information set Ω_{t-1} at time $t-1$, is the conditional probability

$$E_{t-1}(y_{t+5}) = P_{t-1}(y_{t+5} = 1) = \Phi(\pi^y_{t+5}) = p^y_{t+5}.$$

In recession forecasting, an autoregressive probit model

$$\pi^y_{t+5} = c + \phi \pi^y_{t+4} + \boldsymbol{z}'_{t-1}\boldsymbol{b} \tag{4.12}$$

is employed where, according to the findings in the recession forecasting literature, domestic and foreign term spread and lagged nominal stock return are used as predictors. The values of these variables are included in the vector $\boldsymbol{z}_{t-1}$, that is

$$\boldsymbol{z}_{t-1} = \begin{pmatrix} SP^{US}_{t-1} & r^n_{t-1} & SP^{GE}_{t-1} \end{pmatrix}'.$$

The usefulness of the domestic term spread (SP^{US}_t), defined as the spread between the long-term and the short-term interest rates, to predict recession periods is demonstrated in many studies (see, among others, Estrella and Mishkin, 1998; Estrella, 2005). Using dynamic probit models, Nyberg (2010) also suggests that the foreign term spread (SP^{GE}_t, the term spread of Germany) and stock market returns

(r_t^n) can be used to forecast coming recession periods (see also, e.g., Bernard and Gerlach, 1998; Estrella and Mishkin, 1998). [6] In that study, the obtained recession forecasts were quite accurate for at least six months ahead and the probit models with the autoregressive structure (see model (4.2.3)) yield the best out-of-sample forecasts compared with various other probit models. Therefore, in this study, we take model (4.2.3) and the six-month-ahead forecast horizon as a given.

An advantage in model (4.12) is that it does not contain a lagged value of the recession indicator (4.11). It is important to take into account that it takes several months before, for example NBER, can be sure what the state of the economy really is. Hence, the values of the recession indicator are known with a considerable delay, which indicates that it is computationally easier to construct recession forecasts without using the lagged values of the binary recession indicator (4.11) in an estimated model. This is based on the fact that in model (4.12), no multiperiod iterative forecasts (see Kauppi and Saikkonen, 2008) for the recession indicator are made because all predictive power comes from the employed explanatory variables $\boldsymbol{z}_t$. Thus, it is not needed to specify the assumed "publication lag" in the known values of y_t exactly.

4.3 Evaluation of Forecasts

4.3.1 Statistical and Economic Goodness-of-Fit Measures

Both in-sample and out-of-sample performance of predictive models are evaluated with frequently used goodness-of-fit measures. One is Estrella's (1998) pseudo-R^2 measure

$$psR^2 = 1 - (\hat{l}_u/\hat{l}_c)^{-(2/T)\hat{l}_c}, \tag{4.13}$$

where $\hat{l}_u$ is the maximum value of the estimated unconstrained log-likelihood function and $\hat{l}_c$ is its constrained counterpart in a model which only contains a constant term. This measure takes on values between 0 and 1, and it can be interpreted in the same way as the coefficient of determination in linear models. The value of

[6] Further information on the explanatory variables is in Table 4.1 in Section 4.4.1.

the maximized log-likelihood function also enables the comparison of model performance using model selection criteria such as the Schwarz information criterion BIC (Schwarz, 1978).

The binary nature of the dependent variable leads to the question of what the percentage of correct "matches" is of the realized values and the forecasts of the stock indicator. This ratio is denoted by CR. By the hypothesis of no predictability in excess stock return signs, the estimated value of CR should be close to 0.50, which means that the employed model is unable to forecast the future market directions correctly. It is desirable to specify a threshold value that translates the probability forecasts into forecasting signals. The most commonly used and natural threshold choice is 0.50, which is also used in this case.

For financial analysts and investors, the most important model evaluation criterion is the return on their investment. There are many different kinds of trading strategies that can be applied. Here a simple trading simulation similar to that in Leung *et al.* (2000) is used. At the beginning of each month, the investor makes an asset allocation decision. She can shift her assets either into stocks or into risk-free Treasury Bills, and the money that has been invested in either of these alternatives remains there until the next decision date. In this trading strategy, the mentioned 50 percent threshold value is used. Then, the portfolio consists of interest rate investment in Treasury Bills ($I_t^f = 0$), if $p_t^I \leq 0.50$, and stocks ($I_t^f = 1$), if $p_t^I > 0.50$. Here the superscript f refers to forecast.[7]

In this trading simulation, transaction costs are also taken into account. Following Granger and Pesaran (2000), the marginal cost of transaction for asset allocation changes between stocks and interest rates will be denoted by ζ_s and ζ_b, respectively. This means that, every time the asset allocation changes, the amount of the transaction cost is subtracted from the final investment return. In this study, the "low cost scenario" suggested by Pesaran and Timmermann (1995), where $\zeta_s = 0.005$ and $\zeta_b = 0.001$, is applied. For example, when the risk-free

[7] An alternative method for parameter estimation and determining the cutoff threshold, which is assumed to be 50 percent in the above, is for example, the Maximum Utility estimation method proposed by Elliott and Lieli (2007).

interest rate investments are switched to stocks, 0.50% of the whole amount of the portfolio value is lost.

As Granger and Pesaran (2000) have shown, it is also possible to form non-constant "payoff" probability ratios of switches between stocks and interest rates as an alternative for this 50% threshold. However, in this study, these payoff ratios are not very useful because, according to these threshold ratios, the asset allocation decision is to stick to stocks almost all the time. Therefore, probability forecasts, even if accurate, have little economic value.

When considering the predictability of excess stock return signs with different trading rules, one important evaluation criterion is the overall portfolio return, denoted by RET. Nevertheless, as for example, Hong and Chung (2003) emphasize, it is also worth considering risk-adjusted returns because different trading rules involve different levels of risk. In this evaluation, one commonly used measure is the Sharpe ratio (Sharpe, 1966 and 1994)

$$SR = \frac{\overline{RET}^k - \overline{RET}^{rf}}{\hat{\sigma}^k}, \tag{4.14}$$

where $\overline{RET}^k$ is the average portfolio return based on the model and trading rule k, $\overline{RET}^{rf}$ is the average risk-free portfolio return (bond investment strategy), and $\hat{\sigma}^k$ is the sample standard deviation of portfolio returns RET^k. The higher the Sharpe ratio is, the higher the return and the lower the volatility. Portfolios with a high Sharpe ratio are preferable to those with a low Sharpe ratio.

4.3.2 Testing the Statistical Predictability

For the evaluation of the directional forecasting performance and market timing, a test proposed by Pesaran and Timmermann (1992) is available. The null hypothesis of this test is that the value of the correct prediction ratio, CR, does not differ statistically significantly from the ratio that would be obtained in the case of non-predictability, where the forecasts and the realized values of the return indicator I_t are independent. Granger and Pesaran (2000) show that the market timing test can be based on the test statistic

$$PT = \frac{\sqrt{m}\, KS}{\left(\frac{\bar{P}_I(1-\bar{P}_I)}{\bar{I}(1-\bar{I})}\right)^{1/2}}. \tag{4.15}$$

Here KS is the Kuipers score $KS = HR - FR$ between the "hit rate"

$$HR = \frac{\hat{I}^{uu}}{\hat{I}^{uu} + \hat{I}^{du}}$$

and the "false rate"

$$FR = \frac{\hat{I}^{ud}}{\hat{I}^{ud} + \hat{I}^{dd}},$$

where forecast classification is denoted by

$$\begin{aligned} \hat{I}^{uu} &= \sum_{t=1}^{m} \mathbf{1}(I_t^f = 1, I_t = 1),\ \hat{I}^{ud} = \sum_{t=1}^{m} \mathbf{1}(I_t^f = 1, I_t = 0), \\ \hat{I}^{du} &= \sum_{t=1}^{m} \mathbf{1}(I_t^f = 0, I_t = 1),\ \hat{I}^{dd} = \sum_{t=1}^{m} \mathbf{1}(I_t^f = 0, I_t = 0), \end{aligned}$$

where f refers to forecast, u to an "up" signal ($I_t = 1$) and d to a "down" signal ($I_t = 0$), and $\mathbf{1}(\cdot)$ is an indicator function. Furthermore, in the test statistic (4.15), $\bar{I}$ is the sample average of the sign indicator I_t values in the m-month sample period and $\bar{P}_I = \bar{I}HR + (1 - \bar{I})FR$. Under the null hypothesis of non-predictability, the PT test statistic has an asymptotic standard normal distribution.

The directional predictability of an underlying data generation process is, however, not the same thing as a successful trading strategy. To evaluate the forecasts of the best forecasting models, we test the significance of the differences between the investment returns on the best models and trading strategies. This is tested by means of the Diebold-Mariano test (1995). Because the forecast horizon is one month, $h = 1$, the test statistic is

$$DM = \frac{\sqrt{m}\, \bar{d}}{\sqrt{\mathrm{var}(\bar{d})}}, \tag{4.16}$$

where $\bar{d}$ is the average difference between the predicted excess returns of the considered models. As in the PT test, under the null hypothesis of equal forecast accuracy, the DM statistic also has an asymptotic $N(0, 1)$ distribution.

4.4 Empirical Results

4.4.1 Data and Previous Findings

The monthly data set contains financial variables which have been used to predict overall level and the direction of excess stock returns in the previous literature. The data set covers the period from January 1968 to December 2006, and it is obtained from different sources mentioned in Table 4.1. The first 12 observations are used as initial values. The total number of observations, T, is 468. In out-of-sample forecasting, the data set is divided into two subsamples: the estimation and the forecasting sample. The first out-of-sample forecasts will be made for January 1989 and the last for December 2006.

In out-of-sample forecasting, parameters are estimated recursively using an expanding window of observations, where models are estimated using data from the start of the data set through to the present time to obtain a new one-period forecast. This procedure is repeated until the end of the forecasting sample. The use of an alternative rolling estimation window is problematic, because there are not many recession periods in the post-1970 time period and there would be estimation samples with no recession periods at all.

The one-month excess stock return is defined as the continuously compounded return of the price index P_t minus the risk-free interest rate rf_t

$$r_t = 100 \log\left(\frac{P_t}{P_{t-1}}\right) - rf_t. \tag{4.17}$$

Here P_t is the value of the $S\&P500$ stock index and the one-month risk-free return rf_t is approximated by the three-month U.S. Treasury Bill rate i_t. With excess stock returns r_t, the values of the binary stock return indicator described in equation (4.2) can be constructed.

Several explanatory variables to forecast the direction of excess stock returns will be considered. As confirmed by Leung *et al.* (2000), the majority of useful information for forecasting stock returns is contained in interest rates and lagged stock returns. Hence, the financial explanatory variables that are considered in the predictive models are the short-term and long-term interest rates and their first

Table 4.1: Data set of dependent and explanatory variables.

Variable	Description
P_t	Standard&Poors 500 U.S. stock index
P_t^S	CRSP small size firms index, first decile
P_t^L	CRSP large size firms index, tenth decile
r_t, r_t^S, r_t^L	One-month excess return over the risk-free return (see (4.17))
r_t^n	One-month nominal stock return from the S&P 500 index
y_t	U.S. Recession periods (NBER)
i_t	Three-month U.S. Treasury Bill rate, secondary market
R_t	10-year U.S. Treasury Bond rate, constant maturity
$\Delta i_t, \Delta R_t$	First differences of i_t and R_t
SP_t^{US}	U.S. term spread between $R_t - i_t$
SP_t^{GE}	German term spread between German long-term and short-term interest rates
σ_t	Sum of squared daily stock returns in the S&P500 index within one month
DP_t	Dividends over the past year divided by the current stock index value, $DP_t = D_t/P_t$
EP_t	Earnings over the past year divided by the current stock index value, $EP_t = E_t/P_t$

Notes: The sample period is 1968 M1–2006 M12. Monthly and daily S&P500 index series are taken from http://finance.yahoo.com and http://www.econstats.com. Size-sorted CRSP indices are obtained from the Kenneth French Data Library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). Interest rates are from http://www.federalreserve.gov/releases/h15/data.htm. German term spread is constructed as the difference between 10-year Federal security (series WZ9826, the missing values between 1971 M1–1972 M9 are replaced by the OECD 10-year interest rate) and the three-month money market rate (series su0107, see http://www.bundesbank.de/statistik/statistik.). Data for log-dividends D_t and log-earnings E_t were obtained from the homepage of Robert Shiller's book *Irrational exuberance* (http://www.irrationalexuberance.com) [5 January 2009].

differences, the U.S term spread, earnings/price and dividends/price-variables, and the realized volatility (see Table 4.1).

In previous studies, both the lagged excess stock returns and the lagged values of the return indicator (4.2) are used as predictors. Leung *et al.* (2000) use first differences of interest rates and lagged excess stock returns in their comparison between the sign and the overall return forecasting models, and concluded that in probit and logit models, several past returns should be included in the model. If the explanatory power is distributed among many lags of past returns, then the autoregressive models (4.6) and (4.7) could be useful in forecasting. On the other hand, Anatolyev and Gospodinov (2010) use the lagged sign return indicator I_{t-1} in their dynamic logit model for the direction of the future excess stock return. The corresponding estimated regression coefficient was positive, but statistically insignificant.

Interest rate spreads between different maturities may offer information about future expectations in financial markets (see, for instance, Fama and French, 1989). In recession forecasting, the term spread (SP_t^{US}) is expected to transmit the expectations for future monetary policy. The lower is the difference between the long-term and short-term interest rates the more restrictive is the current monetary policy. The term spread could also have its own impact on the stock market, not only on the real economic activity.

Dividends (D_t) and earnings (E_t), divided by the value of the price index (P_t), have been among the most commonly used explanatory variables (see, e.g., Campbell and Shiller, 1988; Cochrane, 1997). The dividend-price (earnings-price) ratio is computed with the dividends (earnings) of $S\&P500$ stock index companies over the past year. Since the monthly data of dividends and earnings are not available, DP_t and EP_t are constructed as sums of dividends and earnings over the past year divided by the current, monthly price level P_t.

Numerous studies have also documented a notable dependence of stock return and stock return volatility with important implications for asset pricing. The realized monthly volatility (σ_t), based on the sum of squared daily observations within one month (see Christoffersen *et al.*, 2007), is examined as a predictor in probit models.

4.4.2 In-Sample Results

Even though our main interest lies in the out-of-sample predictions of the direction of future excess stock returns, first, the in-sample performance of different probit models and combinations of explanatory variables were experimented using the sample period from January 1968 until December 1988. Explanatory variables are included one by one in the model.[8]

The main results and findings are as follows. According to the psR^2 and CR values, and the returns of trading strategies in the chosen in-sample period, the recession forecast and the first difference of the short-term interest rate are the best predictive variables. When employing these variables, there seems to be evidence that the excess stock return signs are predictable in sample. The first difference of the long-term interest rate and the realized volatility also have some predictive power. Interestingly, the corporate earning and dividend variables, used in many previous studies, are not particularly useful predictors.[9] When the recession forecast is employed with different financial explanatory variables in the model equation, the evidence is very much the same as in the above. The first difference of the short-term interest rate appears to be the best predictor with the recession forecast in-sample.

Table 4.2 presents details of parameter estimates in different probit models when the recession forecast (p^y_{t+5}) and the first difference of the short-term interest rate (Δi_{t-1}) are used as explanatory variables. The robust standard errors suggested by Kauppi and Saikkonen (2008) are presented. However, it should be noted that these standard errors may be inaccurate because the estimated recession forecast is employed in the model.[10]

It seems that the lagged stock indicator I_{t-1} has no statistically significant predictive ability for the sign of the stock return. In the error correction model, the autoregressive coefficient π^I_{t-1} is clearly statistically significant, but in other dynamic models, it is not. As expected, the estimated coefficients of the recession

[8] Matlab 7.4.0 and its BFGS optimization routine in the Optimization Toolbox is used in estimation and forecast computation. In models with the autoregressive structure (models (4.6), (4.7) and (4.9)) the initial value π_0 is set to similar way as suggested by Kauppi and Saikkonen (2008). For example, in model (4.7), it is $\pi_0 = (\omega + \delta_1 \bar{I}_{t-1} + \bar{\boldsymbol{x}}_{t-1}\boldsymbol{\beta})/(1-\alpha_1)$, where a bar denotes the sample mean of I_t and $\boldsymbol{x}_{t-1}$.

