IMK
Institut für Makroökonomie
und Konjunkturforschung
Macroeconomic Policy Institute

Januar 2012

89

Working Paper

Thomas Theobald

Combining Recession Probability Forecasts from a Dynamic Probit Indicator

Abstract
This paper analyzes the real-time out-of-sample performance of three kinds of combination schemes. While for each the set of underlying forecasts is slightly modified, all of them are real-time recession probability forecasts generated by a dynamic probit indicator. Among the considered aggregations the most efficient turns out to be one that neglects the correlations between the forecast errors.

Keywords: dynamic probit model, out-of-sample forecasting, real-time econometrics, forecast combinations, correlation matrix, Bayesian average

JEL CLASSIFICATION SYSTEM: C11, C53, E37

January 10, 2012

1 Introduction

Imagine an economic forecasting problem, for which alternative predictions are available. This is, when the question arises, if and how alternative forecasts should be combined to a consensus. Among the first who emphasized the advantages of a combined solution Bates and Granger (1969) have to be mentioned. They discussed that including the ex-ante worser forecast, e.g. measured by the variance of the forecast error, leads to an improvement, if this one contains some independent information contrary to the ex-ante best one. This information for instance can stem from the consideration of other explanatory variables. More recently Timmermann (2006) analyzed the usefulness of forecast combinations and listed four possible reasons: 1. diversification, 2. structural breaks, 3. misspecification in the individual forecasts, 4. systematic differences in the individual loss functions. Among these reasons the presented paper is particularly linked to the first, e.g. by taking into account at least one aggregation method that tries to cover the correlation between the forecast errors. These are needed for the computation of the variance of the consensus error, where in the broadest sense, when following a diversification argument, the consensus forecast represents an equivalent to the market portfolio in classic portfolio theory.

This paper deals with the real-time prediction of business cycle phases. An appropriate estimation procedure is a dynamic probit model, where a binary reference series is regressed on leading indicators as well as on two kinds of autoregressive series in order to determine future recession probabilities. If such a probability exceeds a certain threshold, a turning point of the business cycle will be declared. Such models were estimated by Nyberg (2010) as well as Proaño (2010) [1]. The presented paper uses a modification of the latter model and considers explicitly the usefulness of forecast combinations.

The selection of the presented combination schemes takes into account that because of the time-consuming real-time computation for the underlying forecasts a selected combination scheme should be easy to implement and should not lengthen the running time too much. Consequently, approaches like Switching Regression and Smooth Transition Regression by Deutsch, Granger and Teräsvirta (1994) are excluded, where a single forecast is used as a regressor. Thus in real-time the additional estimations would take a lot of

[1]E-mail: Thomas-Theobald@boeckler.de. I would like to thank Christian Proaño for an inspiring introduction to dynamic probit indicators as well as the whole IMK research team for many helpful comments and suggestions.

running time. In contrast a Bayesian approach along the lines of Winkler (1981) as well as Palm and Zellner (1992) leads to a relatively quick computation after one has clarified that it is possible to apply it to the latent variable of the probit model.

This paper is structured as follows. Section 2 derives the combination schemes theoretically and explains how many underlying probit forecasts are to be considered for each. Section 3 deals with the so called *generators* of the underlying forecasts and shows how another value of a generator changes the recession probability forecast and the explanatory contributions of the selected leading indicators. Section 4 finally reviews the real-time out-of-sample performance of the different combination schemes. Section 5 concludes.

2 Combination Schemes

Whenever a probit model is used for business cycle predictions the binary state variable has to be defined, which here shall be done by

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

Based on current literature, e.g. Fritsche and Stephan (2002), the industrial production y represents the best proxy for monthly German economic activity, while leading indicators can be collected in $\mathbf{x}$. One should keep in mind to separate the reference series y from the latent variable of the probit model denoted by φ. Although a heuristic connection could be established between y and φ, the latter one is really unobservable and just comes from the fact that by definition

$$b_t = \begin{cases} 1 & : \quad \varphi_t > 0 \\ 0 & : \quad \varphi_t \leq 0. \end{cases} \tag{2}$$

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

$$\begin{aligned} \varphi_t = \sum_{j=h+R}^{o} \delta_j b_{t-j} + \sum_{j=h+D_y}^{p} \alpha_j y_{t-j} + \sum_{j=h+D_{\mathbf{x}}}^{q} \mathbf{x}'_{t-j}\beta_j + u_t, \\ u_t \sim N(0,1) \quad \forall t, \quad R > D_y, \end{aligned} \tag{3}$$

