Christian R. Proaño
Thomas Theobald

Predicting German Recessions with a Composite Real-Time Dynamic Probit Indicator

July 2012
Working Paper 05/2012
Department of Economics
The New School for Social Research

Predicting German Recessions with a Composite Real-Time Dynamic Probit Indicator

Christian R. Proaño*
The New School for Social Research
New York, NY

Thomas Theobald
Macroeconomic Policy Institute (IMK)
Düsseldorf, Germany

July 26, 2012

Abstract

In this paper we propose a composite indicator for real-time recession forecasting based on alternative dynamic probit models. For this purpose a large set of monthly macroeconomic and financial leading indicators of the German economy is used. Alternative dynamic probit regressions are specified through automatized general-to-specific as well as specific-to-general lag selection procedures on the basis of slightly different initial sets, and the resulting recession probability forecasts are then combined in order to decrease the volatility of the forecast errors and increase their forecasting accuracy. As it is shown in the paper, this procedure does not only feature good in-sample forecast statistics, but has also good out-of-sample performance, as a real-time evaluation exercise shows.

Keywords: Dynamic probit models, out-of-sample forecasting, yield curve, real-time econometrics

JEL CLASSIFICATION SYSTEM: C25, C53

*E-mail: christian.proano@gmail.com. We would like to thank the IMK research team, and especially Gustav Horn, Peter Hohlfeld, Heike Joebges, Sven Schreiber, Sabine Stephan, Silke Tober and Rudolf Zwiener, as well as Jörg Breitung for many helpful comments and suggestions.

1 Introduction

As is widely acknowledged, the timely and accurate prediction of turning points in the business cycle is one of the most policy-relevant aspects of macroeconomic forecasting. This task is, however, also one of the most challenging, not only due to the many potential nonlinearities at work at the onset of a turning point in economic activity, but also due to the significant uncertainty of macroeconomic data at the end-point and the model uncertainty inherent in applied work, among other things.

As a way to handle the model uncertainty problem, Bates and Granger (1969) were among the first ones who proposed a combinatorial approach, by showing that the inclusion of ex-ante forecasts with the inferior predicting power offered an improvement if they contained some independent information contrary to the ex-ante best forecasts. More recently, Timmermann (2006) also emphasized the usefulness of forecast combinations due to 1) diversification, 2) structural breaks, 3) misspecification of the individual forecasts and 4) systematic differences in the individual loss functions. Concerning the end-point data uncertainty problem, in contrast, Pesaran and Timmermann (2005) have stressed the urgent need to develop robust interactive systems of model specification and evaluation designed explicitly to work in real time, as "by setting out in advance a set of rules for observation windows and variable selection, estimation, and modification of the econometric model, automation provides a way to reduce the effects of data snooping and facilitates learning from the performance of a given model when applied to a historical data set" (Pesaran and Timmermann, 2005, p.212).

As widely known, binary response model class has been used in the prediction of business cycle turning points, see e.g. Estrella and Mishkin (1998), Bernand and Gerlach (1998), Estrella, Rodriguez and Schich (2003), Moneta (2005), Wright (2006), Haltmeier (2008), and Rudebusch and Williams (2009), among others. Along these lines, in this paper we discuss the rationale and structure of a composite indicator for real-time recession forecasting based on alternative dynamic probit models specified through automatized *general-to-specific*, as well as *specific-to-general* variable and lag selection procedures specifically design to work under real time conditions as discussed in Proaño (2010).

The main contribution of this paper to the literature is thus the development of a composite dynamic probit indicator along the lines of recent studies using binary response models such as Kauppi and Saikonnen (2008) and Nyberg (2010) for recession forecasting under real-time conditions on a monthly basis. As it will be discussed in this paper, through the estimation of a variety of alternative dynamic probit regressions and the combination of the resulting recession probability estimates not only a great deal of information is taken into consideration, but also a lower volatility of the recession forecast error is achieved.

The remainder of this paper is organized as follows. In section 2 the modeling of an econometric forecasting tool based on the dynamic binary response approach for recession forecasting in Germany is discussed. Section 3 derives the combination schemes theoretically and explains how many underlying probit forecasts are to be considered for each. Section 4 discusses the real-time in- and out-of-sample performance of the different combination schemes.

Finally, section 5 draws some conclusions from this study and points out possible extensions left for future research.

2 Methodology

As previously mentioned, following the work by Estrella and Hardouvelis (1991), binary response models have been widely used for the estimation and forecasting of recessionary periods during the last twenty years, see e.g. Dueker (1997), Kauppi and Saikonnen (2008), Rudebusch and Williams (2009) and Nyberg (2010). In this strain of the literature, binary recession indicator series representing the state of the economy within the business cycle b_t is set such that

$$b_t = \begin{cases} 1, & \text{if the economy goes through a recessionary phase at time } t \\ 0, & \text{if the economy experiences an expansion at time } t. \end{cases}$$

Let Ω_{t-h} be the information set available at $t-h$, where h represents the forecasting horizon. Assuming a one-period ahead forecast horizon $h = 1$, E_{t-1} and $\text{Prob}_{t-1}(\,\cdot\,)$ represent the conditional expectation and the conditional probability given the information set Ω_{t-1}, where under the assumption that b_t has a Bernoulli distribution

$$b_t |\, \Omega_{t-1} \sim \mathcal{B}(p_t),$$

the conditional probability p_t of b_t taking the value 1 in t is then given by

$$E_{t-1}(b_t) = \text{Prob}_{t-1}(\,b_t = 1) = p_t = \Phi(\varphi_t).$$

φ_t represents the linear model equation of the variables contained in the information set Ω_{t-1} and $\Phi(\,\cdot\,)$ the linking function between φ_t and the conditional probability $\text{Prob}_{t-1}(\,b_t = 1)$ according to the Bernoulli distribution, which in probit models is given by a standard normal distribution function.

As in Proaño (2010) the latent variable of a real-time dynamic probit indicator for the prediction of the business cycle is then given by

$$\begin{aligned} \varphi_t = & \sum_{j=h+R}^{o} \delta_j b_{t-j} + \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_{\mathbf{x}}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t, \\ & u_t \sim N(0,1) \quad \forall t, \quad R > D_y, \end{aligned} \tag{1}$$

where R stands for the recession recognition lag and D_y, D_x for the lagged data availability. As it can be seen in equation (1) the autoregressive explanatory terms consist of both growth rates of the reference series y and the binary state series b generated out of the reference series for the detection of recession and expansion phases. Moreover $\mathbf{x}$ captures a large set of exogenous macroeconomic and leading indicators to be discussed in detail below.