[9] Further information on in-sample performance of the different models is available upon request.

[10] In addition, there are no formal proofs of the asymptotic distributions of the maximum likelihood estimator in models (4.6), (4.7) and (4.9) presently available.

forecast and the first difference of the short-term interest rate are both negative.

Table 4.2: Estimation results of in-sample predictive models.

	static model(4.4)	dynamic model(4.5)	auto. model(4.6)	dyn.auto. model(4.7)	ecm model(4.9)
constant	0.07	0.02	0.07	0.03	-0.07
	(0.10)	(0.12)	(0.09)	(0.13)	(0.03)
π^I_{t-1}			0.04	-0.03	0.85
			(0.24)	(0.28)	(0.05)
I_{t-1}		0.08		0.08	
		(0.11)		(0.19)	
Δi_{t-1}	-0.24	-0.30	-0.30	-0.29	-0.16
	(0.04)	(0.12)	(0.12)	(0.12)	(0.07)
p^y_{t+5}	-0.50	-0.47	-0.49	-0.49	-0.02
	(0.25)	(0.25)	(0.24)	(0.25)	(0.05)
log-L	-161.33	-161.22	-160.72	-160.64	-162.10
psR^2	0.041	0.041	0.046	0.046	0.034
BIC	169.55	172.18	171.68	174.34	173.06
CR	0.580	0.591	0.579	0.579	0.579
RET	10.50	9.98	9.88	9.57	8.12
SR	0.95	0.86	0.79	0.72	0.25
PT	0.003	0.002	0.006	0.006	0.008
DM	0.009	0.017	0.017	0.023	0.123
DM_{ra}	0.000	0.000	0.000	0.000	0.003
$DM^{I_t=0}$	0.000	0.000	0.000	0.000	0.000
$DM^{I_t=0}_{ra}$	0.000	0.000	0.000	0.000	0.000

Notes: The models are estimated using the in-sample data from 1969 M1 to 1988 M12. Robust standard errors, given in parentheses, are computed with procedures suggested by Kauppi and Saikkonen (2008). RET is the average annualized in-sample portfolio return in the considered model and SR is the corresponding Sharpe ratio (4.14). The in-sample return in the B&H strategy RET is 3.60%. The corresponding Sharpe ratio is negative because the average return in the pure risk-free interest rate investment strategy is higher ($\overline{RET}^{rf} = 7.16\%$). The p-values of the market timing test (4.15) and the Diebold and Mariano (1995) test (4.16) are reported. In DM tests, the buy-and-hold trading strategy is the benchmark. Further, ra means the risk-adjusted returns, where returns are standardized with the standard deviation of the returns. The values of test statistics $DM^{I_t=0}$ and $DM^{I_t=0}_{ra}$ are obtained when only months with negative excess stock return ($I_t = 0$) are considered.

In this case the recession forecast is not statistically significant in the error correction model, but the first difference of the short-term interest rate is. In other models, both of these predictors are statistically significant according to presented robust standard errors.

Figure 4.2 depicts the estimated probability of a positive excess stock return in the static model (4.4) and in the error correction model (4.9), whose estimation results are shown in Table 4.2 (the first and the fifth model). Both models seem to give roughly the same in-sample predictions. In recession periods, both models suggest investing in a risk-free interest rate. More or less the only significant difference between the models is the time period between approximately 1976 to 1979. At that time, the probability forecast in the static model is typically above the 0.50 threshold value, while in the error correction model it is below the threshold.

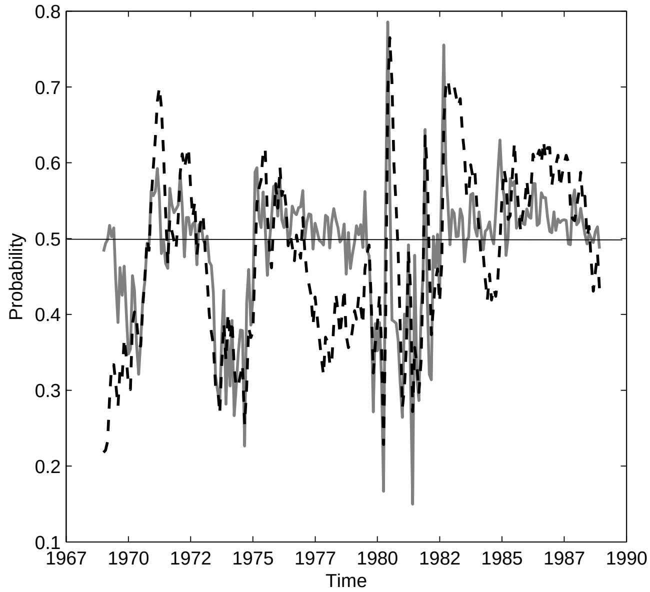

Figure 4.2: In-sample probabilities p_t^I (see (4.3)) of the static model (4.4) (solid line) and the error correction model (4.9) (dashed line). The 50 percent threshold is also depicted.

As the psR^2 values in Table 4.2 indicate, the statistical predictive power for the sign of the excess stock return is, as expected, quite low. Although the statistical predictability is weak, the portfolio investment performance yields evidence about a useful sign predictability in excess stock returns. The average rates of returns for different models and trading strategies are higher than in the "passive" buy-and-hold trading strategy (hereafter B&H strategy), where one is investing only in stocks. This annualized benchmark return is 3.60%. In the best models, the returns, including the transaction costs, are between 9.50% to 10.50%. The presented error correction probit model seems to yield smaller in-sample returns than its counterparts.

The statistical significance of return differences between an examined model and B&H returns was tested. Table 4.2 presents the values of the statistical test statistics introduced in Section 4.3.2. Since we are only interested in cases where the proportion of correctly predicted signs and the portfolio returns in estimated models are higher than under the null hypothesis of no predictability, only the positive and statistically significant values of the PT and DM test statistics (see (4.15) and (4.16)) provide evidence of predictability. The values of the market timing test statistic PT are statistically significant at the 1% level under all experimented models in Table 4.2. Thus the null hypothesis of no predictability is rejected, providing in-sample evidence that excess return signs are predictable. In the DM tests, the null hypothesis of equal performance between the returns in the considered model and the B&H strategy are rejected in all models at a 5% level except in the error correction model where the p-value of the test statistic is 0.123. When risk-adjusted returns are considered, the DM_{ra} test statistics are statistically significant in all models, providing evidence of profitable trading strategies based on the forecasts from the probit models.

Although the unrestricted dynamic autoregressive probit model (4.7) gives a better in-sample fit than the error correction model in the models presented in Table 4.2, in some other models with different explanatory variables, the error correction model gives higher psR^2 and CR values. In contrast to the error correction model (4.9), in many other unrestricted dynamic autoregressive models, the autoregressive coefficient α_1 is typically negative. Thus, if the probability of a positive excess stock return has been high in some period, it tends to be lower in the next period. As an example, consider a model in which the first difference of the short-term interest rate (Δi_{t-1}) is the only explanatory variable in $\boldsymbol{x}_{t-1}$. Figure 4.3 shows the estimated probabilities of positive excess stock returns in the unrestricted dynamic autoregressive model (4.7) and in the error correction model (4.9) in this example case. As the estimate for the autoregressive coefficient α_1 is negative in model (4.7), the probability of excess stock return fluctuates heavily around the threshold value 0.50. On the other hand, in the right panel with a positive and high estimate of α_1, the probability forecasts follow a relatively persistent swing. Therefore, it seems that the error correction model yields less transactions between stocks and bonds, and consequently also less transaction costs, than the unrestricted dynamic autoregressive model. This is particularly striking in models presented in Figure 4.3 and could be an important property in out-of-sample forecasting.

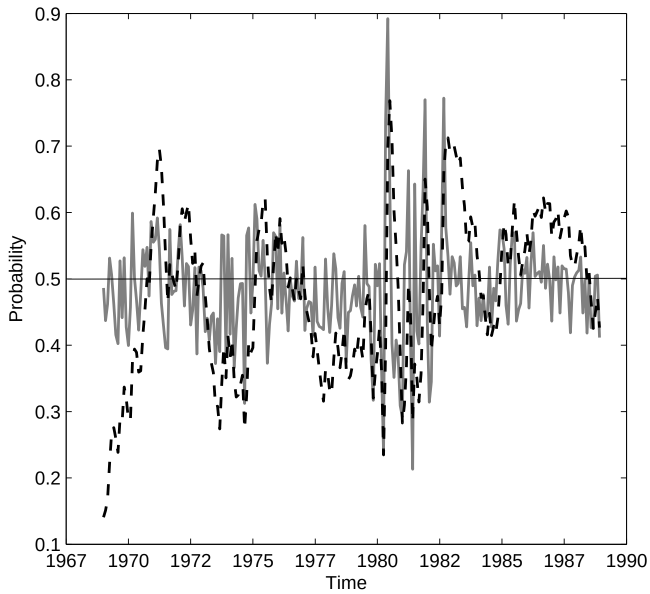

Figure 4.3: In-sample predictions of the dynamic autoregressive model (4.7) (solid line) and the error correction model (4.9) (dashed line) when Δi_{t-1} is the only predictive variable.

4.4.3 Out-of-Sample Results

When forecasting the signs of excess stock returns, it is important to compare the different models out of sample. Previous results on the predictive models for the overall excess stock return suggest that the in-sample predictability does not necessarily imply out-of-sample predictability. For example, Han (2007) finds that a statistically superior predictive VAR-GARCH model in sample does not consistently outperform its competitors in terms of portfolio investment returns out of sample. Goyal and Welch (2008) argue that traditional predictive models for the excess return cannot beat the historical average return out of sample and that there is no single variable that has theoretically meaningful and robust explanatory power. On the other hand, Campbell and Thompson (2008) show that some predictors perform better than the historical average when restrictions on regression coefficients are imposed. They and, for instance, Anatolyev and Gospodinov (2010) have stressed that while the out-of-sample predictive power is small, it can be utilized in market timing decisions to earn economically higher excess stock returns than the B&H strategy even out of sample.

In this study, the out-of-sample period consists of 216 months from January 1989 to December 2006. The out-of-sample recession forecast p^y_{t+5} is constructed before making any sign forecasts for excess stock return signs. As described in Section 4.2.3, the parameters in the sign and in the recession prediction models are estimated recursively. Table 4.3 shows the out-of-sample performance of the best in-sample predictive models and also some other probit models. The idea is to compare the predictive performance of different models when the same combinations of explanatory variables, which turned out to be the best out-of-sample predictive variables, are examined.[11]

According to commonly used statistical model evaluation measures, there is not much out-of-sample predictability in excess stock return signs. The values of psR^2 measures (see (4.13)) are, even in the best models, small or even negative.[12] The

[11] The models with other financial variables, presented in Table 4.1, were also considered and the results of those models are available upon request.

[12] Negative psR^2 value means very poor out-of-sample forecasting performance (Estrella, 1998).

Table 4.3: Out-of-sample performance of different probit models.

model	$\boldsymbol{x}_{t-1}$	psR^2	CR	RET	SR
B&H				8.07	0.79
static (4.4)	p^y_{t+5}	0.015	0.593	7.63	0.75
static (4.4)	ΔR_{t-1} , p^y_{t+5}	0.013	0.588	8.02	0.86
static (4.4)	Δi_{t-1} , p^y_{t+5}	0.014	0.588	7.02	0.61
dynamic (4.5)	–	neg.	0.528	3.92	neg.
dynamic (4.5)	p^y_{t+5}	0.005	0.569	6.16	0.46
dynamic (4.5)	ΔR_{t-1} , p^y_{t+5}	0.002	0.574	7.32	0.76
dynamic (4.5)	Δi_{t-1} , p^y_{t+5}	0.004	0.579	7.46	0.79
auto (4.6)	p^y_{t+5}	neg.	0.514	2.30	neg.
auto (4.6)	ΔR_{t-1} , p^y_{t+5}	0.015	0.583	7.20	0.68
auto (4.6)	Δi_{t-1} , p^y_{t+5}	0.014	0.579	6.53	0.50
dyn.auto (4.7)	–	neg.	0.514	4.11	neg.
dyn.auto (4.7)	p^y_{t+5}	0.008	0.574	5.42	0.28
dyn.auto (4.7)	ΔR_{t-1} , p^y_{t+5}	0.006	0.565	7.03	0.63
dyn.auto (4.7)	Δi_{t-1} , p^y_{t+5}	0.011	0.565	5.46	0.27
ecm (4.9)	–	0.014	0.588	8.62	0.97
ecm (4.9)	p^y_{t+5}	0.018	0.606	10.33	1.46
ecm (4.9)	ΔR_{t-1} , p^y_{t+5}	0.016	0.588	9.09	1.12
ecm (4.9)	Δi_{t-1} , p^y_{t+5}	0.017	0.588	9.78	1.28

Notes: See also notes to Table 4.2. The average return of the buy-and-hold trading strategy (B&H) is 8.07% (annual) with the corresponding Sharpe ratio SR= 0.79. The risk-free return on interest rate investments is 4.21%. The note "neg" means a negative psR^2 value and "–" in $\boldsymbol{x}_{t-1}$ indicates that there are no explanatory variables in the model. Note that the transaction costs are also taken into account in RET.

percentage of correct forecasts, CR, vary between 0.51 and 0.61. Contrary to the employed statistical measures, the results of portfolio returns RET and Sharpe ratios SR exhibit evidence of useful predictability for asset allocation decisions even though average portfolio returns vary strongly between different models. As in the in-sample evidence, the models with recession forecasts generate better sign forecasts than models without these forecasts. It is worth noting that the sign prediction models containing the constructed recession forecast outperform the models including the variables used in recession forecasting, especially out of sample.

It is interesting that the error correction model (4.9) clearly outperforms the corresponding unrestricted dynamic autoregressive model (4.7) out of sample. As seen in the best in-sample models in Table 4.2, the autoregressive model (4.6) and the dynamic autoregressive model (4.7) outperform the error correction model (4.9), but the out-of-sample evidence seems to be very different. In Table 4.3, the psR^2 values of the error correction models are clearly positive and the ratios of correct predictions, CR, are higher than in the other probit models considered. Above all, error correction models can generate more profitable trading strategies than the other probit models. Perhaps the most striking finding is the performance of the model with no explanatory variables ("–" in Table 4.3). The psR^2 values, CR ratio, average excess returns, and Sharpe ratios are clearly higher in the error correction model. A potential identification problem in the dynamic autoregressive model (4.7) discussed in Section 4.2.2 is a possible explanation for this superior out-of-sample performance of the error correction model.

Overall, compared with the dynamic models (4.5)–(4.7), the static probit model (4.4), without the autoregressive model structure π^I_{t-1} or the lagged I_{t-1}, seems to be an adequate model for the excess stock return sign. The error correction model (4.9) appears to be the only dynamic model which yields better forecasts in this data set than the static model.

The recession forecast is the main predictive variable in different models. The first differences of the short-term and the long-term interest rates are also fairly good predictors almost in all probit models and perform consistently better than the other financial explanatory variables examined. For instance, the realized volatility σ_t was a quite good predictive variable in sample, but its out-of-sample performance is very poor in probit models.

Figure 4.4 depicts the out-of-sample probability forecasts for the positive excess stock return in two models presented in Table 4.4. The most notable difference is that in 2001–2003 the error correction model gives a signal to invest in a risk-free interest rate when the monthly stock returns are most of the time negative. Further, in Table 4.4, the values of the PT and DM test statistics in these best error correction and static models, in terms of investment return (RET), are shown. In the error correction model, the p-value of the PT test statistic is 0.053 and the p-value of the DM test statistic is 0.121. The risk-adjusted returns are statistically significantly higher than the returns in the B&H strategy (p-value 0.000). On the other hand, the p-values of test statistics in the best static probit model shows that the excess stock return signs are not predictable with this model.