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 (3) the autoregressive explanatory terms consist of both growth rates of the reference series y and the binary state series b, which corresponds to what Kauppi and Saikkonen (2008) call the 'dynamic autoregressive' specification. The binary series reflects the recession and expansion phases of the reference series respectively. Moreover $\mathbf{x}$ captures a large set of exogenous macroeconomic, survey and financial indicators. Among these there are different yield spreads ranging from short-term to long-term maturity, which if all included in an equation of type (3) would automatically induce high multicollinearity. On the other hand all remaining explanatory variables should be included before a Likelihood Ratio (LR) test is used for model specification in order to avoid an omitted variable bias. Considering the large data set, it is obvious that the model selection strategy (from general to specific or the other way round) matters, where it shall be bounded by a maximum lag of five per each of the leading indicators. We therefore recommend to apply the LR test based on a general-to-specific (G) and on a specific-to-general (S) approach. Thus, in order to consider several maturities of the yield curve and to use both model selection procedures, pooling the forecasts of different specifications of (3) seems to be appealing.

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

$$\begin{aligned} \varphi_{t+h}^i = \mathbf{z}_t'^i \beta + u_{t+h}^i, \quad u_{t+h}^i \sim N(0,1), \quad i \in I, \\ b_{t+h}^i = \begin{cases} 1 & : \quad \varphi_{t+h}^i > 0 \\ 0 & : \quad \varphi_{t+h}^i \leq 0 \end{cases} \end{aligned} \tag{4}$$

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

$$\begin{aligned} E\left(b_{t+h}^i | \mathbf{z}_t^i, \beta\right) = \mu_{t+h|t}^i = P\left(b_{t+h}^i = 1 | \mathbf{z}_t^i, \beta\right) \\ = \Phi\left(\mathbf{z}_t'^i \beta\right) = \Phi\left(E\left(\varphi_{t+h|t}^i\right)\right). \end{aligned} \tag{5}$$

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. The next section motivates this in detail and adds another component, which will be predictions of the same future value from different forecast horizons.

This paper deals with three different kinds of combination schemes, where for each a short derivation is provided with I being slightly modified. Let $\mu_{t+h|t} = \left(\mu_{t+h|t}^1, \ldots, \mu_{t+h|t}^{|I|}\right)'$ denote the vector of single forecasts and $\theta = C\left(\mu_{t+h|t}; w_c\right)$ the consensus forecast aggregating the underlying forecasts by means of the combination weights. Within such a setting

Timmermann (2006) analyzes that based on a straightforward diversification argument simple combinations of forecasts often lead to better results then the ex-ante best one. Thus the *first* case to be considered here will be the one of equal weights (simple average), i.e.

$$\theta = \frac{1}{|I|} \sum_{i=1}^{|I|} \mu^i_{t+h|t} \\ |I| = \#\{\text{interest rate spreads}\} \times \#\{G, S\}. \tag{6}$$

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. After studying the preservation properties of a linear pooling operator Genest and Zidek (1986) as well as Genest and Wagner (1987) proposed to introduce an additional probability distribution for the decision maker of the consensus and even extended it to a logarithmic form. One consequence of this is the possibility of single forecasts balancing each other by $w_i \in [-1, 1]$. At this point one could enter the stage of developing highly sophisticated pooling operators for the presented probit model. However, according to the appealing simplicity we trust the idea of preservation to some extent and center the consensus even more to the majority of single forecasts arising from different interest rate spreads and specification orders. But as a consequence of the analysis in the next section we will see that there is also a difference in forecast results when instead of using the most recent information the last available observation is omitted from the information set. In contrast to the estimation of the previous month with this estimation the revisions are taken into account. The motivation for such a procedure is to stabilize the forecast based on more recent data by one, which on the one hand is based on a longer horizon, but on the other on data that will be less revised. In a nutshell this means to add different forecast horizons for the same future value as an extra source of generating the underlying forecasts. Thus the *second* aggregation method to be considered here will be a two-stage procedure, i.e.

