

Theobald, Thomas

Conference Paper

Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Forecast

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Forecasting, No. E20-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Theobald, Thomas (2013) : Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Forecast, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Forecasting, No. E20-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/79911>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Markov Switching with Endogenous Number of Regimes and Leading Indicators in a Real-Time Business Cycle Forecast

February 25, 2013

Abstract

This paper uses several macroeconomic and financial indicators within a Markov Switching (MS) framework to predict the turning points of the business cycle. The presented model is applied to monthly German real-time data covering the recession and the recovery after the financial crisis. We show how to take advantage of combining single MSARX forecasts with the adjusting of the number of regimes on the real-time path, which both lead to higher forecast accuracy through the non-linearity of the underlying data-generating process. Adjusting the number of regimes implies distinguishing between recessions which are either normal or extraordinary, i.e. specifically determining as early as possible the point in time from which the recession in the aftermath of the financial crisis structurally exceeded previous ones. In fact it turns out that the Markov Switching model can signal quite early whether a conventional recession will occur or whether an economic downturn will be more pronounced.

Keywords: Markov Switching, Business Cycle, Leading Indicators, Macro-Forecast

JEL CLASSIFICATION SYSTEM: C24,C53,E37

1 Introduction

As a consequence of the financial crisis in the years 2008 and early 2009, Germany suffered its strongest decline in GDP since the global economic crisis of 1929. More than ever it seems worthwhile to attempt to predict the business cycle and its turning points in order to react properly to recessions through counter-cyclical economic policy. This paper delivers real-time predictions within a Markov Switching (MS) framework. In this class of models it is also possible to analyze the impact of additional regressors by Markov Switching Autoregressive Models with Exogenous Variables (MSARX). Among others Lee, Liang and Chou (2009) used such an approach for regressing a proxy variable of the real estate cycle on its lags as well as on a composite leading index. Still, most of the MS business cycle literature concentrates on purely autoregressive estimations following the famous Markov Switching Mean Model (MSM) by Hamilton (1989), or as stated by Boldin (1996): ‘Because the estimated parameters of relatively simple MSM specifications match many stylized facts about the business cycle, this framework has become an important alternative to linear, autoregressive structures.’ Contrary to the linear case, a straightforward set of specification tests for MS models - in particular covering a highly parameterized design - with results that clearly support the MSM, are not available. Although some test procedures have been developed, see Breunig, Najarian and Pagan (2003), it is difficult to align them with a more parameterized design. But in contrast to a purely autoregressive MSM, the inclusion of leading indicators as explanatory variables can be appealing due to promising additional information for policy makers. The paper at hand considers this additional information in univariate MS regressions with lagged dependent and lagged leading variables. In particular we draw conclusions about the significance and prediction performance of leading indicators according to their (real-time) characteristics. These characteristics are comprised of two dimensions: whether an indicator is subject to a publication lag or not, and whether the indicator belongs to the class of financial or real economy variables. Apart from such an analysis of the leading indicators, the forecast of business cycle turning points still represents the central feature of the model. With respect to this aim, different MSARX specifications allow us to take advantage of averaging effects and the fact that we endogenize the number of regimes depending on available real-time data. Indeed, it turns out that both of our extensions, the combination of different MSARX specifications as well as the real-time change of the number of regimes can significantly improve the levels of forecast accuracy.

The paper is structured as follows: Section 2 provides a short overview of the literature related to real-time business cycle forecasts. Section 3 reviews elements of the filter introduced by Hamilton (1989, 1990) in order to stress the fact that exogenous variables have a numerical impact on the state probabilities when using an expectation-maximization (EM) algorithm for the MS model. Section 4 describes data input and in-sample results of the model. With respect to both the dependent variable (reference series) and the independent variables (indicator series), real-time characteristics are provided. In subsection 4.2 an ex-post-dating algorithm - based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002) - is presented to serve as an evaluation tool for the MS results. Afterwards model specification is discussed in detail, illustrating how the structure of the model results from a compromise between information needs and available real-time data records. While the former suggest a super highly parameterized design, the latter clearly restricts the parameter space to some extent. Next, in-sample results for a certain publication are discussed. In the final part of section 4, we deal with goodness-of-fit measures in the case of more than two regimes. Based on these measures we assess if in general those leading indicators which are not subject to a publication lag perform better. Section 5 turns to the out-of-sample results. There we introduce the notion of how to change the number of regimes in real-time. Then we analyze the sensitivities of financial versus real economy indicators to see if they had been generally higher at the outset of the financial crisis. Next, we present real-time out-of-sample results which are improved by averaging the different MSARX specifications and by changing the number of regimes in real-time. Finally the procedure is repeated for German data of the OECD Composite Index of Leading Indicators (CLI). Section 6 concludes.

2 The Literature

Back in the 1970s, research efforts were mainly focused on exact dating of business cycle turning points represented by the seminal book from Bry and Boschan (1971), where they developed a solid working non-parametric dating algorithm. Nowadays the focus has turned to real-time business cycle predictions ahead of the publication point in time. Therefore it is crucial to deal with two questions: Firstly, what estimation procedure to use, and secondly what indicators related to the business cycle are to be included. *Inter alia* the development of estimation procedures was fostered by Chauvet and Potter (2005)

using different specifications of the probit model, by Stock and Watson (1989) introducing dynamic factor models for the business cycle, and by Hamilton (1989) proposing the Markov Switching model which this paper extends to incorporate a combination of different MSARX specifications. The resulting model is then applied to monthly German real-time data. Simultaneously, within the development of different prediction procedures the set of leading real economy indicators was extended to include financial ones such as spreads from the term structure of interest rates, e.g. by Estrella and Hardouvelis (1991) and the spread between corporate and public issuers, e.g. by Friedman and Kuttner (1992). Most recently the connection between the corporate spread and economic development has been analyzed by Gilchrist, Yankov and Zakrajsek (2009) as well as by Meeks (2011). As the rapid expansion of the credit derivative market is seen as one of the reasons for the extent of the last crisis, credit growth may also emerge as a predictor of the business cycle. The general relationship is dealt with by Biggs, Mayer and Pick (2009), whereas the paper at hand concentrates on a specific proxy for credit growth - namely as it arises in the balance sheets of monetary financial institutions within ECB reporting.

Disparity between the characteristics of financial and real economy indicators becomes particularly essential with real-time forecasts. While financial data, at least with monthly frequency, is provided immediately and is not subject to revisions, this is not the case for most of the real economy variables as they are subject to a publication lag and data revisions. As Diebold and Rudebusch (1991) pointed out for the U.S. Composite Leading Index, revisions and the lagged data availability substantially affect the predictive power of leading indicators. That is why this paper considers real-time data and additionally contributes to the literature by adapting the MS model while proceeding on the real-time path. For this purpose, proposals by Hamilton (2011) and Timmermann (2006) are elaborated upon by averaging the inference from alternative specifications. Whereas this has been done for other methods, see e.g. Proaño (2010), to the best of our knowledge no Markov Switching literature follows such an approach. Indeed it is the basis for showing how a data sample-dependent change in the number of regimes can significantly improve real-time forecasts.

Over the years several papers contributed to the literature by attempting to refine the initial Hamilton (1989) model, see Chauvet and Hamilton (2005) and Chauvet and Piger (2008) among others. Chauvet and Hamilton (2005)¹ were among the first to apply Markov

¹The authors use the NBER announcements as a benchmark. Unlike in the U.S., there is no official

Switching estimations systematically to U.S. real-time data including monthly frequency. Unfortunately, unlike in the U.S., real-time matrices available for Germany are relatively low dimensional in terms of publications and are additionally restricted by the structural break in terms of observations due to reunification ², so that highly parameterized specifications only recently became possible to estimate. Moreover, although smoothed recession probabilities in Chauvet and Hamilton (2005) behave very well in general, some of their figures reveal single months in which the business cycle state is misrepresented in the form of a recession probability above 0.5. The extensions in this paper also aiming at eliminating such mistakenly declared business cycle phases.

3 Markov Switching and Exogenous Variables

When starting business cycle modeling, it is useful to look for a well-defined and generally acknowledged borderline between recessions and expansions such as the one given by the National Bureau of Economic Research, NBER (2011), for the U.S.:

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. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion.

At first glance this definition may suggest a five-dimensional MSVAR model, as introduced by Krolzig (1997). But MSVAR specifications would additionally enlarge the parameter space beyond all the extensions we are introducing in this paper. Given the currently available German real-time data records such an approach is likely to fail. For the same reason one-dimensional equations are arranged to include only two kinds of regressors, lags of the dependent variable and one exogenous variable (leading indicator) including its lags. Moreover, only the coefficient of the most recent lag is chosen to switch in order to minimize the number of parameters that have to be estimated. This leads to the following

business cycle turning point dating in Germany, compare Schirwitz (2009).

²Explicitly, this means that the observation period can start at the earliest on January 1991.

form

$$\begin{aligned}
y_t = & \beta_0^{S_t} + \beta_{1,y}^{S_t} y_{t-f-D_y} + \sum_{j=2}^p \beta_{j,y} y_{t-j+1-f-D_y} \\
& + \beta_{1,x}^{S_t} x_{t-f-D_x} + \sum_{j=2}^q \beta_{j,x} x_{t-j+1-f-D_x} + u_t, u_t \sim N(0, \sigma^{S_t}), \quad t = 1, \dots, T,
\end{aligned} \tag{1}$$

where f represents the forecasting horizon and D_y, D_x the data availability lag of the dependent and independent variable. S_t stands for the latent states that generate the total process of the observed time series y_t . For the following review of Hamilton (1989, 1990)'s filter consider a publication lag being of 0, a forecasting horizon of 1 and vector notation so that (1) is rewritten as

$$\begin{aligned}
y_t = & z_t' \beta^{S_t} + u_t, \quad u_t \sim N(0, \sigma^{S_t}) \quad \text{with} \\
z_t' = & (1, y_{t-1}, y_{t-2}, \dots, y_{t-p}, x_{t-1}, x_{t-2}, \dots, x_{t-q}), \\
\beta^{S_t} = & (\beta_0^{S_t}, \beta_{1,y}^{S_t}, \beta_{2,y}, \dots, \beta_{p,y}, \beta_{1,x}^{S_t}, \beta_{2,x}, \dots, \beta_{q,x}), \\
\theta' = & (\beta^{S_t}, \sigma^{S_t}).
\end{aligned} \tag{2}$$

The central characteristic of the Markov Switching model is the fact that the hidden states of the dependent variable are generated by a first order Markov chain, whose transition matrix for a two regime setting will look like

$$P = (p_{ij}) = \begin{pmatrix} \mathcal{P}(S_t = 1 | S_{t-1} = 1) & \mathcal{P}(S_t = 1 | S_{t-1} = 2) \\ \mathcal{P}(S_t = 2 | S_{t-1} = 1) & \mathcal{P}(S_t = 2 | S_{t-1} = 2) \end{pmatrix}. \tag{3}$$

Later on, the model will be extended to four regimes so that for the transition matrix alone 12 parameters have to be estimated. This is where the available real-time data records become especially relevant. For reasons of simplicity let us stick to the two regime case where the second element in identifying the Markov chain, the starting distribution, will look like

$$\hat{\xi}_{1|0} = \begin{pmatrix} \mathcal{P}(S_1 = 1 | y_0) \\ \mathcal{P}(S_1 = 2 | y_0) \end{pmatrix}. \tag{4}$$

y_0 cannot be observed, so the $\hat{\cdot}$ -notation hints at an initial guess for the starting distribution. In fact at the beginning of the maximization algorithm, the entries for both the transition matrix and starting distribution are chosen uniformly, which later plays a role when deciding which regimes are related to recessions and which to expansions of the business cycle. The same applies to the entries for θ .³ The normality assumption delivers the

³Matlab code for all estimations is available on request.

following density vector

$$\eta_t = \begin{pmatrix} f(y_t|S_t = 1, z_t, \theta) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\left(\frac{y_t - z_t'\beta^1}{2\sigma_1^2}\right)^2\right) \\ f(y_t|S_t = 2, z_t, \theta) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\left(\frac{y_t - z_t'\beta^2}{2\sigma_2^2}\right)^2\right) \end{pmatrix}. \quad (5)$$