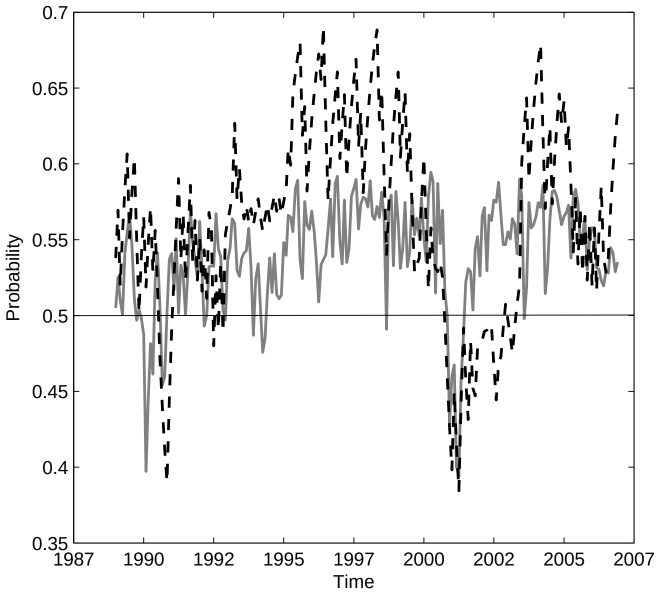

Figure 4.4: Out-of-sample predictions of the static (4.4) model with ΔR_{t-1} and p^y_{t+5} (RET= 8.02%, solid line). The dashed line is the error correction model (4.9) with p^y_{t+5} (RET=10.33%).

It should be pointed out that all models mostly suggest investing in stocks. For instance, in Figure 4.4, the conditional probabilities of a positive excess stock return are typically above the 0.50 threshold. Thus, the return differences between probit models and the B&H trading strategy are zero in most months. In addition, because the probability of recession is principally almost zero when the economy is in an expansionary state, the recession forecast should be a particularly useful predictor of negative excess stock return months when the economic activity is declining. Hence, the values of the DM test statistics are also calculated based on only those months when the excess stock returns have been non-positive (that is $I_t = 0$). There are 86 months with a negative excess return in the out-of-sample period. In Table 4.4, the values of test statistics $DM^{I_t=0}$, and $DM_{ra}^{I_t=0}$ in the case of risk-adjusted returns, are strongly statistically significant in the best models. As seen, the in-sample results in Table 4.2 are similar.

Table 4.4: Statistical tests for the best error correction probit model and the best static probit model.

model	$\boldsymbol{x}_{t-1}$	CR	RET	SR	PT	DM	DM_{ra}	$DM^{I_t=0}$	$DM_{ra}^{I_t=0}$
ecm (4.9)	p_{t+5}^y	0.61	10.33	1.46	0.053	0.121	0.011	0.000	0.000
static (4.4)	p_{t+5}^y, ΔR_{t-1}	0.59	8.02	0.86	0.326	0.616	0.184	0.000	0.000

Notes: The best error correction model (4.9) (ecm) and the best static model (4.4), reported in Table 4.3 and depicted in Figure 4.4, are presented. The p-values of the PT and DM tests are reported. In the DM tests the B&H trading strategy is the alternative asset allocation strategy. In the table, ra means the risk-adjusted returns and the test statistics $DM^{I_t=0}$ and $DM_{ra}^{I_t=0}$ are obtained when only months with negative excess stock return ($I_t = 0$) are taken into account.

Furthermore, when the investment returns of the best error correction model and the best static model are compared, the p-value of the DM test statistic is 0.111 and 0.048, respectively, when the risk-adjusted returns are considered. Thus, the error correction model yields higher returns, but the statistical significance between return differences is relatively weak in the considered out-of-sample period. However, according to the "asymmetric" DM test statistics discussed above ($DM^{I_t=0}$ and $DM_{ra}^{I_t=0}$), the best error correction model outperforms the best static model on all traditional statistical significance levels.

4.4.4 Comparison Between Probit and Alternative Predictive Models

It is interesting to make some comparisons between probit models and alternative models, such as ARMAX models and models where forecasts for the asset return volatility are employed to produce sign forecasts for excess stock returns. In fact, there are not many previous studies that compare the predictive performance of these models. Leung *et al.* (2000) find some evidence that qualitative response models, including logit and probit models, outperform models for the continuous dependent variables in their out-of-sample forecasting. They considered a sample of U.S., U.K. and Japanese stock indices from January 1991 to December 1995. In their study, the ratios of correct sign predictions and the investment returns from qualitative dependent models are higher than in models for continuous variables.

In ARMAX models the same explanatory variables as in probit models are considered. The dependent variable is the excess stock return r_t and it is assumed that a positive forecast gives the signal to buy stocks (i.e. $I_t^f = 1$). This is consistent with the definition of the stock return indicator (4.2). As in probit models, the in-sample predictive performance of different ARMAX models is first analyzed.[13] The estimated values of the BIC model selection criterion (Schwarz, 1978) suggested an ARMAX(2,0)[14] model with the first difference of the short-term interest rate (Δi_{t-1}) and the recession forecast (p_{t+5}^y) as explanatory variables. An ARMAX(1,0) model with the recession forecast and the U.S. term spread (SP_t^{US}) generates the highest in-sample investment return.

Out-of-sample forecasting performance of the best in-sample models and some other ARMAX models are shown in Table 4.5. As a whole, the percentage of the correct forecasts among the best probit models appears to be somewhat higher than in the best ARMAX models, but the investment return performance in particular is clearly better among the best probit models. In Table 4.5, only the best two ARMAX models in terms of RET values yield considerably higher returns than other ARMAX models. These two models are also depicted in Figure 4.5. When the return differences between the best error correction probit model and the best ARMAX models are tested, the return differences are also statistically significant at a 5% level based on the all considered DM test statistics shown in Table 4.6. Therefore, the error correction probit model seems to be a superior predictive model also against the alternative ARMAX models.

[13] In the estimation of ARMAX models, the UCSD_GARCH toolbox package for Matlab is used.

[14] For instance, ARMAX(2,0) is the same as the AR(2) model with explanatory variables.

Table 4.5: Out-of-sample performance of ARMAX and volatility models.

model	$\boldsymbol{x}_{t-1}$	CR	RET	SR
B&H			8.07	0.79
ARMAX(1,0)	p^y_{t+5}	0.565	7.12	0.67
ARMAX(1,0)	p^y_{t+5} , SP^{US}_{t-1}	0.542	6.07	0.45
ARMAX(1,0)	p^y_{t+5} , ΔR_{t-1}	0.556	4.98	0.19
ARMAX(1,0)	p^y_{t+5} , Δi_{t-1}	0.569	6.16	0.45
ARMAX(2,0)	p^y_{t+5}	0.576	5.83	0.37
ARMAX(2,0)	p^y_{t+5} , SP^{US}_{t-1}	0.514	4.36	0.04
ARMAX(2,0)	p^y_{t+5} , ΔR_{t-1}	0.532	4.76	0.13
ARMAX(2,0)	p^y_{t+5} , Δi_{t-1}	0.588	6.28	0.47
Volatility models				
Non-parametric		0.481	5.57	0.65
Extended		0.514	4.93	0.16

Notes: The ARMAX(p, 0) model for r_t is $r_t = a + \sum_{i=1}^{p} b_i \, r_{t-i} + \boldsymbol{x}'_{t-1}\boldsymbol{d}$. "Non-parametric" and "Extended" refer to the predictive models proposed by Christoffersen *et al.* (2007), which are based on the volatility forecasts $\hat{\sigma}_{t|t-1}$. See also notes to Table 4.3.

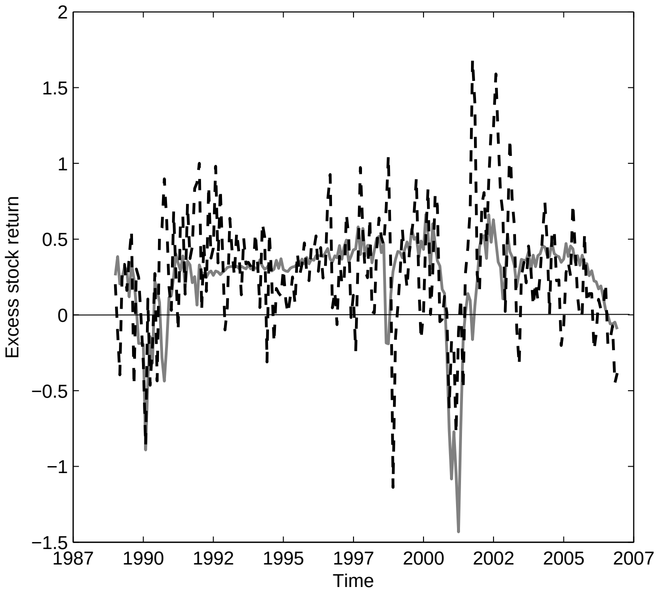

Figure 4.5: Out-of-sample predictions of the ARMAX(1,0) model with p^y_{t+5} (solid line) and ARMAX(2,0) model (dashed line) with p^y_{t+5} and Δi_{t-1} are depicted.

Table 4.6: Diebold-Mariano tests between the best error-correction probit model and the best ARMAX models.

model	$\boldsymbol{x}_{t-1}$	DM	DM_{ra}	$DM^{I_t=0}$	$DM_{ra}^{I_t=0}$
ARMAX(1,0)	p_{t+5}^{y}	0.046	0.019	0.000	0.000
ARMAX(2,0)	p_{t+5}^{y} , Δi_{t-1}	0.015	0.006	0.000	0.000

Notes: The p-values of the Diebold and Mariano (1995) tests between the investment returns from the error correction probit model presented in Table 4.4 and the ARMAX models mentioned in the first column are reported.

As presented in equation (4.1), if the volatility σ_t, conditional on the information at time $t-1$, is predictable, then the signs of stock returns should be predictable as well, provided that $\mu_t \neq 0$, although the conditional mean μ_t could be unpredictable. Using the same terminology as Christoffersen *et al.* (2007) for predictive models based on volatility forecasts, a "non-parametric" model indicates a model where the one-step-ahead volatility forecast is also used to compute the conditional mean forecast, which then together with the volatility forecast determine the probability forecast for positive excess return. In an "extended" model, the skewness and kurtosis of excess returns are also taken into account in the model. The percentage of correct forecasts CR shown in Table 4.5 shows that the volatility models do not produce out-of-sample sign predictability, and when compared to the best error correction probit model, the latter produces higher value of CR and higher investment returns. The p-values of DM test statistics between the models are 0.012 (non-parametric model) and 0.008 (extended model), indicating that the return differences are statistically significant at a 5% level.

4.4.5 Sign Predictability of Small and Large Size Firms' Returns

Finally, we extend our analysis by considering the sign predictability of small and large size firms. Perez-Quiros and Timmermann (2000), among others, find that there is a close link between stock returns of different size of firms and the state of the economy. They propose that recession, which indicates, for instance, worsening credit market conditions, is expected to affect the expected returns of small firms more strongly than large size firms' returns. In their model, Perez-Quiros and Timmermann (2000) employed a Markov switching model where the continuous excess stock return is modeled as a function of lagged Treasury Bill rate, default spread, changes in the money stock growth and dividend yield.

In our limited study, we employ the presented error correction probit model (4.9) with the six-month recession forecast employed for both return indicator series. The values of the return indicators (4.2) are constructed by using the excess stock returns r_t^S and r_t^L (see details in Table 4.1) from the size-sorted CRSP decile portfolios. It is worth noting that there is a significant correspondence with the binary values of stock indicator series between S&P500 and these size-sorted stock indices, as expected. Especially in the case of large size firms the correspondence is about 96%, as expected. In the returns from small size firms there is more variation between the values of return indicators. Overall, the mean and the volatility of small size firms' returns are higher than in the case of large size firms.

Table 4.7 presents the out-of-sample forecasting performance of considered error correction probit models. The p-values of the PT market timing test (4.15) are statistically significant, which shows that the signs are predictable out of sample in both cases. The percentage of correct forecasts, CR, is even higher than in the case of S&P500 returns. Further, the sign predictability does also convert to the higher investment returns in our simple trading simulation compared with B&H strategy. However, the differences are statistically significant only in terms of risk-adjusted returns where the standard deviation of the portfolio returns are taken into account.

Table 4.7: Out-of-sample sign predictions for small and large size firms' returns.

Firm size	model	$\boldsymbol{x}_{t-1}$	psR^2	CR	RET	SR	PT	DM	DM_{ra}	$DM^{I_t=0}$	$DM_{ra}^{I_t=0}$
Small	B&H				13.76	1.37					
	ecm (4.9)	p_{t+5}^y	0.029	0.625	15.72	2.24	0.002	0.267	0.014	0.000	0.000
Large	B&H				10.89	1.36					
	ecm (4.9)	p_{t+5}^y	0.007	0.634	12.91	2.07	0.014	0.116	0.009	0.000	0.000

Notes: As in Tables 4.3 and 4.5, the average of annual risk-free interest rate return is 4.21% in both series. See also notes to Table 4.3.

It seems that there is not much difference between the results of large and small size firms in terms of predictive accuracy. For small size firms the error correction model produces a higher value of psR^2 measure. On the other hand, the percentage of correct forecasts are higher in the case of large size firms. However, one significant difference between the models can be seen in Figure 4.6. There the conditional probability forecast for small firms fluctuates much more than in the case of large size firms. In that model the estimated value for coefficient α_1 is between 0.30–0.50 for the out-of-sample period, whereas for large size firms it is about the same magnitude as in the analysis of the S&P500 index ($\alpha_1 \approx 0.90$). Despite the fact that this fluctuation around the 50% threshold indicates more transaction costs (about a 2 percent deficit in investment returns compared with the returns without transaction costs), as discussed in Section 4.2.2, the investment returns from our trading simulation are higher than in the B&H strategy.

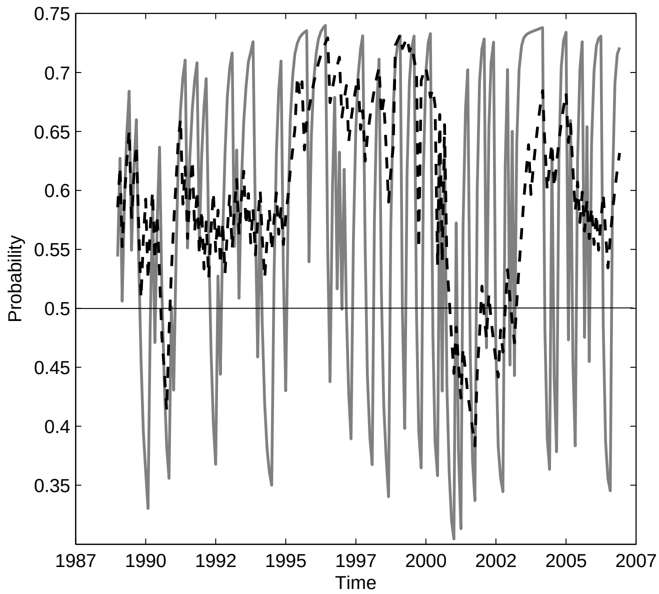

Figure 4.6: Out-of-sample predictions of the error correction probit models (4.9) with the recession forecast p^y_{t+5} for small (solid line) and large (dashed line) size firms.

4.5 Conclusions

We examine the predictability of the U.S. excess stock return signs by using dynamic binary probit models. The proposed forecasting method, where the six-month recession forecast for the recession indicator is used as an explanatory variable, seems to outperform other predictive models. Using the S&P500 stock index, the direction of the excess stock return is predictable and it is possible to earn statistically significant higher investment returns compared with the buy-and-hold trading strategy in sample. However, out-of-sample predictability turns out to be weaker. This is in line with previous findings related to stock return forecasting. In fact, in out-of-sample forecasting, the best dynamic probit model appears to be the error correction model proposed in this chapter. Using this model, the number of correct sign predictions and investment returns are higher than in other probit models, ARMAX models, or predictive models based on volatility forecasts. In the best error correction model the average investment returns are also higher than in the buy-and-hold trading strategy. Compared with the evidence from S&P500 returns, it appears that the out-of-sample sign predictability is higher in small and large size sorted firms' returns using the error correction probit model with recession forecasts.

