CESifo WORKING PAPERS

6457
2017

April 2017

Predicting Ordinary and Severe Recessions with a Three-State Markov-Switching Dynamic Factor Model. An Application to the German Business Cycle

Kai Carstensen, Markus Heinrich, Magnus Reif, Maik H. Wolters

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6457*
Category 6: Fiscal Policy, Macroeconomics and Growth

Predicting Ordinary and Severe Recessions with a Three-state Markov-Switching Dynamic Factor Model.

An Application to the German Business Cycle

Abstract

We estimate a Markow-switching dynamic factor model with three states based on six leading business cycle indicators for Germany preselected from a broader set using the Elastic Net soft-thresholding rule. The three states represent expansions, normal recessions and severe recessions. We show that a two-state model is not sensitive enough to reliably detect relatively mild recessions when the Great Recession of 2008/2009 is included in the sample. Adding a third state helps to clearly distinguish normal and severe recessions, so that the model identifies reliably all business cycle turning points in our sample. In a real-time exercise the model detects recessions timely. Combining the estimated factor and the recession probabilities with a simple GDP forecasting model yields an accurate nowcast for the steepest decline in GDP in 2009Q1 and a correct prediction of the timing of the Great Recession and its recovery one quarter in advance.

JEL-Codes: C530, E320, E370.

Keywords: Markov-Switching Dynamic Factor Model, business cycles, Great Recession, leading indicators, turning points, GDP-nowcasting, GDP-forecasting.

Kai Carstensen
University of Kiel / Germany
carstensen@stat-econ.uni-kiel.de

Markus Heinrich
University of Kiel / Germany
m.heinrich@stat-econ.uni-kiel.de

Magnus Reif
ifo Institute / Munich / Germany
reif@ifo.de

Maik H. Wolters
University of Jena / Germany
maik.wolters@uni-jena.de

20th April 2017
We thank Rudi Bachman, Christian Schumacher, and participants of the Ifo macro seminar for helpful comments.

1 Introduction

The failure of macroeconomists to predict the Great Recession of 2008 and 2009 has evoked much public criticism. While the debate mostly focuses on the state of macroeconomic modeling, it has also raised the question why professional forecasters even at the onset of the Great Recession did not foresee the steep output contraction that loomed around the corner. The case of Germany illustrates this failure. It was not until November 2008 that professional forecasters started predicting a recession despite clear warning signals accumulating throughout the year 2008.[1] For example, the expectation component of the Ifo business climate index—viewed by professional forecasters as one of the most important early indicators for German GDP—began its descent already in June 2007 and plunged heavily in July 2008, well before GDP plummeted in the fourth quarter of 2008 and the first quarter of 2009.

In this paper we take up the debate and ask whether it is possible to reliably predict in real time both business cycle turning points and GDP growth rates around these turning points, particularly during the Great Recession episode. We focus on Germany as a representative of the group of countries that show little persistence in GDP growth (other countries with this characteristic are, *inter alia*, Japan, Australia, and Norway). The lack of persistence is important because the usual approach to predict GDP growth by augmenting an autoregressive distributed lag model with a business cycle measure derived from coincidence indicators (see, e.g., Chauvet and Potter, 2013) works well only for countries like the US that exhibit significant sample autocorrelations.[2] As a more promising approach for low-persistence countries we suggest to directly exploit the information of leading indicators for future GDP. For Germany we show that this yields very competitive one-quarter ahead forecasts of business cycle turning points and GDP growth.

To extract information from leading indicators of the German business cycle, we use the Markov-switching dynamic factor model (MS-DFM) proposed by Diebold and Rudebusch (1996) and Kim and Yoo (1995) because it has been shown to be a valuable device for assessing the state of an economy (Chauvet, 1998; Kim and Nelson, 1998; Camacho et al., 2014) and its results are much more timely available than those of simple benchmark approaches such as the Bry-Boschan algorithm. However, unlike the previous literature we specify the MS-DFM with three states. Specifically, we add to the conventional expansion and (ordinary) recession states a third state which reflects a severe recession.[3] This is motivated both by the general perception that the Great Recession was different from previous post-war recessions and may thus require a special

[1] See Drechsel and Scheufele (2012) for an analysis of the performance of leading indicators during the financial crisis and Heilemann and Schnorr-Bäcker (2016) for a detailed documentation of the chronological sequence of data releases and publications of professional forecasts in 2008.

[2] The large difference in the persistence of GDP growth across countries, explanations for these differences and its implications for forecasting are hardly discussed in the literature at all. One exception is Stock and Watson (2005) who document differences in persistence of GDP growth across countries.

[3] Three-state Markov-switching models have been applied mainly to the US (Boldin, 1996; Layton and Smith, 2000; Krolzig and Toro, 2001; Ferrara, 2003; Nalewaik, 2011; Ho and Yetman, 2012) but also to the euro area (McAdam, 2003; Artis et al., 2004; Anas et al., 2008). However, they have been implemented in univariate and vectorautoregressive contexts but not in a dynamic factor model. In addition, these papers intend to identify a recession, a normal growth regime, and a high growth regime, the latter being typically interpreted as a recovery in the line of Sichel (1994) and Morley and Piger (2012). The only exception is Hamilton (2005) who identifies a severe recession regime in a univariate model of the US.

econometric treatment, and by our empirical finding documented below that a MS-DFM with two states becomes instable in 2008.

We also address the question of how to determine the number of states in real time. This is highly relevant as the severe recession state is only weakly identified before the Great Recession which is probably why studies analyzing pre-2008 data report that the German business cycle can well be represented with two states (Bandholz and Funke, 2003; Artis et al., 2004; Kholodilin and Siliverstovs, 2006). We propose to choose—at each point in real time—the number of states that optimizes the quadratic probability score which measures how well the MS-DFM fits the Bry-Boschan algorithm. Thereby, we effectively train the MS-DFM to yield results close to a simple benchmark but at the same time exploit its advantage to detect turning points instantaneously at the sample end.