Applying the proposition of total probability and denoting an element-wise multiplication by \odot we obtain Hamilton (1994, p.692)'s result

$$\hat{\xi}_{t|t} = \frac{\left(\hat{\xi}_{t|t-1} \odot \eta_t\right)}{\mathbf{1}'\left(\hat{\xi}_{t|t-1} \odot \eta_t\right)}, \quad (6)$$

which for the beginning of the series means nothing else than that filtered probabilities

$$\hat{\xi}_{1|1} = \begin{pmatrix} \mathcal{P}\left(S_1 = 1|y_1, \hat{\theta}\right) \\ \mathcal{P}\left(S_1 = 2|y_1, \hat{\theta}\right) \end{pmatrix} \quad (7)$$

have been calculated. The important implication for this paper is that whenever estimates of θ change with different exogenous variables, $\hat{\xi}_{1|1}$ and in general all filtered probabilities will also change. Given the equivalence of $S_t = i \Leftrightarrow \xi_t = e_i$, where the last one represents the i -th unity vector, the preliminary probability for the second observation can be computed by

$$\hat{\xi}_{2|1} = \hat{P}\hat{\xi}_{1|1}. \quad (8)$$

Here P stands for the transition matrix. Using Kim (1994)'s smoothed state probabilities we obtain

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} \odot \left(P' \left(\hat{\xi}_{t+1|T} \oslash \hat{\xi}_{t+1|t}\right)\right), \quad (9)$$

where \oslash denotes an element-wise division. Repeating the procedure recursively finally delivers (smoothed) state probabilities for all observations. What remains is to concretize the maximization process. As a manipulation of the standard maximum likelihood approach, Hamilton (1994, p.696)'s result for a target function in the case of equation (2) will be

$$\sum_{t=1}^T \sum_{i=1}^{r=2} \mathcal{P}(S_t = i|Z_T) \log f(Y_t|S_t = i, Z_t, \theta). \quad (10)$$

Thus, the first order conditions are

$$\sum_{t=1}^T \left(\frac{\partial \log \eta_t}{\partial \theta'}\right)' \hat{\xi}_{t|T} = 0. \quad (11)$$

For simplification restrict θ' to $(\beta_0^{S_t}, \beta_{1,y}^{S_t}, \beta_{1,x}^{S_t}, \sigma^{S_t})$. Then (11) is equivalent to the following non-linear system of equations

$$\begin{aligned}
\sum_{t=1}^T \mathcal{P}(S_t = 1|z_T) \frac{\partial \log \eta_t^1}{\partial \beta_0^1} + \mathcal{P}(S_t = 2|z_T) \frac{\partial \log \eta_t^2}{\partial \beta_0^2} &= 0 \\
\sum_{t=1}^T \mathcal{P}(S_t = 1|z_T) \frac{\partial \log \eta_t^1}{\partial \beta_{1,y}^1} + \mathcal{P}(S_t = 2|z_T) \frac{\partial \log \eta_t^2}{\partial \beta_{1,y}^2} &= 0 \\
\sum_{t=1}^T \mathcal{P}(S_t = 1|z_T) \frac{\partial \log \eta_t^1}{\partial \beta_{1,x}^1} + \mathcal{P}(S_t = 2|z_T) \frac{\partial \log \eta_t^2}{\partial \beta_{1,x}^2} &= 0 \\
\sum_{t=1}^T \mathcal{P}(S_t = 1|z_T) \frac{\partial \log \eta_t^1}{\partial \sigma^1} + \mathcal{P}(S_t = 2|z_T) \frac{\partial \log \eta_t^2}{\partial \sigma^2} &= 0.
\end{aligned} \tag{12}$$

Its non-linearity arises from the derivative with respect to σ^{S_t} . Perlin (2012) uses a gradient ascent method to obtain $\hat{\theta}$ as a solution of (12). Moreover, Hamilton (1990) derives estimates for the transition probabilities by

$$\hat{p}_{ij} = \frac{\sum_{t=2}^T \mathcal{P}(S_t = j, S_{t-1} = i|z_T, \hat{\theta})}{\sum_{t=2}^T \mathcal{P}(S_{t-1} = i|z_T, \hat{\theta})}. \tag{13}$$

With $\hat{\theta}$, $\hat{\xi}_{1,T}$ and \hat{p}_{ij} replacing the randomly chosen initial values, the process reaches its second phase. Again state probabilities are computed and the whole procedure iterates up to the convergence of the likelihood function. Through this detailed description, we want to stress the fact that with different exogenous explanatory variables the solution of the non-linear system of equations in (12) also changes. Thus it is a double effect² of different entries in z'_t and in β^{S_t} during the whole maximization process that ultimately generates different state probabilities for the different MSARX specifications.

At first sight this may sound trivial, but the impact becomes evident when considering that the model does not comprise one equation of type (1). Instead, the whole model is constructed in the spirit of Timmermann (2006) assuming that each equation of type (1) may be subject to a ‘misspecification bias of unknown form’ and that simple averaging

¹One could include (\hat{p}_{ij}) in $\hat{\theta}$ but - to facilitate some remarks in this section as well as in the section related to the LR test - transition probabilities are separated by means of the selected notation.

²The most intuitive term in which to observe this is the conditional expectation $z'_t \beta^{S_t}$ in the density vector (5).

can lower the effect of such a bias. As already mentioned, one reason for such a bias could be the limit of the parameter space given the available real-time data records. In the following, each equation (or in other words specification) will be represented by its exogenous variable. Against the background of trying to achieve proper forecasts, this variable is a carefully chosen leading indicator of the business cycle.

4 Model Specification and In-sample Evaluation

4.1 Data Selection

When it comes to the question of frequency, the paper at hand selects monthly data in order to be ready for forecasts between the quarterly publications of GDP in Germany. Fritsche and Stephan (2002) point out that the highest-correlated monthly proxy of overall German economic activity is industrial production, which captures the volatile parts of GDP such as investments and exports. Hence we select industrial production to serve as the dependent variable in equations of type (1). In order to evaluate the combined forecast later on, a non-parametric ex-post-dating algorithm based on the work of Bry and Boschan (1971) and Harding and Pagan (2002) will be employed. In the tradition of this kind of business cycle literature, industrial production is also referred to as the reference series. All financial and real economy data stem from Deutsche Bundesbank (2011*a,b,c*). As a survey variable we use data provided by ifo Institut für Wirtschaftsforschung (2011) and we take OECD (2011) as the relevant data source for a German composite leading index. In order to avoid the structural break linked to the German reunification, all data starts in early 1994. At the time of writing, the corresponding publication lag of industrial production spans two months.

Figure 1 shows the growth rates of the reference series. The high volatility of monthly data delivers several months with a relatively high positive value even in recessions and a relatively low negative value even in expansions, e.g. $\pm 2\%$. As it turns out, this short-term contrary dynamic cannot be captured by the autoregressive terms, which results in the MS model switching inconsistently with the business cycle phases. In this case the filtered regimes seem to reflect the asymmetry between positive and negative rates exceeding a certain absolute value.

Growth rates of backwards smoothed industrial production

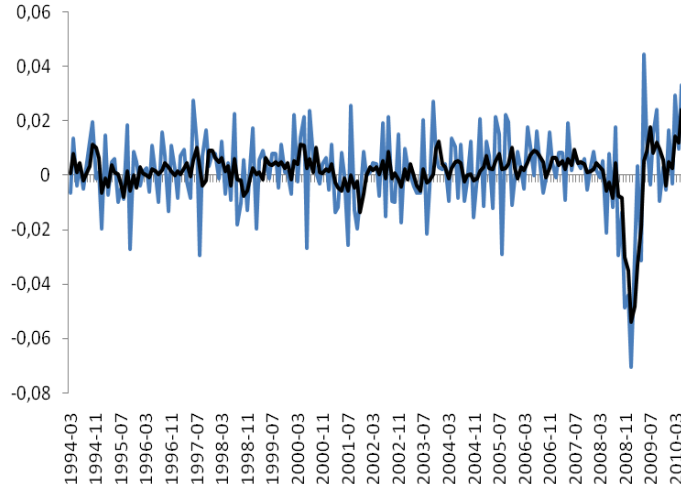


Figure 1: The blue line illustrates the growth rates of monthly industrial production on publication on November 2010 for observations from March 1994 to September 2010, which lead to the MS model switching inconsistently with the business cycle phases. This changes when taking a slightly backwards smoothed version of the industrial production (black line).

BDS Test for growth rates of smoothed industrial production

Dimension	BDS Statistic	Std. Error	z-Statistic	Normal Prob.	Bootstrap Prob
2	0.051300	0.007325	7.003095	0.0000	0.0000
3	0.082194	0.011679	7.037683	0.0000	0.0000
4	0.097266	0.013956	6.969353	0.0000	0.0000
5	0.095695	0.014599	6.554701	0.0000	0.0000
Raw epsilon		0.007572			
Pairs within epsilon		25952.00	V-Statistic	0.703993	
Triples within epsilon		3866120.	V-Statistic	0.546225	
Dimension	C(m,n)	c(m,n)	C(1,n-(m-1))	c(1,n-(m-1))	c(1,n-(m-1))^k
2	9845.000	0.542574	12718.00	0.700909	0.491274
3	7616.000	0.424172	12556.00	0.699304	0.341978
4	6100.000	0.343352	12513.00	0.704323	0.246086
5	4885.000	0.277904	12505.00	0.711401	0.182210

Figure 2: The test results in a clear rejection of the i.i.d - hypothesis confirms the appropriateness of a non-linear model. When computing ϵ/σ this lies in the interval of $[0.5, 2]$. This, as well as a maximum embedded dimension of $m = 5$, represent the relevant range for i.i.d - series according to Brock et al. (1996). Since N/m is not large, additional bootstrapped p-values are calculated.

But, as is stated in the NBER definition, this does not automatically correspond to recessions and expansions, since dynamics in the environment, in particular over the duration of more than a few months and the significance of changes, should be taken into account. One way to handle this is by slightly smoothing the reference series backwards - in this case by a moving average of order 3. In doing so we obtain the coincidence between iterated regimes and business cycle phases also found by Hamilton (1989), who runs his model on quarterly data. As comprehensively discussed by Krolzig (1997, p.20) MS in general (and in particular if highly parameterized as in this paper) represents a non-linear model ⁴. Thus, before starting to estimate, it is reasonable to apply a nonlinearity test to the reference series. A widely used test for nonlinearity is the one developed by Brock et al. (1996).⁵ Thereby it is analyzed whether the residuals of a well-specified ARMA⁶ estimation follow an i.i.d process:

$$H_0 : \text{The residual time series is i.i.d.}$$

As figure 2 shows, H_0 is rejected on every regular level of significance, which in fact suggests the need to apply a non-linear method to the data.

Having dealt with the reference series, we now turn to the leading indicator series. We selected foreign and domestic orders, construction permits, CDAX stock index, the spread between corporate and public issuers' current yield, the 3-month EURIBOR interest rate, the ifo business climate index, credit growth according to the European System of Central Banks (ESCB) statistics, the maturity spread between 10-year federal bonds and 3-month EURIBOR as well as job vacancies. Apart from the corporate spread, the credit growth and job vacancies, which were already mentioned in section 2 because of their capability for early signaling, a similar information set contributes to the U.S. Composite Index of Leading Indicators.

The lead of most of the selected indicators is obvious since they reflect pre-stages to the production process, such as orders or construction permits, or expectations of economic development, such as business climate or job vacancies. As it turns out later on, job vacancy coefficients are only weakly significant. Nevertheless its purpose is also

⁴Apart from special cases it is generally not possible to find a MA_∞ -representation.

⁵Maheu and McCurdy (2000) chose exactly the same procedure to test for the appropriateness of their MS framework. In general the BDS tests for independence, but the procedure here analyzes whether there is any non-linear dependence in the series after the linear ARMA model has been fitted.

⁶The ARMA model is specified due to information criteria which in our case led to an ARMA (2,7).