The analysis can be extended in various ways. A system analysis in which the recession and sign forecasts are determined simultaneously in the same model is of particular interest. In this chapter, the six-month ahead recession forecast is taken as given, but this selection need not be optimal in terms of predictive power in sign predictions. It could also be interesting to form a somewhat more complicated trading strategy rule that could take the sign predictability, and perhaps the risk related to different models, even better into account in investment allocation decisions.

References

Anatolyev S, Gospodinov N. 2010. Modeling financial return dynamics via decomposition. *Journal of Business and Economic Statistics* **28**: 232–245.

Bernard H, Gerlach S. 1998. Does the term structure predict recessions? The international evidence. *International Journal of Finance and Economics* **3**: 195–215.

Breen W, Glosten LR, Jagannathan R. 1989. Economic significance of predictable variations in stock index returns. *Journal of Finance* **44**: 1177–1189.

Campbell JY, Shiller RJ. 1988. Stock Prices, earnings, and expected dividends. *Journal of Finance* **43**: 661–676.

Campbell JY, Thompson SB. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average. *Review of Financial Studies* **21**: 1509–1532.

Chauvet M, Potter S. 2000. Coincident and leading indicators of the stock market. *Journal of Empirical Finance* **7**: 87–111.

Chen N. 1991. Financial investment opportunities and the macroeconomy. *Journal of Finance* **46**: 529–554.

Cochrane JH. 1997. Where is the market going? Uncertain facts and novel theories. *Economic Perspectives*, Federal Reserve Bank of Chicago **21**: 3–37.

Christoffersen PF, Diebold FX. 2006. Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science* **52**: 1273–1287.

Christoffersen PF, Diebold FX, Mariano RS, Tay AS, Tse YK. 2007. Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: International evidence. *Journal of Financial Forecasting* **1**: 3–24.

de Jong RM, Woutersen TM. In press. Dynamic time series binary choice. *Econometric Theory*, forthcoming.

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–263.

Elliott G, Lieli RP. 2007. Predicting binary outcomes. Unpublished manuscript, University of Texas.

Engle RF, Bollerslev T. 1986. Modeling the persistence of conditional variances. *Econometric Reviews* **5**: 1–50.

Estrella A. 1998. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics* **16**: 198–205.

Estrella A. 2005. The yield curve as a leading indicator: Frequently asked questions. Federal Reserve Bank of New York. http://www.newyorkfed.org/research/capital_markets/ycfaq.pdf. [5 January 2009].

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45–61.

Fama EF. 1990. Stock returns, expected returns, and real activity. *Journal of Finance* **45**: 1089–1108.

Fama EF, French KR. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* **25**: 23–49.

Goyal A, Welch I. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* **21**: 1455–1508.

Granger CWJ, Pesaran HM. 2000. Economic and statistical measures of forecast accuracy. *Journal of Forecasting* **19**: 537–560.

Han Y. 2007. Return predictability, economic profits, and model mis-specification: How important are the better specified models? Tulane University. http://ssrn.com/abstract=967564. [2 July 2009].

Hong Y, Chung J. 2003. Are the directions of stock price changes predictable? Statistical theory and evidence. Manuscript, Cornell University.

Kauppi H, Saikkonen P. 2008. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics* **90**: 777–791.

Leitch G, Tanner J. 1991. Economic forecast evaluation: Profits versus the conventional error measures. *American Economic Review* **81**: 580–590.

Leung MT, Daouk H, Chen AS. 2000. Forecasting stock indices: A comparison of classification and level estimation models. *International Journal of Forecasting* **16**: 173–190.

Maddala GS. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, New York.

Nyberg H. 2010. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting* **29**: 215–230.

Pagan A. 1984. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review* **25**: 221–247.

Perez-Quiros G, Timmermann A. 2000. Firm size and cyclical variations in stock returns. *Journal of Finance* **55**: 1229–1262.

Pesaran HM, Timmermann A. 1992. A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics* **10**: 461–465.

Pesaran HM, Timmermann A. 1995. Predictability of stock returns: Robustness and economic significance. *Journal of Finance* **50**: 1201–1228.

Russell JR, Engle RF. 2005. A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial-autoregressive conditional duration model. *Journal of Business and Economic Statistics* **23**: 166–180.

Rydberg T, Shephard N. 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.

Sharpe WF. 1966. Mutual fund performance. *Journal of Business* **39**: 119–138.

Sharpe WF. 1994. The Sharpe ratio. *Journal of Portfolio Management* **21**: 49–58.

Chapter 5

A Bivariate Autoregressive Probit Model: Predicting U.S. Business Cycle and Growth Rate Cycle Recessions

Abstract[1]

We propose a new bivariate autoregressive probit model for binary time series. The model nests various special cases, such as two separate univariate models, for which a LM test against the unrestricted bivariate model is developed. The parameters of the model are estimated by the method of maximum likelihood and forecasts can be computed by using explicit formulae. The model is applied to predict the current state of the U.S. business cycle and growth rate cycle recession periods. Evidence of predictability of both recession periods is obtained by using financial variables as predictors. The bivariate model is found to outperform the univariate models built separately for each cycle indicator.

[1] An earlier version of this chapter has been published in HECER Discussion Papers, No. 272, 2009.

5.1 Introduction

In the previous literature on time series models for binary dependent variables, the models have typically been univariate. Given the importance of vector autoregressive models for continuous dependent variables, it is of interest to study multivariate binary time series models, where the probabilities of different binary outcomes are modeled jointly.

In this chapter, we present a bivariate autoregressive probit model as an extension to the univariate autoregressive probit model of Kauppi and Saikkonen (2008). The model can also be seen as an extension of the "static" bivariate probit model of Ashford and Sowden (1970), where the dependence between two binary time series is modeled by using a bivariate cumulative normal distribution function. In our bivariate model, the static model is extended by the inclusion of the autoregressive model structure.

In the previous literature, only few bivariate and multivariate models have been considered. Those models have mainly been based on the latent variable formulation, where the values of binary time series are realizations of corresponding continuous latent variables (see, e.g., Chib and Greenberg, 1998; Mosconi and Seri, 2006). In this chapter, the latent variable approach is not used. An advantage of our model is that parameter estimation can conveniently be carried out by the method of maximum likelihood and forecasts can be computed using explicit formulae. This is not typically the case in dynamic models based on the latent variables, such as the dynamic univariate model by Chauvet and Potter (2005) and the Qual VAR model of Dueker (2005). Our bivariate model is closely related to the model proposed by Anatolyev (2009), but we model the dependence between the two binary time series in a different way.

As an empirical application, we consider several alternative specifications of the proposed bivariate autoregressive probit model to nowcast the current state of the U.S. economy. We measure the state of the economy in terms of recession periods defined by the business cycle and the growth rate cycle indicators. Predicting business cycle recession periods with univariate probit models has attracted considerable attention in the literature (see, e.g., Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Chauvet and Potter, 2005), where the growth rate cycle indicator has hardly been considered at all. Given the fact that some economic slowdown periods do not turn into business cycle recessions, it is also of interest to consider the binary growth rate cycle indicator. To the best of our knowledge, this type of bivariate framework of two cycle indicators has not been considered in the previous literature.

A growth rate cycle is defined in terms of periods of increasing and decreasing growth rate in economic activity (see details in Banerji and Hiris, 2001; Osborn, Sensier and van Dijk, 2004). While "classical" business cycle recession periods are associated with the level of economic activity, a growth rate recession may occur without a decline in the level of economic activity. Therefore, growth rate cycle recessions are more numerous than classical business cycle recessions, but from the viewpoint of economic policy, they may be at least equally important and informative. For example, monetary policy decisions made by central banks are based on the real time assessment of the current, and also expected future, economic conditions using the data available at the time the decision is made. As Osborn *et al.* (2004) point out, growth rate cycles are closely related to the estimated output gap, which is supposedly an important variable affecting monetary policy decisions.

We concentrate on the predictive power of financial variables for business cycle and growth rate cycle recessions. The advantage of those variables is that they are available on a continuous basis without revisions. A difficulty with macroeconomic predictive variables, such as initial estimates of the real GDP or the estimated output gap, in contrast, is that they face substantial revisions during subsequent months and observations of some variables are not even available on a monthly basis. These properties of macroeconomic variables, which ultimately determine the values of both cycle indicators, also mean that the real-time state of the economy is always uncertain to some extent. Therefore, nowcasting the business cycle and growth rate cycle indicators is of interest, and the real-time availability supports financial variables as predictors.

Our results demonstrate the advantages of modeling the probabilities of business cycle and growth rate cycle recessions jointly. As a matter of fact, among the considered univariate and bivariate specifications, the proposed unrestricted bivariate autoregressive probit model yields the best in-sample, but also out-of-sample predictions. The lagged first difference of the Federal funds rate and monthly stock market returns turn out to be the best predictive variables for the U.S. growth rate cycle. As suggested in many previous studies, the U.S. term spread is an important predictive variable for predicting business cycle recessions, but its predictive power for growth rate cycle periods is limited.

The remainder of this chapter is organized as follows. The bivariate autoregressive probit model is introduced in Section 5.2. Issues of parameter estimation, testing, and forecasting are discussed in Section 5.3. Section 5.4 presents the empirical results. Section 5.5 concludes.

5.2 Bivariate Autoregressive Probit Model

Consider two binary time series, y_{1t} and y_{2t}, $t = 1, 2, ..., T$. Let us assume that conditional on information set Ω_{t-1} the random vector (y_{1t}, y_{2t}) follows a bivariate Bernoulli distribution,

$$(y_{1t}, y_{2t})|\Omega_{t-1} \sim B_2(P_{11,t}, P_{10,t}, P_{01,t}, P_{00,t}), \tag{5.1}$$

where

$$P_{ij,t} = P_{t-1}(y_{1t} = i, y_{2t} = j), \quad i, j = 0, 1, \tag{5.2}$$

and

$$P_{11,t} + P_{10,t} + P_{01,t} + P_{00,t} = 1. \tag{5.3}$$

Hence, the conditional marginal probabilities of the separate outcomes $y_{1t} = 1$ and $y_{2t} = 1$ are equal to

$$P_{1t} = P_{11,t} + P_{10,t}, \tag{5.4}$$

and

$$P_{2t} = P_{11,t} + P_{01,t}, \tag{5.5}$$

respectively.

A bivariate probit model was first proposed by Ashford and Sowden (1970) for analyzing cross-sectional data. In their model the joint probabilities for different outcomes of the vector (y_{1t}, y_{2t}) are determined as

$$\begin{aligned} P_{11,t} &= P_{t-1}(y_{1t}=1, y_{2t}=1) = \Phi_2(\pi_{1t}, \pi_{2t}, \rho), \\ P_{10,t} &= P_{t-1}(y_{1t}=1, y_{2t}=0) = \Phi_2(\pi_{1t}, -\pi_{2t}, -\rho), \\ P_{01,t} &= P_{t-1}(y_{1t}=0, y_{2t}=1) = \Phi_2(-\pi_{1t}, \pi_{2t}, -\rho), \\ P_{00,t} &= P_{t-1}(y_{1t}=0, y_{2t}=0) = \Phi_2(-\pi_{1t}, -\pi_{2t}, \rho), \end{aligned} \tag{5.6}$$

where $\Phi_2(\cdot)$ is the cumulative distribution function of the bivariate normal distribution with zero means, unit variances and correlation coefficient ρ, $|\rho| < 1$, and π_{1t} and π_{2t} are assumed to be linear functions of variables $\boldsymbol{x}_{1,t-k}$ and $\boldsymbol{x}_{2,t-k}$ included in the information set Ω_{t-1}, respectively. The sign changes in the arguments of the bivariate cumulative normal distribution function are needed to guarantee that condition (5.3) holds (see, for example, Greene, 2000, 849–850).

To complete the bivariate probit model, a parametrization for π_{1t} and π_{2t} needs to be specified. Ashford and Sowden (1970) introduced the following parametrization,

$$\begin{bmatrix} \pi_{1t} \\ \pi_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} \boldsymbol{x}'_{1,t-k} & 0 \\ 0 & \boldsymbol{x}'_{2,t-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}, \tag{5.7}$$

where ω_1 and ω_2 are constant terms and $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are coefficient vectors of the lagged explanatory variables included in $\boldsymbol{x}'_{1,t-k}$ and $\boldsymbol{x}'_{2,t-k}$, respectively. Note that using the same lag k in all explanatory variables is only for notational convenience and can easily be relaxed in practice. Equations (5.6) and (5.7) together define the static bivariate probit model.[2]

Dynamic extensions of the static model (5.7) can be obtained in various ways. We propose the following "bivariate autoregressive probit model",

$$\begin{bmatrix} \pi_{1t} \\ \pi_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \pi_{1,t-1} \\ \pi_{2,t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{x}'_{1,t-k} & 0 \\ 0 & \boldsymbol{x}'_{2,t-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}, \tag{5.8}$$

[2] The corresponding multivariate model is considered by Ashford and Sowden (1970), and Chib and Greenberg (1998), among others.

where π_{1t} and π_{2t} are specified as linear functions of their lags and the lagged values of the explanatory variables included in $\boldsymbol{x}'_{1,t-k}$ and $\boldsymbol{x}'_{2,t-k}$. Model (5.8) can compactly be written as

$$\boldsymbol{\pi}_t = \boldsymbol{\omega} + \mathbf{A}\boldsymbol{\pi}_{t-1} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}, \tag{5.9}$$

where $\boldsymbol{\pi}_t = \begin{pmatrix}\pi_{1t} & \pi_{2t}\end{pmatrix}'$, $\boldsymbol{x}_{t-k} = \operatorname{diag}\begin{pmatrix}\boldsymbol{x}'_{1,t-k} & \boldsymbol{x}'_{2,t-k}\end{pmatrix}'$, $\boldsymbol{\omega} = \begin{pmatrix}\omega_1 & \omega_2\end{pmatrix}'$, $\boldsymbol{\beta} = \begin{pmatrix}\boldsymbol{\beta}'_1 & \boldsymbol{\beta}'_2\end{pmatrix}'$, and

$$\mathbf{A} = \begin{pmatrix}\alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22}\end{pmatrix}.$$

In this bivariate autoregressive model we explicitly allow for the possibility that different explanatory variables can be included in $\boldsymbol{x}'_{1,t-k}$ and $\boldsymbol{x}'_{2,t-k}$. If the parameter matrix $\mathbf{A}$ is unrestricted, both π_{1t} and π_{2t} can depend on the lagged values of π_{1t} and π_{2t}. Thus, even if $\rho = 0$ in (5.6), the coefficients α_{12} and α_{21} provide a linkage between the variables π_{1t} and π_{2t} in model (5.8). Note that the static model is obtained from (5.9) with the restriction $\mathbf{A} = \mathbf{0}$. Furthermore, it is only in the special case where $\rho = 0$ and $\alpha_{12} = \alpha_{21} = 0$ that our bivariate autoregressive probit model reduces to two independent univariate autoregressive probit models.

Model (5.8) is somewhat similar to the multivariate dynamic binary model of Anatolyev (2009). The main difference is that Anatolyev (2009) suggests using the so called "dependence ratios" (cf. Ekholm, Smith, and McDonald, 1995) between the dependent variables to construct the conditional joint probabilities of the different outcomes of (y_{1t}, y_{2t}). In parameter estimation the dependence ratios and marginal probabilities for the variables y_{1t} and y_{2t} are handled separately by using a logistic function. In our model, the dependence between y_{1t} and y_{2t} is instead modeled by using the autoregressive specification (5.8) and the bivariate cumulative normal distribution function, where the correlation coefficient ρ is allowed to be nonzero. In addition, in the bivariate autoregressive probit model, parameter estimation can be carried out within the same system without dependence ratios.

Note that if the roots of $\det(I_2 - \mathbf{A}z)$ lie outside the unit circle, we obtain by recursive substitution of (5.9) the following representation,

$$\boldsymbol{\pi}_t = \sum_{j=1}^{\infty} \mathbf{A}^{j-1}\boldsymbol{\omega} + \sum_{j=1}^{\infty} \mathbf{A}^{j-1}\boldsymbol{x}'_{t-k-j+1}\boldsymbol{\beta}. \tag{5.10}$$

This shows that in the bivariate autoregressive probit model (5.8) π_{1t} and π_{2t} depend on the whole infinite history of the explanatory variables in a parsimonious way and, therefore, the model can be interpreted as an "infinite order" extension of the static model (5.7). Furthermore, assuming that the explanatory variables included in $\boldsymbol{x}_{t-k}$ are stationary, also $\boldsymbol{\pi}_t$ is stationary.