$$\theta^*_k = \sum_{i=1}^{|I^*|} \frac{\left(\sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|\right) - |\mu^{i,k}_{t+h|t} - \mu^{med}_{t+h|t}|}{(|I^*| - 1) \sum_{j=1}^{|I^*|} |\mu^{j,k}_{t+h|t} - \mu^{med}_{t+h|t}|}, \; i \in I^* \subset I, \; k \in I \backslash I^*, \tag{7}$$

$$\theta = \sum_{i=1}^{\#\{\text{horizons}\}} \frac{2^{h-1}}{\sum_{j=0}^{\#\{\text{horizons}\}-1} 2^j} \theta_h^*, \;\; h \in I \backslash I^*, \tag{8}$$
$$|I| = \#\{\text{interest rate spreads}\} \times \#\{G, S\} \times \#\{\text{horizons}\},$$

where *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 (6). Note again that taking into account the horizon as an additional generator of the underlying forecasts in (8) means nothing else than balancing the actual prediction by one which is generated with a longer horizon and thus considers less actual information. In such a constellation it 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 largest 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 *horizon average* approach.

Finally as a *third* aggregation method we consider a *Bayesian average* approach that is based on the correlations between the forecast errors. We first study the general form along the lines of derivation in Zellner (1971) as well as Palm and Zellner (1992) in order to apply it then to the forecasts provided by the dynamic probit indicator. Let us assume unbiased forecast errors observed for T periods in a d-dimensional combination space, i.e.

$$\mathbf{U} = (u_1, \ldots, u_T)' = (u_{ti}), \qquad t = 1, \ldots, T, \quad i = 1, \ldots, d, \quad \text{where}$$
$$\mathbf{u_t} = (u_{t1}, \ldots, u_{td})' \sim N(0, \mathbf{\Sigma}), \quad u_{ti} := \varphi_t - f_{ti}. \tag{9}$$

Here φ_t denotes the true value and f_{ti} the i-th forecast, which shall be summarized in the forecast vector $\mathbf{f}_t$. Then for instance one row of future forecast errors can be summarized in $\mathbf{W}$ which is $N(0, \mathbf{\Sigma})$ -distributed just as it was assumed for the rows of $\mathbf{U}$. In order to derive the predictive probability density function conditional on the observed forecast errors it is possible to integrate out the covariance matrix, i.e.

$$\begin{aligned} p(\mathbf{W}|\mathbf{U}) &= \int p\left(\mathbf{W}, \mathbf{\Sigma}^{-1}|\mathbf{U}\right) d\mathbf{\Sigma}^{-1} \\ &= \int p\left(\mathbf{\Sigma}^{-1}|\mathbf{U}\right) p\left(\mathbf{W}|\Sigma^{-1}, \mathbf{U}\right) d\mathbf{\Sigma}^{-1} \\ &= \int p\left(\mathbf{\Sigma}^{-1}|\mathbf{U}\right) p\left(\mathbf{W}|\Sigma^{-1}\right) d\mathbf{\Sigma}^{-1}, \end{aligned} \tag{10}$$

where the last equation reflects the assumption that given the covariance matrix future error terms do not depend on the past ones. Since the covariance matrix is assumed to be unknown, the reason why to turn to the inverse is that the Inverse-Wishart distribution can be shown to be a conjugate prior with respect to Gaussian data, e.g. see Sawyer (2007). This means nothing else than that the posterior distribution

$$p\left(\mathbf{\Sigma}^{-1}|\mathbf{U}\right) \propto p\left(\mathbf{\Sigma}^{-1}\right) p\left(\mathbf{U}|\mathbf{\Sigma}^{-1}\right) \tag{11}$$

and the prior distribution $p\left(\mathbf{\Sigma}^{-1}\right)$ are in the same family. According to the multivariate normal distribution the likelihood is given by

$$p\left(\mathbf{U}|\mathbf{\Sigma}^{-1}\right) \propto \det\left(\mathbf{\Sigma}\right)^{-T/2} \exp\left(-\text{trace}\left(\mathbf{U}'\mathbf{U}\mathbf{\Sigma}^{-1}\right)/2\right). \tag{12}$$

Thus $p\left(\mathbf{W}|\mathbf{U}\right)$ can be calculated by plugging in the likelihood and the probability density function of the Inverse-Wishart distribution in (11) and following (10). Then, after adding the forecast vector, $p\left(\varphi_{T+h}|\mathbf{f}_{T+h},\mathbf{U}\right)$ occurs as the marginal predictive probability density function of $p\left(\mathbf{W}|\mathbf{U}\right)$. Winkler (1981) obtained the result to be an univariate Student-t probability density function