Another methodological contribution to the literature is to prepend a flexible indicator selection procedure to the MS-DFM. This is important because there are many potentially useful business cycle indicators available for an economy to be fed into the MS-DFM, while the nonlinear one-step estimation approach by Kim and Yoo (1995), which simultaneously determines the factor and the state probabilities, is subject to numerical problems if the number of parameters is large.[4] We use a soft thresholding procedure that accounts for multivariate correlations among the variables to extract a small number of variables from a medium-sized set of pre-selected indicators because Bai and Ng (2008) show that hard thresholding, i.e., using statistical tests to ensure that a predictor is significant irrespective of other predictors, might be inadequate in such situations. Specifically, we use the elastic net (EN) algorithm of Zou and Hastie (2005), which is a convex combination of a ridge regression and a Least Absolute Shrinkage and Selection Operator (LASSO). It is suited particularly for datasets with highly correlated variables like business cycle indicators.

We structure our empirical analysis in two parts. We first study whether the MS-DFM reasonably describes the German business cycle ex post using revised data for the period January 1991 to June 2016. Subsequently, we examine how well the MS-DFM is suited to timely detect, and predict, business cycle turning points in real time. In both parts, we compare the properties of models with two and three states, emphasizing the Great Recession period.

In the ex-post analysis presented in Section 4, we apply the EN algorithm to select three out of 16 hard indicators such as new orders and three out of 19 survey indicators, all of which have been considered as early indicators in the literature on German business cycle dynamics. We then feed these six indicators in one-factor MS-DFMs with two and three states, estimate the parameters, and smooth out the factors, which can be interpreted as composite leading indicators, and the conditional state probabilities. It turns out that the three-state model is superior in several dimensions. Its factor correlates more strongly with GDP growth (if aggregated to the quarterly frequency) and its states can be interpreted nicely as expansion, ordinary recession, and severe

[4] While this problem can be circumvented by a two-step approach which first extracts a linear factor from the dataset and subsequently uses this factor to estimate a univariate Markov-switching model, Camacho et al. (2015) argue that the one-step method is—although it involves a higher computational burden—more robust against misspecification. Furthermore, Doz and Petronevich (2016) compare the performance of both methods on dating French business cycle turning points and find that one-step estimation is more precise in indicating the beginning and end of recessions.

recession, while the two-state model seems to identify a low-growth regime and a medium-severe recession that is too fierce for any pre-2008 downturn and to mild for the Great Recession. The three-state model also dates recessions in general, and the Great Recession in particular, much more in line with conventional wisdom and the Bry-Boschan algorithm.[5] In contrast, the two-state model is less sensitive and thus typically comes a bit late because the business cycle needs to deteriorate considerably before it is classified as medium-severe recession.

In Section 5 we present the second part of our empirical analysis. We ask whether the superiority of the three-state model carries over to a forecasting situation in real time in which the data exhibit a ragged-edge structure and the Bry-Boschan algorithm is not suited because its standard version requires a lag of at least 5 months until it is able to signal a turning point. To this end, we set up a recursive nowcasting exercise from January 2001 to June 2016 that in each month selects six indicators by means of the EN algorithm and estimates one-factor MS-DFMs with two and three states. We find that the two-state model signals turning points fairly well but becomes instable during the Great Recession, while the three-state model appears poorly identified before the Great Recession but works properly thereafter. These results suggest that a forecaster would have dismissed the two-state model after the Great Recession and moved towards the three-state model. To operationalize this, we use real-time model selection based on the quadratic probability score and the BIC which yields a combined two-state/three-state model. It produces precise and timely nowcasts of business cycle turning points.

Using a recursive out-of-sample forecasting exercise we even demonstrate that the combined model is able to provide excellent 3-month ahead turning point predictions that would have been extremely useful for policy makers during the Great Recession. In particular, it predicts an upcoming recession with almost 100 percent probability already in July 2008 and thus four months ahead of most professional forecasters. Moreover, in March 2009 it correctly predicts that the recession comes to an end soon, one month before the German public started to discuss a third stimulus package.

Finally, we assess whether point forecasts of German GDP growth rates benefit from including the information provided by the MS-DFMs. Specifically, we augment an autoregressive forecasting model with the dynamic factor and the recession probabilities extracted from the early indicators. Specifically, augmenting an autoregressive forecasting model with the dynamic factor and the recession probabilities extracted from the early indicators considerably improves nowcasts and short-term forecasts, especially during recessions. In particular, it yields an accurate nowcast for the steepest decline in GDP in 2009Q1.

This paper adds to the literature that applies Markov-switching models to the German business cycle. Ivanova et al. (2000) estimate univariate Markov-switching models for various interest rate spreads and examine their predictive power for business cycle turning points. Bandholz and

[5]We apply the Bry-Boschan algorithm because there is no widely accepted monthly business cycle chronology for the German economy available against which we can assess the results of our MS-DFM. The chronology published by the Economic Cycle Research Institute (ECRI) is based on both an unknown data set and an unknown method and is provided with a lag of approximately one year. The business cycle dates published by the OECD are determined by applying the Bry-Boschan algorithm on the OECD's composite leading indicator on a quarterly basis. A useful proposal is made by Schirwitz (2009) who suggests a consensus business cycle chronology based on the results of different methods. However, it is again on the quarterly frequency. Hence, we use the Bry-Boschan algorithm applied to monthly industrial production as a benchmark.

Funke (2003) use a MS-DFM model with a bivariate dataset to construct a leading indicator for the German business cycle. Kholodilin (2005) augment that model with a second factor and interpret it as a coincidence indicator. Abberger and Nierhaus (2010) demonstrate the predictive power of the Ifo business climate index with regard to business cycle turning points in a univariate framework. None of these contributions considers a flexible data selection approach based on a large dataset or a distinction between severe and ordinary recessions. Moreover, they are based on revised data, while we analyse the predictive ability of the model in a real-time setting.

The remainder of the paper is structured as follows: Section 2 outlines our baseline MS-DFM model and the estimation method. In section 3 we describe our dataset and the variable selection procedure. Section 4 and 5 presents our estimation results as described above. Finally, section 6 concludes.