Sample: 1994:03 – 2010:09	publication lag	revision	ADF test result
Industrial Production	2 months	yes	I(1)
Foreign Orders	2 months	yes	I(1)
Domestic Orders	2 months	yes	I(1)
Construction Permits	2 months	no	I(0) trend-stationary
CDAX	0 month	no	I(1)
Corporate Spread	0 month	no	I(1)
Euribor - 3M	0 month	no	I(1)
ifo Business Climate	0 month	no	I(0)
Credit Growth	1 month	no	I(1)
Maturity Spread	0 month	no	I(1)
Job Vacancies	0 month	no	I(1)

Table 1: Real-time characteristics and stationarity properties of selected variables.

to include at least one variable from the labor market. The reasoning behind the corporate spread is that whenever a recession approaches, this will lead to higher default rates of companies, whereas federal bonds remain a safe haven. Since short-term interest rates react more sensitively to the current economic situation, the spread between long-term and short-term maturity implies predictive potential - sometimes even ending up with an inverse yield curve. The role of credit growth will be dealt with in detail later on.

All data is calendar and seasonally adjusted. In general, financial and survey variables are not subject to revisions and to lagged data availability, whereas real economy variables are. For the real economy variables used here, the publication lag is taken to be two months - except for job vacancies, which are provided immediately. Credit growth is the only financial variable in the ESCB reports with a publication lag (one month). Finally, table 1 displays Augmented Dickey Fuller (ADF) test results ⁷ which suggest that one should difference most of the variables including the reference series. Similar non-stationarity properties have been found by Levanon (2010). For some of the series such as the term spread, there is a controversial literature whether to difference or not. In these cases we decided to potentially over-difference which is usually appraised as the minor mistake.

4.2 The Benchmark Model

In order to evaluate the results of our MS model, an ex-post-dating algorithm based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002) was employed.

⁷Similar results are also displayed in Schreiber, Theobald, Proano, Stephan, Rietzler and Detzer (2012).

Coming back to the definition in section 3, this algorithm should find recessions between peaks and troughs of the reference series (and vice versa expansions). Rewriting the reference series in levels, in a first step, candidates for the turning points of the business cycle have to be recognized. Local extrema can be found by

$$\{y_t : y_t > y_{t\pm k}, k = 1, \dots, 5\} \cup \{y_t : y_t < y_{t\pm k}, k = 1, \dots, 5\}. \quad (14)$$

Not each local extremum automatically represents a turning point since one might be confronted with the phenomenon of extra cycles or non-alternating extrema. That is why (14) only reflects the necessary condition for the dating algorithm. The sufficient condition is introduced according to the triangle approximation of a recession, Harding and Pagan (2002). It consists of the product of duration and relative amplitude exceeding a certain limit, i.e. for recessions

$$\Delta = \frac{1}{2} (t_{peak} - t_{trough}) \frac{y_{peak} - y_{trough}}{y_{peak}} > 0.025. \quad (15)$$

There is no unique value in the literature for the right hand side of the inequality, so it has been chosen to be quite selective. Imagine a recession lasting 5 months. Then the decline of the overall economic downturn must add up to 1% of the base level, i.e. the value at the very last peak. If the duration of the recession is 10 months, a decline of 0.5% of the base level will be sufficient.

The resulting binary series (benchmark series) generated by this procedure detects six recessions (or periods of stagnation) between 1994 and 2010 which is compatible with results from both Schirwitz (2009) as well as the Economic Cycle Research Institute, ECRI (2012). See figure 3 ⁸. Due to the number of forward-looking months in (14) the dating algorithm can only be used for ex-post analysis, namely after 5 months plus the publication lag have expired. For comparison, the results of the ex-post-dating algorithm are colored gray in figures 4, 5 and 6. Whenever MSARX specifications match the results of the ex-post-dating algorithm, MSARX forecasts are favorable out of construction since they deliver the recession information several months in advance.

⁸Note that a tendency towards labeling a stagnation phase as a recession is fully intended. If one does not agree to such an approach, the model offers enough adjusting possibilities to lower the number of detected recessions by simultaneously increasing the probability of a potentially missed recession.

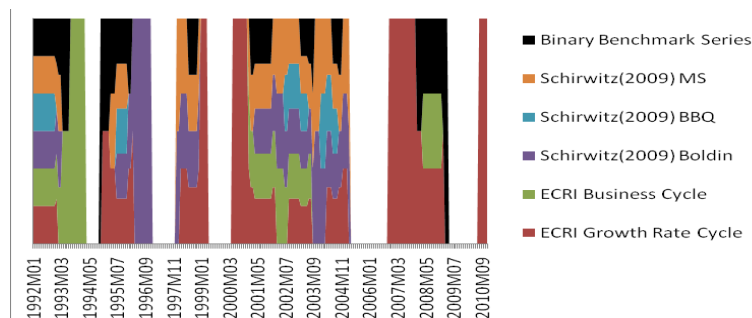


Figure 3: The binary benchmark series (BBS) and its compatibility to other recession detection methods from the literature. ECRI distinguishes between growth rate and business cycle where the latter represents a subset of the former. Schirwitz (2009) follows a similar approach by employing different methods and allowing only for recession detection when all agree. We proceed differently since we consider the missing of a recession as the more serious error. Consequently, our recession definition embeds to some extent periods of economic stagnation. But for each detected recession of our favored approach, the BBS, at least two other methods can be found which totally agree with the detected period. Periods where none of the approaches detects a recession are colored white.

4.3 Model Specification

The usual trade-off between improving the overall fit by additional significant regressors and making it worse by over-specification arises particularly with MS models. Whereas in a standard linear estimation an equation with only two kinds of regressors and restricted lag selection may lead to an omitted variable bias, in a combination of MSARX specifications this is only obvious when the bias occurs in each of the regime dependent equations. However, extending the equations with additional variables or switching parameters in such a way that the optimization described in section 3 only finds local maxima seems to be the greater danger, see Boldin (1996). Therefore, as it becomes obvious by means of the degree of freedom in table 2, we accepted at most four switching coefficients in a single equation in order to guarantee the numerical robustness of the approach. Nevertheless, in general restricted real-time data records still remains a problem.

In the following we explain which coefficients we chose to switch and why. The first MS

business cycle model introduced by Hamilton (1989):

$$y_t - \beta_{S_t} = \sum_{j=1}^4 \phi_j (y_{t-j} - \beta_{S_{t-j}}) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2), \quad S_t = 1, 2.$$

helps in deciding whether the intercept in each equation of type (1) should switch or not. In particular, Perlin (2012) showed that estimating the following version of the Hamilton model

$$y_t = \beta_{S_t} + \epsilon_{a,t}, \quad \epsilon_{a,t} \sim N(0, \sigma_a^2), \quad S_t = 1, 2$$

$$\epsilon_{a,t} = \sum_{j=1}^4 \phi_j \epsilon_{a,t-j} + \epsilon_{b,t}, \quad \epsilon_{b,t} \sim N(0, \sigma_b^2)$$

delivers similar values for the state probabilities and switching coefficients. Indeed, this two-stage estimation identifies the switching intercept as the most relevant part for the iteration of the state probabilities and explains the inclusion of a switching intercept in our model.

In addition, we include more switching coefficients in each of the equations. The reasons for that have been touched upon in the introduction. Technically, they are as follows: Firstly, a switching variance of the error term allows for applying the Welch test to identify different normal distributions when turning to a higher regime setting, see section 5.1. Hence, the selected design implicitly delivers a way of testing the change in the number of regimes. Secondly, at least one switching coefficient of the embedded leading indicator (and of the lagged dependent variable for consistency) allows for measuring how much the coefficients change between the regimes representing different intensities of the same business cycle phase (weak and strong recession or weak and strong expansion), see table 5 in section 5.1.

Table 2 illustrates the final MSARX specifications according to a general-to-specific lag choice by information criteria. Lags can be selected up to a maximum of 5 when in addition a minimum of state probabilities agree with the time-consistent benchmark series generated by the algorithm of section 4.2. In a former version, lag choice was implemented to be renewed for each publication on the real-time path. But because of an exploding running time the lag structure in table 2 was fixed for all real-time estimations.⁴

⁴First of all, lag choice is based on the publication from January 2007, i.e. the last before real-time forecasts of section 5.2 start, but at this point in time a higher regime design is only hypothetically possible and not a result attributable to the criteria (19) which will be introduced in section 5.1. That is why in-sample results presented in section 4.4 help to check if the selected lag structure still works for a publication (November 2010) after the number of regimes has really changed due to the criteria (19).

4.4 In-sample Evaluation

Tables 2 and 3 as well as figure 4 present in-sample results for the different MSARX specifications. Table 2 already deals with a four regime setting and provides standard errors for each of the coefficients in parentheses. Standard errors have been calculated in line with the method from Perlin (2012) to be robust to heteroskedasticity and serial correlation according to Newey and West (1987). Most of the coefficients are highly significant. The only instance of systematic insignificance is the rejection of third regime's parameters when running the specification with job vacancies. In this case it cannot be excluded that each of the regime's coefficients and with this the contribution to the fit could be equal to 0. With respect to the consistency of the whole system and to the intention to include at least one variable from the labor market these results were accepted anyway. Minor instances of insignificance arise in the specification with credit growth. That is why we qualify our conclusion, in that credit growth should not be categorically neglected as a predictor of the business cycle.

Generally, one links an extended credit flow, which reflects better access to finance investments, with an upturn of the economic situation. Surprisingly Biggs et al. (2009) found as 'a stylized fact that after financial crises economic activity recovers without a rebound in credit', but they do not exclude this relation for the flow of credit. When considering credit growth rates extracted from the balance sheets of monetary financial institutions, this leaves a totally different impression. See the corresponding sub-figure of figure 4. The mid-points of the last recessions as detected by the benchmark series seem to coincide with the peaks of credit growth. This coherency becomes especially evident when including credit flows to other monetary financial institutions. In fact, this might reflect the unhealthy credit growth linked to previous credit bubbles. As pointed out earlier, the MS model is based on growth rates, which means including the growth of credit growth as a regressor. However, results are good enough not to categorically reject credit growth as a predictor of the business cycle. ⁹

⁹Uncertainty also arises from two issues: The series of credit growth is only available from the start of the ESCB, i.e. the end of 1998. Additionally most credit derivatives, which could serve as an easy explanation for the results, were declared as off-balance transactions in the past.

Sample: 1994:03 – 2010:09 Publication: 2010:11	Switching Intercept				Switching Endogenous Lag				Additional Endo. Lag
	r1	r2	r3	r4	r1	r2	r3	r4	
Purely Autoregressive	0.010*** (0.0018)	0.004*** (0.0011)	-0.002** (0.0011)	-0.040*** (0.0008)	- (-)	- (-)	- (-)	- (-)	-0.113*** (0.0090)
Foreign Orders	0.005*** (0.0003)	0.000 (0.0003)	-0.020*** (0.0002)	-0.034*** (0.0000)	-0.092*** (0.0000)	-0.132*** (0.0021)	-1.281*** (0.0141)	0.349*** (0.0000)	- (-)
Domestic Orders	0.013*** (0.0001)	0.003*** (0.0002)	-0.002*** (0.0006)	-0.033*** (0.0009)	- (-)	- (-)	- (-)	- (-)	- (-)
Construction Permits ⁵	0.005*** (0.0005)	0.005*** (0.0009)	-0.002*** (0.0005)	-0.039*** (0.0006)	-0.100*** (0.0146)	0.455*** (0.0101)	-0.121*** (0.0173)	-0.069*** (0.0156)	-0.017 (0.0126)
CDAX	0.006*** (0.0003)	-0.001** (0.0003)	-0.010*** (0.0001)	-0.031*** (0.0008)	-0.101*** (0.0019)	-0.094*** (0.0026)	0.220*** (0.0018)	0.977*** (0.0022)	- (-)
Corporate Spread	0.010*** (0.0002)	0.003*** (0.0007)	-0.003 (0.0022)	-0.054*** (0.0031)	-0.057*** (0.0068)	-0.089*** (0.0158)	-0.122*** (0.0002)	-0.420*** (0.0700)	- (-)
Euribor - 3M	0.012*** (0.0032)	0.004*** (0.0058)	-0.003*** (0.0017)	-0.043*** (0.0010)	- (-)	- (-)	- (-)	- (-)	- (-)
ifo Business Climate	0.010*** (0.0008)	0.009*** (0.0001)	0.006*** (0.0003)	0.000 (0.0007)	0.274*** (0.0431)	1.087*** (0.0677)	-0.085** (0.0411)	-0.127*** (0.0264)	- (-)
Credit Growth ⁶	0.010*** (0.0014)	0.005 (0.0264)	-0.003 (0.0021)	-0.043*** (0.0071)	-0.117* (0.0663)	-0.122*** (0.0273)	-0.128** (0.0648)	-0.232 (0.2028)	- (-)
Maturity Spread	0.010*** (0.0010)	0.003*** (0.0006)	-0.002*** (0.0005)	-0.027*** (0.0043)	- (-)	- (-)	- (-)	- (-)	- (-)
Job Vacancies	0.012** (0.0046)	0.003** (0.0013)	-0.002 (0.0065)	-0.016*** (0.0024)	-0.123** (0.0481)	-0.095** (0.0417)	-0.175 (0.1692)	-0.306*** (0.0233)	- (-)