$$\begin{aligned} p\left(\varphi_{T+h}|\mathbf{f}_{T+h},\mathbf{U}\right) &\propto \left(\frac{1+\left(\varphi_{T+h}-m\right)^2}{\left(T+d-1\right)s^2}\right)^{-(T+d)/2}, \\ m = \frac{\mathbf{1}'\mathbf{\Sigma}_0^{-1}\mathbf{f}_{T+h}}{\mathbf{1}'\mathbf{\Sigma}_0^{-1}\mathbf{1}}, &\quad s^2 = \frac{T+\left(m\mathbf{1}-\mathbf{f}_{T+h}\right)'\mathbf{\Sigma}_0^{-1}\mathbf{f}_{T+h}}{\left(T+d-1\right)\mathbf{1}'\mathbf{\Sigma}_0^{-1}\mathbf{1}}. \end{aligned} \tag{13}$$

Looking back at (4) it is clear enough that $\mathbf{u}_{t+h} = \left(u_{t+h}^1,\dots,u_{t+h}^d\right)$ will be $N\left(0,\mathbf{\Sigma}\right)$ - distributed with unknown covariance matrix $\mathbf{\Sigma}$. According to (5) it is obvious that when applying the Bayesian procedure to the latent variable of the probit model the forecast vector will be given by

$$\mathbf{f}_{t+h} = \left(E\left(\varphi_{t+h|t}^1\right),\dots,E\left(\varphi_{t+h|t}^d\right)\right). \tag{14}$$

Thus as a consensus forecast we suggest to consider the mean m from the distribution above, i.e.

$$\theta = \Phi\left(E\left(\varphi_{t+h}|\mathbf{f}_{t+h},\mathbf{U}\right)\right) = \Phi\left(m\right). \tag{15}$$

For calculating the mean in (13) it is necessary to know the elements of Σ_0^{-1}, which arises as a parameter matrix from the Inverse - Wishart prior distribution.[2] Winkler (1981)

[2]Using the spectral theorem of linear algebra the maximum likelihood estimator of the covariance matrix of a multivariate normal sample can be shown to be the normalized scatter matrix following a Whishart distribution. This matrix is hidden behind Σ_0; see also Sawyer (2007).

provides an illustration of his model, thereby suggesting that if no other information about Σ is available, the elements of Σ_0 should be the covariances between the residuals of the single forecast. Here a problem arises from the probit model since in the latent model the residuals cannot be calculated by the difference between actual and fitted values. One way to handle this case is to generate residuals that show similar properties to those of the familiar linear models. This was proposed by Gourieroux, Monfort, Renault and Trognon (1987). In particular their *generalized residuals* are derived by the condition that the resulting residuals are orthogonal to all of the explanatory variables, which leads to

$$\tilde{u}_i(\beta) = \frac{\phi(z_t'\beta)}{\Phi(z_t\beta)(1-\Phi(z_t'\beta))}\left(b_t - \Phi(z_t'\beta)\right), \tag{16}$$

where ϕ, Φ stand for the standard normal probability density function and cumulative distribution function. As always when using generalized residuals in combination with the probit model the missing normal distribution of the resulting residuals can be mentioned as a weak point, but together with the Bayesian approach this represents a computable way to consider the correlations between the forecast errors within this class of models.

Obviously another assumption for the Bayesian method is the non-singularity of $\mathbf{\Sigma}$. In order to avoid similar covariances for contiguous interest rate spreads we concentrate on the one and ten years maturity. This fits a result by Ang, Piazzesi and Wei (2006), who find the maximum among the maturity differences to work best in forecasting GDP growth. Thus one obtains $|I| = 8$ for the Bayesian average approach, where again

$$|I| = \#\{\text{interest rate spreads}\} \times \#\{G, S\} \times \#\{\text{horizons}\}. \tag{17}$$

3 Dimension of the Combinations

As mentioned in section 1 one of the reasons for combining forecasts follows a diversification argument from portfolio theory; see also Timmermann (2006). This requires keeping two conditions of an additional generator for the underlying forecasts: Firstly, the resulting forecasts should differ to some extent. Otherwise it is not worthwhile to combine since the information is already available, Clemen (1987). Secondly, the resulting forecasts should be near the efficient frontier. Whereas the first condition is trivial to check, the second is certainly not. At least standard specification tests applied to the probit model from section 2 suggest the presented forecasts, but this clearly does not mean that these are the efficient