Summarizing all available explanatory variables and lags in $\mathbf{z}_t^i$ then the i-th specification of a h-step ahead recession forecast with the probit model is given by

$$\begin{aligned} \varphi_{t+h}^i = \mathbf{z^i}'_t\beta + u_{t+h}^i, \quad u_{t+h}^i \sim N(0,1), \quad i \in I, \\ b_{t+h}^i = \begin{cases} 1 & : \quad \varphi_{t+h}^i > 0 \\ 0 & : \quad \varphi_{t+h}^i \leq 0 \end{cases} \end{aligned} \tag{2}$$

and in terms of the expected future value conditional on current information this leads to

$$\begin{aligned} E\left(b_{t+h}^i|\mathbf{z}_t^i,\beta\right) &= \mu_{t+h|t}^i = P\left(b_{t+h}^i = 1|\mathbf{z}_t^i,\beta\right) \\ &= \Phi\left(\mathbf{z}_t'^i\beta\right) = \Phi\left(E\left(\varphi_{t+h|t}^i\right)\right). \end{aligned} \tag{3}$$

In there the size of I is equal to the dimension of the combination space times the elements in each of its components. For instance with five different interest rate spreads and two different kinds of lag choice ten specifications can be taken into account. In this context it should be clear that the inclusion of all yield spreads in the same set of regressors would automatically give rise to a multicollinearity problem. The a-priori choice of just one of the available yield spreads, on the contrary, may also turn out to be wrong if that yield spread has low predictive power.

In order to avoid the latent problem of choosing an arbitrary model specification based on an ad-hoc selection of lagged values – and of the explaining variables in general –, following Proaño (2010) each alternative dynamic probit specification was estimated using a *general-to-specific* as well as a *specific-to-general* approach. In the *general-to-specific* selection procedure (see Campos, Ericsson and Hendry (2005)), the explanatory contribution of each lag of each explanatory variable was tested using a redundant variables Likelihood Ratio (LR) test, with the LR statistic computed as

$$LR = -2(\mathcal{L}_R - \mathcal{L}_U)$$

where $\mathcal{L}_{\mathcal{R}}$ and $\mathcal{L}_{\mathcal{U}}$ are the maximized values of the (Gaussian) log likelihood function of the unrestricted and restricted regressions.[1] In the *specific-to-general* selection procedure, in contrast, the added explanatory value of an additional lag of each explaining variable was tested using an omitted variables Likelihood Ratio test, where under the H_o the coefficient of the additionally added variable (lag) is not significant.

3 Combination Scheme

Given the uncertainty linked with the use of macroeconomic data, as well as the potential misspecification of some/all of the ten dynamic probit regressions, however, it is impossible both a-priori and a-posteriori to select one particular specification as “the one” representing

[1]Under the H_o of this asymptotically χ^2 distributed test with one degree of freedom, the coefficient of a redundant variable (lag) is zero. A rejection of this test results in the conservation of the tested variable (lag) in the model specification.

in the best way the data-generating process. In contrast, it seems advantageous to pursue a combinatorial approach where the information of each regression is incorporated while its eventual bias is relativized.

To express such a combinatorial approach in a more exact way, following Theobald (2012) let

$$\mu_{t+h|t} = \left(\mu^1_{t+h|t}, \ldots, \mu^{|I|}_{t+h|t}\right)'$$

denote the vector of single forecasts and

$$\theta = C\left(\mu_{t+h|t}; w_c\right)$$

the combinatorial forecast resulting from the aggregation of the underlying forecasts by means of determinate combination weights. In the simple case of equally weighted recession probability forecasts the combinatorial forecast is then given by

$$\theta \quad = \quad \frac{1}{|I|}\sum_{i=1}^{|I|} \mu^i_{t+h|t}, \qquad \text{with } |I| = \#\{\text{interest rate spreads}\} \times \#\{G, S\}. \tag{4}$$

Because of considering indeed five long-term maturities (1, 2, 3, 5 and 10 years), where for each the corresponding spread is calculated by subtracting the 3 month euribor interest rate, one obtains $|I| = 10$ for the *simple average* approach. Obviously this is a special case of the *linear opinion pool* with non-negative weights summing up to one.

At this point we could obviously formulate more sophisticated pooling operators for the presented probit models. As a consequence of the analysis in Theobald (2012), however, we formulate a combination scheme which centers the forecast combinations of single forecasts arising from different interest rate spreads and specification order around its median and adds the different forecast horizons for the same future value as an extra source of generating the underlying forecasts. This combination scheme can be thought of as a two-stage procedure, where

$$\begin{aligned} \theta &= \sum_{i=1}^{\#\{\text{horizons}\}} \frac{2^{h-1}}{\sum_{j=0}^{\#\{\text{horizons}\}-1} 2^j} \theta^*_h, \;\; h \in I \backslash I^*, \\ |I| &= \#\{\text{interest rate spreads}\} \times \#\{G, S\} \times \#\{\text{horizons}\}, \end{aligned} \tag{5}$$

and

$$\theta^*_k = \sum_{i=1}^{|I^*|} \frac{\left(\sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|\right) - |\mu^{i,k}_{t+h|t} - \mu^{med}_{t+h|t}|}{(|I^*| - 1)\sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|}, \;\; i \in I^* \subset I, \;\; k \in I \backslash I^*. \tag{6}$$

med denotes the median of the forecast vector and $I \backslash I^*$ the well-defined set after aggregating over the spreads and specification order similar to eq.(4). Note that in real-time applications the horizon as an additional generator of the underlying forecasts in eq.(5) means nothing else than balancing the actual prediction by one which is generated with a longer horizon out of the first revisions and thus considers less actual information. In such a constellation it still seems to be preferable to put more weight on the forecasts using the last available observations although these of course are also subject to the strongest revisions. Obviously the number of horizons that can be taken into account is limited by future uncertainty. For instance if aiming at a parsimonious running time and taking I^* with the same size as in the case of the simple average approach it is reasonable just to choose $\#\{\text{horizons}\} = 2$. This altogether leads to $|I| = 20$ underlying forecasts for each prediction generated by what is called here the *weighted average* approach.[2]

The following sections describe the application of this approach to recession forecasting using German macroeconomic data.

4 Empirical Results

4.1 Data Description

For the following exercise of recession forecasting, a wide dataset of macroeconomic indicators of the German economy was employed. All financial and real economy variables stem from the Bundesbank database (`www.bundesbank.de/statistik/`), with the exception of the orders, which stem from the GENESIS-Online database from the German Statistical Office (`https://www-genesis.destatis.de/genesis/online`), and the ifo business cycle climate index (`http://www.cesifo-group.de`) The estimation sample comprises monthly observations from 1991:1 to 2011:8.

As previously discussed and as already pointed out by Burns and Mitchell (1946), an economic recession is characterized by a *widespread* and *synchronized* downturn in overall economic activity observable on a broad set of economic variables. The proper dating of economic expansions and recessions should therefore result from a multivariate approach which takes into account this fact. For the sake of expositional simplicity and in order to assess the occurrence of turning points on a monthly basis and thus in a more timely fashion, however, in this paper a univariate business cycle dating approach is preferred, with the index of industrial production as the business cycle reference series, as also done in a large number of previous studies, see e.g. Anas, Billio, Ferrara and Mazzi (2008) and Darné and Ferrara (2009). Indeed, as discussed in Fritsche and Stephan (2002, p.291), the use of the index of industrial production as a proxy for business cycle movements can be justified by the fact that industrial production is "much closer to the 'volatile' aggregates of GDP like investment and exports – which are at the heart of most business cycle theories". Furthermore, besides the fact that the index of industrial production is published on a monthly basis, what, as mentioned before, greatly enhances the timely account of the business cycle, this series is also less prone to revisions than the (quarterly) GDP figures.