Sample: 1994:03 – 2010:09 Publication: 2010:11	Switching Exogenous Lag				Additional Exogenous Lags		Degree of Freedom
	r1	r2	r3	r4			
Purely Autoregressive	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	162
Foreign Orders	0.015** (0.0051)	0.045*** (0.0057)	0.737*** (0.0050)	-0.078*** (0.0000)	0.032*** (0.0000)	- (-)	154
Domestic Orders	0.159*** (0.0072)	0.052*** (0.0073)	0.050** (0.0242)	0.053*** (0.0053)	0.036** (0.0141)	- (-)	158
Construction Permits ⁵	-0.003 (0.0034)	0.159*** (0.0159)	-0.0028* (0.0017)	-0.209*** (0.0033)	-0.002 (0.0031)	-0.004* (0.0022)	148
CDAX	-0.007** (0.0023)	-0.005*** (0.0003)	0.136*** (0.0009)	0.289*** (0.0024)	-0.005*** (0.0011)	- (-)	154
Corporate Spread	0.041*** (0.0091)	0.018** (0.0089)	-0.135*** (0.0109)	0.197*** (0.0262)	- (-)	- (-)	155
Euribor - 3M	0.654*** (0.0497)	0.043*** (0.0103)	-0.102*** (0.0282)	1.389*** (0.0279)	- (-)	- (-)	159
ifo Business Climate	1.569*** (0.0506)	0.206*** (0.0546)	0.088** (0.0378)	0.103*** (0.0216)	- (-)	- (-)	155
Credit Growth ⁶	0.025** (0.0253)	0.009* (0.0046)	0.012 (0.0111)	0.059 (0.0617)	0.001 (0.0029)	-0.010*** (0.0028)	93
Maturity Spread	-0.043 (0.0310)	0.018 (0.0130)	0.006 (0.0052)	-0.204*** (0.0537)	- (-)	- (-)	159
Job Vacancies	0.068*** (0.0210)	0.020 (0.0156)	0.040 (0.1057)	0.862*** (0.0821)	0.048** (0.0217)	0.019 (0.1211)	153

Table 2: MSARX regressions with 4 regimes (r1 - r4) for publication on November 2010. Each specification is represented by its embedded leading indicator. The upper table deals with the intercept and lagged dependent variable of the equations and the lower with the (lagged) exogenous variables (leading indicators). *p<0.10, **p<0.05, ***p<0.01. Standard errors are in parentheses.

When presenting in-sample results of real-time estimations, it must be mentioned that statements about significance are linked to a certain publication - in this case to the publication on November 2010. Indeed p-values may change (slightly) for each and every publication, but it goes beyond the scope of this paper to consider each of the publications for the in-sample analysis. Figure 4 illustrates the state probabilities iterated out in each of the MSARX specifications. Indeed they do change, as was theoretically anticipated in section 3. Moreover, it is reasonable to assume that the predictive power of a single indicator, and thus of the corresponding specification, varies in different recessions so that the ex-ante best forecast need not be the same for the next publication. In a nutshell, the changing regime probabilities with different MSARX specifications make it advisable to rely on a combined forecast.

Whereas the ex-post-dating algorithm of section 4.2 detects six recessions (or periods of stagnation), not each MSARX specification ends up with the same number (with respect to a recession probability above 50%). But most of the MSARX specifications interpret the same periods as recessions, so that there are also six recessions when averaging the state probabilities. Each MSARX specification identifies the recession linked to the financial crisis as an additional regime, see figure 4. Thus, when turning to the out-of-sample forecasts, the task will be to allow for a flexible number of regimes. But how can different regimes, in particular in the case of more than two, be identified as recessionary or expansionary states? In general, it is natural to relate regimes to recessions or expansions. The latter determine the business cycle and with the MS model the iterated regimes generate the reference series which again is a proxy for the business cycle. Under certain conditions the identification can focus on the switching intercept. Each iteration starts with uniformly distributed initial values, so it is not necessary that the same regime label (e.g. r1) always represents the same business cycle state for each and every publication. Hence, in a two regime setting, the higher intercept will identify the expansionary regime and the lower the recessionary. In section 5.1 we will discuss the more sophisticated four regime case, where different intensities of expansion and recession come into play.

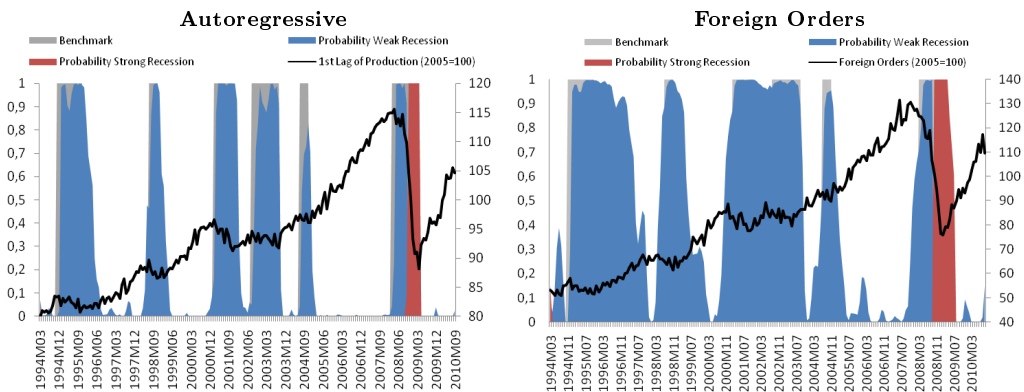
Another interesting question arising with different MSARX specifications is whether a leading indicator can really bequeath its predictive power to the corresponding specification. Hints of such a relation can be found in figure 4. For instance the presumed continuous recession between February 2001 and September 2003 in the specifications of CDAX and the

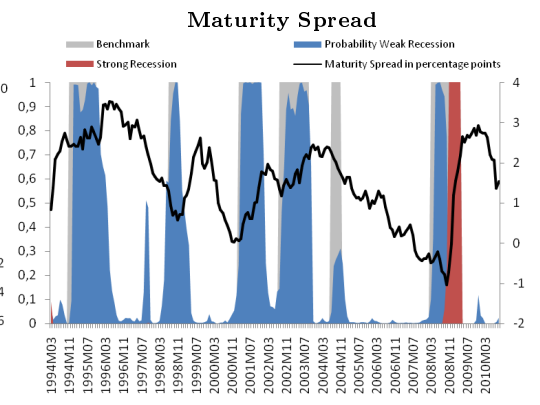
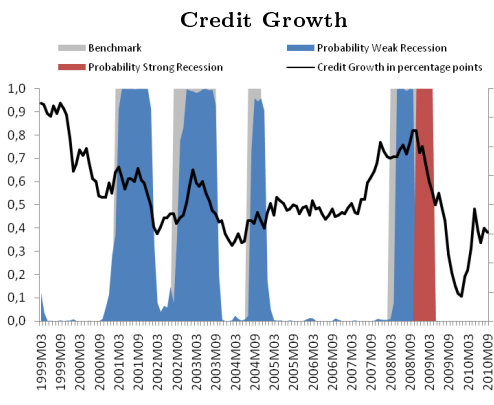
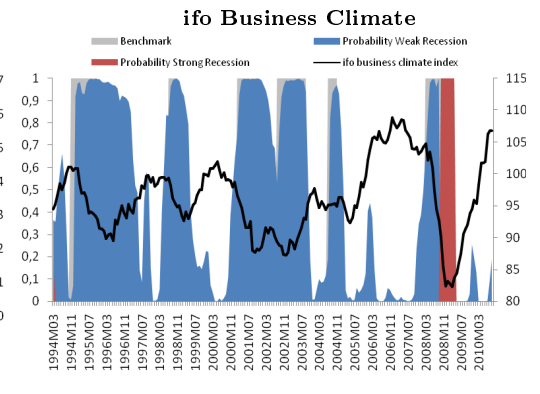
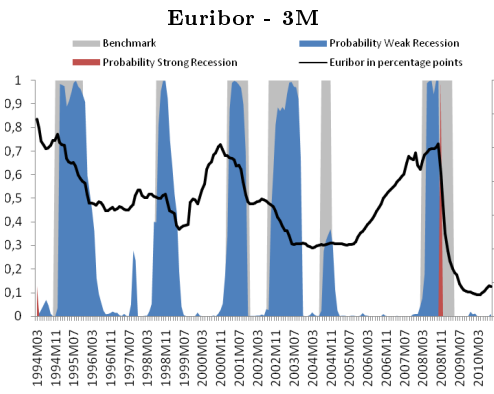
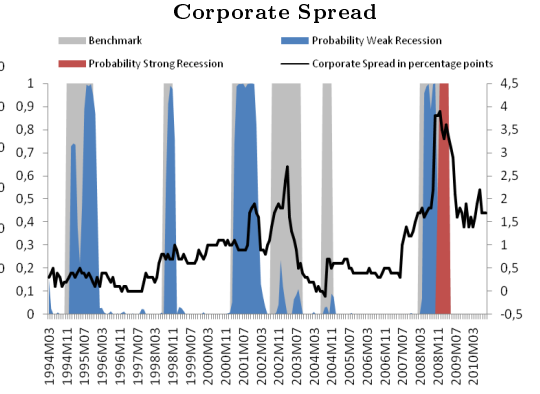
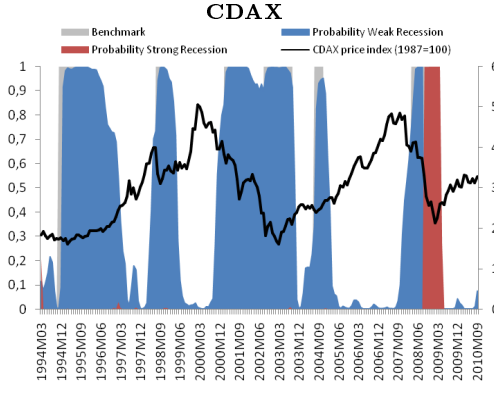
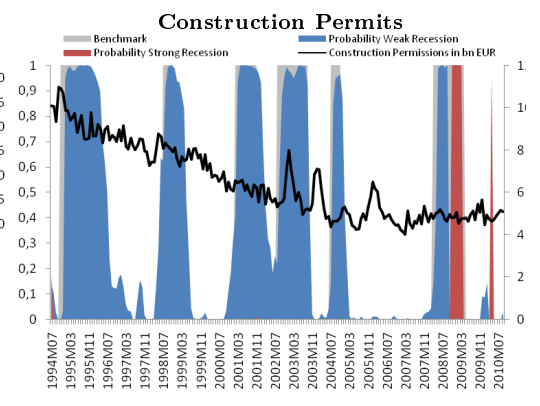
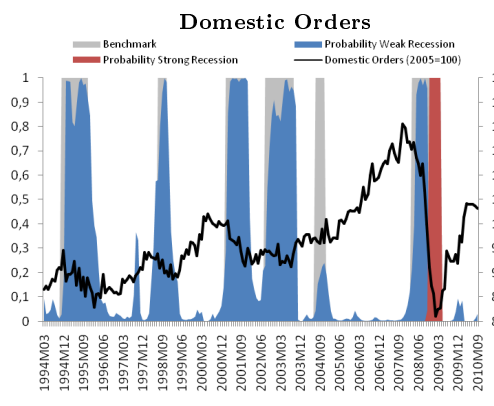
⁵Sample 1994:07 - 2010:09 for table 2, 3, 4 and sample 1994:07 - 2008:06 for table 5.