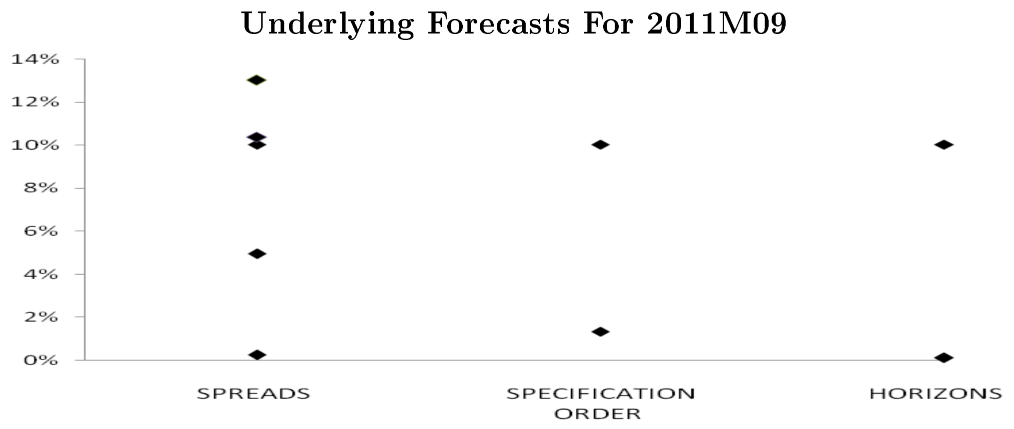


Figure 1: All underlying forecasts for 2011M09 are based on the publication in 2011M08. This in particular means that with the longer forecast horizon (2 months) the first revisions represent the last observations to be considered. As it can be seen the underlying forecasts differ in a certain range independent of looking at the specifications generated by different interest rate spreads, specification methods or forecast horizons.

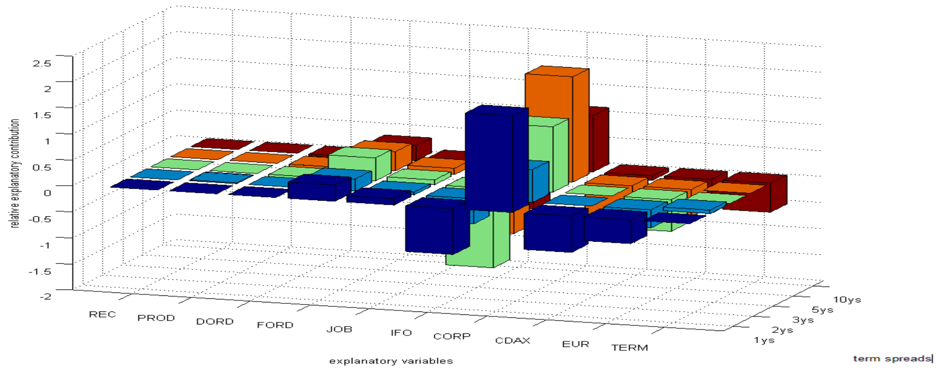
among all the possible ones. However, this section concentrates on showing that the forecasts from the above mentioned generators - interest rate spreads, specification order and forecast horizons - vary enough to justify a combination.

When looking at figure (1) it becomes obvious that the underlying forecasts differ in a certain range for all the selected generators. As always when dealing with real-time forecasts the illustration is linked to a certain publication - here 2011M08. At the time of work this also represents the last observation for (survey and financial) indicators which are not subject to a data availability lag. Thus forecasts for 2011M09 are not pseudo out-of-sample predictions. Of course this does not exclude that for other publications the dispersion can be higher or lower but a complete analysis would go beyond the scope of this paper.

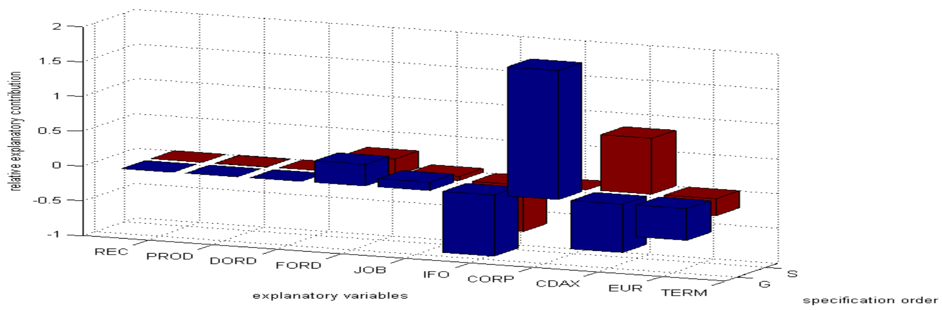
In this context an interesting question is how the influence of the explanatory variables change with changing values of the generator. Note that such changes can either stem from the fact that certain variables respectively their lags are excluded or from the fact that the coefficients and in the case of the horizons the indicator values change. After all these changes lead to the forecast differences presented in figure (1).