[2]Theobald (2012) also considers a *Bayesian average* approach that is based on the correlations between the forecast errors. However, as found in many empirical studies before, e.g. see Clemen (1989) and Timmermann (2006), the results discussed in Theobald (2012) indicate that the simpler approaches neglecting the correlations even work better.

Specifically, in this paper the Bry and Boschan (1971) algorithm was employed, according to which a peak in the business cycle is identified when

$$\{y_{t-k} < y_t > y_{t+k}, \quad k = 1, \ldots, 5\}$$

while, analogously, a trough is assumed to take place when

$$\{y_{t-k} > y_t < y_{t+k}, \quad k = 1, \ldots, 5\},$$

where y_t is the two-month moving average of the German index of industrial production – the business cycle reference series.[3] Furthermore, as an additional censoring rule for the identification of recessionary periods and thus for the generation of the binary recession indicator series b_t, following Harding and Pagan (2002) a *triangle approximation to the cumulative movements* was pursued in order to measure the "severity" of an economic downturn j – and by extension the eventual occurrence of a recession –, defined as

$$S_j = 0.5 \times Deepness_j \times Duration_j,$$

where the *duration* is equal to the number of months between peak and trough (according to the NBER definition of a recession as a *significant decline in economic activity* [of] *more than a few months*), and the deepness is defined as the percentage decline,

$$Deepness_j = (y_p - y_t) / y_p,$$

where y_p and y_t are the respective values of the index of industrial production at the corresponding peak and trough, see Anas et al. (2008). A recessionary period was identified selectively when $S_j > 0.025$, as there is no consensus on the reference minimum duration and deepness of recessions (Darné and Ferrara, 2009, p.5). For instance this reflects a decline of 1% of the peak level of economic activity coinciding with a duration of at least 5 months.

Figure 1 illustrates the relationship between the underlying industrial production series and the resulting binary recession indicator series generated by the Bry and Boschan (1971) algorithm.

For the empirical analysis of this paper a variety of macroeconomic and financial variables were considered. Concerning the subset of variables which are supposed to reflect the real economy development, besides the index of industrial production, the following indicators were chosen: the open vacancies in the productive sector, the domestic and foreign orders received by the industrial sector, as well as the ifo business sentiment indicator (all variables as month-to-month % changes).

The financial indicators were selected as to represent a broad dimension of the financial markets. On the one hand, following Bernanke (1990) and Friedman and Kuttner (1992) the spread between average corporate bond yields of all maturities traded and the average yield of

[3]Given the high volatility of monthly data, it is usual in the turning points dating literature to "smooth" the underlying business cycle reference series among other things to avoid potential outliers biases.

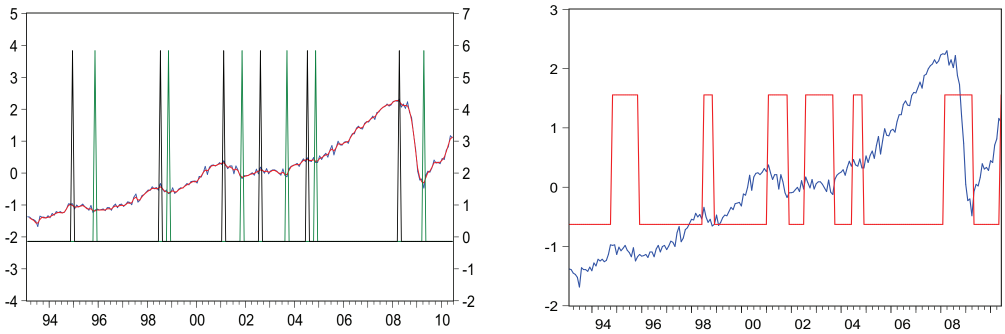


Figure 1: German industrial production, business cycle peaks and troughs calculated on the basis of the BB algorithm and related binary recession indicator. Normalized scaling method.

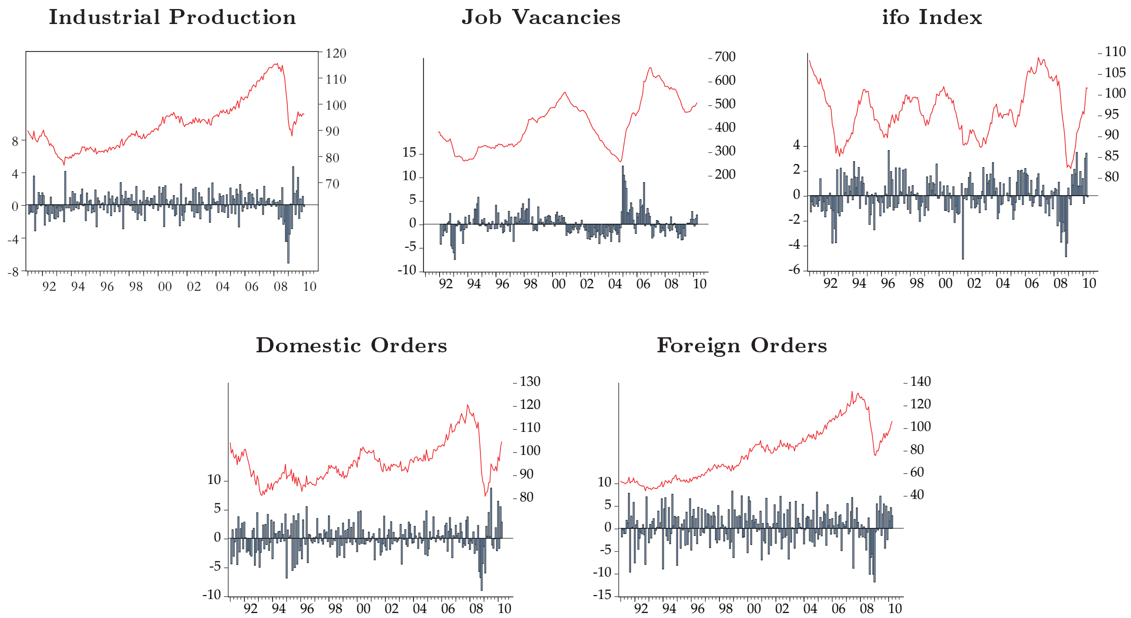


Figure 2: Macroeconomic Indicators: Industrial production index, job vacancies, foreign and domestic orders received by the productive sector, and ifo business sentiment index. Sources: Deutsche Bundesbank, DESTATIS, ifo Institute.

public securities was used, as well as the the growth rate of the CDAX price index in order to incorporate the German stock markets developments. Furthermore, along the lines of Stock and Watson (1989), Estrella and Hardouvelis (1991), Estrella and Mishkin (1998), and more recently Kauppi and Saikonnen (2008) and Nyberg (2010), the yield spread between the long-term and the short-term interest rate – was included in the general set of regressors. More specifically, alternative dynamic probit specifications using the 1-, 2-, 3-, 5- and 10-year yield