⁶Sample 1999:03 - 2010:09 for table 2, 3, 4 and sample 1999:03 - 2008:06 for table 5.

Ifo Business Climate reflects the course of the corresponding indicators. Another example is the leading start of the last recession, reflecting the predictive power of foreign orders. Nevertheless there are also counterexamples, e.g. the development of corporate spread and job vacancies clearly fits the downturn from August 2002 to September 2003, whereas the corresponding specifications do not report a recession. But even if such an intuitive relation is not consistently apparent, it still makes sense to use different leading indicators with respect to different MSARX specifications. The reasons are two-fold: On the one hand, the ex-ante worse specifications can still improve the consensus forecast whenever they contain some independent information from the ex-ante better ones, see Bates and Granger (1969). On the other hand, for each new publication a single specification can be subject to a misspecification bias of unknown form, see Timmermann (2006). To a certain extent, this potential weakness can be balanced out in a combined forecast. Later on, the OECD Composite Index of Leading Indicators (CLI) will be used as the explained variable in the MS regression. One could consider using the CLI as an explanatory variable, but as can be seen in the last sub-figure of figure 4, we do not find persistent regimes in such a regression. Thus this approach is left out whenever we take industrial production as the dependent variable.

Finally, table 3 summarizes the number of correct recession and expansion detections for each of the MSARX specifications in comparison to the ex-post-dating algorithm presented in section 4.2. As pointed out in section 4.3 the purely autoregressive estimation is most comparable with the Hamilton model. Both specifications with a higher number of correct recessions (e.g. CDAX) and with a higher number of correct expansions (e.g. corporate spread) as well as specifications with a higher degree of total correctness (domestic orders) can be found.





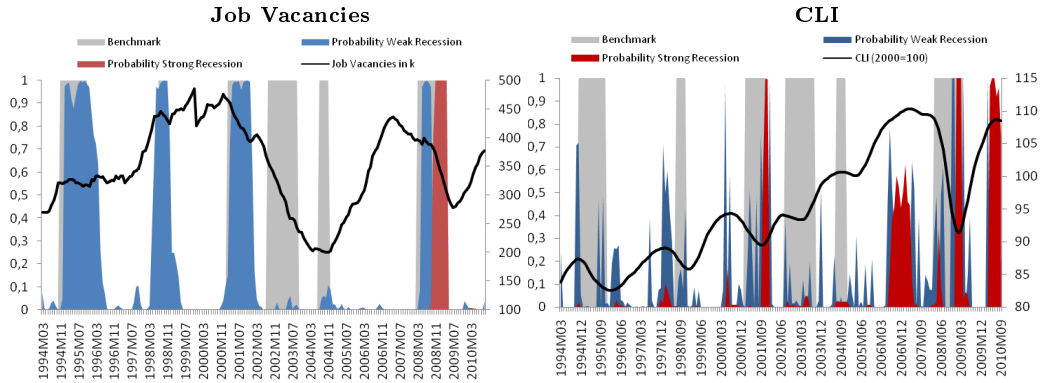


Figure 4: Recession probabilities of MSARX specifications 1 - 12 (left axis) and the embedded leading indicators (right axis). In the regression, growth rates of the reference series are regressed on the growth rates (differences) of the leading indicators. But for illustration purposes levels are provided in the sub-figures. All data is calendar and seasonally adjusted. Weak and strong expansion regimes are colored white, but in most cases the strong expansion regime fits the recovery after the financial crisis. Not all MSARX specifications recognize six recessions (or periods of stagnation) in the observation period, whereas the ex-post dating algorithm does. But most of the MSARX specifications do produce six, so the number of recessions is the same when the state probabilities are averaged. A regression with the CLI as an explanatory variable does not deliver persistent regimes (last graph), so this specification is not considered for the consensus forecast.

Sample: 1994:03 – 2010:09	recession		expansion		total		\bar{R}^2	RMSE	SIC	LR_T
	#	%	#	%	#	%				
Purely Autoregressive	52	85.25	127	92.03	179	89.95	0.8234	0.0033	-1467.39	260.54
Foreign Orders	59	96.72	89	44.72	148	74.37	0.7664	0.0037	-1440.76	228.32
Domestic Orders	50	81.97	133	96.38	183	91.96	0.7435	0.0039	-1437.91	187.09
Construction Permits ⁵	52	85.25	115	85.82	167	85.64	0.8131	0.0033	-1369.53	228.73
CDAX	58	95.08	98	71.01	156	78.39	0.7825	0.0035	-1440.37	250.77
Corporate Spread	33	54.10	131	94.93	164	82.41	0.7803	0.0036	-1438.46	265.84
Euribor - 3M	42	68.85	127	92.03	169	84.92	0.7528	0.0038	-1446.56	192.57
ifo Business Climate	58	95.08	91	65.94	149	74.87	0.0980	0.0072	-1447.86	247.49
Credit Growth ⁶	37	86.05	90	93.75	127	91.37	0.8023	0.0037	- 957.64	219.76
Maturity Spread	49	80.33	124	89.86	173	86.93	0.8409	0.0031	-1432.48	231.48
Job Vacancies	37	60.66	125	90.58	162	81.41	0.7944	0.0034	-1452.31	268.24
AR(1)	-	-	-	-	-	-	0.5916	0.0055	-	-

Table 3: Goodness of fit for different MSARX specifications. First six columns are linked to MSARX regime probabilities and compare their allocation to recession and expansion regimes with the results of the ex-post dating algorithm. The adjusted R^2 as well as the root mean squared error (RMSE) are calculated with respect to the fitted growth rates. These values are also compared with results from a simple AR(1) estimation. Information criteria as well as the likelihood ratio test statistic are linked to the log likelihood function as it is described in section 3 (10). In order to compute the test statistic, for each of the MSARX specifications a linear version without switching coefficients is estimated and the corresponding log likelihood is taken into account, see (16).

In addition to measures related to the regime probabilities table, 3 also reports standard measures linked to the fitted growth rates of the reference series. These are the adjusted R^2 and the root mean squared error. After regime probabilities are already filtered and the state dependent coefficients are estimated, fitted values can be computed from the different state equations. Therefore each state equation is weighted by the corresponding regime probability. Except for the specification with the Ifo Business Climate Index, each of the MS regressions reaches values above 0.7 for the adjusted R^2 . In this way they outperform a simple AR(1) benchmark. In fact the low value in the regression with the Ifo Business Climate might be based on a local maximum found by the expectation-maximization algorithm (section 3) for the specific publication on November 2010. Thus, it might reveal that one of the eleven regressions is subject to a misspecification bias for this certain publication. Again, the instrument to handle this problem is averaging single forecasts. Timmermann (2006) finds that for such cases particularly simple combinations often reach better results than the ex-ante best one. A similar impression arises after computing the root mean squared error. Again, it is the regression with the Ifo Business Climate Index which reaches a far higher value than the others. Moreover, in table 3 there are also measures linked to the log likelihood function of the single MSARX estimations. Here none of the results are conspicuously out of range. More details for the interpretation of the results are provided in the next sections.

4.5 Likelihood Ratio Test in a 4 Regime MS Model

The last column of table 3 contains the test statistic for a likelihood ratio test between the unconstrained MS model and the linear constrained versions of each of its equations, where there is no parameter switch, i.e.

$$LR_T = -2 \left(Q_T^{\text{unconstrained}} \left(\hat{\theta}, (\hat{p}_{ij}) \right) - Q_T^{\text{constrained}} \left(\tilde{\theta} \right) \right). \quad (16)$$

Here Q_T represents the log likelihood, where T in this special case stands for the last observation of the publication on November 2010. A careful reading of the notation makes it obvious that there are parameters, namely the entries of the transition matrix, which are arbitrary under the null hypothesis of a true linear model. This is what Hansen (1996) calls ‘inference when a nuisance parameter is not identified under the null’. Choosing the

following linear and unique decomposition of the switching parameters in equation (1),

$$\begin{aligned}
y_t &= (\alpha_0^0 + (\alpha_0^1 + \alpha_0^2 + \alpha_0^3) S_t) + (\alpha_{1,y}^0 + (\alpha_{1,y}^1 + \alpha_{1,y}^2 + \alpha_{1,y}^3) S_t) y_{t-f-D_y} \\
&+ \sum_{j=2}^p \beta_{j,y} y_{t-j+1-f-D_y} + (\alpha_{1,x}^0 + (\alpha_{1,x}^1 + \alpha_{1,x}^2 + \alpha_{1,x}^3) S_t) x_{t-f-D_x} \\
&+ \sum_{j=2}^q \beta_{j,x} x_{t-j+1-f-D_x} + (\sigma^0 + (\sigma^1 + \sigma^2 + \sigma^3) S_t) \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad t = 1, \dots, T,
\end{aligned}$$

the null hypothesis can be written as

$$H_0 : (\alpha_0^1, \alpha_0^2, \alpha_0^3, \alpha_{1,y}^1, \alpha_{1,y}^2, \alpha_{1,y}^3, \alpha_{1,x}^1, \alpha_{1,x}^2, \alpha_{1,x}^3, \sigma^1, \sigma^2, \sigma^3)' = 0. \quad (17)$$

we obtain a nested design. Because of the nuisance parameters, which are only available under H_1 , the (asymptotic) LR-distribution will not conform to the standard χ_{d-c}^2 -distribution, where the degree of freedom is usually equal to the difference in the number of estimated parameters. Although Garcia (1998) as well as Carrasco, Hu and Ploberger (2005) develop asymptotic null distributions for similar test problems, their results do not exactly fit our requirements. Both only consider the simple case of 2 regimes⁷ and a lower number of switching parameters.⁸ It is left to future research to extend test procedures explicitly to the case of 4 regimes and to a higher parameterized design. Nevertheless the test statistics in table 3 provide some heuristics without simulating the distribution of LR_T explicitly.

Firstly, when creating a ranking of the leading indicators in the next section, test statistics can be used comparatively. This should be feasible since, apart from the influence of the leading indicators, the asymptotic null distribution for each of the MSARX specifications only differs because of including a switching coefficient with the most recent autoregressive lag or not.⁹ Secondly, analyzing critical values, as they are provided by Garcia (1998), shows that they are higher than in the standard case and that they increase with a higher

⁷This is the comfortable case where $E(S_t = i) = \frac{1 - \mathcal{P}(S_t=j|S_{t-1}=j)}{2 - \mathcal{P}(S_t=i|S_{t-1}=i) - \mathcal{P}(S_t=j|S_{t-1}=j)}$, $i, j = 1, 2$, which is used as an entry in all covariance matrices for the asymptotic χ^2 -process developed by Garcia (1998).

⁸Another problem with the approach of Carrasco et al. (2005) is that test results are obtained without estimating the model under the alternative. In doing so, it is difficult to capture the differences between the single MSARX specifications, which are at heart of the paper at hand.

⁹Compared to the impact of the three other switching coefficients in all equations, i.e. the switching intercept, the switching exogenous lag and the switching variance of the error term, this is assumed to be subordinated.

parameterized design. But still, these values are below 40 on a one percent level of significance - even in the case of a model with at least a switching intercept and a switching variance of the error term. Hence, the distance to the test statistics of table 3 makes a rejection of the null - i.e. the appropriateness of switching coefficients - likely for all of the specifications.¹⁰

4.6 Ranking Leading Indicators

As pointed out earlier, the MS framework at hand is constructed in a way which includes all the information given by all the selected leading indicators. Due to the notion that the predictive power of a certain indicator (and implicitly the impact of a potential misspecification bias) may change for different publications, a simple but effective way to include all the information and to deal with time-variant behavior is the averaging of the single MSARX specifications (see also Timmermann (2006)).