Figure (2) answers the question above. For each of the explanatory variables the relative contribution to the fit of the latent variable is provided, i.e. the product of the coefficient

Explanatory Contributions - Interest Rate Spreads



Explanatory Contributions - Specification Order



Explanatory Contributions - Forecast Horizons

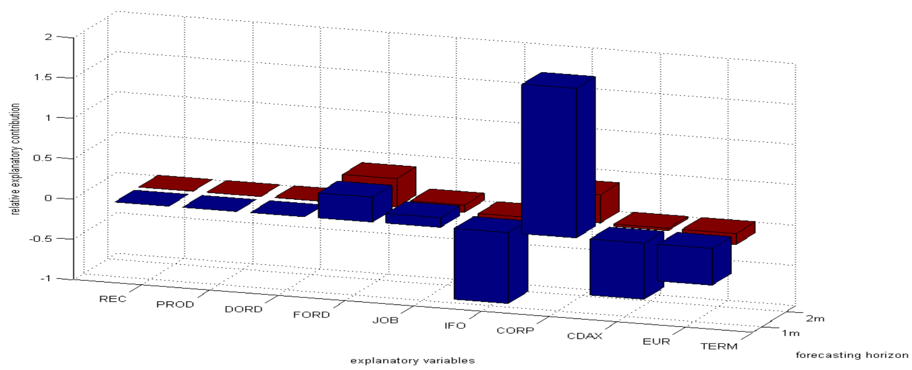


Figure 2: All (relative) explanatory contributions are ratios between the contribution of a certain explanatory variable to the latent fit and the latent fit itself. In each case - spreads, specification order or horizons - partial results differ significantly. This is why with two of the aggregation methods all three possible ways of generating the underlying forecasts are taken into account.

and indicator value for all the selected lags of a certain indicator divided by the sum of these terms over all of the indicators. As a third dimension the generator appears, i.e. different interest rate spreads (1,2,3,5,10 years term spread) in the upper sub-figure, the difference between a general-to-specific and a specific-to-general lag choice in the central sub-figure and the difference between an one month and a two month ahead forecast for 2011M09 (both based on publication 2011M08) in the lower sub-figure. The influence of the generator is analyzed individually. This means that in the case of the term spreads we stick to one month ahead general-to-specific forecasts. In the case of the specification order we only consider one month ahead forecasts, where the one year term spread can be selected as an explanatory variable. Finally in the case of the forecast horizons only general-to-specific forecasts with the one year term spread are considered. This explains the number of illustrated forecasts and represents a reasonable reduction of the maximum number of underlying forecasts ($|I| = 20$).

A more detailed description about the involved regressors is given by Proaño (2010), but for a complete comprehension of figure (2) we list them here. REC stands for lags of the binary recession indicator, PROD for lags of the industrial production as the underlying monthly reference series, DORD for domestic orders, FORD for foreign orders, JOB for job vacancies, IFO for the ifo business climate index, CORP for the credit spread between corporate and public issuer's current yield, CDAX for the corresponding stock price index, EUR for the 3 month euribor interest rate and TERM for one of the possible term spreads.

For the first generator, the term spreads, one finds high consistency with respect to the sign of the explanatory contribution, where a minus corresponds to a recession contribution and a plus to an expansion contribution. Nevertheless there are also significant differences, in particular the high dispersion of the contributions of ifo business climate index, the credit spread and the CDAX. Moreover the specification with the 10 years term spread is the only which finds a significant contribution from the term spread. The consistency with respect to the sign of the contribution changes when turning to the specification order. Here one finds a highly positive contribution of the credit spread, whereas with the converse specification order there is no contribution. The contribution of CDAX changes from positive to negative. In the case of the forecast horizon consistency is given again, but the contributions of ifo business climate index, the credit spread, the CDAX and the euribor interest rate are much higher when using the shorter forecast horizon. To sum up the underlying forecasts and explanatory contributions change enough to justify a combination when assuming that each of the generators provides additional information.

4 Real-time Combination Results

Figure (3) and table (1) show the real-time out-of-sample performance of the combination schemes. For the graphical comparison a non-parametric dating procedure based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002) is used as an ex-post benchmark. This approach 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 (3) represent forecasts which are all generated by the probit model at the date of the first publication.