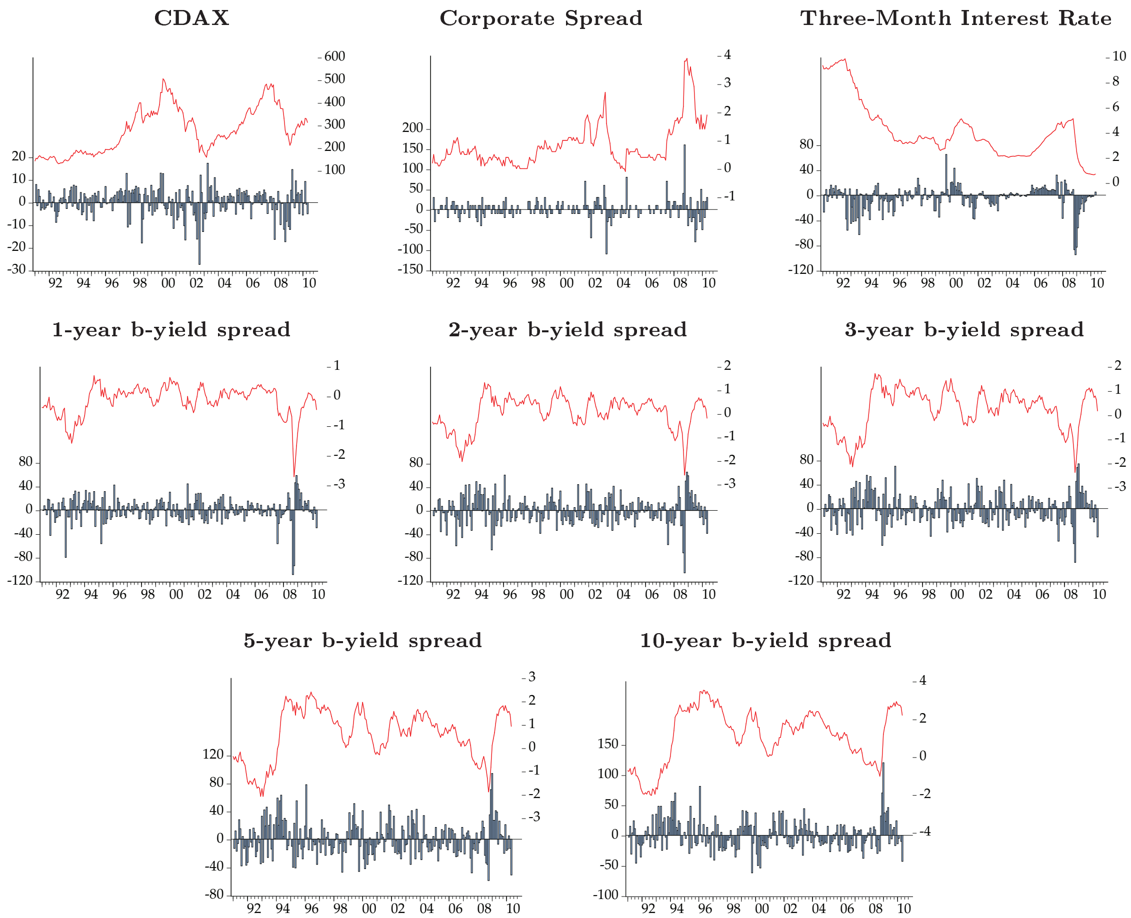


Figure 3: Financial Indicators: CDAX, Corporate Spread, Three-Month Euribor and yield spreads of different maturities. Source: Deutsche Bundesbank

(calculated by the Svensson's method) spreads to the three-month EURIBOR were estimated in order to address the uncertainty about which yield spread has the "best" prediction power, among other reasons which will be discussed in detail below. Furthermore, the short-term interest rate was also included in the set of regressors. In this respect, Ang, Piazzesi and Wei (2006) show, using a dynamic factor model, that the two principal factors of the term structure at all traded maturities, which in their study account for 90% of the variation of the whole term structure, are highly correlated with the short-term interest rate and the 10-year yield spread. Additionally, Wright (2006) shows that probit models with the yield spread of the 10-year T-bond to the three-month T-bill *and* the short-term three-month T-bill interest rate outperform probit specifications using only the yield spread in the MSE sense. Table 1 summarizes the descriptive statistics of the whole set of explanatory variables.

As previously discussed, in order to avoid eventual multicollinearity problems which may arise due to the strong correlation between the yield spreads of different maturities, as pre-

Table 1: Descriptive Statistics of Macroeconomic Indicators

Sample: 1991m1 – 2010m2, Obs: 217									
	Mean	Median	Max.	Min.	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	Prob.
IPIDX	92.38	92	115.3	76.4	9.74	0.69	2.75	17.89	0.00
JOBVAC	418.82	415	659	259	107.75	0.35	2.03	13.04	0.00
DOM. ORDERS	84.55	83.6	135.5	53.6	18.84	0.65	2.72	16.11	0.00
FOR. ORDERS	84.55	83.6	135.5	53.6	18.84	0.65	2.72	16.11	0.00
B-SPRD1Y	3.87	3.61	9.47	0.72	1.77	1.06	4.52	61.31	0.00
B-SPRD2Y	4.05	3.89	9.11	1.27	1.65	1.03	4.39	56.20	0.00
B-SPRD3Y	4.26	4.15	8.88	1.66	1.56	0.94	3.97	40.78	0.00
B-SPRD5Y	4.63	4.48	8.47	2.29	1.44	0.74	3.07	19.65	0.00
B-SPRD10Y	5.21	4.96	7.96	3.21	1.33	0.51	2.09	16.92	0.00
CDAX	84.06	75.69	06.08	31.29	98.53	0.33	2.06	11.89	0.00
CRP-SPRD	0.84	0.6	3.9	-0.1	0.76	1.91	7.12	285.24	0.00

viously mentioned, alternative dynamic probit models underlying the (invariant) set of explaining variables given by the industrial production index, the job vacancies, the ifo business sentiment index, the CDAX price index (all in % month-to-month changes), the corporate spread and the three-month euribor, and alternatively the 1-, 2-, 3-, 5- and 10-year bond yield spreads (to the three-month euribor) were specified and estimated. In the following the empirical results of such an automatized model specification procedure are discussed.

4.2 In-Sample Evaluation

In the following we discuss the estimation results of the dynamic probit specifications obtained by the *general-to-specific* (denoted by a G) as well as the *specific-to-general* (denoted by a S) approaches for 1-, 2- and 3-month ahead forecasts for the estimation sample 1991:1–2010:5. It should be pointed out, however, that these estimation results just represent an arbitrary "snap-shot" of the performance of the composite indicator, as all regressions underlying it are re-estimated in each and every month based on the newly available information through the automatized real-time specification procedure previously discussed.[4] The estimation results are summarized in Tables 2 – 4.