It is worth noting that because of the characteristics of our empirical application (see Section 5.4.2), the lagged values of y_{1t} and y_{2t} included in Anatolyev's (2009) model are excluded. However, that would be a possible extension of model (5.8). This extension can be based on the univariate model of Kauppi and Saikkonen (2008), where the lag y_{t-1} is also included in the right hand side of the model.

5.3 Parameter Estimation, Testing and Forecasting

5.3.1 Maximum Likelihood Estimation

As in corresponding univariate models, parameter estimation in the bivariate autoregressive model defined by (5.6) and (5.8), as well as its special cases, can conveniently be carried out by the method of maximum likelihood (ML). Using the conditional probabilities in (5.6), one can write the likelihood function and obtain the maximum likelihood estimate by using numerical methods.

Following Greene's (2000, 849–850) notation, the log-likelihood function can be constructed as follows. Define $q_{jt} = 2y_{jt} - 1$ and $\mu_{jt} = q_{jt}\pi_{jt}$, $j = 1, 2$, so that

$$q_{jt} = \begin{cases} 1 & \text{if } y_{jt} = 1, \\ -1 & \text{if } y_{jt} = 0, \end{cases}$$

and

$$\mu_{jt} = \begin{cases} \pi_{jt} & \text{if } y_{jt} = 1, \\ -\pi_{jt} & \text{if } y_{jt} = 0. \end{cases}$$

Furthermore, set

$$\rho_t^* = q_{1t}q_{2t}\rho.$$

The conditional probabilities of the different outcomes of (y_{1t}, y_{2t}) can be expressed as (cf. (5.6))

$$P_{t-1}(y_{1t}, y_{2t}) = \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*).$$

Let $\boldsymbol{\theta} = \Big(\mathrm{vec}(\mathbf{A})' \quad \boldsymbol{\omega}' \quad \boldsymbol{\beta}' \quad \rho\Big)'$ denote the vector of the parameters of the bivariate autoregressive probit model. The log-likelihood function, conditional on initial values, is the sum of the individual log-likelihood functions $l_t(\boldsymbol{\theta})$[3],

$$\begin{aligned} l(\boldsymbol{\theta}) &= \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \log\Big(\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)\Big) \\ &= \sum_{t=1}^{T} \Big(y_{1t}y_{2t}\log(P_{11,t}) + y_{1t}(1-y_{2t})\log(P_{10,t}) \\ &\qquad + (1-y_{1t})y_{2t}\log(P_{01,t}) + (1-y_{1t})(1-y_{2t})\log(P_{00,t})\Big). \end{aligned} \tag{5.11}$$

It is worth noting that if the correlation coefficient ρ in (5.6) is zero, the conditional probabilities in (5.6) are products of the marginal probabilities (5.4) and (5.5). For instance, in that case the conditional probability of the outcome $(y_{1t} = 1, y_{2t} = 1)$ is

$$P_{11,t} = P_{t-1}(y_{1t} = 1, y_{2t} = 1) = P_{t-1}(y_{1t})P_{t-1}(y_{2t}) = \Phi(\pi_{1t})\Phi(\pi_{2t}). \tag{5.12}$$

The score vector of the log-likelihood function (5.11) is

$$s(\boldsymbol{\theta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^{T} s_t(\boldsymbol{\theta}) = \sum_{t=1}^{T} \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \tag{5.13}$$

where

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{1}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} \frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \boldsymbol{\theta}}.$$

[3] In the bivariate autoregressive model (5.8) the selection of the initial value $\boldsymbol{\pi}_0 = (\pi_{01} \quad \pi_{02})'$ is also needed to construct the log-likelihood function (5.11). Following Kauppi and Saikkonen (2008) one way to obtain the initial values $\pi_{0i}, i = 1, 2$, is to select $\pi_{0i} = (\omega_i + \bar{\boldsymbol{x}}_{i,t-k}\boldsymbol{\beta}_i)/(1-\alpha_{ii})$, where a bar indicates the sample mean of variables included in $\boldsymbol{x}_{i,t-k}$.

At this point it is convenient to split the parameter vector into three disjoint components, namely $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^{'} \quad \boldsymbol{\theta}_2^{'} \quad \rho)^{'}$, where the parameters in $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are related to the specifications of π_{1t} and π_{2t}, respectively. The score vector can be partitioned accordingly as

$$s_t(\boldsymbol{\theta}) = \Big(s_{1t}(\boldsymbol{\theta}_1)^{'} \quad s_{2t}(\boldsymbol{\theta}_2)^{'} \quad s_{3t}(\rho)\Big)^{'}. \tag{5.14}$$

The first component of $s_t(\boldsymbol{\theta})$ can be written as

$$\begin{aligned} s_{1t}(\boldsymbol{\theta}_1) = \frac{\partial l_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_1} &= \frac{1}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} \frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \boldsymbol{\theta}_1} \\ &= \frac{1}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} \frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \mu_{1t}} \frac{\partial \mu_{1t}}{\partial \boldsymbol{\theta}_1} \\ &= \frac{1}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} \phi(\mu_{1t}) \Phi\Big(\frac{\mu_{2t} - \mu_{1t}\rho_t^*}{\sqrt{1-\rho_t^{*2}}}\Big) q_{1t} \frac{\partial \pi_{1t}}{\partial \boldsymbol{\theta}_1}, \end{aligned}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density function and the cumulative distribution function of the standard normal distribution, respectively. The value of $s_{1t}(\boldsymbol{\theta}_1)$ depends on the realized values of y_{1t} and y_{2t}. For instance, if $(y_{1t} = 1,\ y_{2t} = 1)$, then by the definitions of μ_{jt} and q_{1t},

$$s_{1t}(\boldsymbol{\theta}_1) \quad = \quad \frac{1}{\Phi_2(\pi_{1t}, \pi_{2t}, \rho)} \phi(\pi_{1t}) \Phi\Big(\frac{\pi_{2t} - \pi_{1t}\rho}{\sqrt{1-\rho^2}}\Big) \frac{\partial \pi_{1t}}{\partial \boldsymbol{\theta}_1}.$$

It can be seen that the main difference between the score vector of the static model ($\mathbf{A} = \mathbf{0}$) and model (5.8) is in the derivative term $\partial \pi_{1t}/\partial \boldsymbol{\theta}_1$. In model (5.8),

$$\frac{\partial \pi_{1t}}{\partial \boldsymbol{\theta}_1} = \begin{pmatrix} \frac{\partial \pi_{1t}}{\partial \omega_1} \\ \frac{\partial \pi_{1t}}{\partial \alpha_{11}} \\ \frac{\partial \pi_{1t}}{\partial \alpha_{12}} \\ \frac{\partial \pi_{1t}}{\partial \boldsymbol{\beta}_1} \end{pmatrix} = \begin{pmatrix} 1 + \alpha_{11}\frac{\partial \pi_{1,t-1}}{\partial \omega_1} + \alpha_{12}\alpha_{21}\frac{\partial \pi_{1,t-2}}{\partial \omega_1} \\ \pi_{1,t-1} + \alpha_{11}\frac{\partial \pi_{1,t-1}}{\partial \alpha_{11}} + \alpha_{12}\alpha_{21}\frac{\partial \pi_{1,t-2}}{\partial \alpha_{11}} \\ \pi_{2,t-1} + \alpha_{11}\frac{\partial \pi_{1,t-1}}{\partial \alpha_{12}} + \alpha_{12}\alpha_{21}\frac{\partial \pi_{1,t-2}}{\partial \alpha_{12}} \\ \boldsymbol{x}_{1,t-k} + \alpha_{11}\frac{\partial \pi_{1,t-1}}{\partial \boldsymbol{\beta}_1} + \alpha_{12}\alpha_{21}\frac{\partial \pi_{1,t-2}}{\partial \boldsymbol{\beta}_1} \end{pmatrix},$$

whereas, in the static model, it reduces to $\Big(1 \quad \boldsymbol{x}_{1,t-k}\Big)^{'}$. The derivative $\partial l_t(\boldsymbol{\theta})/\partial \boldsymbol{\theta}_2$ is obtained in the same way by replacing $\partial \pi_{1t}/\partial \boldsymbol{\theta}_1$ in the definition of $s_t(\boldsymbol{\theta})$ by

$$\frac{\partial \pi_{2t}}{\partial \boldsymbol{\theta}_2} = \Big(\frac{\partial \pi_{2t}}{\partial \omega_2} \quad \frac{\partial \pi_{2t}}{\partial \alpha_{22}} \quad \frac{\partial \pi_{2t}}{\partial \alpha_{21}} \quad \frac{\partial \pi_{2t}}{\partial \boldsymbol{\beta}_2}\Big)^{'}.$$

Given the result

$$\frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \rho_t^*} = \phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*),$$

the derivative with respect ρ is

$$\frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \rho} = \frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \rho_t^*} \frac{\partial \rho_t^*}{\partial \rho} = \phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*) q_{1t} q_{2t}.$$

Therefore, the score of the correlation coefficient ρ becomes

$$s_{3t}(\rho) = \frac{\partial l_t(\boldsymbol{\theta})}{\partial \rho} = \frac{1}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} \frac{\partial \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\partial \rho_t^*} \frac{\partial \rho_t^*}{\partial \rho} = \frac{\phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)}{\Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*)} q_{1t} q_{2t}.$$

The value of $s_{3t}(\rho)$ depends on realized values of the dependent variables. For example, if $y_{1t} = 1$ and $y_{2t} = 1$,

$$s_{3t}(\rho) = \frac{\phi_2(\pi_{1t}, \pi_{2t}, \rho)}{\Phi_2(\pi_{1t}, \pi_{2t}, \rho)},$$

and if $y_{1t} = 1$ and $y_{2t} = 0$,

$$s_{3t}(\rho) = -\frac{\phi_2(\pi_{1t}, -\pi_{2t}, -\rho)}{\Phi_2(\pi_{1t}, -\pi_{2t}, -\rho)}.$$

Maximization of the log-likelihood function (5.11) yields the maximum likelihood estimate $\hat{\boldsymbol{\theta}}$, which solves the first order condition $s(\hat{\boldsymbol{\theta}}) = 0$. Under appropriate regularity conditions, including the stationarity of $\boldsymbol{\pi}_t$ and explanatory variables and the correctness of the probit model specification, the conventional large sample theory of ML estimation gives the usual asymptotic distribution

$$T^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{L} N(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})^{-1}), \tag{5.15}$$

where $\mathcal{I}(\boldsymbol{\theta}) = \operatorname{plim} T^{-1} \partial^2 l(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'$ exists and is positive definite.

A practical difficulty with the bivariate autoregressive probit model (5.8) is that the number of parameters can become large if many explanatory variables are included. ML estimation is considerably simplified if the correlation coefficient ρ is restricted to zero because then the bivariate probabilities in the log-likelihood function (5.11) factor into products of marginal probabilities, as in (5.12). Thus, it is of interest to test for the hypothesis $\rho = 0$. In the next section, a LM test for this purpose is developed.

5.3.2 LM Test for the Correlation Coefficient

For testing the significance of the correlation coefficient, the Lagrange Multiplier test is attractive because it only requires ML estimation under the null hypothesis $\rho = 0$. Kiefer (1982) has proposed a corresponding LM test for the static bivariate probit model (5.7). In this section, the test is extended to the bivariate autoregressive model (5.8).

Let $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_1^{\prime} \quad \tilde{\boldsymbol{\theta}}_2^{\prime} \quad 0)^{\prime}$ be the restricted ML estimate of $\boldsymbol{\theta}$ obtained by assuming

$$H_0 : \rho = 0. \tag{5.16}$$

The general form of the LM test statistic (see, for example, Engle, 1984) is

$$LM = s(\tilde{\boldsymbol{\theta}})^{\prime}\tilde{\mathcal{I}}(\tilde{\boldsymbol{\theta}})^{-1}s(\tilde{\boldsymbol{\theta}}), \tag{5.17}$$

where $\tilde{\mathcal{I}}(\tilde{\boldsymbol{\theta}})$ is a consistent estimate of the information matrix $\mathcal{I}(\boldsymbol{\theta})$ and $s(\tilde{\boldsymbol{\theta}})$ is the score vector (5.13) evaluated at the restricted ML estimates $\tilde{\boldsymbol{\theta}}$. Under the null hypothesis (5.16) the test statistic has an asymptotic χ_1^2 distribution.

Due to the complexity of the second derivatives of the log-likelihood function (5.11), the outer-product of the score is an attractive estimator of the information matrix $\mathcal{I}(\boldsymbol{\theta})$. The resulting test statistic is

$$LM^{\rho} = \boldsymbol{\iota}^{\prime} S(\tilde{\boldsymbol{\theta}})\left(S(\tilde{\boldsymbol{\theta}})^{\prime} S(\tilde{\boldsymbol{\theta}})\right)^{-1} S(\tilde{\boldsymbol{\theta}})^{\prime}\boldsymbol{\iota}, \tag{5.18}$$

where $\boldsymbol{\iota}$ is a vector of ones and the matrix $S(\tilde{\boldsymbol{\theta}})$ is given by

$$S(\tilde{\boldsymbol{\theta}}) = \left(s_1(\tilde{\boldsymbol{\theta}}) \quad s_2(\tilde{\boldsymbol{\theta}}) \ldots s_T(\tilde{\boldsymbol{\theta}})\right)^{\prime}.$$

As in (5.14), the score vector $s_t(\tilde{\boldsymbol{\theta}})$, evaluated at $\tilde{\boldsymbol{\theta}}$, consists of three components. The bivariate densities and probabilities factor into products of marginals and, consequently, the components of the score reduce to

$$\begin{aligned} s_{1t}(\tilde{\boldsymbol{\theta}}_1) &= \frac{\phi(\tilde{\mu}_{1t})}{\Phi(\tilde{\mu}_{1t})} q_{1t} \frac{\partial \tilde{\pi}_{1t}}{\partial \boldsymbol{\theta}_1}, \\ s_{2t}(\tilde{\boldsymbol{\theta}}_2) &= \frac{\phi(\tilde{\mu}_{2t})}{\Phi(\tilde{\mu}_{2t})} q_{2t} \frac{\partial \tilde{\pi}_{2t}}{\partial \boldsymbol{\theta}_2}, \end{aligned}$$

and

$$s_{3t}(0) = \frac{\phi(\tilde{\mu}_{1t})\phi(\tilde{\mu}_{2t})}{\Phi(\tilde{\mu}_{1t})\Phi(\tilde{\mu}_{2t})} q_{1t} q_{2t},$$

where "$\sim$" on the right hand side means that the corresponding quantities are evaluated at $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$. The derivatives $\partial\pi_{1t}/\partial\boldsymbol{\theta}_1$ and $\partial\pi_{2t}/\partial\boldsymbol{\theta}_2$ depend on the considered specifications of π_{1t} and π_{2t} (see Section 5.3.1).

5.3.3 Forecasting

As shown by Kauppi and Saikkonen (2008), explicit formulae can be used to obtain one-period and multiperiod forecasts in the case of the univariate model. The obtained forecasts are probability forecasts for different outcomes of (y_{1t}, y_{2t}). In the following we show that the same principles can also be applied in the proposed bivariate model.