When comparing the different aggregation methods, a large congruency between the simple and the horizon 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 (1). But with the help of the measures of forecast accuracy provided in table (1) it can be seen that for two of the three horizons the horizon average delivers better results. This also becomes obvious when looking at the outliers in figure (3), which correspond to months, where the probit forecasts exceed 50% recession probability although the benchmark method does not signal a recession here (and vice versa). Both the number of these outliers (simple 2, horizon 1, Bayesian 5) and their levels (simple and Bayesian in part over 90%, horizon around 60%) advise a policy maker to prefer the horizon average approach. This confirms that combining different forecasting horizons for the same future value stabilizes the predictions.

According to figure (3) as well as table (1) the Bayesian average performs worst among the considered combinations although from a theoretical point of view it seems to be privileged because of considering the correlation between the forecast errors. On the one hand this confirms the findings by Timmermann (2006) that in empirical studies 'simple combinations that ignore correlations between the forecast errors often dominate more refined combination schemes aimed at estimating the theoretically optimal combination weights' (p.1). On the other hand this constraints the appropriateness of Gourieroux et al. (1987)'s generalized residuals for appropriately capturing the correlation structure.

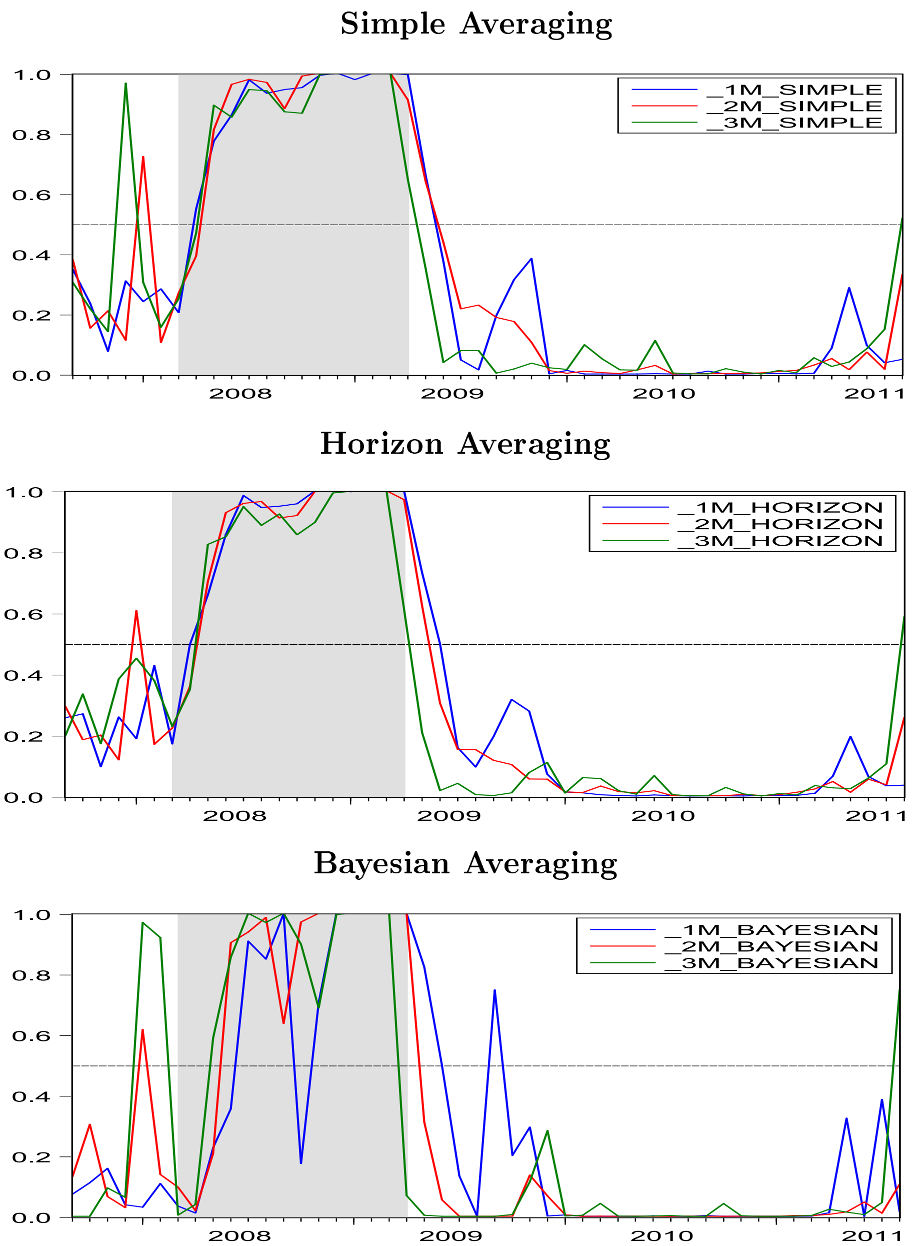


Figure 3: Real-time recession probabilities according to different combination schemes. The time axis is linked to the publications between 2007M09 and 2011M08, which means that the last observation of an involved series is given for the date of publication minus the data availability lag. The different lines represent the forecast horizons starting from the date of publication. A non-parametric benchmark method is used for ex-post dating of the business cycle turning points (grey area). In doing so seven months later the beginning and the end of a recession can be announced.

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
horizon average	2M	0.1296	0.2346	0.2080	2008M5
	3M	0.1289	0.2223	0.2028	2008M5
	1M	0.1942	0.3633	0.3376	2008M7
bayesian average	2M	0.1244	0.2790	0.2547	2008M6
	3M	0.1570	0.3485	0.3149	2008M5

Table 1: Measures of forecast accuracy and time of the recession signal for the different combination schemes. For one month ahead forecasts the simple average reaches best values, where as the horizon average does for the remaining horizons. Both of them lead to identical signaling times.

5 Conclusions

This paper reviews the real-time out-of-sample performance of different combination schemes - each pooling business cycle predictions generated by a dynamic probit model. It can be shown for all the considered generators - interest rate spreads, specification order and forecasting 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). 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 < 0.34).