A variety of issues are worth to be highlighted. In the first place, at a more general level, the heterogeneity of the dynamic probit model estimations at all three analyzed forecast horizons corroborates the combinatorial approach pursued in this paper. Indeed, as it can be clearly observed in Tables 2 – 4, the significance level of the majority of variables (lags) is affected by the specific yield spread included in the respective regression sets, on the one hand, as well as by the lag selection procedure (*general-to-specific* or *specific-to-general*), on the other. There is, however, a certain "constancy" in the significance level of some variables (lags), which depends on the underlying forecast horizon of the respective regressions. At

[4]Note that a complete analysis of the in-sample properties on the whole real-time path would go beyond the scope of this paper.

the one-month-ahead forecast horizon, for example, the third lag (relative to the end-point) of the job vacancies series and of the foreign orders, as well as the seventh lag of the CDAX monthly growth rate are significant across all probit specifications.

In the same sense, the ifo business sentiment index does not seem to have any statistical significance at the one-month ahead horizon when included among the sets of indicators employed in this paper, as well as the binary recession indicator series and the short-term interest rate. Furthermore, the inclusion of (the growth rate of) the business cycle reference series (the index of industrial production) seems to be valid from the statistical point of view, at least in some probit specifications. Also worth highlighting is the fact that the statistical significance of the different yield curves also seems to be affected by the variables (lag) selection procedures, as well as the corporate spread series and the series of the domestic orders received by the industrial sector.

In contrast to the one-month-ahead forecast specifications, in the two-month-ahead forecast regressions the binary recession indicator series (at the ninth lag) is statistically significant in all specifications, as well as (various lags of) the CDAX price index. In contrast, while both job vacancies and foreign orders (the variables with the highest statistical “constancy” in the previous case) does not seem to have any predictive power at the two-month-ahead forecast horizon, the opposite seems to hold for the corporate spread, which coefficients are statistically significant in eight of ten specifications. Also interesting is the fact that, in contrast to the previous case summarized in Table 2, in the two-month-ahead forecast regressions the ifo business sentiment index is statistically significant on four out of ten specifications.

Finally, the most remarkable fact concerning the estimation results of the three-month-ahead forecast specifications is that in only one out of ten specifications the yield spread (specifically, the spread of the 5-year federal security rate to the three-month euribor) seems to have a statistically significant predictive power for recessions.

When compared with the outcomes of previous related empirical studies, among the just discussed estimation results a particularly interesting one is not only the corroboration of the predictive power of stock price developments respecting future economic activity already documented by Harvey (1989), Stock and Watson (1999), and recently by Haltmeier (2008), but also that in contrast the predictive power of the yield spread (irrespective the underlying maturity) does not seem to be as statistically significant as commonly thought.

Table 2: Summary of Dynamic Probit Regressions, One-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YG	EQ.B-SPRD1YS	EQ.B-SPRD2YG	EQ.B-SPRD2YS	EQ.B-SPRD3YG	EQ.B-SPRD3YS	EQ.B-SPRD5YG	EQ.B-SPRD5YS	EQ.B-SPRD10YG	EQ.B-SPRD10YS
RECES_IND	-	-	-	-	-	-	-	-	-	-
IPIDX	-	3, 4, 5	-	3, 4	-	3, 4	-	-	-	3, 4, 5, 6
DOM.ORDERS	-	3, 4, 5	-	7	-	-	-	-	-	3, 4, 5
for.orders	3, 4	3	3	3, 4, 5	3	3, 4	3	3, 4	3	3, 4
JOB VAC	3	3	3	3	3	3	3	3	3	3
IFO-IDX	-	-	-	-	-	-	-	-	-	-
CRP_SPRD	-	-	3	-	3	7	3	-	-	-
CDAX	7	4, 6, 7	3, 4, 6, 7	7	3, 4, 6, 7	3, 6, 7	3, 4,6	3, 6, 7	4, 6, 7	4, 6, 7, 8
EURIBOR 3M	-	-	-	-	-	-	-	-	-	-
B-SPRD1Y	3, 4	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	3	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	3	3	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	3	-	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	3	8
Pseudo-R^2	0.291	0.345	0.304	0.284	0.308	0.309	0.3148	0.279	0.303	0.399
SSR	30.993	28.750	29.716	30.274	29.571	29.834	29.345	30.701	30.390	26.085
Avg. Log-Likelihood	-0.441	-0.407	-0.432	-0.444	-0.430	-0.429	-0.426	-0.448	-0.432	-0.373
AIC	0.945	0.924	0.946	0.971	0.942	0.958	0.934	0.960	0.929	0.892
SC	1.054	1.109	1.085	1.110	1.081	1.128	1.072	1.068	1.037	1.139
HQC	0.989	0.999	1.002	1.027	0.998	1.027	0.990	1.004	0.973	0.992

Table 3: Summary of Dynamic Probit Regressions, Two-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YG	EQ.B-SPRD1YS	EQ.B-SPRD2YG	EQ.B-SPRD2YS	EQ.B-SPRD3YG	EQ.B-SPRD3YS	EQ.B-SPRD5YG	EQ.B-SPRD5YS	EQ.B-SPRD10YG	EQ.B-SPRD10YS
RECES_IND	9	9	9	9	9	9	9	9	9	9
IPIDX	-	-	-	-	-	-	-	-	-	-
DOM.ORDERS	4, 5	-	4, 5	-	4, 5	-	4, 5	-	-	7
FOR.ORDERS	-	-	-	-	-	-	-	-	-	-
JOB VAC.	-	-	-	-	-	-	-	-	-	-
IFO-IDX	-	4, 5	-	5	-	-	-	-	4, 5	4, 5
CRP_SPRD	4	7, 9	-	9	4	9	7, 9	9	7, 9	-
CDAX	4,. . . ,9	4, 6,. . . ,9	6,. . . ,9	4,. . . ,9	4,. . . ,9	6,. . . ,9	4, 6,. . . ,9	4, 6, 7	4, 6,. . . ,9	4, 7, 9
EURIBOR 3M	-	-	4	-	-	-	-	9	-	9
B-SPRD1Y	4	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	-	8	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	4	8, 9	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	4	-
Pseudo-R^2	0.296	0.301	0.343	0.284	0.298	0.348	0.311	0.358	0.335	0.348
SSR	30.559	31.645	29.253	30.274	30.650	28.314	30.425	28.938	29.654	28.670
Avg. Log-Likelihood	-0.433	-0.429	-0.403	-0.444	-0.431	-0.398	-0.423	-0.393	-0.408	-0.400
AIC	0.972	0.955	0.887	0.971	0.969	0.876	0.953	0.865	0.921	0.879
SC	1.153	1.121	1.023	1.110	1.151	1.011	1.135	1.000	1.103	1.015
HQC	1.045	1.022	0.942	1.027	1.043	0.930	1.027	0.920	0.995	0.934