In the mean-square sense, the optimal h-period forecast based on the given information available at time $t-h$, $h \geq 1$, is the conditional expectation

$$E_{t-h}(y_{1t}, y_{2t}) = \Phi_2(\mu_{1t}, \mu_{2t}, \rho_t^*). \tag{5.19}$$

For example, the forecast for outcome $(y_{1t} = 1, \, y_{2t} = 1)$ is given by

$$E_{t-h}(y_{1t} = 1, y_{2t} = 1) = \Phi_2(\pi_{1t}^{(h)}, \pi_{2t}^{(h)}, \rho),$$

where, as shown in (5.10), by recursive substitution the bivariate system (5.9) can be written as

$$\boldsymbol{\pi}_t^{(h)} = \mathbf{A}^h \boldsymbol{\pi}_{t-h} + \sum_{j=1}^{h} \mathbf{A}^{j-1} \Big(\boldsymbol{\omega} + \boldsymbol{x}_{t-k-j+1}^{'} \boldsymbol{\beta} \Big), \tag{5.20}$$

where $\boldsymbol{\pi}_t^{(h)} = \begin{pmatrix} \pi_{1t}^{(h)} & \pi_{2t}^{(h)} \end{pmatrix}^{'}$ and the vector $\boldsymbol{\pi}_{t-h}$ is a function of values of the explanatory variables and the initial values of $\pi_{1,0}$ and $\pi_{2,0}$. In addition, the condition $k \geq h$ for all predictors included in $\boldsymbol{x}_{t-k}$ must hold indicating that the employed lags of the predictive variables are always tailored to match the information available at the time of forecasting. The usual case is obtained by selecting $k = h$. The right hand side of (5.20) gives the h step forecast for the outcome $(y_{1t} = 1, y_{2t} = 1)$ "directly" using the information up to the forecast time $t-h$. Forecasts for other outcomes of vector (y_{1t}, y_{2t}) are obtained by imposing necessary sign changes in bivariate normal cumulative distribution function (see (5.6)).

Both in-sample and out-of-sample predictive performance of the employed models can be evaluated with goodness-of-fit measures commonly used for binary dependent variables. Comparisons between different models can be based on the value of the maximized log-likelihood function (5.11), denoted by logL below. It can also be used to compute values of model selection criteria, such as the Schwarz information criterion (Schwarz, 1978) defined as

$$BIC = -\text{logL} + K\,\frac{\log(T)}{2}, \tag{5.21}$$

where K is the number of parameters in $\boldsymbol{\theta}$ and T is the number of observations. Another goodness-of-fit measure is the quadratic probability score, QPS, suggested by Diebold and Rudebusch (1989). Using the marginal conditional probability forecasts P_{1t} and P_{2t} (see (5.4) and (5.5)), the quadratic probability score for variable y_{jt} is

$$QPS_j = \frac{1}{T}\sum_{t=1}^{T} 2\Big(y_{jt} - P_{jt})\Big)^2, \tag{5.22}$$

where $j = 1, 2$. The values of the QPS_j lie on the interval [0,2] with the value 0 indicating a perfect fit. It can be seen as a counterpart of the mean square error used with models for continuous variables.

Because of the binary nature of the dependent variable, the percentage of correct predictions (CR) is a natural measure of predictive performance. However, a threshold value must be specified that translates the probability forecasts into signal forecasts ($y_{jt} = 1$ or $y_{jt} = 0$, $j = 1, 2$). The most commonly used and natural threshold value is 0.50, which is also used in this study. When the signal forecasts are constructed, a test proposed by Pesaran and Timmermann (1992) is available for the evaluation of the directional predictive performance of a model. The null hypothesis of the test is that the value of the correct prediction ratio does not differ significantly from the ratio that would be obtained in the case of no predictability, where the forecasts and realized values of y_{jt}, $j = 1, 2$, are independent. Under the null hypothesis of no predictability, the test statistic has an asymptotic standard normal distribution.

5.4 Empirical Application: Predicting the Current State of the U.S. Economy

We apply different bivariate probit model specifications to predict the state of the U.S. economy. In this application, we predict, or more specifically "nowcast", values of the dependent U.S. business cycle and growth cycle indicators to be discussed in more detail in Section 5.4.1 using the real-time information on financial predictive variables. The monthly sample size covers the period from January 1971 to December 2005.

Knowledge of the current state of aggregate economic activity is important for many economic agents in business and finance, as well as for policymakers, such as central banks and government organizations. However, because of informational lags and revisions of important macroeconomic variables, such as the real GDP, the current state of the economy is always uncertain to some extent. In our nowcasting exercise, the forecast horizon will therefore be one month, $h = 1$. Thus, we are interested in predicting the probabilities of business cycle and growth cycle recessions for month t using the information up to the end of the previous month $t-1$. In other words, the nowcasts are constructed at the beginning of month t.[4]

5.4.1 Binary Indicators for the Business and Growth Rate Cycles

Forecasting the recession periods of the economy with various univariate binary time series models has attracted considerable attention in many previous studies (see, among others, Estrella and Mishkin, 1998; Chauvet and Potter, 2005; Kauppi and Saikkonen, 2008; Nyberg, 2010). These recession periods are related to business cycle fluctuations defined in terms of the level of economic activity. Thus, our

[4] Matlab 7.8.0. and its BFGS optimization routine in the Optimization Toolbox is employed in estimation.

first binary recession indicator is

$$y_{1t} = \begin{cases} 1, & \text{if the economy is in a recession at time } t, \\ 0, & \text{if the economy is in an expansion at time } t. \end{cases} \tag{5.23}$$

The best-known indicator for the U.S. is that one provided by the National Bureau of Economic Research (NBER). It is based on the definition in which a recession is "a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales."[5] It is important to note that the NBER uses a broader array of economic indicators than just, say, the real GDP, to determine the recession periods.

In this chapter, we are mainly interested in predicting the growth rate cycles of the U.S. economy. To the best of our knowledge, only Osborn *et al.* (2004) have so far studied these cycle periods by means of binary time series models. In contrast to classical business cycles characterized by the recession indicator (5.23), growth rate cycles are related to the growth rate of aggregate economic activity. We adopt the growth rate cycle periods defined by the Economic Cycle Research Institute (ECRI). Based on their definition of "periods of cyclical upswings and downswings in growth", we introduce the binary indicator

$$y_{2t} = \begin{cases} 1, & \text{if the growth rate cycle is in a downswing state at time } t, \\ 0, & \text{if the growth rate cycle is in an upswing state at time } t. \end{cases} \tag{5.24}$$

In this chapter, the "downswing state" of the growth rate indicator ($y_{2t} = 1$) is referred to as a "growth rate recession".[6]

ECRI determines the turning points in the growth rate cycle in a way analogous to the "NBER approach", where the co-movements and cyclical turns in various measures of the aggregate macroeconomic activity are taken into account. Banerji and Hiris (2001) provide a more detailed discussion on the business cycle and growth rate cycle periods (see also Layton and Moore, 1989). It is expected that the growth rate cycle indicator y_{2t} exhibits more regime switches than the business

[5] See details on http://www.nber.org/cycles/recessions.html [2 July 2009].

[6] Osborn *et al.* (2004) refer these periods as "growth regimes".

cycle indicator y_{1t}. The reason is that a period with a lower growth rate may be classified as a growth rate recession, but it is not necessarily defined as a business cycle recession.

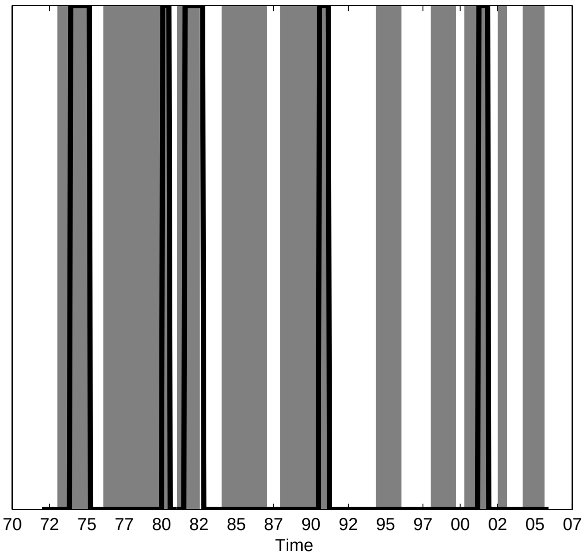

Figure 5.1: U.S. business cycle recession periods ($y_{1t} = 1$, line) since January 1972 until December 2005. Shaded areas indicate growth rate recessions (downswing periods in the growth rate cycle, i.e. $y_{2t} = 1$).

Figure 5.1 depicts the ECRI growth rate recession periods along with the NBER recession periods. Table 5.1 shows the cross-tabulation of the realized values of y_{1t} and y_{2t} defined in terms of these periods. As expected, growth rate cycles are more numerous than "classical" business cycles. All slowdowns in the growth of economic activity do not involve business cycle recessions. On the other hand, the growth rate cycle recessions seem to lead the business cycle recessions: the growth rate recession has typically started a few months before a business cycle recession period. Furthermore, it should be pointed out that the rare outcome ($y_{1t} = 1, y_{2t} = 0$), i.e. the economy is in a business cycle recession, but at the same time in a growth rate expansion, is also possible. Table 5.1 shows that this outcome has been occurred in five months in our data set. When taking a closer look at the turning point chronologies of the NBER and ECRI, it can be seen that those periods have been related, as expected, to the endpoints of the business cycle recessions ($y_{1t} = 1$), where the growth rate expansion ($y_{2t} = 0$) has started before the business cycle expansion ($y_{1t} = 0$).

Table 5.1: Dependent variables and the cross-tabulation of realized values.

		y_{2t}	
		0	1
y_{1t}	0	162	205
	1	5	48

Notes: U.S. business cycle periods y_{1t} (recession/expansion) and growth rate cycle periods y_{2t} (growth rate recession/growth rate expansion) are obtained from http://www.nber.org/cycles/cyclesmain and http://www.businesscycle.com. The sample period is 1972 M1–2005 M12. [30 April 2009]

Osborn *et al.* (2004) emphasize the relationship between the output gap (i.e. the difference between the realized level and potential of output) and growth rate cycle recessions and expansions periods (see Figure 5.1 and Section 1.3.2). The output gap is of interest for many economic agents, especially central banks in their setting of monetary policy. However, information on the value of the current output gap is not available in real time (see, e.g., Orphanides and van Norden, 2002) because of the informational delays and revisions in real-GDP which is often employed to determine the output gap. Thus, because of this above-mentioned relationship, accurate predictions for the growth rate cycles can be very useful, for example, in monetary policy decision making.

5.4.2 Data Set and Predictive Models

In addition to y_{1t} and y_{2t}, our data set consists of a number of financial variables, such as interest rates and stock market returns, which are used as predictors in bivariate probit models.[7] Both levels and first differences of various interest rates are considered. Assuming that monetary policy has an impact on real economic activity and its growth rate, it is of interest to study which interest rate variable

[7] The variables are described in more detail in Table 5.2.

is the most informative predictor. The Federal funds rate (FF_t) is closely related to the monetary policy in the U.S., so that it is a natural candidate variable (see, e.g., Bernanke and Blinder, 1992).

Table 5.2: Explanatory variables.

r_t	Stock market return, log-difference of the S&P500 index
FF_t	Federal funds rate
i_t	Three-month Treasury bill rate, secondary market
R_t	10-year Treasury bond yield rate, constant maturity
SP_t	Term spread, $R_t - i_t$
ΔFF_t	First difference in Federal funds rate
Δi_t	First difference in three-month Treasury bill rate
ΔR_t	First difference in 10-year Treasury bond yield
SP_t^{GE}	German term spread between the long-term and short-term interest rate.

Notes: Interest rates are from Federal Reserve Statistical Release Historical Data set (http://www.federalreserve.gov/releases/h15/data.htm). S&P500 stock market index is taken from Yahoo Finance (http://finance.yahoo.com) and from http://www.econstats.com. German term spread is constructed as the difference between 10-year Federal security (series WZ9826, the missing values between 1971 M1-1972 M9 are replaced by the OECD 10-year interest rate) and the three-month money market rate (series su0107, see http://www.bundesbank.de/statistik/statistik) [30 April 2009].

The term spread (SP_t) between the long-term interest rate and the short-term interest rate has often been found the most important predictor of business cycle recessions (see, e.g., Estrella and Mishkin, 1998). Hence, it is of interest to examine whether the term spread is also an important predictor for the growth rate cycle periods. Furthermore, as a forward-looking variable and incorporating expectations of future dividends and profitability of firms, stock market returns (r_t) should also have predictive power.

We concentrate on the potential predictive information of financial variables, which are available with no revisions or informational lags at the monthly frequency. For example, the output gap, which Osborn *et al.* (2004) find the best predictor of growth rate cycle recessions with the long-term and short-term interest rates and stock market returns for European countries, is not available on a real-time basis. In addition to the real-time availability of predictive variables, another issue that should be taken into account in the specification of the predictive model is the fact that the values of the NBER business cycle phases (y_{1t}), and apparently also the growth cycle periods defined by the ECRI (y_{2t}), become available with very long delays. We call these delays as "publication lags". Without explicit assumptions concerning the publication lags it is difficult to use lagged values of the cycle indicators in the predictive model. Overall, the publication lags are typically so long that it is likely that the lagged values of the indicators are statistically insignificant in estimated models (see the evidence on univariate models in Nyberg, 2010).[8] Therefore, we only consider models excluding the lagged values of the indicators y_{1t} and y_{2t}.

We consider four different model specifications obtained from the bivariate autoregressive probit model given in (5.6) and (5.8). The models are defined as follows:

$$\begin{aligned}
\text{Model 1} &: \alpha_{12} = 0, \alpha_{21} = 0, \rho = 0,\\
\text{Model 2} &: \alpha_{12} = 0, \alpha_{21} = 0,\\
\text{Model 3} &: \rho = 0,\\
\text{Model 4} &: \text{unrestricted model.}
\end{aligned}$$

Model 1 consist of two independent autoregressive probit models. Model 2 is obtained from Model 1 by removing the restriction that the correlation coefficient ρ is zero. Note that Models 1 and 2 are already extensions of the static bivariate model (5.7) because both π_{1t} and π_{2t} follow univariate autoregressive models. Model 4 is the bivariate autoregressive probit model (5.8) without any restrictions, whereas only the correlation coefficient ρ in (5.6) is restricted to zero in Model 3.

5.4.3 Model Selection and In-Sample Results

Model selection considered in this section is based on models estimated over the entire sample period from January 1971 to December 2005. The first 12 observa-

[8] The most recent publication lags of the NBER have varied from five up to twenty months (see http://www.nber.org/cycles/cyclesmain.html) [2 July 2009].

tions are used as initial values in estimation. This section is also a starting point for out-of-sample forecasts for the U.S. at time period 2006–2008 considered in Section 5.4.5.

As mentioned in Section 5.4.1, unlike the business cycle recession periods, there are few previous results on the predictability of growth rate cycle periods. In model selection, therefore, we concentrate on predicting the growth rate cycle recessions with various financial variables. For simplicity, we employ the same explanatory variables as Nyberg (2010) in the predictive models of the U.S. business cycle recession periods, i.e.,

$$\boldsymbol{x}_{1,t-k} = \begin{pmatrix} SP_{t-6} & r_{t-1} & SP^{GE}_{t-6} \end{pmatrix}', \tag{5.25}$$

where SP_t is the U.S. term spread, r_t is the monthly stock market return, and SP^{GE}_t is the German term spread (see also the evidence in Estrella and Mishkin, 1998; Bernard and Gerlach, 1998). The German term spread is used as a "representative" foreign term spread reflecting the state of the economy in the euro area. Here the employed lags are the same as in Nyberg (2010) and, because the forecast horizon is one month ($h = 1$), the condition $k \geq h$ (see Section 5.3.3) is satisfied.

We apply the following model selection procedure concerning the variables included in $\boldsymbol{x}_{2,t-k}$. First, we estimate univariate autoregressive probit models with one predictor. The considered predictive variables are listed in Table 5.2. When the best single predictor based on BIC is found, it will be retained in the model and different models with two predictors are estimated. We mainly restrict ourselves to models with two predictors, but make some experiments with models containing the third predictor as well. Finally, we consider bivariate models with the predictors selected at the first stage.

Table 5.3 shows values of the Schwarz information criterion (5.21) for Model 1 with different explanatory variables, when the employed lag varies from one ($k = 1$) to six ($k = 6$).[9] Especially in more general bivariate probit models, such as Model 4, the number of parameters is quite large, indicating that ML estimation may

[9] It appears that the evidence of predictive power of different explanatory variables is the same when goodness-of-fit measures other than the BIC are used.

become difficult. Therefore, we prefer parsimonious models which makes BIC a suitable model selection criterion.