Nevertheless, in this framework the need can also arise to rank leading indicators depending on a certain hypothesis. For instance, we want to examine whether, in general, indicators do better when they are not subject to a data availability lag. Again we restrict ourselves to the publication on November 2010. As the criteria to rank the indicators, the aforementioned standard measures can be used. Obviously these measures are linked to the MSARX specifications. Hence, whenever a certain specification is outperforming the others, we presume that the embedded leading indicator is responsible. Or in other words we take the leading indicators as being representative of the corresponding MSARX specifications. Table 4 lists the first three indicators that reach the best values according to the measures of table 3. The consistency between the adjusted R^2 and $RMSE$ can be explained by them both depending on the sum of squared residuals. The adjusted R^2 additionally considers the number of estimated parameters. According to the overview in table 4 the best performer in the set of indicators is the maturity spread - occurring three times and being available

¹⁰This statement is based on the fact that in the case of 4 regimes a supremum-type test (Hansen (1996) theorem 3, Garcia (1998) theorem 1), i.e. $LR_T = \sup_{p_{ij} \in \Gamma} LR_T(p_{ij})$, is still feasible and the test statistics converge to some χ^2 -process. But for an explicit computation, the corresponding covariance matrices have to be extended to some degree, which goes beyond the scope of this paper. Extensions to the dimension of the vector in (17) requires at least 18 derivatives to be computed resulting in an 18×18 covariance matrix for most of the specifications. Given such a computation, it would be possible to simulate the whole range of transition probabilities.

Sample: 1994:03 – 2010:09	1. Position	2.Position	3. Position
Total Correctness	Domestic Orders	Maturity Spread	Euribor
Adjusted R^2	Maturity Spread	Construction Permits ⁵	Credit Growth ⁶
Root Mean Squared Error	Maturity Spread	Construction Permits ⁵	Job Vacancies
LR Test Statistic	Job Vacancies	Corporate Spread	CDAX

Table 4: A ranking of leading indicators due to the measures of table 3. Taking the indicators as representative of the corresponding MSARX specifications, the first three indicators are listed that reach the best values according to each of the measures. Results from the purely autoregressive estimation are neglected here.

without a publication lag. But, although there are more indicators which are subject to a data availability lag, 8 versus 4, the results generated are not clear enough ¹⁰ to conclude that in general they will perform better in a business cycle analysis.

5 Out-of-sample Evaluation

5.1 Real-time Forecast Methodology

In general, forecasts with MS models are produced similarly to those from equation (8), i.e.

$$\hat{\xi}_{T+f|T} = \hat{P}^f \hat{\xi}_{T|T}, \quad (18)$$

where the exponent represents the real forecast horizon ¹¹. According to equation 18, the future state probability of being in a certain regime comes from the regime probabilities of the last observation. These are weighted by the probability of changing to a certain regime. Note that the model at hand follows a standard approach insofar as transition probabilities do not depend on the amplitude or the duration of the last regime, as for instance introduced by Durland and McCurdy (1994). Obviously, the idea of extending the model to additional regimes, explicitly to a third and to a fourth one, arose during the economic downturn of the financial crisis and the recovery thereafter. But real-time forecasts cannot be based on information received later on. That is why a criterion has been developed which determines when to introduce new regimes on the real-time path.

¹⁰For this statement we find particularly important the non-parametric rank-sum tests, which are appropriate in the scenario above.

¹¹Given a data availability lag of 2 months the real forecast horizon for a one-month-ahead forecast is 3 months.

In fact, the in-sample results presented in figure 4 show that the probability of the additional regimes has only been allocated since the financial crisis. But these are in-sample results for a certain publication - here on November 2010. With real-time forecasts it is clear enough that whenever new regimes are introduced some probability will be allocated to them. The essential question is how great a probability is allocated. The answer to this question lies at the heart of the above-mentioned criterion. With an increasing number of regimes, more extreme events can be reproduced. This enables the model to distinguish between a strong and a weak intensity of the same kind of business cycle phase. Considering a first month as representing a potential outlier, it makes sense to change the number of regimes whenever the probability of the strong intensity exceeds the one of the weak intensity for two consecutive months, i.e.

$$\begin{aligned} & \mathcal{P}(S_{T+t+1} = \text{strong}|y_{T+t}) > \mathcal{P}(S_{T+t+1} = \text{weak}|y_{T+t}) \\ \wedge & \mathcal{P}(S_{T+t+2} = \text{strong}|y_{T+t+1}) > \mathcal{P}(S_{T+t+2} = \text{weak}|y_{T+t+1}), \quad t = 0, 1, \dots, \end{aligned} \tag{19}$$

where $T + 1$ stands for the beginning of the out-of-sample forecasts. From an operational point of view, in order to apply this criterion, it is necessary to run both in parallel - the setting with less regimes and the one with more regimes.

At this point the question arises whether to increase the number of regimes by one or by two. Here we take a symmetric approach, which in times of financial crisis means that the total number of regimes can only change from two to four. The reason is quite simple. When introducing a third regime in real-time it will not be clear without laborious computation whether to allocate its probability to a recession or an expansion. Certainly, this advantage of a symmetric approach has to be weighed against the potential instability arising from the additional parameter. At least in our test runs the four-regime model did not provide any evidence of over-parameterization.

In the case of two more regimes, two of the four will lead to higher growth rates in absolute values. Hence the regime with the most positive intercept will be allocated to strong expansion, the one with the most negative intercept to strong recession. Then the remaining regimes form the weak expansion and weak recession states ¹². Naturally this approximation is only feasible whenever the switching intercept is identified as the most relevant part for the iteration of the state probabilities, see section 4.3. Moreover Welch test results

¹²In doing so a possible misallocation cannot be excluded categorically since with a single specification there might occur one expansion and three recession regimes or vice versa, but a symmetric approach fits the main empirical findings of section 4.4.

given in table 5 confirm the symmetric approach except for one case (the specification using construction permits as the embedded leading indicator). Changing the number of regimes due to criterion (19) the Welch test hypothesis

$$\begin{aligned}
H_0 : \quad & \text{Regimes chosen to stand for different intensities of a recession or an expansion} \\
& \text{can have the same (normal) distribution.} \\
\Leftrightarrow \quad & \mu_{\text{expansion}}^{\text{strong}} | \sigma_{\text{expansion}}^{\text{strong}} = \mu_{\text{expansion}}^{\text{weak}} | \sigma_{\text{expansion}}^{\text{weak}} \\
\wedge \quad & \mu_{\text{recession}}^{\text{strong}} | \sigma_{\text{recession}}^{\text{strong}} = \mu_{\text{recession}}^{\text{weak}} | \sigma_{\text{recession}}^{\text{weak}}
\end{aligned} \tag{20}$$

can be rejected for the publication on August 2008^{13 14}. Hence Welch test results support a change to four regimes, although at this point in time Germany's recovery had not started yet.

After the decision has been made in favor of changing from two to four regimes, another issue can be analyzed. Obviously the strong downturn of the last recession is linked to turbulences that occurred on financial markets. Thus, one could suggest that given such an origination of recessionary phases, financial variables will exhibit a higher predictive power for the business cycle than real economy ones do. Relating this statement to the model, one would expect the (relative) change between the coefficients linked to weak and strong recession regimes to be higher with financial than with real economy variables. Table 5 contains the corresponding results. Although, indeed, the largest change occurs with a financial variable (3-month EURIBOR interest rate), the hypothesis cannot be confirmed on a general level.¹¹

¹³The hypothesis assesses whether the means can be equal under the given differences in empirical variances. Because the normal distribution is involved, a rejection of equal expectations is taken as a rejection of that same distribution.

¹⁴Taking S_t as a filtration calculus of the conditional expectation shows how to approximate parameter μ by the intercept and the exogenous parts of the equation:

$$\begin{aligned}
E(y_t | S_t) &= E(E(y_t | S_t) | S_{t-1}) \\
&= E\left(\beta_0^{S_t} + \beta_{1,y}^{S_t} E(y_{t-1} | S_t) + \beta_{1,x}^{S_t} x_{t-1} | S_{t-1}\right) \\
&= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t} E(y_{t-1} | S_{t-1}) \\
&= \beta_0^{S_t} + \beta_{1,x}^{S_t} x_{t-1} + \beta_{1,y}^{S_t} \left(\beta_0^{S_{t-1}} + \beta_{1,x}^{S_{t-1}} x_{t-1} + \beta_{1,y}^{S_{t-1}} E(y_{t-2} | S_{t-2})\right).
\end{aligned}$$

Since for each β -term one has $\beta \ll 1$, the last term consisting of products of β s can be neglected. A similar assessment identifies the error's σ^2 as the essential part of the conditional variance.

¹¹For instance, large changes also occur with domestic and foreign orders. Again, the statement is formally based on non-parametric rank sum tests.

Sample: 1994:03 – 2008:06 Publication: 2008:08	Regime	Switching Intercept	Switching Exogenous Lag	Error's σ	Change of the Intercept	Change of the Exogenous	Welch Test Result
Autoregressive	r1	0.0060	-	0.0026		-	T=8.44
	r2	0.0032	-	0.0010	45.80%		rejected
	r3	-0.0001	-	0.0012	97.50%	-	T=24.73
	r4	-0.0054	-	0.0019			rejected
Foreign Orders	r1	0.0065	0.0041	0.0023	64.89%	-701.24%	T=18.57
	r2	0.0023	0.0327	0.0018			rejected
	r3	-0.0009	0.0801	0.0002	70.17%	-230.28%	T=1.97
	r4	-0.0029	0.0242	0.0028			rejected
Domestic Orders	r1	0.0099	0.0423	0.0013	69.07%	73.26%	T=21.76
	r2	0.0031	0.0113	0.0027			rejected
	r3	0.0001	0.1650	0.0015	101.58%	239.02%	T=21.13
	r4	-0.0041	-0.1187	0.0024			rejected
Construction Permits ⁵	r1	0.0113	-0.0889	0.0000	54.33%	94.34%	T=17.92
	r2	0.0052	-0.0050	0.0025			rejected
	r3	-0.0015	-0.0317	0.0011	45.07%	259.60%	T=2.20
	r4	-0.0027	0.0198	0.0033			not rejected
CDAX	r1	0.0064	-0.0118	0.0025	38.44%	16.06%	T=9.63
	r2	0.0039	-0.0099	0.0010			rejected
	r3	0.0002	-0.0042	0.0013	103.26%	30.05%	T=19.19
	r4	-0.0052	-0.0061	0.0021			rejected
Corporate Spread	r1	0.0050	0.0263	0.0028	71.20%	-61.62%	T=12.73
	r2	0.0014	0.0425	0.0017			rejected
	r3	-0.0019	0.0487	0.0024	67.75%	187.37%	T=14.72
	r4	-0.0059	-0.0558	0.0012			rejected
Euribor - 3M	r1	0.0065	0.2096	0.0034	47.22%	93.89%	T=7.51
	r2	0.0034	0.0128	0.0027			rejected
	r3	0.0026	0.4194	0.0011	199.79%	619.07%	T=14.49
	r4	-0.0026	-0.0808	0.0034			rejected
ifo Business Climate	r1	0.0057	0.0805	0.0025	26.09%	227.80%	T=15.68
	r2	0.0042	-0.1028	0.0007			rejected
	r3	0.0006	0.1260	0.0015	112.68%	-233.53%	T=8.07
	r4	-0.0044	0.0378	0.0025			rejected
Credit Growth ⁶	r1	0.0075	-0.0485	0.0006	28.99%	122.55%	T=4.37
	r2	0.0053	0.0109	0.0021			rejected
	r3	0.0002	0.0029	0.0021	102.95%	84.84%	T=29.81
	r4	-0.0073	0.0193	0.0010			rejected
Maturity Spread	r1	0.0041	0.0062	0.0032	101.85%	3107.29%	T=3.81
	r2	-0.0001	-0.1861	0.0002			rejected
	r3	-0.0018	0.0017	0.0035	86.00%	98.97%	T=41.57
	r4	-0.0128	0.1674	0.0001			rejected
Job Vacancies	r1	0.0052	0.0232	0.0031	34.37%	-39.98%	T=4.56
	r2	0.0034	0.0325	0.0012			rejected
	r3	-0.0006	0.0765	0.0005	84.27%	9.70%	T=12.24
	r4	-0.0036	0.0847	0.0022			rejected

Table 5: Change of coefficients and Welch test results for the publication on August 2008. At this point in time the model changes from 2 to 4 regimes (r1 - r4) due to criterion (19) so that the relative change between the coefficients linked to the weak and to the strong intensity of an expansion or a recession can be measured for each MSARX specification. Welch test results are also related to the comparison of weak and strong recession or expansion regimes, namely to the difference in the normal distributions.