Table 4: Summary of Dynamic Probit Regressions, Three-Month Forecast Horizon

	Sample: 1991:1 – 2010:5									
	EQ.B-SPRD1YG	EQ.B-SPRD1YS	EQ.B-SPRD2YG	EQ.B-SPRD2YS	EQ.B-SPRD3YG	EQ.B-SPRD3YS	EQ.B-SPRD5YG	EQ.B-SPRD5YS	EQ.B-SPRD10YG	EQ.B-SPRD10YS
RECES_IND	8	9	8	10	8	10	8	10	8	10
IPIDX	-	-	-	-	-	-	-	-	-	-
DOM.ORDERS	9, 10	-	7,. . . ,10	-	7,. . . ,10	-	7,. . . ,10	-	7,. . . ,10	-
FOR.ORDERS	-	-	7,. . . ,10	-	7,. . . ,10	-	7,. . . ,10	-	7,. . . ,10	-
JOB VAC	-	-	-	-	-	-	-	-	-	-
IFO-IDX	-	5	-	5	-	5	-	5	-	5
CRP-SPRD	7, 9	7, 10	9	-	9	7	9	10	9	10
CDAX	5,. . . ,10	6,. . . ,10	5,. . . ,8	6, 7, 10	5,. . . ,8	6, 7, 10	5,. . . ,8	10	5,. . . ,8	10
EURIBOR 3M	5	10	5	10	5	10	5	10	5	10
B-SPRD1Y	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD2Y	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B-SPRD3Y	n.a.	n.a.	n.a.	n.a.	-	-	n.a.	n.a.	n.a.	n.a.
B-SPRD5Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	8	n.a.	n.a.
B-SPRD10Y	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	-	-
Pseudo-R^2	0.356	0.366	0.362	0.328	0.362	0.328	0.362	0.351	0.362	0.327
SSR	26.757	26.996	28.783	29.314	28.783	29.314	29.345	28.160	28.783	29.530
Avg. Log-Likelihood	-0.395	-0.388	-0.391	-0.412	-0.391	-0.412	-0.391	-0.398	-0.391	-0.412
AIC	0.905	0.874	0.923	0.886	0.923	0.886	0.923	0.858	0.923	0.878
SC	1.101	1.040	1.165	0.991	1.165	0.991	1.165	0.964	1.165	0.968
HQC	0.984	0.941	1.021	0.928	1.021	0.928	1.021	0.901	1.021	0.914

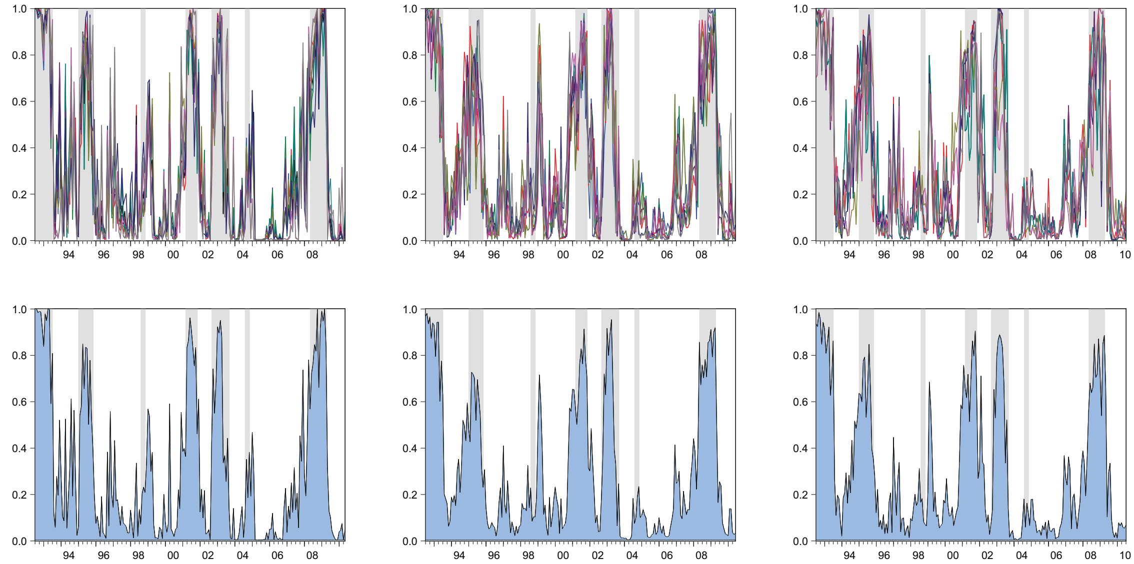


Figure 4: In-Sample Fit of Estimated and Average Recession Probabilities, One- Two- and Three-Month Forecast Horizon

Let us know focus on the advantage of the combination of the different estimated probabilities at the one- two- and three-month-ahead forecast horizon illustrated in Figure 4.

As it is clearly observable, while by and large the estimated recession probabilities of all probit specifications feature a similar pattern, there are some periods where the range of estimated probabilities becomes particularly high. This is especially important in middle ranges of the interval [0 – 1], where the signal threshold of a recession might be set.

In order to assess in a more formal manner the capability of the probit regressions to deliver accurate signals for the occurrence of a recession, the percentage of Type I and Type II errors for a success cut-off value of 0.5 are computed, see table 5. This indirectly highlights the value-added of combining the estimated probabilities of the alternative probit specifications since the specifications cover a certain range, which will even dynamically change over time.

Indeed, as the summary statistics in Table 5 clearly show, the accuracy in predicting especially the recessionary periods vary from correctly predicting 36 out of 69 recessionary periods (52.17 %) by EQ.B-SPRD1YG to 47 out of 69 (68.12 %) by EQ.B-SPRD5YG. Furthermore, it is also interesting to note that the forecast accuracy in predicting recessions of the different probit specifications varies across the forecast horizon: At the one-month forecast horizon EQ.B-SPRD5YG has the highest forecast accuracy, at the two-month the specification with the best performance is EQ.B-SPRD3YS, and at the three-month horizon EQ.B-SPRD1YG and EQ.B-SPRD1YS deliver best values.

Table 5: Expectation-Prediction Evaluations of Probit Regressions, Sample: 1991:1–2010:5