2 The Markov-switching dynamic factor model

We use a Markov-switching dynamic factor model (MS-DFM) to extract common nonlinear business cycle dynamics from a set of leading indicators. We distinguish between n_h hard indicators, $y_{it}^{(h)}$, such as new orders, interest rates, and oil prices, which typically account for rather short-term fluctuations, and n_s survey indicators, $y_{it}^{(s)}$, such as the Ifo business climate index and the ISM purchasing managers index which capture primarily medium-term business cycle dynamics. The distinction is important because quarterly growth rates of hard indicators generally correlate well with quarter-on-quarter GDP growth (and monthly growth rates of hard indicators correlate with monthly business cycle indicators like industrial production), while business surveys typically rather fit year-on-year GDP growth. We model these differences along the lines of Camacho et al. (2014): For the hard indicators we assume a standard factor structure,

$$y_{i,t-l_{h,i}}^{(h)} = \gamma_i^{(h)} f_t + z_{it}^{(h)}, \qquad i = 1, \ldots, n_h, \tag{1}$$

where $z_{it}^{(h)}$ is an idiosyncratic component, f_t is a scalar dynamic factor that leads the month-on-month business cycle dynamics by three months, and $l_{h,i}$ is the lag with which leading indicator $y_{it}^{(h)}$ enters the model. For the survey indicators we assume a slightly different specification,

$$y_{i,t-l_{s,i}}^{(s)} = \gamma_i^{(s)} \sum_{k=0}^{11} f_{t-k} + z_{it}^{(s)}, \qquad i = 1, \ldots, n_s, \tag{2}$$

where $z_{it}^{(s)}$ is an idiosyncratic component and $l_{s,i}$ is the lag with which leading indicator $y_{it}^{(s)}$ enters the model. We include the sum of lags 0 to 11 of the factor as a parsimonious way to incorporate the phase shift associated with a year-on-year growth cycle that correlates with the survey indicators.

For all indicators, we take into account that they lead the cycle to different extents and thus should enter the factor model with different lags $l_{h,i}$ and $l_{s,i}$. To make the factor lead the business cycle by 3 months, we include indicators that lead GDP by 1, 2, and 3 quarters with a lag of 0, 3, and 6 months, respectively (in Section 3 below we describe in detail how we choose the indicators and their lags).

Following Doz and Petronevich (2016), we model the vector of idiosyncratic components, $z_t = [z_{1t}^{(h)}, \dots, z_{n_h t}^{(h)}, z_{1t}^{(s)}, \dots, z_{n_s t}^{(s)}]'$, as a diagonal VAR process of lag order q,

$$z_t = \psi_1 z_{t-1} + \cdots + \psi_q z_{t-q} + \varepsilon_t, \qquad \varepsilon_t \sim i.i.d\, N(0, \Sigma_z), \tag{3}$$

where $\psi_1, \dots, \psi_q$ and Σ_z are diagonal matrices, and ε_t is a vector of independent Gaussian shocks. We specify the common factor as an autoregressive process of lag order p with regime-dependent intercept,

$$f_t = \beta_{S_t} + \phi_1 f_{t-1} + \cdots + \phi_p f_{t-p} + \eta_t \qquad \eta_t \sim i.i.d\, N(0, 1), \tag{4}$$

where η_t is an independent Gaussian shock. The intercept, β_{S_t}, depends on the state variable $S_t \in \{1, ...m\}$ as follows:

$$\beta_{S_t} = \beta_1 S_{1,t} + \beta_2 S_{2,t} + \cdots + \beta_m S_{m,t},$$

where $S_{m,t}$ is equal to unity if S_t equals m and zero otherwise. We assume that S_t follows a first-order ergodic Markov chain. The corresponding $m \times m$ transition matrix, Π, has elements p_{ij} defining the probability to switch from regime i to regime j, with $\sum_{j=1}^{m} p_{ij} = 1$ for every $i = 1, \dots, m$. We do not impose restrictions on the duration of any regime. We consider models with two regimes ($m = 2$) that represent expansions and recessions and with three regimes ($m = 3$) with the aim to distinguish in addition between ordinary and severe recessions.

Defining the vector $y_t = [y_{1,t-l_{h,1}}^{(h)}, \dots, y_{n_h,t-l_{h,n_h}}^{(h)}, y_{1,t-l_{s,1}}^{(s)}, \dots, y_{n_s,t-l_{s,n_s}}^{(s)}]'$ of dimension $n = n_h + n_s$, we cast the model into state-space form,

$$y_t = B a_t \tag{5}$$

$$a_t = \mu_{S_t} + F a_{t-1} + R \omega_t, \tag{6}$$

where a_t is the state vector, ω_t is a vector of independent Gaussian shocks with mean zero and covariance matrix Q, B, F and R are coefficient matrices, and μ_{S_t} is a state-dependent intercept. For details, we refer to Appendix 7.1.

We estimate the MS-DFM by maximizing the likelihood function.[6] To this end, we employ the filter proposed by Kim (1994), see Appendix 7.2 for details. It yields the latent dynamic regime dependent factor as well as the Markov-switching probabilities.

We use the following starting values. In a first step, we approximate f_t by a static principal components analysis and plug it into (4) with invariant intercept to estimate starting values for ϕ_1 to ϕ_p by OLS. We also plug f_t into (1) and (2), and run OLS regressions to obtain starting values for $\gamma^{(h)}$ and $\gamma^{(s)}$. The residuals of these regressions approximate the idiosyncratic components z_t. We use them to estimate a diagonal VAR model of lag order q to find starting values for ψ_1 to ψ_q, and Σ_z. In the next step, we take all these values to initialize and estimate a dynamic factor model with a single regime. This yields starting values for $\gamma^{(h)}$, $\gamma^{(s)}$, $\phi_1, \dots, \phi_p$, $\psi_1, \dots, \psi_q$ and Σ_z. Finally, combining the results of the single-regime model with starting values for the

[6]We use the Matlab globalsearch class based on the routine fmincon to obtain a global maximum.

transition matrix and the regime dependent means completes the set of required parameters. Specifically, we initialize the transition matrix by assuming persistent regimes (high values on the main diagonal and small values on the off-diagonal). We construct starting values for the regime dependent means as follows. In case of the two-state model we take the average over all positive factor values and the average of all negative factor values for the expansion and recession regime, respectively. For the three state model we use the same approach and take in addition the smallest factor value in the sample as starting value for the mean of the severe recession regime.