Table 5.3: BIC values for Model 1 with different explanatory variables for the growth rate cycle indicator.

k	r_{t-k}	FF_{t-k}	i_{t-k}	R_{t-k}	SP_{t-k}	ΔFF_{t-k}	Δi_{t-k}	ΔR_{t-k}
1	308.77	312.76	315.26	331.25	300.52	335.83	335.83	335.41
2	314.75	312.20	314.66	330.95	300.51	280.41	286.39	335.48
3	319.02	314.11	314.92	330.27	304.30	280.34	282.19	336.05
4	321.64	317.55	318.36	330.21	313.58	289.94	286.34	335.90
5	327.01	321.22	320.82	330.55	318.26	298.25	286.87	319.66
6	331.01	323.71	323.53	330.80	323.52	304.70	308.30	325.52

Notes: Explanatory variable included in $\boldsymbol{x}_{2,t-k}$ and its lag k are mentioned in the first row and the first column of the table. Schwarz information criterion, BIC, is defined in (5.21).

Table 5.4: BIC values for Model 1 with the lagged first difference of the Federal funds rate and other explanatory variables for the growth rate cycle indicator.

	k	r_{t-k}	FF_{t-k}	i_{t-k}	R_{t-k}	SP_{t-k}	Δi_{t-k}	ΔR_{t-k}
ΔFF_{t-2}	1	269.75	277.56	277.49	276.91	283.39	281.06	278.26
	2	273.18	277.49	277.31	276.54	303.51	283.08	280.49
	3	278.27	277.49	277.25	276.26	283.28	282.00	282.94
	4	277.05	277.64	277.35	276.11	283.20	280.28	283.30
	5	278.38	277.88	277.57	276.12	283.11	276.31	282.87
	6	282.23	278.19	277.97	276.23	282.94	281.82	282.66
ΔFF_{t-3}	1	270.27	278.18	277.90	277.49	283.32	282.57	280.85
	2	274.09	278.43	278.08	277.22	303.44	281.84	281.95
	3	278.64	278.62	278.73	276.84	307.16	282.85	282.64
	4	278.61	278.62	278.30	276.83	282.92	283.23	283.24
	5	279.22	278.63	278.30	276.55	282.87	280.26	283.34
	6	282.07	278.78	278.51	276.53	282.72	282.28	282.93

Notes: The employed predictive variables in $\boldsymbol{x}_{2,t-k}$ are the second, or the third, lag of the first difference of the Federal funds rate and variable mentioned in the first row of the table. See also notes to Table 5.3.

According to Table 5.3, the best single predictive variable seems to be the first difference of the Federal funds rate lagged by two or three months (ΔFF_{t-2} or

ΔFF_{t-3}). The first difference of the short-term interest rate (Δi_{t-k}) also performs quite well. Furthermore, as seen from Table 5.4, when models with two predictive variables are considered, the lagged stock market return r_{t-1} combined with ΔFF_{t-2} provides the lowest value of the BIC. Thus the vector $\boldsymbol{x}_{2,t-k}$ is selected as

$$\boldsymbol{x}_{2,t-k} = \begin{pmatrix} \Delta FF_{t-2} & r_{t-1} \end{pmatrix}'. \tag{5.26}$$

This selection is also meaningful from the viewpoint of the predictive ability of financial markets because variables reflecting the effect of both the monetary policy (ΔFF_{t-2}) and the stock market returns (r_{t-1}) are now included in the model.

Based on the evidence in Tables 5.3 and 5.4, the U.S. term spread has some ability to predict the growth rate cycle periods. However, it is outperformed by several other variables. The term spread is also found to be a statistically insignificant predictor when used a third predictor in Models 1–4 together with the explanatory variables given in (5.26).

Table 5.5 shows the estimation results for Models 1–4 with the explanatory variables given in (5.25) and (5.26). The signs of the estimated coefficients are as expected. In the case of the growth rate cycle periods, increasing values of the differenced Federal funds rate and negative stock market returns increase the probability of growth rate recession. The results thus indicate that the U.S. monetary policy has a statistically significant predictive impact for the current state of the growth rate cycle via the first difference of the Federal funds rate. Stock market returns also have predictive power for the growth rate cycles as well as for business cycle recession periods. In addition, the U.S. term spread (SP_{t-6}) and the German term spread (SP^{GE}_{t-6}) are also statistically significant predictors.

In Model 2, the estimate of the correlation coefficient ρ is statistically significant. The LM^{ρ} test based on the restricted Model 1 yields the same conclusion. A positive estimate of ρ is obtained, as expected, since the correlation between the dependent variables is positive. Further, in Model 3, the estimates of the off-diagonal elements of the matrix $\mathbf{A}$, α_{12} and α_{21}, for $\pi_{2,t-1}$ and $\pi_{1,t-1}$ are statistically significant, and according to the values of the BIC, Model 3 outperforms

Table 5.5: Estimation results of different bivariate models.

model	variable	Model 1	Model 2	Model 3	Model 4
π_{1t}	constant$_1$	0.06	0.06	-0.07	-0.16
		(0.03)	(0.03)	(0.05)	(0.05)
	$\pi_{1,t-1}$	0.85	0.86	0.88	0.86
		(0.02)	(0.02)	(0.01)	(0.01)
	$\pi_{2,t-1}$			0.17	0.18
				(0.05)	(0.05)
	SP_{t-6}	-0.19	-0.16	-0.08	-0.06
		(0.03)	(0.04)	(0.03)	(0.02)
	r_{t-1}	-0.11	-0.10	-0.10	-0.10
		(0.02)	(0.02)	(0.02)	(0.02)
	SP_{t-6}^{GE}	-0.08	-0.07	-0.13	-0.11
		(0.03)	(0.02)	(0.03)	(0.03)
π_{2t}	constant$_2$	0.07	0.07	-0.02	-0.05
		(0.01)	(0.01)	(0.01)	(0.01)
	$\pi_{2,t-1}$	0.91	0.91	0.92	0.96
		(0.02)	(0.01)	(0.01)	(0.01)
	$\pi_{1,t-1}$			-0.03	-0.04
				(0.01)	(0.01)
	ΔFF_{t-2}	0.51	0.51	0.34	0.31
		(0.18)	(0.06)	(0.06)	(0.05)
	r_{t-1}	-0.03	-0.03	-0.06	-0.06
		(0.01)	(0.01)	(0.01)	(0.01)
	ρ		0.58		0.53
			(0.21)		(0.21)
logL		-242.70	-240.31	-223.38	-217.94
BIC		269.75	270.36	256.44	254.01
QPS_1		0.061	0.063	0.061	0.062
QPS_2		0.340	0.330	0.302	0.284
$CR_1^{50\%}$		0.963	0.961	0.961	0.963
$CR_2^{50\%}$		0.713	0.723	0.770	0.789
LM^ρ		8.58		27.21	
p-value		0.000		0.000	

Notes: The models are estimated using the data from 1971 M1 to 2005 M12. The first 12 observations are used as initial values. Standard errors of the estimated coefficients are given in parentheses. Estimated values of the log-likelihood function (5.11), logL, and Schwarz (1978) information criteria, BIC, are reported, as well as the values of quadratic probability scores, QPS_j, where $j = 1, 2$. Further, $CR_j^{50\%}$ indicate the ratio of correct prediction with using the 50% threshold value in the classification of probability forecasts. Lagrange Multiplier test statistics LM^ρ (see (5.18)) for the null hypothesis (5.16) and the corresponding p-values are also reported.

Model 2 as an extension of Model 1. However, both the LM^ρ test based on Model 3 and the Wald test based on Model 4 point to a nonzero value of the correlation coefficient ρ. Thus, Model 4 clearly yields the best in-sample fit. This is especially the case when the models for the growth rate cycle periods are compared with the values of QPS_2 and CR_2.

The fact that Model 4 outperforms alternative bivariate probit models reflects the fact that recession probabilities of the two cycle indicators are dependent on both $\pi_{1,t-1}$ and $\pi_{2,t-1}$. For instance, a positive estimate of α_{12} indicates that the probability of a business cycle recession is high when the lagged probability of growth rate recession is high (high value of π_{2t}). This is in line with the fact that the growth rate recession appears to precede occurred business cycle recession periods. Note that in Models 3 and 4, the U.S. term spread, which is a statistically significant predictor for business cycle recession periods, has also an effect on the growth rate cycle recession probability via the coefficient α_{21} for $\pi_{1,t-1}$ in π_{2t}.

As an extension of the results presented in Table 5.5, we consider the possibility that there has been a structural break in the data generating process. In the previous literature, it has been suggested that there might have been a structural change in the U.S. economic activity in the mid-1980s. For example, McConnell and Perez-Quiros (2000) and Blanchard and Simon (2001) have documented that the variability of output, and also the variability of inflation, has declined after the mid-1980s.[10] Sensier and van Dijk (2004) also provide evidence that there have been structural breaks in the unconditional volatility of many U.S. macroeconomic time series around the years 1984 to 1986.

In our application, a parsimonious way to allow for the potential effect of a structural break is the inclusion of an additional dummy variable in the model. The dummy variable, denoted by $\mathbf{1}_{71-84}$, takes the value one before the beginning of the year 1985. The results for the augmented models are presented in Table 5.6, where the dummy variable turns out to be a statistically significant predictor. The models are also superior to their counterparts in Table 5.5 according to BIC values. Model 4 augmented with the dummy variable $\mathbf{1}_{71-84}$ yields clearly the best in-sample predictions, especially for growth rate cycle recessions.

[10] This time period after the mid-1980s is often referred to the "Great Moderation" period.

Table 5.6: Estimation results of different models with the additional dummy variable $\mathbf{1}_{71-84}$.

model	variable	Model 1	Model 2	Model 3	Model 4
π_{1t}	constant$_1$	0.09	0.07	0.08	0.03
		(0.05)	(0.05)	(0.04)	(0.04)
	$\mathbf{1}_{71-84}$	-0.04	-0.00	-0.25	-0.21
		(0.06)	(0.05)	(0.09)	(0.08)
	$\pi_{1,t-1}$	0.85	0.86	0.89	0.88
		(0.02)	(0.02)	(0.01)	(0.01)
	$\pi_{2,t-1}$			0.11	0.11
				(0.04)	(0.03)
	SP_{t-6}	-0.20	-0.16	-0.11	-0.10
		(0.04)	(0.04)	(0.03)	(0.02)
	r_{t-1}	-0.11	-0.09	-0.09	-0.09
		(0.04)	(0.02)	(0.02)	(0.02)
	SP_{t-6}^{GE}	-0.08	-0.08	-0.12	-0.11
		(0.03)	(0.03)	(0.03)	(0.02)
π_{2t}	constant$_2$	0.06	0.06	-0.07	-0.09
		(0.01)	(0.01)	(0.02)	(0.02)
	$\mathbf{1}_{71-84}$	0.06	0.06	0.09	0.08
		(0.02)	(0.02)	(0.03)	(0.02)
	$\pi_{2,t-1}$	0.91	0.91	0.93	0.94
		(0.01)	(0.01)	(0.01)	(0.01)
	$\pi_{1,t-1}$			-0.04	-0.04
				(0.01)	(0.01)
	ΔFF_{t-2}	0.55	0.55	0.40	0.38
		(0.07)	(0.06)	(0.08)	(0.07)
	r_{t-1}	-0.03	-0.03	-0.08	-0.07
		(0.01)	(0.01)	(0.01)	(0.01)
	ρ		0.59		0.77
			(0.22)		(0.15)
logL		-235.15	-232.76	-198.63	-192.46
BIC		268.21	268.83	237.71	234.54
QPS_1		0.061	0.063	0.058	0.059
QPS_2		0.327	0.310	0.256	0.254
$CR_1^{50\%}$		0.963	0.961	0.958	0.958
$CR_2^{50\%}$		0.725	0.740	0.816	0.826
LM^{ρ}		8.14		26.21	
p-value		0.000		0.000	

Notes: See notes to Table 5.5. Variable $\mathbf{1}_{71-84}$ indicates a variable which takes value one at period from 1971 M1 to 1984 M12, and zero otherwise.

Although Model 4 seems to outperform its special cases, the estimated coefficients of the explanatory variables are almost equal in all models in Tables 5.5

and 5.6. These findings confirm our earlier results that the changes in the Federal funds rate and stock market returns appear to be the main predictive variables for growth rate cycles also in the more general bivarite probit models.

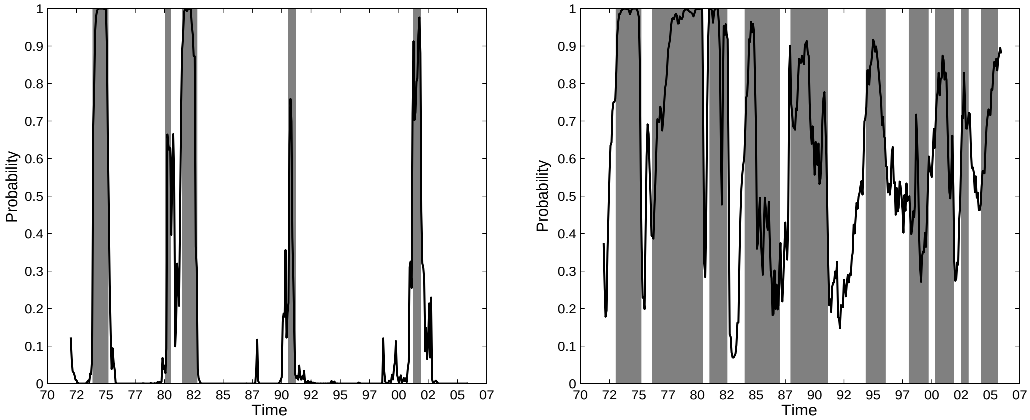

Figure 5.2: In-sample fitted values from the Model 1 presented in Table 5.6. Recession periods are depicted with shaded areas. Business cycle recession probabilities are presented in the left panel, probabilities for growth rate cycle periods in the right panel.

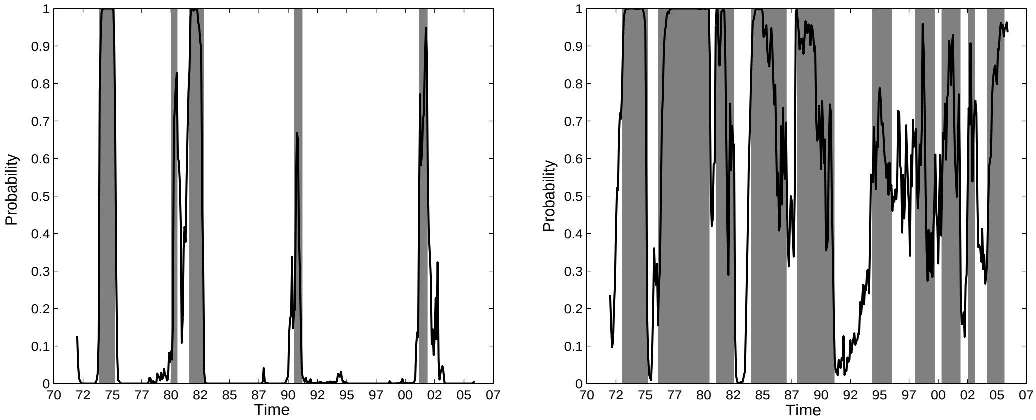

Figure 5.3: In-sample fitted values from the Model 4 presented in Table 5.6. Recession periods are depicted with shaded areas. Business cycle recession probabilities are presented in the left panel, probabilities for growth rate cycle periods in the right panel.

Figures 5.2 and 5.3 depict the in-sample fitted values from Models 1 and 4 shown in Table 5.6. As the estimation results in Tables 5.5 and 5.6 suggest, the business cycle recession periods predicted by these two models are almost identical. For the growth rate recession periods, Model 4 gives somewhat more precise signals around the years from 1984 to 1987 and before the year 2000.

In conclusion, especially the general bivariate autoregressive model (Model 4), but also its special cases (Models 2 and 3), are superior to the independent univariate autoregressive models for both cycle indicators (Model 1). These findings indicate that superior predictive power can be found by using bivariate models instead of univariate models.