	Dep=0	Correct	% Correct	% Incorrect	Dep=1	Correct	% Correct	% Incorrect
One-Month-Ahead Forecast Horizon								
EQ.B-SPRD1YG	151	138	91.39	8.61	69	36	52.17	47.83
EQ.B-SPRD1YS	151	139	92.05	7.95	69	45	65.22	34.78
EQ.B-SPRD2YG	151	141	93.38	6.62	69	43	62.32	37.68
EQ.B-SPRD2YS	151	139	92.05	7.95	69	38	55.07	44.93
EQ.B-SPRD3YG	151	141	93.38	6.62	69	45	65.22	34.78
EQ.B-SPRD3YS	151	139	92.05	7.95	69	42	60.87	39.13
EQ.B-SPRD5YG	151	138	91.39	8.61	69	47	68.12	31.88
EQ.B-SPRD5YS	151	141	93.38	6.62	69	36	52.47	47.83
EQ.B-SPRD10YG	151	136	90.07	9.93	69	45	65.22	34.78
EQ.B-SPRD10YG	151	135	89.40	10.60	69	46	66.67	33.33
Two-Month-Ahead Forecast Horizon								
	Dep=0	Correct	% Correct	% Incorrect	Dep=1	Correct	% Correct	% Incorrect
EQ.B-SPRD1YG	157	145	92.36	7.64	69	40	57.97	42.03
EQ.B-SPRD1YS	158	143	90.51	9.49	69	36	52.17	47.83
EQ.B-SPRD2YG	157	144	91.72	8.28	69	42	60.87	39.13
EQ.B-SPRD2YS	158	143	90.51	9.49	69	40	57.97	42.03
EQ.B-SPRD3YG	157	144	91.72	8.28	69	41	59.42	40.58
EQ.B-SPRD3YS	160	144	90.00	10.00	69	47	68.12	83.41
EQ.B-SPRD5YG	157	143	91.08	8.92	69	39	56.52	43.48
EQ.B-SPRD5YS	159	141	88.68	11.32	69	45	65.22	34.78
EQ.B-SPRD10YG	158	144	91.14	8.86	69	41	59.42	40.58
EQ.B-SPRD10YS	158	144	91.14	8.86	69	42	60.87	39.13
Three-Month-Ahead Forecast Horizon								
	Dep=0	Correct	% Correct	% Incorrect	Dep=1	Correct	% Correct	% Incorrect
EQ.B-SPRD1YG	158	145	91.77	8.23	69	50	72.46	27.54
EQ.B-SPRD1YS	158	149	94.30	5.70	69	50	72.46	27.54
EQ.B-SPRD2YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD2YS	158	146	92.41	7.59	69	43	62.32	37.68
EQ.B-SPRD3YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD3YS	158	146	92.41	7.59	69	43	62.32	37.68
EQ.B-SPRD5YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD5YS	158	141	89.24	10.76	69	44	63.77	36.23
EQ.B-SPRD10YG	158	141	89.24	10.76	69	45	65.22	34.78
EQ.B-SPRD10YS	158	146	92.41	7.59	69	40	57.97	42.03

4.2.1 Out-of-Sample Evaluation

In order to assess the out-of-sample forecasting performance of the estimated probit models, following Moneta (2005), the out-of-sample recession probability forecasts were computed under real-time conditions by performing the following steps: First, the different probit regressions were estimated over the 1991:1 to 2007:8 period in order to have a good starting estimation of the parameters. Then, the probability of recession at a given month ahead was forecasted and its value recorded. After adding one more month to the revised estimation period and dynamically re-estimating each time the different probit regressions, the procedure was repeated. At the end a series of out-of-sample estimated probabilities over the publications 2007:9 to 2011:8 was obtained, while for each publication the 1-, 2- and 3-month-ahead forecasts were recorded.

To evaluate the out-of-sample forecasting performance of an estimated probit model $\mathcal{M}$, three common measures of forecast accuracy (see e.g. Rudebusch and Williams (2009)) were employed: the mean absolute error (MAE)

$$MAE(\mathcal{M}, h) = \frac{1}{T}\sum_{t=1}^{T} |P_{t|t-h}^{\mathcal{M}} - b_t|,$$

the root mean squared error (RMSE)

$$RMSE(\mathcal{M}, h) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \left(P_{t|t-h}^{\mathcal{M}} - b_t\right)^2},$$

and the Theil Inequality Coefficient

$$Theil = \frac{\sqrt{\sum_{t=T+1}^{T+h}(P_{t|t-h}^{\mathcal{M}} - b_t)^2/h}}{\sqrt{\sum_{t=T+1}^{T+h}(P_{t|t-h}^{\mathcal{M}})^2/h} + \sqrt{\sum_{t=T+1}^{T+h} b_t^2/h}},$$

which, as it is widely known, lies in the interval $[0, 1]$, where 0 represents a perfect fit and 1 no explanation whatsoever.

As Table 6 clearly summarizes, the estimated probability series resulting from the combination of the underlying forecasts seem to deliver not only statistically meaningful results and significant predictive power, but also feature good out-of-sample properties.[5]

Figure (5) and table (6) show the real-time out-of-sample performance of the combination schemes. For the graphical comparison a non parametric dating based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002) is used as an ex-post benchmark method, see section 4.1. This leads to a dating of the financial crisis between 2008M03 and 2009M04 (grey area). In addition this algorithm works behind the dependent variable of the probit model since the industrial production has to be transferred to a binary reference series.

Note that the benchmark method can only decide several months after the first publication if a recession has started, where as the different lines in figure (5) represent forecasts all generated by the probit at the date of the first publication.

When comparing the different aggregation methods, a large congruency between the simple and the weighted average approach can be realized. In particular the time of the recession signal is for all horizons the same, if this signal is based on a recession probability above 50%, see table (6). But with the help of the measures of forecast accuracy provided in table (6) it can be seen that for two of the three horizons the intuitive average delivers better results. This also becomes obvious when looking at the outliers in figure (6), which correspond to months,

[5]Theobald (2012) shows that there is enough variation among the dynamic regression variables to justify a combinatorial approach from the statistical perspective.

Table 6: Statistical Evaluation Measures for Combined Real-Time Recession Probability Forecasts (Starting estimation: 1991:1 – 2007:8, real-time out-of-sample path: 2007:9 – 2011:8)

Combination	Horizon	MAE	RMSE	Theil	Time of Signal$_{\geq 0.5}$
	1M	0.1302	0.2315	0.2036	2008M4
simple average	2M	0.1363	0.2469	0.2158	2008M5
	3M	0.1296	0.2429	0.2163	2008M5
	1M	0.1429	0.2496	0.2193	2008M4
weighted average	2M	0.1296	0.2346	0.2080	2008M5
	3M	0.1289	0.2223	0.2028	2008M5

where the probit forecasts exceed 50% recession probability although the benchmark method later on does not recognize a recession here (and vice versa). Both the number of these outliers (simple 2, weighted 1) and the level of them (simple in part over 90%, weighted around 60%) advise a policy maker to prefer the weighted average approach. This confirms the benefits from a combination of real-time predictions with different forecast horizons, while aiming at the same future month. The reason for such benefits is the fact that forecasts based on the most recent, but also most uncertain information (zero revisions as the youngest considered data) can be stabilized by ones based on less recent, but also less uncertain information (first revisions as the youngest considered data).

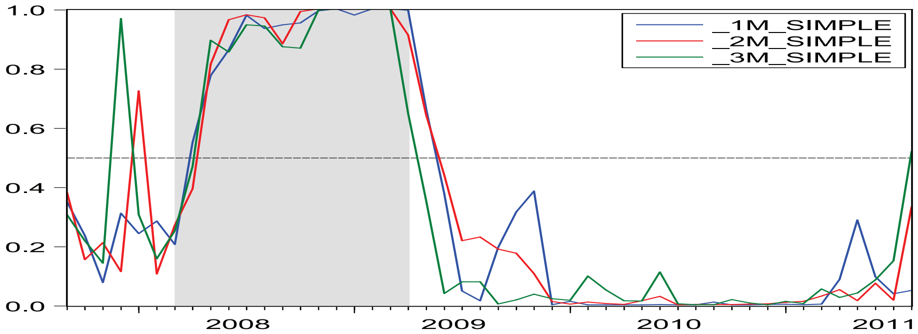
5 Concluding Remarks

As previously pointed out, the timely and accurate recognition of turning points in the business cycle is one of the most important, but also one of the most difficult tasks in macroeconomic forecasting. As a contribution along these lines of research, in this paper a practical econometric approach to forecast recessions under real time conditions based on the combination of alternative dynamic probit regressions was presented.