3 Indicator selection

While there are many business cycle indicators available for the German economy, the challenge is to reduce their number such that they carry all necessary cyclical information without overburdening the nonlinear maximum likelihood technique described above with estimating too many parameters. Boivin and Ng (2006) demonstrate that even linear factor models do not always benefit from adding more and more variables in particular in the context of forecasting. Camacho et al. (2015) focus specifically on Markov-Switching Dynamic Factor models and show that once a small number of high quality indicators is included that adding more indicators yields only small improvements in terms of the identification of business cycle turning points. Finally, Schumacher (2010) shows in application specifically for factor-based forecasting German GDP that preselecting a set of targeted predictors can improve prediction accuracy. Hence, we first pre-select a medium-sized set of potentially useful indicators based on previous results in the literature and then apply to it a variable selection algorithm that chooses only a few final indicators to be fed into the MS-DFM.

Our pre-selection is primarily based on previous results of the literature (Fritsche and Stephan, 2002; Kholodilin and Siliverstovs, 2006; Drechsel and Scheufele, 2012; Lehmann and Wohlrabe, 2016) on the German business cycle. As hard indicators we choose 6 industrial order inflow series, 2 commodity prices, 3 interest rates, the German contribution to the EMU M2, and the DAX index which have all been found to give early business cycle signals. To take into account Germany's dependence on foreign markets, we also include US industrial production as a simple indicator for world market fluctuations. We finally add German consumer prices and employment as important economic state variables even though they are typically thought to lag the business cycle. We leave it to the indicator selection algorithm below to decide whether they are promising candidates. As survey indicators we pre-select 9 series published by the European commission and 7 series published by the Ifo institute. These series cover a broad range of economic activity, with a specific focus on expectations. We add the purchasing manager index for the US, the Belgium business confidence indicator—which is sometimes found to lead the EU cycle, see Vanhaelen et al. (2000)—and the Euro-coin index to reflect the importance of major foreign markets.[7] Altogether, we pre-select a set of 35 monthly business cycle indicators, of

[7] Although it consists of both hard and survey indicators, the Euro-coin index is assigned to the survey category because it exhibits, as the other survey indicators, the highest correlation with year-on-year GDP growth (Altissimo et al., 2010).

which 16 are categorized as hard and 19 as survey indicators. To ensure stationarity, we apply log differencing to all hard indicators while the survey indicators are stationary by construction. The indicators are then standardized to mean zero and variance one. A complete description is provided in Tables 7 and 8 in the Appendix. Our sample starts in January 1991 in order to avoid any issues associated with the German reunification break, and runs until June 2016.

Based on the pre-selected data set, we employ an automatic indicator selection algorithm. As our goal is to provide early signals for business cycle turning points, the algorithm should select only those hard indicators that exhibit a strong lead correlation with quarter-on-quarter GDP growth rates, $\Delta \log(GDP_t)$, and only those survey indicators that exhibit a strong lead correlation with year-on-year GDP growth rates, $\Delta_4 \log(GDP_t)$. To this end, we transform our monthly indicators to quarterly frequency by averaging over the respective quarter and estimate the predictive regressions

$$\Delta \log(GDP_t) = \sum_{i=1}^{16} \sum_{l=1}^{3} b_{i,l}^{(h)} y_{i,t-l}^{(h)} + u_t^{(h)} \tag{7}$$

for hard indicators and

$$\Delta_4 \log(GDP_t) = \sum_{i=1}^{19} \sum_{l=1}^{3} b_{i,l}^{(s)} y_{i,t-l}^{(s)} + u_t^{(s)} \tag{8}$$

for survey indicators, where l denotes quarterly lags. The parameters $b^{(h)}$ and $b^{(s)}$ are estimated using the elastic net (EN) proposed by Zou and Hastie (2005) and successfully used by Bai and Ng (2008) for indicator selection. The elastic net is a convex combination of a ridge regression and a LASSO and yields nonzero parameter estimates only for a few important indicators. It solves the following optimization problem:

$$\mathsf{L} = (\lambda_1, \lambda_2, \mathsf{b}) = |\mathsf{y} - \mathsf{X}\mathsf{b}|^2 + \lambda_1 |\mathsf{b}|_1 + \lambda_2 |\mathsf{b}|^2 \tag{9}$$

where

$$|\mathsf{b}|_1 = \sum_j |\mathsf{b}_j| \quad \text{and} \quad |\mathsf{b}|^2 = \sum_j \mathsf{b}_j^2.$$

$\mathsf{y} = (\mathsf{y}_1, \ldots, \mathsf{y}_T)$ denotes a centered response variable—in our setting either $\Delta \log(GDP_t)$ or $\Delta_4 \log(GDP_t)$—and $\mathsf{X} = (\mathsf{x}_1, \ldots, \mathsf{x}_N)$ is a set of N standardized predictors $\mathsf{x}_i = (\mathsf{x}_{1i}, \ldots, \mathsf{x}_{Ti})'$—in our setting either the hard indicators $y_{i,t-l}^{(h)}$, $i = 1, \ldots, 16$, $l = 1, \ldots, 3$, or the survey indicators $y_{i,t-l}^{(s)}$, $i = 1, \ldots, 19$, $l = 1, \ldots, 3$.

The tuning parameters λ_1 and λ_2 control the weight on the L_1 and L_2-norm penalty, respectively. For increasing relative weight λ_1 the EN approaches the LASSO which is known to shrink coefficients to zero due to the non-smoothness of its objective function, while for increasing relative weight λ_2 the EN approaches the ridge regression which is capable of handling highly correlated predictors. Zou and Hastie (2005) show that the EN inherits both properties and is thus particularly suited for our purpose. They also demonstrate that the EN can be transformed into a LASSO problem which can be estimated by the Least Angel Regression (LARS) of Efron et al. (2004). This algorithm, called LARS-EN, is a forward stepwise additive fitting procedure.