Certainly financial variables are important for business cycle predictions. But the results in table 5 illustrate that when one runs different specifications, the impact of financial variables should not be overestimated relative to real economy variables.¹²

5.2 Real-time Forecasts with the Industrial Production

This section deals with one-month-ahead real-time forecasts with industrial production as the dependent variable. The methodology for how to produce these forecasts including the real-time change in the number of regimes was described in the previous section. Table 6 contains measures of forecast accuracy for the different MSARX specifications and for their average. MAE stands for the mean absolute value, i.e.

$$MAE = \frac{1}{h} \sum_{t=T+1}^{T+h} |\mathcal{P}_{t|t-1} - b_t|, \quad (21)$$

RMSE for the root mean squared error, i.e.

$$RMSE = \sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1} - b_t)^2} \quad (22)$$

and Theil for the Theil coefficient, i.e.

$$Theil = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1} - b_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} (\mathcal{P}_{t|t-1})^2 / h} + \sqrt{\sum_{t=T+1}^{T+h} b_t^2 / h}}. \quad (23)$$

Theil coefficients are normalized to the unit interval with 0 representing a perfect fit. b_t is the binary variable reporting the state of the business cycle with the benchmark series, while $\mathcal{P}_{t|t-1}$ stands for the predicted state probabilities of the Markov Switching model. Table 6 shows that the average outperforms each of the single forecasts. In this context, Timmermann (2006, p.1) mentions that ‘simple combinations that ignore correlations between forecast errors often dominate more refined’ ones. Yet, the average, as listed in table 6, does not only achieve the best values because of the averaging effect, but also because of the fact that the model changes the number of regimes due to criterion (19). This change takes place with the forecast for September 2008¹⁵.

¹²For instance, mixed and high frequency models often rely strongly on financial data.

¹⁵The title of the Gemeinschaftsdiagnose, the professional opinion of important German economic research institutes, in fall 2008, was ‘Germany on the edge of a recession’. In contrast to this report, the real-time introduction of new regimes for September 2008 could be interpreted as Germany no longer being on the edge, but in the middle of a recession becoming deeper than all previous recessions.

Sample: 2007:02 – 2010:12	MAE	RMSE	Theil
Average	0.2948	0.3898	0.4060
Autoregressive	0.3070	0.4439	0.4309
Foreign Orders	0.3651	0.5007	0.4340
Domestic Orders	0.3141	0.4385	0.4203
Construction Permits	0.3050	0.4394	0.4112
CDAX	0.3234	0.4569	0.4421
Corporate Spread	0.3609	0.4958	0.4921
Euribor - 3M	0.3132	0.4485	0.4537
ifo Business Climate	0.3564	0.4991	0.4860
Credit Growth	0.2559	0.4395	0.4128
Maturity Spread	0.2987	0.4585	0.4279
Job Vacancies	0.3774	0.5038	0.4564

Table 6: Measures of forecast accuracy for the different MSARX specifications. Averaging the forecasts and including a change of the embedded regimes due to criterion (19) leads to the best values. But with the slightly delayed recession recognition, the Theil coefficient does not reach values below 40%.

One reason for the Theil coefficient not reaching lower values is the slightly delayed recession detection, see figure 5¹⁶. Compared to the recession start in March 2008, as it is reported by the benchmark model in October 2008 (7 months later), the MS recession probability exceeds the 0.5 threshold in August 2008. On the one hand, this represents a delay of 5 months. On the other hand, the recession is recognized significantly earlier by the MS model than by the benchmark method. As the forecast for August is made in July (one month forecasting horizon), the time in advance between the MS and the benchmark model stands at 3 months. Additionally, the recession probability forecast for July 2008 is above 30%. Such an indication of a recession cannot be provided by the nonparametric benchmark method since this focuses on a binary decision (1=recession, 0=expansion). Both MS and ex-post dating algorithms continuously announce the last recession between August 2008 and April 2009. Compared to the end of the recession in April 2009, as it is reported by the benchmark series in November 2009, the MS forecast for July 2009 is the last above the 0.5 threshold (delay of 3 months). Again considering the point in time when the information about the end of the recession was provided, the MS model is 4 months earlier than the benchmark series.¹⁷ Thus the forerun of the MS is even longer in the case of the

¹⁶Another reason is that, compared to the binary series, the averaged MS state probabilities are likely to cover only a certain range between 0 and 1. Thus, the measures of forecast accuracy could only be improved by generating a binary variable out of the MS state probabilities according to the 0.5 threshold.

1M - Real-time Forecast for Regime Probabilities of Production

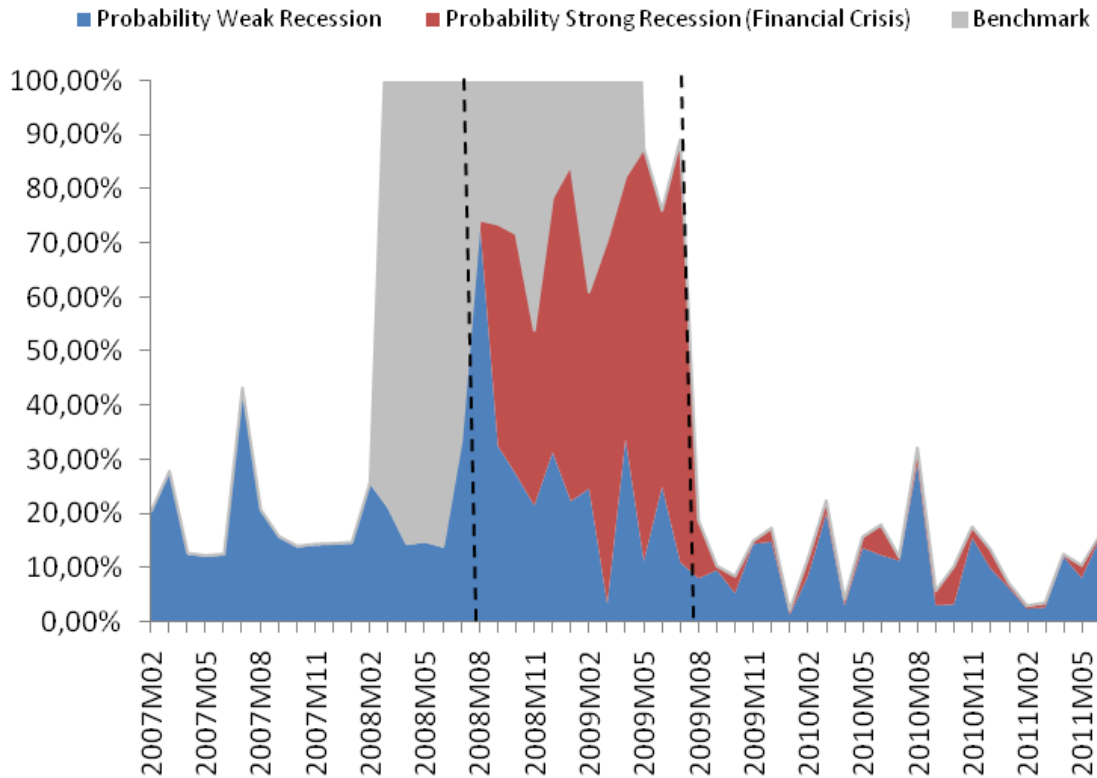


Figure 5: The regime probabilities are averaged one month ahead real-time forecasts over the different MSARX specifications. Probabilities of weak and strong recession intensity are added up to a total recession probability. For September 2008, the model changes from 2 to 4 regimes due to the criteria described in section 5.1, where the new regime clearly points to the magnitude of the economic decline. When defining a recession on the 0.5 threshold, the downturn is predicted continuously between August 2008 and July 2009 (dashed lines). Actually this represents a delay to the recession start, as it is reported later on by the ex-post dating algorithm. But considering the point in time when the recession is recognized, the MS model is 3 months earlier than the benchmark method.

end of the recession than the beginning. Together with the aggregated recession probability not reaching values above 90%, this reveals a certain restraint towards an erroneous recession declaration, which seems to be functional with respect to the forecast accuracy. Indeed

¹⁷The end of the recession has to be seen technically. For example this does not mean that at this point in time the former output level was already again reached.

there is no extra period in the out-of-sample evaluation where the economic situation is misinterpreted as a recession.

5.3 Real-time Forecasts with the CLI

Although the MS model outperforms the ex-post dating algorithm, the question arises as to whether there was any alternative in recognizing the recession in advance. Lahiri and Wang (1994) showed how to apply the MS model to the Composite Index of Leading Economic Indicators (CLI). Such monthly data for Germany is provided by OECD (2011). The idea behind the CLI is to generate a synthetic series that represents a lead to the business cycle and anticipates its turning points. To achieve this, leading indicators - similar to the way they are used as regressors in the previous estimations - are aggregated. That is why the CLI is also subject to revisions. Before aggregation, the data is seasonally adjusted, outliers are eliminated, trends are removed, and filters for smoothing and normalization are applied in order to obtain homogenized cyclical amplitudes for each of the component series. It is not the topic of this paper to discuss the OECD methods in detail, but it turns out that the procedure above leads to the CLI often exhibiting relative undecidedness between up- and downturns on the current edge. Nevertheless with the Hamilton filter generating the state probabilities endogenously out of the observations, and with the result of lagged recession recognition in the case of the industrial production, it is quite appealing to run a specification where the above-mentioned reference series is substituted by the CLI.

In doing so, some differences to the previous MS regressions have to be considered. Firstly, smoothing backwards by a moving average is no longer necessary since the series is already smoothed. Secondly, the lag choice, described in section 4.3, only makes sense for a purely autoregressive specification since there must be a bias with leading indicators standing on both sides of the equation in a different manner (aggregated versus disaggregated). In fact the lag choice results in favoring no autoregressive terms so that the right hand side of the equation only consists of a switching intercept and an error term. As a consequence of this parsimonious design it is not sufficient to choose the regimes with the lowest intercepts to stand for recessions, but to request that these intercepts be negative. In figure 6 the MS regression with the CLI delivers an early signal for the recession linked to the financial crisis, but forecasts are very volatile. Among the real-time out-of-sample predictions between February 2006 and June 2011 there are three periods (8 months), in which the economic situation is misinterpreted as a recession. The early signal is given six months in advance,

1M - Real-time Forecast for Regime Probabilities of CLI

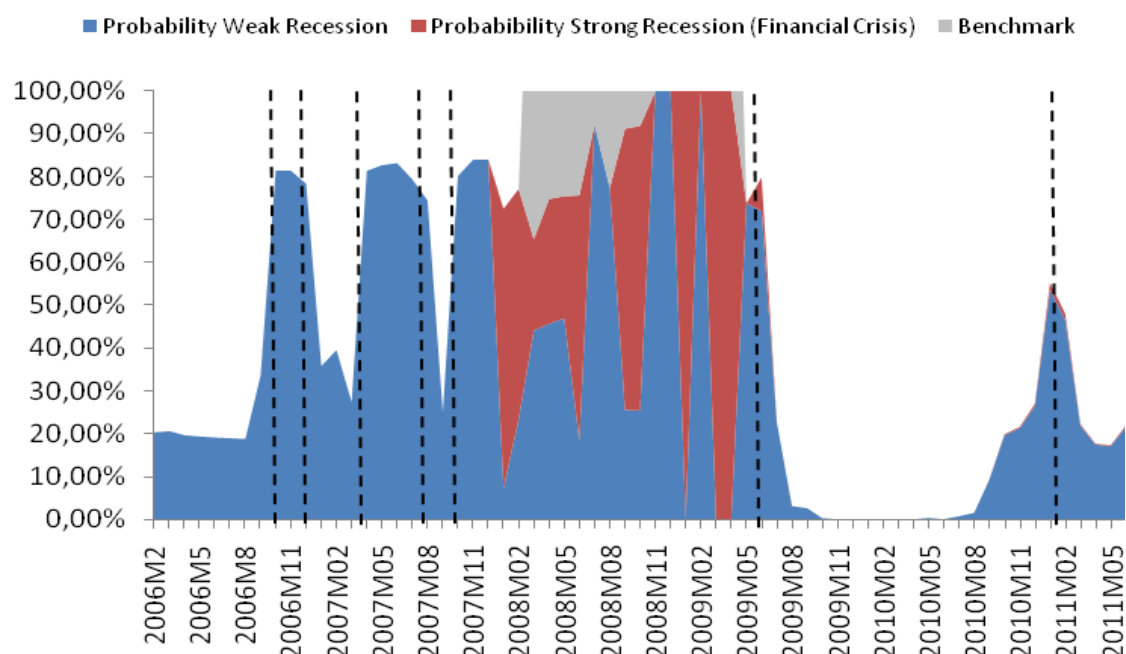


Figure 6: The regime probabilities are one-month-ahead real-time forecasts with no autoregressive terms. Probabilities of weak and strong recession intensity can be added up to a total recession probability. For December 2007 the model changes from 2 to 4 regimes according to the criteria described in section 5.1, where the new regime clearly points to the magnitude of the economic decline. When defining a recession on the 0.5 threshold, there are three periods (October 2006 - December 2006, April 2007 - August 2007 and January 2011), in which the economic situation is misinterpreted as a recession. Among the very volatile forecast results an early indication of the approaching recession can be identified in October 2007. Compared to the beginning of the recession in March 2008 as is reported by the benchmark series, this early signal would represent a lead of 6 months. The recession is then predicted continuously between October 2007 and June 2009 (dashed lines).

whereas a timely signal in accordance with the forecast horizon would have to come one month ahead of the publication point. This reveals a general problem with CLI data: Given the high number of misinterpreted recessionary phases it is not clear whether the forerun is stable and always equal to 6 months as is claimed by the OECD. Nevertheless - from an operational point of view - MS regressions on the CLI and on the industrial production

can be complementary. While the former might signal the recession in advance, the latter can confirm it accurately a short time after its onset.