While two of the considered aggregation methods neglect the correlations between the forecast errors, a Bayesian approach along the lines of Winkler (1981) as well as Palm and Zellner (1992) taking the correlation structure into account is shown to work on the basis of the underlying probit model. But as found in many empirical studies before, e.g. see

Clemen (1989) and Timmermann (2006), the results indicate that the simpler approaches neglecting the correlations even work better. Among these the one, which also covers different horizons as an additional generator of the underlying forecast, delivers the best values.

References

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.

Bry, G. and Boschan, C. (1971), *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, National Bureau of Economic Research, Inc.

Clemen, R. T. (1987), 'Combining overlapping information', *Management Science* **33**, 373–380.

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

Deutsch, M., Granger, C. and Teräsvirta, T. (1994), 'The combination of forecasts using changing weights', *International Journal of Forecasting* **10**, 47–57.

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

Genest, C. and Wagner, C. G. (1987), 'Further evidence against independence preservation in expert judgement synthesis', *Aequationes Mathematicae* **32**, 74–86.

Genest, C. and Zidek, J. (1986), 'Combining Probability Distributions: A Critique and an Annotated Bibliography', *Statistical Science* **1**(1), 114–148.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987), 'Generalised Residuals', *Journal of Econometrics* **34**, 5–32.

Harding, D. and Pagan, A. (2002), 'Dissecting the cycle: a methodological investigation', *Journal of Monetary Economics* **49**(2), 365–381.

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

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

Palm, F. and Zellner, A. (1992), 'To combine or not to combine ? Issues of combining forecasts', *Journal of Forecasting* **11**(8), 687–701.

Proaño, C. (2010), Recession forecasting with dynamic probit models under real time conditions, IMK Working Paper 10-2010.

Sawyer, S. (2007), Wishart Distributions and Inverse Wishart Sampling.
URL: *www.math.wustl.edu/ sawyer/hmhandouts/Whishart.pdf*

Timmermann, A. (2006), Forecast combinations, *in* 'Handbook of Economic Forecasting', Vol. 1, G. Elliot, C. Granger and A. Timmermann, eds., pp. 135–196. North-Holland, Amsterdam.

Winkler, R. L. (1981), 'Combining Probability Distributions from Dependent Information Sources', *Management Science* **27**(4), 479–488.

Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

Publisher: Hans-Böckler-Stiftung, Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
Phone: +49-211-7778-331, IMK@boeckler.de, http://www.imk-boeckler.de

IMK Working Paper is an online publication series available at:
http://www.boeckler.de/imk_5016.htm

ISSN: 1861-2199

Hans **Böckler Stiftung**

Fakten für eine faire Arbeitswelt.