The number of steps, k, equal the number of included variables and corresponds, for given λ_2, to a specific value of λ_1. Hence, instead of choosing λ_1 and λ_2, one may equivalently choose λ_2 and k which is what we do in the following.[8]

We apply the LARS-EN algorithm to both (7) and (8). We select $n_h = 3$ hard indicators and $n_s = 3$ survey indicators in order to avoid predominance of one of the categories.[9] In both cases, we thus set $k = 3$.[10] In addition, we choose $\lambda_2 = 100$ which is a fairly large value and allows high correlation between the selected indicators.[11]

4 Ex post business cycle dating for Germany

In the following, we apply our dynamic factor model combined with the LARS-EN indicator selection to identify the German business cycle turning points in the full sample. Such an ex post business cycle dating based on revised data is of its own interest as it complements simple but purely univariate dating algorithms like Bry-Boschan and undisclosed multivariate procedures like the one published by the Economic Cycle Research Institute (ECRI). Our main interest is, however, to show that our empirical approach produces reasonable results in-sample before we subsequently use it to predict turning points out-of-sample in a real-time forecasting setting.

4.1 Selected indicators

We first apply the LARS-EN algorithm with the aforementioned settings to the pre-selected set of indicators. We obtain the following results. The selected hard indicators comprise—in the order of selection—new foreign orders of capital goods, new domestic orders of intermediate goods, and new domestic orders of capital goods. The selected survey indicators include—again in the order of selection—overall industry production expectations, overall business expectations, and export expectations published by the Ifo institute. All six indicators are selected with a lag of one quarter implying that they lead the business cycle by three months. To obtain a factor with the same lead property, we include the indicators contemporaneously in the monthly factor model, i.e., set $l_{h,i} = l_{s,i} = 0$ in equations (1) and (2) for all i. Altogether the selection reflects common knowledge that orders of production inputs and business expectations are valuable early indicators. It also highlights the openness of the German economy as foreign trade plays a role in both indicator sets.

[8] For given λ_2 this works as follows. Since LASSO shrinks coefficients to zero, start with a sufficiently large λ_1 (which yields zero estimates of all coefficients) and iteratively lower λ_1 until the prespecified number, k, of nonzero coefficient estimates is obtained.

[9] We also tried to select more indicators and different numbers of hard and survey indicators without being able to improve in-sample business cycle dating and out-of-sample recession forecasts. Therefore, we do not report those results.

[10] In some instances, the LARS-EN algorithm selects two different lags of the same indicator. In such a case, we include in our factor model the lag selected first and increase k by one to select another indicator to be included.

[11] Higher values for λ_2 do not change the selection. Smaller values for λ_2 cause LARS-EN to select only one of a set of correlated indicators which is problematic in our setting because we rather select similar indicators with high correlation and good forecasting power for GDP than very different indicators of which some are only loosely related to GDP. We also tried to choose λ_2 according to cross validation based on MSE but this method leads to inferior results and we do not report them here.

4.2 Factor estimate for MS(2)-DFM

Based on the selected indicators, we first estimate a "classical" two-state model, MS(2)-DFM, that distinguishes between expansions and recessions. Before estimation, we have to determine the lag orders of the factor and the idiosyncratic components. Camacho and Perez-Quiros (2007) and Aastveit et al. (2016) argue that the main dynamics of a business cycle can be captured solely by a switching intercept, and Boldin (1996) shows for univariate Markov-switching models that overparameterization can lead to severe problems. Therefore, we set the lag order, p, of the factor to zero.[12] This allows us to treat our intercept as a switching mean. The autocorrelation functions of the idiosyncratic components indicate a lag order of $q = 2$.

The estimated parameters of the MS(2)-DFM are reported in Table 1. State 1 features a positive mean and a high persistence probability and occurs 87 percent of the time unconditionally. It can thus be interpreted as an expansionary regime. State 2 exhibits a negative mean, is less persistent, and takes place 13 percent of the time which is why it appears like a standard recession regime. However, the estimated means have a strong implication. To see this, recall that the factor is constructed from standardized indicators and has a sample mean of approximately zero. Therefore, the expansionary (recessionary) mean describes the average positive (negative) deviation from "normal times". While the scale is arbitrary, the relative sizes are not. Hence, the estimates imply that a recession is, in absolute terms, about 6.5 times stronger than an expansion. This appears very large and is a consequence of effectively treating the Great Recession as a normal recession.

Table 1: Estimated parameters of the MS(2)-DFM

Parameter	β_1	β_2	p_{11}	p_{22}	P_1	P_2
Estimate	0.32 (0.10)	−2.12 (0.34)	0.97 (0.02)	0.79 (0.13)	0.87	0.13

Notes: Estimated standard errors are reported in parentheses below the estimates. $P_1 = P(S_t = 1)$ and $P_2 = P(S_t = 2)$ are the unconditional probabilities of being in the expansionary and recessionary states, respectively.