5.4.4 Out-of-Sample Performance

In this section, we examine the out-of-sample predictive performance of different models. The first out-of-sample nowcasts are made for January 2000 and the last ones for December 2005. Notice that the out-of-sample period contains three growth rate cycle recession periods, but only one business cycle recession in 2001 (see Figure 5.1). In addition to these out-of-sample nowcasts, in Section 5.4.5, the best bivariate probit models, according to model selection criteria and in-sample predictions provided in Section 5.4.3, are used to assess the state of U.S. economy from 2006 to 2008.

Different predictive models are estimated using the data up to December 1999 assuming that the state of the economy is known at that time. To emulate real-time forecasting, we should take into account the fact that the latest values of the dependent variables are unknown at the time the forecast is made (see Section 5.4.2). Thus the out-of-sample exercise is carried out without updating the parameter estimates using the sample period up to December 1999. The same framework is also applied in Section 5.4.5. It turns out that the model selection procedure employed in Section 5.4.3, using the data set up to December 1999, yields the same conclusions as obtained when using the whole sample in estimation. Therefore, the differenced Federal funds rate and the stock market returns have predictive power for the growth rate cycles also in this estimation period.

Table 5.7 shows the out-of-sample performance of the employed models. Out-of-sample forecast accuracy is evaluated by the QPS and the percentage of correct forecasts (CR). The first four models also include the dummy variable $\mathbf{1}_{71-84}$. As in Section 5.4.3, the bivariate autoregressive probit model (Model 4) yields the best out-of-sample predictions although Model 3 also produces good forecasts. Based on the predictability test of Pesaran and Timmermann (1992) the reported percentages of correct forecasts are statistically significant at the 1% level in models including the $\mathbf{1}_{71-84}$ variable. Without the additional dummy variable the percentages of correct forecasts are lower and, consequently, the p-values are higher and statistically insignificant at 5% level.

Table 5.7: Out-of-sample performance of different models.

	QPS_1	CR_1	QPS_2	CR_2
Model 1 with $\mathbf{1}_{71-84}$	0.395	0.722	0.395	0.708
Model 2 with $\mathbf{1}_{71-84}$	0.435	0.722	0.395	0.708
Model 3 with $\mathbf{1}_{71-84}$	0.167	0.889	0.303	0.833
Model 4 with $\mathbf{1}_{71-84}$	0.130	0.903	0.307	0.847
Model 1	0.179	0.833	0.446	0.625
Model 2	0.228	0.819	0.446	0.625
Model 3	0.094	0.931	0.446	0.625
Model 4	0.108	0.944	0.608	0.611

Notes: The out-of-sample values of QPS_j and CR_j, $j = 1, 2$, based on models mentioned on the left. Explanatory variables included in the model are the same as in the estimation results presented in Tables 5.5 and 5.6. Nowcasts from Model 4 with the dummy variable $\mathbf{1}_{71-84}$ (the fourth model) are depicted in Figure 5.4.

Although the dummy variable $\mathbf{1}_{71-84}$ is useful in predicting the growth rate cycle recessions, this is not the case for nowcasts of the business cycle recession periods. However, the more general Models 3 and 4 outperform Models 1 and 2 even in this case. Note that the percentages of correct forecasts for business cycle periods are statistically significant in all models presented in Table 5.7.

Figure 5.4 illustrates the out-of-sample performance of the unrestricted Model 4. It appears that the recession probabilities match well the realized values of both cycle indicators. As depicted in the left panel, Model 4 predicts the business cycle recession in 2001 really well, but the increase in the recession probability in

2002 weakens the performance of the model in terms of the values of the statistical goodness-of-fit measures reported. The out-of-sample nowcasts depicted in the right panel match the growth rate cycle periods well. For instance, when the 50% threshold value for probabilitity forecasts to construct signal forecasts for growth rate recessions and expansions is used, Model 4 gives the correct signal forecast with approximately 85% accuracy ($CR = 0.847$).

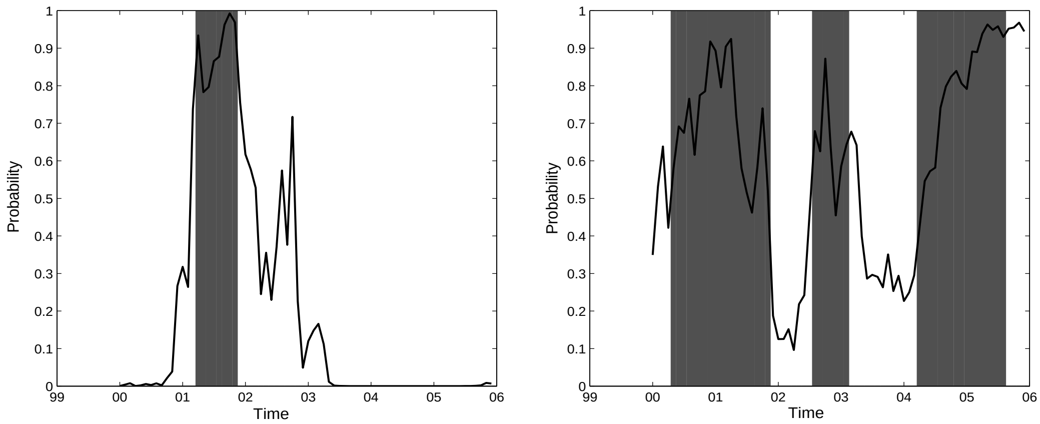

Figure 5.4: Out-of-sample nowcasts for business cycle recession (left panel) and growth rate cycle (right panel) recession periods (shaded areas) using the bivariate autoregressive probit model (Model 4) which contains also the dummy variable $\mathbf{1}_{71-84}$ given in Section 5.4.3.

In summary, the results confirm that the proposed general bivariate autoregressive model (5.8) outperforms its restricted versions and the values of both cycle indicators appear to be predictable also out of sample.

5.4.5 Predictions for 2006–2008

There has been great uncertainty about the state of the economy in the United States during the last couple of years. Therefore, it is of great interest to consider the probabilities of the business and growth rate cycle periods during 2006–2008. Business cycle recession probabilities are of special interest due to the fact that the NBER Business Cycle Dating Committee declared that a peak in the U.S. economic activity occurred in December 2007 indicating that the value of the recession indicator (5.23) has been one, at least in some months from January 2008.

In this section, we examine nowcasts of the state of the economy using the best models found in Sections 5.4.3 and 5.4.4. The first predictions are made for January 2006 and the last ones for November 2008. According to the recent announcement of the NBER, the U.S. economy has been in business cycle recession since the beginning of the year 2008. In the case of the growth rate cycle periods, it is, however, not evident that there has been a constant downswing state (i.e. a growth recession, $y_{2t} = 1$) after January 2006. Thus, at the time this chapter is written, the realized values of y_{2t} are not known after January 2006.

Figures 5.5 and 5.6 depict the nowcasts from Models 1 and 4. The probabilities in Figure 5.6 are based on the models including the additional dummy variable $\mathbf{1}_{71-84}$. The evidence appears to be ambiguous between different models. It seems that Model 1 produces greater business cycle recession probabilities than Model 4 at the beginning of the recession in January 2008 (see Figure 5.5). On the other hand, in both cases, Model 1 gives higher "false" recession risks than model 4 for some months before the recession started. Overall, the employed models seem to nowcast the beginning of the recession period at 2008 reasonably well.

As discussed above, the latest values of the growth rate cycle indicator are unknown and, thus, it is difficult to make comparisons between different models with the currently available information. All in all, it seems that the growth rate recession probability have been decreasing from mid-2006. As the business cycle recession started at the beginning of 2008, it is likely that the U.S. economy has been in a growth recession some months before. For those months, the predicted probabilities in Figure 5.5 are quite high and exceed the 50% threshold value, indicating a growth rate cycle recession. Note also that a decreasing probability of growth rate recession affects the business cycle recession probability in Model 4. This is a potential reason why the business cycle recession probabilities in the left panels of Figures 5.5 and 5.6 are lower than in the independent univariate autoregressive model (Model 1).

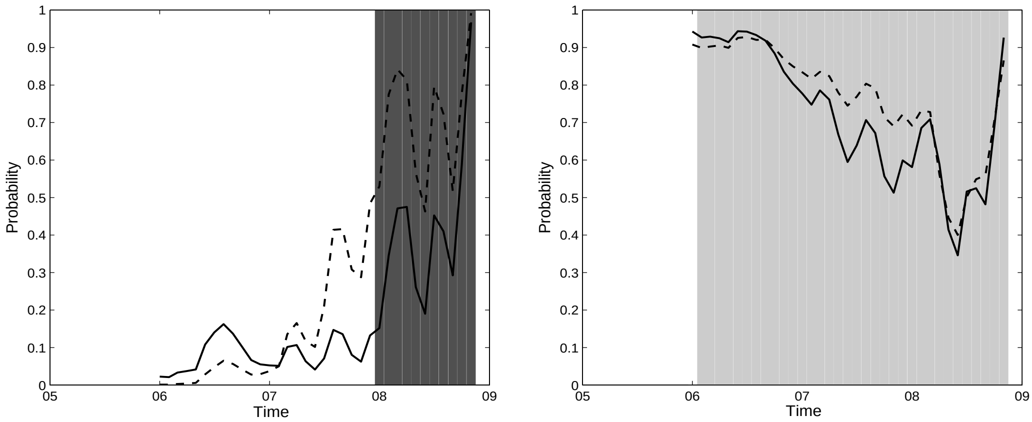

Figure 5.5: Real-time predictive probabilities from Model 1 (dashed line) and from Model 4 for the business cycle (left panel) and the growth rate cycle periods (right panel) using the models described in Table 5.5. In the left panel, the business cycle recession that began at 2008 is depicted with shaded area. In the right panel, the shaded area corresponds to the time period of unknown values of the growth rate cycle indicator y_{2t}.

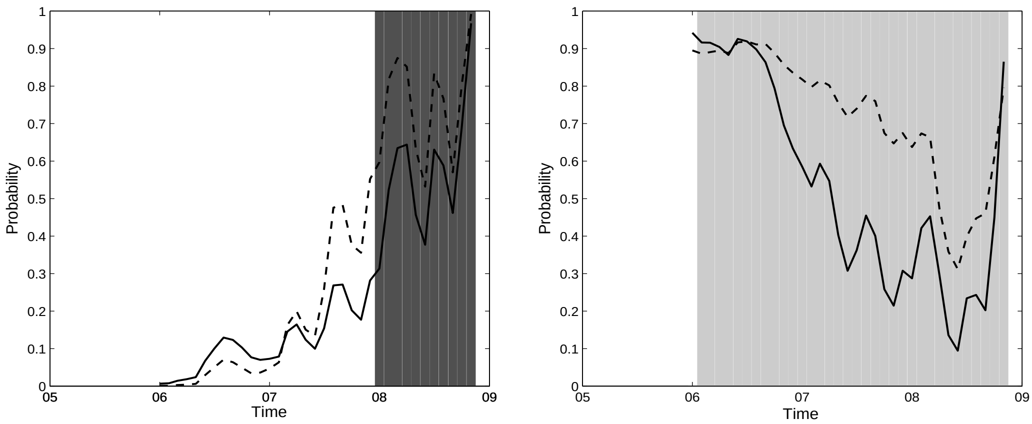

Figure 5.6: Real-time recession probabilities from Model 1 (dashed line) and from Model 4 for the business cycle (left panel) and the growth rate cycle periods (right panel) using the models described in Table 5.6. In the left panel the business cycle recession period that began at 2008 is depicted with shaded area. In the right panel, the shaded area corresponds to the time period of unknown values of the growth rate cycle indicator y_{2t}.

5.5 Conclusions

We introduce a new bivariate time series model for binary dependent variables. The bivariate autoregressive probit model is a bivariate extension of the univariate autoregressive model of Kauppi and Saikkonen (2008), but it can also be considered a dynamic extension of the static bivariate probit model of Ashford and Sowden (1970).

The bivariate autoregressive probit model, and its special cases, are applied to predict the current state of the U.S. economy using binary indicator variables for the level and the growth rate of the U.S. economic activity. The proposed bivariate model framework extends the traditional univariate analysis of business cycle recession periods examined in the previous literature, where only business cycle recessions have been considered.

We found strong in-sample and out-of-sample evidence in favor of the proposed bivariate autoregressive probit model. The results suggest that it is possible to gain additional predictive power by modeling the recession probabilities of the two cycle indicators jointly instead of considering two independent univariate autoregressive probit models. The lagged first difference of the Federal funds rate is the most useful single predictor of the state of the growth rate cycle, but also monthly stock market returns turn out to be statistically significant predictors for both cycle indicators. As suggested in previous studies, a term spread between the long-term and short-term interest rates is an important explanatory variable for business cycle recession periods, but it turned out not to be the best predictive variable for the growth rate cycle periods. We also found evidence that the probability of growth rate recession was systematically higher in the 1971–1984 period than after the mid-1980s.

References

Anatolyev S. 2009. Multi-market direction-of-change modeling using dependence ratios. *Studies in Nonlinear Dynamics and Econometrics* **13**, article 5.

Ashford JR, Sowden RR. 1970. Multivariate probit analysis. *Biometrics* **26**: 535–546.

Banerji A, Hiris L. 2001. A framework for measuring business cycles. *International Journal of Forecasting* **17**: 333–348.

Bernanke BS, Blinder AS. 1992. The Federal funds rate and the channels of monetary transmission. *American Economic Review* **82**: 901–921.

Bernard H, Gerlach S. 1998. Does the term structure predict recessions? The international evidence. *International Journal of Finance and Economics* **3**: 195–215.

Blanchard O, Simon J. 2001. The long and large decline in U.S. output volatility. *Brooking Papers on Economic Activity* **1**: 135–164.

Chauvet M, Potter S. 2005. Forecasting recession using the yield curve. *Journal of Forecasting* **24**: 77–103.

Chib S, Greenberg E. 1998. Analysis of multivariate probit models. *Biometrika* **85**: 347–361.

Diebold FX, Rudebusch GD. 1989. Scoring the leading indicators. *Journal of Business* **62**: 369–391.

Dueker MJ. 2005. Dynamic forecasts of qualitative variables: A qual VAR model of U.S. recessions. *Journal of Business and Economic Statistics* **23**: 96–104.

Ekholm A, Smith PWJ, McDonald JW. 1995. Marginal regression analysis of a multivariate binary response. *Biometrika* **82**: 847–854.

Engle RF. 1984. Wald, likelihood ratio and Lagrange Multiplier tests in econometrics, in Griliches Z. and Intriligator MD. (eds.), *Handbook of Econometrics*, Vol. II. Amsterdam, North-Holland.

Estrella A, Hardouvelis GA. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* **46**: 555–576.

Estrella A, Mishkin FS. 1998. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics* **80**: 45–61.

Greene WH. 2000. *Econometric Analysis*. Fourth edition. Prentice-Hall International, London.

Kauppi H, Saikkonen P. 2008. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics* **90**: 777–791.

Kiefer NM. 1982. Testing for dependence in multivariate probit models. *Biometrika* **69**: 161–166.

Layton AP, Moore GH. 1989. Leading indicators for the service sector. *Journal of Business and Economic Statistics* **7**: 379–386.

McConnell MM, Perez-Quiros P. 2000. Output fluctuations in the United States: What has changed since the early 1980's? *American Economic Review* **90**: 1464–1476.

Mosconi R, Seri R. 2006. Non-causality in bivariate binary time series. *Journal of Econometrics* **132**: 379–407.

Nyberg H. 2010. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting* **29**: 215–230.

Osborn DR, Sensier M, van Dijk D. 2004. Predicting growth regimes for European countries, in Lucrezia Reichlin (eds.), *The Euro Area Business Cycle: Stylized Facts and Measurement Issues*. Centre for Economic Policy Research.

Orphanides A, van Norden S. 2002. The unreliability of output-gap estimates in real time. *Review of Economics and Statistics* **84**: 569–583.

Pesaran HM, Timmermann A. 1992. A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics* **10**: 461–465.

Schwarz G. 1978. Estimating the dimension of a model. *Annals of Statistics* **6**: 461–464.

Sensier M, van Dijk D. 2004. Testing for volatility changes in U.S. macroeconomic time series. *Review of Economics and Statistics* **86**: 833–839.