6 Conclusion

This paper uses a Markov Switching framework applied to German monthly real-time data. While the appropriateness of the method for business cycle applications is well-known since Hamilton's innovation in 1989, based on current literature there are some new insights which can be fruitfully applied to monthly German real-time data and which are helpful from an operational point of view:

Given limited data records, it is appealing to connect Timmermann (2006)'s idea of a single forecast being subject to a misspecification bias with the Markov Switching model generating each of the single forecasts. In order to reduce the bias, forecasting results are averaged. When generating the forecasts as above, several macroeconomic and financial leading indicators serve as exogenous variables in univariate MSARX regressions. This design also opens up room for the evaluation of the predictive properties of single indicators or a group of indicators. In the paper at hand such statements are:

In the MSARX specification with credit growth, as it is reported in the ESCB statistics from the balance sheets of monetary financial institutions including interbank deals, most of the regime and publication dependent coefficients turn out to be significant. Thus, this variable deserves further consideration as a potential predictor of the business cycle.

In general, we do not find that leading indicators which are available immediately perform better than those which are subject to a publication lag. Although intuition tells us that the last recession had its origin on financial markets, in real-time predictions we do not find financial variables in general reacting more sensitively than real economy variables. This stresses the fact that in a business cycle model the role of financial variables should not be overestimated, for instance by including a similar number of financial and real economy time series. This finding may also be relevant for the increasing number of mixed-frequency models, often influenced by the greater availability of financial market data.

Allowing the MS model to change the number of embedded regimes in real-time stabilizes forecasting results. By introducing a criterion for the real-time regime change it is also possible to determine the point in time from which the recession after the financial crisis structurally exceeded the previous ones. In our analysis this turns out to be September 2008, where the forecast is made in August - one month before the investment bank Lehman

Brothers declared bankruptcy.

When selecting the industrial production as a proxy of overall economic activity, six recessions (or periods of stagnation) can be found in the observation period from March 1994 to September 2010. All in all this fits and extends the suggestion for a German business cycle chronology by Schirwitz (2009), when considering her disaggregated results and accepting small time shifts since she used quarterly GDP instead of monthly industrial production. The MS real-time forecast from February 2007 to Mai 2011 for industrial production outperforms a non-parametric ex-post-dating method based on the work of Bry and Boschan (1971) as well as Harding and Pagan (2002), while revealing similar characteristics: On the one hand recession start and end are recognized too late, while the delay for the end of the recession is considerably shorter. On the other hand, at least when considering different specifications and changing the number of regimes, no business cycle phase is misinterpreted as a recession. This fits the fact that Hamilton (2011) compares his MS results with the time of the NBER announcements, which are usually also made several months after the beginning and end of the recession in order to provide the official dating as accurately as possible.

In order to counterbalance the aforementioned inertia of the MS model in the case of industrial production, it is appealing to apply it to the OECD Composite Index of Leading Indicators, which seeks to be a leading proxy for the business cycle. In doing so the findings for the combination of the MS model and a leading index are more ambivalent. On the one hand, we can confirm Lahiri and Wang (1994)'s result of obtaining an early signal. On the other hand, predictions are of low quality since in several instances recessions are mistakenly declared. Finally several extensions of our MS framework are possible, such as the inclusion of different forecast horizons.

References

- Bates, J. and Granger, C. (1969), ‘The combination of forecasts’, *Operations Research Quarterly* **20**, 451–468.
- Biggs, M., Mayer, T. and Pick, A. (2009), Credit and economic recovery, Working Paper 218, De Nederlandsche Bank.
- Boldin, M. (1996), ‘A Check on the Robustness of Hamilton’s Markov Switching Model Approach to the Economic Analysis of the Business Cycle’, *Studies in Nonlinear Dynamics and Econometrics* **1**, 35–46.
- Breunig, R., Najarian, S. and Pagan, A. (2003), ‘Specification Testing of Markov Switching Models’, *Oxford Bulletin of Economics and Statistics* **65**, 703–725.
- Brock, W., Dechert, W., Scheinkman, J. and LeBaron, B. (1996), ‘A test for independence based on the correlation dimension’, *Econometric Reviews* **15**, 197–235.
- Bry, G. and Boschan, C. (1971), *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, National Bureau of Economic Research, Inc.
- Carrasco, M., Hu, L. and Ploberger, W. (2005), ‘Optimal Test for Markov Switching’, <http://www.econ.sinica.edu.tw/upload/file/20060725.pdf>. Accessed on 18 July 2011.
- Chauvet, M. and Hamilton, J. D. (2005), Dating business cycle turning points, in ‘Non-linear Time Series Analysis of Business Cycles’, Van Dijk, Milas and Rothman, ed., pp. 1–54. Elsevier’s Contributions to Economic Analysis Series.
- Chauvet, M. and Piger, J. (2008), ‘A comparison of the real time performance of business cycle dating methods’, *Journal of Business and Economic Statistics* **26**, 42–49.
- Chauvet, M. and Potter, S. (2005), ‘Forecasting recessions using the yield curve’, *Journal of Forecasting* **24**, 77–103.
- Deutsche Bundesbank (2011a), ‘Financial Data Deutsche Bundesbank’, http://www.bundesbank.de/statistik/statistik_zinsen.php. Accessed on 11 June 2011.
- Deutsche Bundesbank (2011b), ‘MFI Balances Sheets Deutsche Bundesbank’, http://www.bundesbank.de/statistik/statistik_eszb_tabellen.php. Accessed on 11 June 2011.

- Deutsche Bundesbank (2011c), ‘Realtime Data Deutsche Bundesbank’, http://www.bundesbank.de/statistik/statistik_realtime.php. Accessed on 11 June 2011.
- Diebold, F. X. and Rudebusch, G. D. (1991), ‘Forecasting output with the composite leading index: A real-time analysis’, *Journal of the American Statistical Association* **86**, 603–610.
- Durland, J. M. and McCurdy, T. H. (1994), ‘Duration-Dependent Transitions in Markov Model of U.S. GNP Growth’, *Journal of Business and Economic Statistics* **1**, 279–288.
- ECRI (2012), ‘International Business Cycle Dates’, <http://www.businesscycle.com>. Accessed on 18 December 2012.
- Estrella, A. and Hardouvelis, G. A. (1991), ‘The term structure as a predictor of real economic activity’, *Journal of Finance* **46**, 555–576.
- Friedman, B. M. and Kuttner, K. N. (1992), ‘Money, income, prices and interest rates’, *American Economic Review* **82**, 472–492.
- Fritsche, U. and Stephan, S. (2002), ‘Leading indicators of German business cycles. an assessment of properties’, *Jahrbuecher fuer Nationaloekonomie und Statistik* **222**, 289–315.
- Garcia, R. (1998), ‘Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models’, *International Economic Review* **39**, 763–788.
- Gilchrist, S., Yankov, V. and Zakrajsek, E. (2009), ‘Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets’, *Journal of Monetary Economics* **56**, 471–493.
- Hamilton, J. D. (1989), ‘A new approach to the economic analysis of nonstationary time series and the business cycle’, *Econometrica* **57**, 357–384.
- Hamilton, J. D. (1990), ‘Analysis of time series subject to changes in regime’, *Journal of Econometrics* **45**, 39–70.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.

- Hamilton, J. D. (2011), ‘Calling Recessions in Real Time’, *International Journal of Forecasting* **27**, 1006 – 1026.
- Hansen, B. E. (1996), ‘Inference when a nuisance parameter is not identified under the null hypothesis’, *Econometrica* **64**, 413–430.
- Harding, D. and Pagan, A. (2002), ‘Dissecting the cycle: a methodological investigation’, *Journal of Monetary Economics* **49**, 365–381.
- ifo Institut für Wirtschaftsforschung (2011), ‘ifo Geschäftsklimaindex’, <http://www.cesifo-group.de>. Accessed on 11 June 2011.
- Kim, C. (1994), ‘Dynamic linear models with Markov - switching’, *Journal of Econometrics* **60**, 1–22.
- Krolzig, H. M. (1997), *Markov Switching Vector Autoregressions. Statistical Inference and Application to Business Cycle Analysis*, Springer Science+Business Media.
- Lahiri, K. and Wang, J. (1994), ‘Predicting Cyclical Turning Points with Leading Index in a Markov Switching Model’, *Journal of Forecasting* **13**, 245–263.
- Lee, C., Liang, C. and Chou, H. (2009), ‘Identifying Turning Points of Real Estate Cycles in Taiwan’, *Journal of Real Estate Portfolio Management* **15**, 75–85.
- Levanon, G. (2010), ‘Evaluating and comparing leading and coincident economic indicators’, *Business Economics* **45**, 16–27.
- Maheu, J. M. and McCurdy, T. H. (2000), ‘Identifying bull and bear markets in stock returns’, *Journal of Business and Economic Statistics* pp. 100–112.
- Meeks, R. (2011), Do credit market shocks drive output fluctuations ? Evidence from corporate spreads and defaults, Working Paper 0906, Federal Reserve Bank of Dallas.
- NBER (2011), ‘Determination of the 2007 Peak in Economic Activity’, <http://www.nber.org/cycles/dec2008.html>. Accessed on 10 January 2011.
- Newey, B. and West, D. (1987), ‘A Simple, Positive semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix’, *Econometrica* **55**, 347–370.
- OECD (2011), ‘Realtime Data OECD’, <http://stats.oecd.org/mei/default.asp?rev=1>. Accessed on 11 June 2011.

- Perlin, M. (2012), MS_Regress - A Package for Markov Regime Switching Models in Matlab, SSRN Discussion Paper 1714016. <http://ssrn.com/abstract=1714016>.
- Proaño, C. (2010), Recession forecasting with dynamic probit models under real time conditions, IMK Working Paper 10-2010, Macroeconomic Policy Institute.
- Schirwitz, B. (2009), 'A comprehensive German Business Cycle Chronology', *Empirical Economics* **37**, 287–301.
- Schreiber, S., Theobald, T., Proano, C., Stephan, S., Rietzler, K. and Detzer, D. (2012), Verfahren der konjunkturellen Wendepunktbestimmung unter Berücksichtigung der Echtzeitproblematik, IMK Study 27, Macroeconomic Policy Institute.
- Stock, J. H. and Watson, M. W. (1989), New indexes of coincident and leading economic indicators, *in* 'NBER Macroeconomics Annual', Vol. 4, O. Blanchard and S. Fischer (Eds.), pp. 351–409. Cambridge: The MIT Press.
- Timmermann, A. (2006), Forecast combinations, *in* 'Handbook of Economic Forecasting', Vol. 1, G. Elliot, C. Granger and A. Timmermann (Eds.), pp. 135–196. Amsterdam: North-Holland.