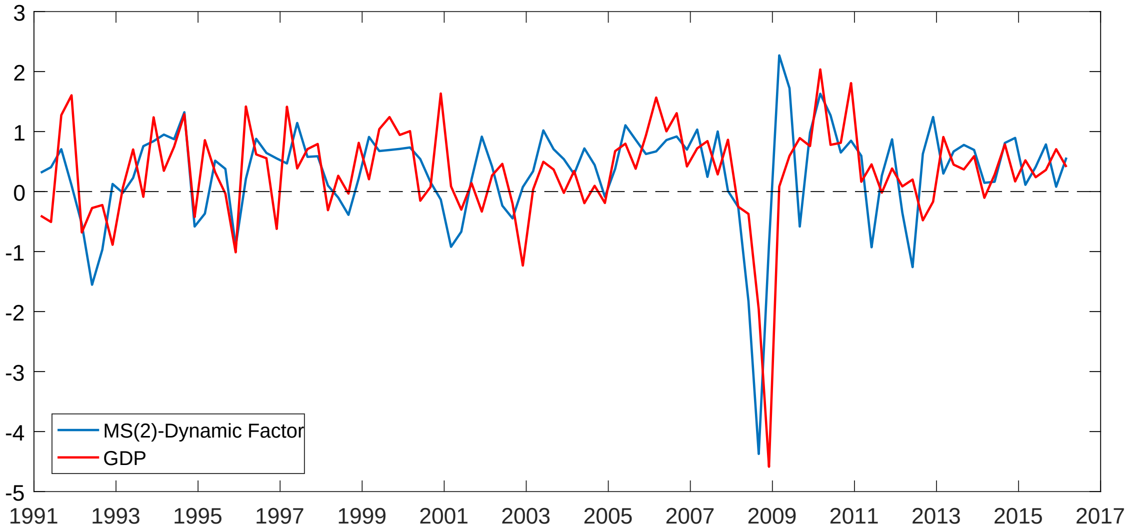
Nevertheless, the factor corresponds closely to GDP growth, see Figure 1 where we display quarterly averages of the filtered factor along with quarterly German GDP growth rates. Even though the factor solely summarizes the fluctuations of the six leading indicators identified above, it tracks GDP growth remarkably well. In several instances it appears to lead GDP growth as intended by construction. In fact, it exhibits the strongest correlation of 0.64 to GDP growth with a lead of one quarter which suggests that already the MS(2)-DFM may be well suited to forecast business cycle turning points.

4.3 Factor estimate for MS(3)-DFM

Now we introduce a third state. The idea is to account for, and predict, extraordinary strong output contractions like the Great Recession. The majority of the literature only considers two

[12]We also estimated models with $p = 1$ and $p = 2$ but obtained inferior results for the in-sample fit, thereby confirming the results of Camacho and Perez-Quiros (2007) and Aastveit et al. (2016).

Figure 1: Filtered factor of the MS(2)-DFM and GDP growth



Notes: The figure displays quarterly averages of the filtered factor estimated from a MS(2)-DFM, and quarterly German GDP growth rates, 1991Q1-2016Q2. The factor is re-scaled to fit mean and variance of GDP.

regimes. The few exceptions that consider three regimes rather aim at identifying weak growth phases (sometimes called stall phases) in addition to recessions and expansions (Boldin, 1996; Ferrara, 2003; Artis et al., 2004; Nalewaik, 2011). Instead, we aim at identifying regime 1 as expansionary, regime 2 as ordinary recession and regime 3 as severe recession as in Hamilton (2005).

To identify the three regimes and obtain numerically stable results of the numerical estimation procedure, we impose two economically sensible restrictions on the 3×3 transition matrix. Specifically, as in Hamilton (2005) we do not allow to directly switch from regime 1 (expansion) to regime 3 (severe recession) or vice versa. This is motivated by the observation that the Great Recession started off like an ordinary recession at the beginning of 2008, became severe after the Lehman collapse (industrial production dropped by more than 3 percent in each of the four months between November 2008 and February 2009), and phased out in the subsequent months.[13] The restricted transition matrix reads as follows:

$$\Pi_r = \begin{bmatrix} p_{11} & (1-p_{11}) & 0 \\ p_{21} & p_{22} & (1-p_{21}-p_{22}) \\ 0 & (1-p_{33}) & p_{33} \end{bmatrix}. \tag{10}$$

Except for adding a third state, we apply the same specification choices as before. In particular, we include the same six indicators as in the two-state model, set the lag order, p, of the factor to zero and the lag order, q, of the idiosyncratic components to two. The estimated parameters of the MS(3)-DFM are reported in Table 2. They compare favorably to the results of the two-state approach because the relative size of the means is more in line with what one would

[13] A likelihood ratio test of the two restrictions was not rejected with a p-value of almost 1 which indicates that the two transition probabilities are not well identified by the data anyway.

Table 2: Estimated parameters of the MS(3)-DFM

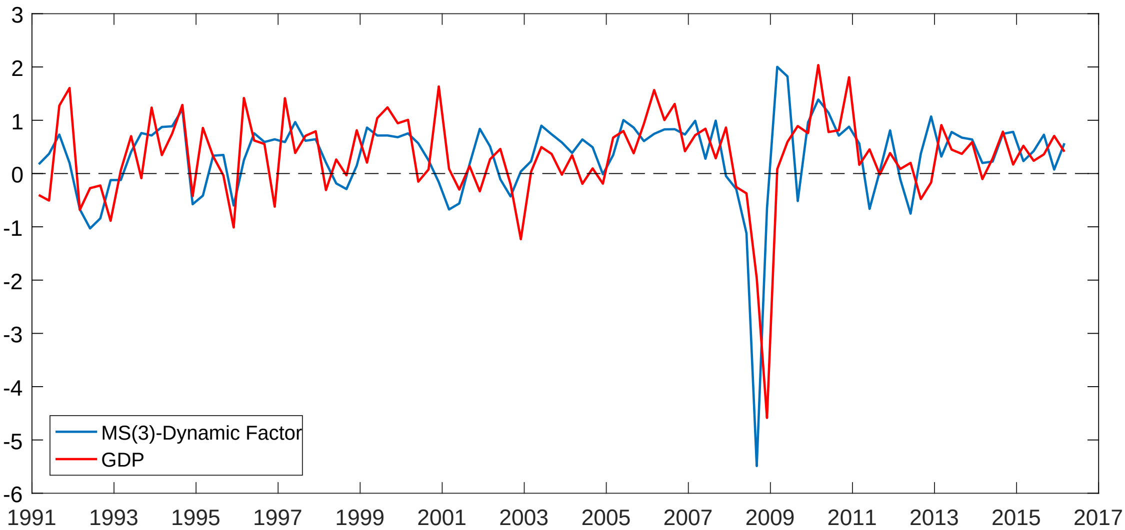
Parameter	β_1	β_2	β_3	p_{11}	p_{22}	p_{33}	p_{21}	P_1	P_2	P_3
Estimate	0.61 (0.12)	−1.42 (0.29)	−7.93 (1.08)	0.94 (0.03)	0.83 (0.10)	0.66 (0.51)	0.16 (0.09)	0.73	0.26	0.01