While for each of the underlying regressions the presented automatized variable and lag selection procedure is employed, they all differ either in the yield spread potentially included as a regressor or in the order of the lag selection or in the length of the forecast horizon (and implicitly in the grade of data revision that is used). It can be shown for all these generators - yield spreads, specification order and forecast horizons - that the partial independent information sets in the sense of Bates and Granger (1969) lead to sufficiently different underlying forecasts, which justifies combining them. Moreover this is the condition for benefiting from the combination according to a portfolio diversification argument, see Timmermann (2006).

This paper also reviews the real-time out-of-sample performance of different combination schemes for business cycle predictions with a dynamic probit model. Although we do not provide an intensive analysis of the total space of combination schemes and thus cannot explicitly determine the efficient frontier, all of the considered schemes reveal a minimum size of forecast accuracy (Theil coefficient $<$ 0.22).

Simple Averaging



Weighted Averaging

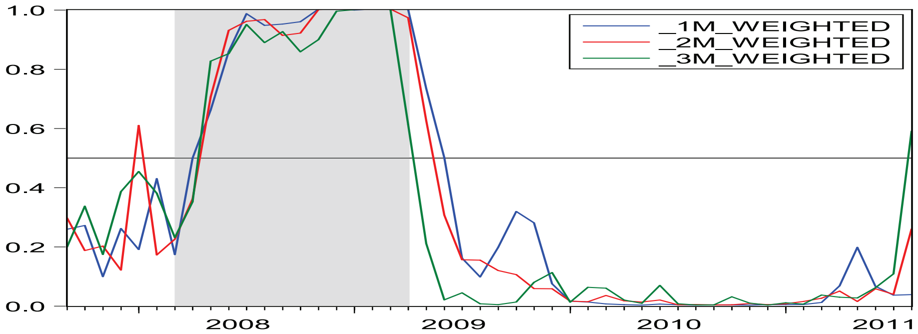


Figure 5: Real-time recession probabilities. The time axis is linked to the publications between 2007M09 and 2011M08, which means that the last observation of an involved series is given for the date of publication minus the data availability lag. The different lines represent the forecast horizons starting from the date of publication.

References

Anas, J., Billio, M., Ferrara, L. and Mazzi, G. L. (2008), 'A system for dating and detecting turning points in the euro area', *The Manchester School* **76**, 549 – 577.

Ang, A., Piazzesi, M. and Wei, M. (2006), 'What does the yield curve tell us about gdp growth?', *Journal of Econometrics* **131**, 359 – 403.

Bates, J. M. and Granger, C. W. J. (1969), 'The Combination of Forecasts', *Operations Research Quarterly* **20**, 451–468.

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

Bernanke, B. (1990), 'On the predictive power of interest rates and interest rate spreads', *New England Economic Review* **Nov./Dec.**, 51 – 68.

Bry, G. and Boschan, C. (1971), *Cyclical Analysis of Times Series: Selected Procedures and Computer Programs*, NBER.

Burns, A. F. and Mitchell, W. C. (1946), *Measuring Business Cycles*, NBER.

Campos, J., Ericsson, N. R. and Hendry, D. (2005), General-to-specific modeling: An overview and selected bibliography, *in* 'General-to-specific Modeling', Edwar-Elgar.

Clemen, R. T. (1989), 'Combining forecasts: A review and annotated bibliography', *International Journal of Forecasting* **5**, 559–581.

Darné, O. and Ferrara, L. (2009), Identification of slowdowns and accelerations for the Euro Area economy, Working Paper 42/2009, CEPR/EABCN.

Dueker, M. (1997), Strengthening the case for the yield curve as a predictor of U.S. recessions, Review, Federal Reserve Bank of St. Louis.

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

Estrella, A. and Mishkin, F. S. (1998), 'Predicting U.S. recessions: Financial variables as leading indicators', *Review of Economics and Statistics* **80**(1), 45 – 61.

Estrella, A., Rodriguez, A. P. and 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.

Friedman, B. M. and Kuttner, K. N. (1992), 'Money, income and prices, and interest rates', *American Economic Review* **82**(3), 472 – 492.

Fritsche, U. and Stephan, S. (2002), 'Leading indicators of German business cycles. an assessment of properties', *Jahrbücher für Nationalökonomie und Statistik* **222**, 289 – 315.

Haltmeier, J. (2008), Predicting cycles in economic activity, International Finance Discussion Papers 926, Board of Governors of the Federal Reserve System, Washington D.C.

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

Harvey, C. (1989), 'Forecasts of economic growth from the bond and stock markets', *Financial Analysts Journal* pp. 38 – 45.

Kauppi, H. and Saikonnen, P. (2008), 'Predicting U.S. recessions with dynamic binary response models', *Review of Economics and Statistics* **90**(4), 777 – 791.

Moneta, F. (2005), 'Does the yield spread predict recessions in the euro area', *International Finance* **8**(2), 263 – 301.

Nyberg, H. (2010), 'Dynamic probit models and financial variables in recession forecasting', *Journal of Forecasting* **29**(1-2), 215–230.

Pesaran, H. and Timmermann, A. (2005), 'Real time econometrics', *Econometric Theory* **21**, 212–231.

Proaño, C. (2010), Recession forecasting with dynamic probit models under real time conditions, IMK Working Paper 10-2010, IMK at the Hans Boeckler Foundation, Macroeconomic Policy Institute.

Rudebusch, G. D. and Williams, J. C. (2009), 'Forecasting recessions: The puzzle of the enduring power of the yield curve', *Journal of Business & Economic Statistics* **27**(4), 492 – 503.

Stock, J. and Watson, J. (1999), Business cycle fluctuations in U.S. macroeconomic time series, *in* J. B. Taylor and M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1A, Elsevier, pp. 3–64.

Stock, J. and Watson, M. (1989), New indexes of coincident and leading economic indicators, *in* O. Blanchard and S. Fischer, eds, 'NBER Macroeconomics Annual', Vol. 4, The MIT Press, Cambridge, MA.

Theobald, T. (2012), Combining recession probability forecasts from a dynamic probit indicator, Working Paper 89, Macroeconomic Policy Institute (IMK), Düsseldorf, Germany.

Timmermann, A. (2006), Forecast combinations, *in* 'Handbook of Economic Forecasting', North-Holland, Amsterdam, pp. 135 – 196.

Wright, J. H. (2006), The yield curve and predicting recessions, Technical report, Division of Monetary Affairs, Federal Reserve Board.