Notes: Estimated standard errors are reported in parentheses below the estimates. $P_1 = P(S_t = 1)$, $P_2 = P(S_t = 2)$, and $P_3 = P(S_t = 3)$ are the unconditional probabilities of being in the states of expansion, recession, and severe recession, respectively.

expect. The first regime has a positive mean implying that an expansion is characterized by a positive deviation from average times. The second regime has a negative mean of an absolute size that is 2.5 times the mean of the first regime. Hence, a normal recession is characterized by a negative deviation from average times, and it is 2.5 times as strong as an expansion. The third regime has a much lower mean and can thus safely be interpreted as a severe recession. The estimate implies that a severe recession is more than five times worse than a normal recession. Not much surprisingly given the development of the Great Recession, a severe recession is estimated to be much less persistent than normal recessions and expansions. In addition, the probability to switch from the ordinary recession to the severe recession is much lower ($1 - \hat{p}_{21} - \hat{p}_{22} = 0.01$) than to switch back ($1 - \hat{p}_{33} = 0.34$), and the unconditional probability of being in a severe recession is much lower than that of being in an ordinary recession.

The factor of the MS(3)-DFM corresponds closely to GDP growth, see Figure 2. Again it appears to lead GDP growth. It exhibits the strongest correlation of 0.68 to GDP growth with a lead of one quarter. This correlation is slightly larger than the one of the two-state factor which indicates that the three-state model might be better suited to predict German business cycle turning points.

Figure 2: Filtered factor of the MS(3)-DFM and GDP growth



Notes: The figure displays quarterly averages of the filtered factor estimated from a MS(3)-DFM, and quarterly German GDP growth rates, 1991Q1-2016Q2. The factor is re-scaled to fit mean and variance of GDP.

4.4 Which model gives a more realistic characterization of the German business cycle?

In the following we present the smoothed recession probabilities of the two-state and three-state models and assess whether they give a realistic picture of the German business cycle phases. As a benchmark we would ideally use a generally accepted monthly business cycle chronology for Germany comparably to the one of the NBER for the US. Since this is not available, we construct our own benchmark. To this end, we apply the Bry-Boschan business cycle dating algorithm because it is an often-used method and easily replicable. Given a monthly benchmark series, x_t, the algorithm defines peaks by

$$\wedge_t = \{(x_{t-d}, \cdots, x_{t-1}) < x_t > (x_{t+1}, \cdots, x_{t+d})\}$$

and troughs by

$$\vee_t = \{(x_{t-d}, \cdots x_{t-1}) > x_t < (x_{t+1}, \cdots, x_{t+d})\},$$

where d is the minimum duration which also implies that peak and trough must be at least d periods apart. The definition reveals the major drawback of the algorithm. To identify a turning point it requires at least d subsequent observations. Throughout the literature it has become standard to assume $d = 5$ months (and additionally a minimum length of a full cycle of 15 months). We follow this convention. Thus the algorithm exhibits a lag of at least 5 months until it signals that the state of the business cycle has changed, while the MS-DFM is—if it is applied in a real-time situation—designed to identify turning points instantaneously.

As benchmark series, x_t, to be fed into the Bry-Boschan algorithm we choose industrial production excluding construction.[14] This is motivated by the stylized fact that German industrial and overall activity are so strongly correlated that the industry sector, which exhibits a much more pronounced cyclical behavior than GDP, is generally thought of as the driver of the German business cycle. As industrial production is available at a monthly frequency, it enables us to determine the state of the economy on a monthly basis.

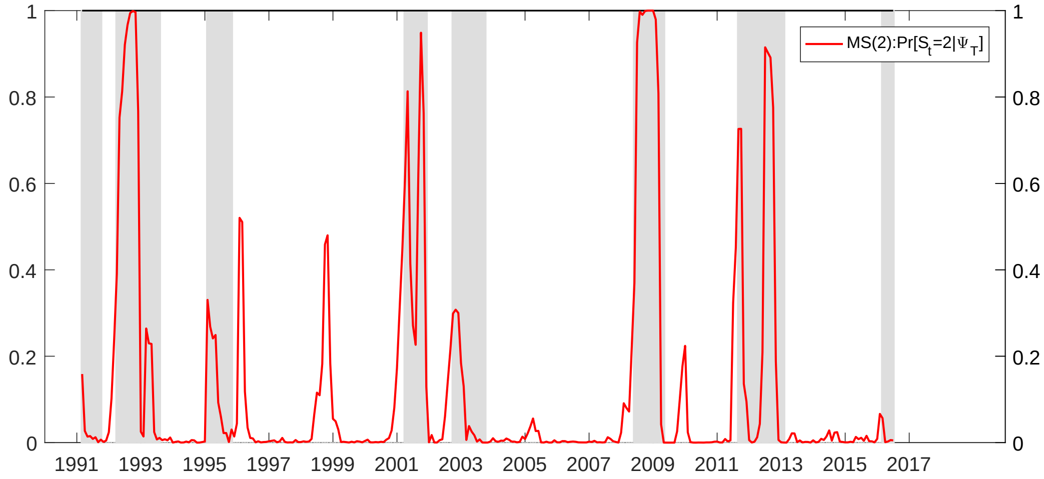
Figure 3 presents the smoothed recession probabilities of the MS(2)-DFM (panel a) and the MS(3)-DFM (panel b). Generally, the probabilities match the Bry-Boschan classification (indicated by shaded areas) quite well. In particular, the recession probabilities start rising slightly before, or at the beginning of, all benchmark recessions. Further, the MS(3)-DFM model identifies the steepest contraction of GDP during the Great Recession as a severe recession regime, while the probability of a severe recession is close to zero for the rest of the sample. There are some important differences between the recession probabilities of the two models. The two-state model detects the Bry-Boschan recessions starting in January 1995 and September 2002 with probabilities of less than 0.4, while the three-state model identifies them with probabilities of more than 0.9.[15] This finding indicates that the three-state model is much more sensitive than

[14] We exclude construction because particularly in the 1990s after German reunification the construction cycle was decoupled from the overall business cycle in Germany.

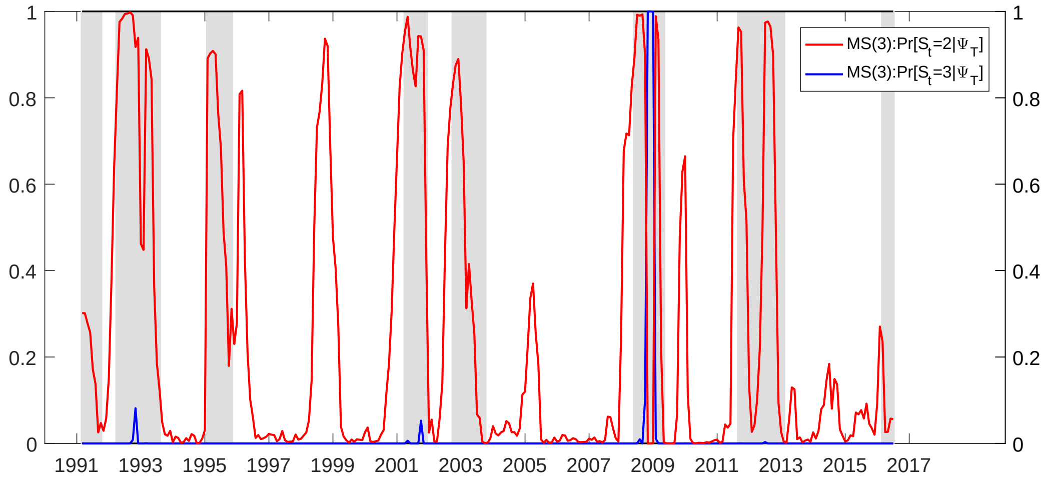
[15] A similar episode is the Bry-Boschan recession of February to June 2016. We do not take it too seriously, however, because the data are still relatively preliminary which may induce divergences in the information content of the early indicators and industrial production that may vanish after future data revisions.

Figure 3: Recession probabilities of MS(2)-DFM and MS(3)-DFM

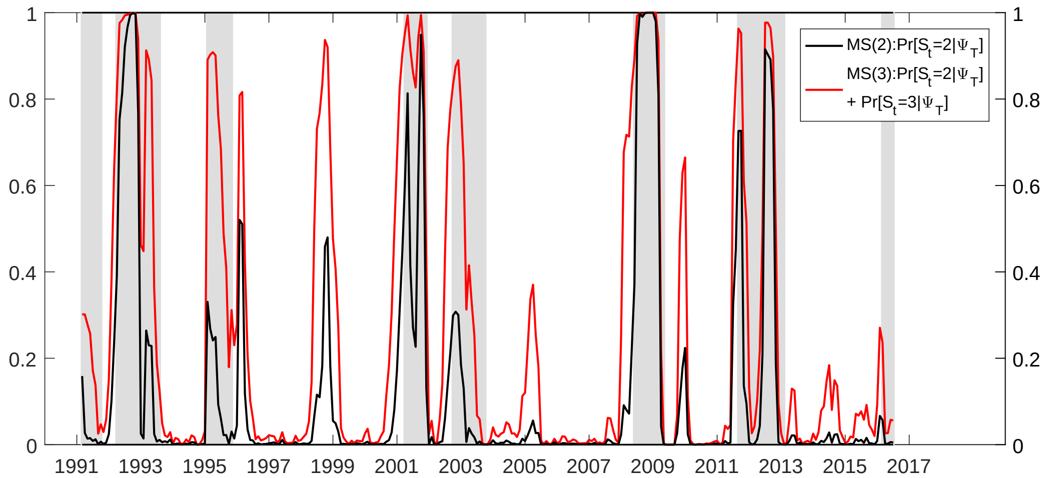
(a) Probability of a recession estimated from an MS(2)-DFM



(b) Probability of a normal and severe recession estimated from an MS(3)-DFM



(c) Comparison of the recession probabilities



Notes: Panel (a) displays smoothed recession probabilities of the MS(2)-DFM. Panel (b) displays smoothed probabilities of an ordinary recession (red line) and severe recession (blue line) of the MS(3)-DFM. Panel (c) compares the probability of a recession from the MS(2)-DFM (red line) with the joint probability of an ordinary or severe recession from the MS(3)-DFM (blue line). Shaded areas correspond to the recessions dated by the Bry-Boschan algorithm.

its two-state counterpart because the distinction between ordinary and severe recessions allows it to assign already mildly weak times as (ordinary) recessions. The increased sensitivity can also be inferred from panel (c) of Figure 3 which displays the recession probabilities of the two-state model and the joint probabilities of a normal or severe recession of the three-state model. Clearly, the latter are always higher than the former.

As a potential drawback, an increased sensitivity may go hand in hand with a higher risk of false alarms. In fact, the three-state model indicates the existence of a recession in a few cases when both the two-state model and the Bry-Boschan algorithm do not. It is instructive to examine one example in more detail. In September 1998 the recession probability of the three-state model exceeds 0.5 for 7 months in a row while the two-state probability remains slightly below 0.5 and the benchmark does not indicate a recession at all. At that time the German business cycle was temporarily fragile as indicated by a majority of the selected indicators. After a peak in July 1998, industrial production exhibited a weak period of more than 6 months before it picked up again. However, the trough was already in November 1998 which is not more than five months away from the peak. Hence, the Bry-Boschan algorithm neglects this episode. The example demonstrates that it is ultimately a matter of definition whether an episode should be classified as a recession. It also shows that what might appear as oversensitivity at first sight, may carry useful information that is more precise than a 0-1 rule.

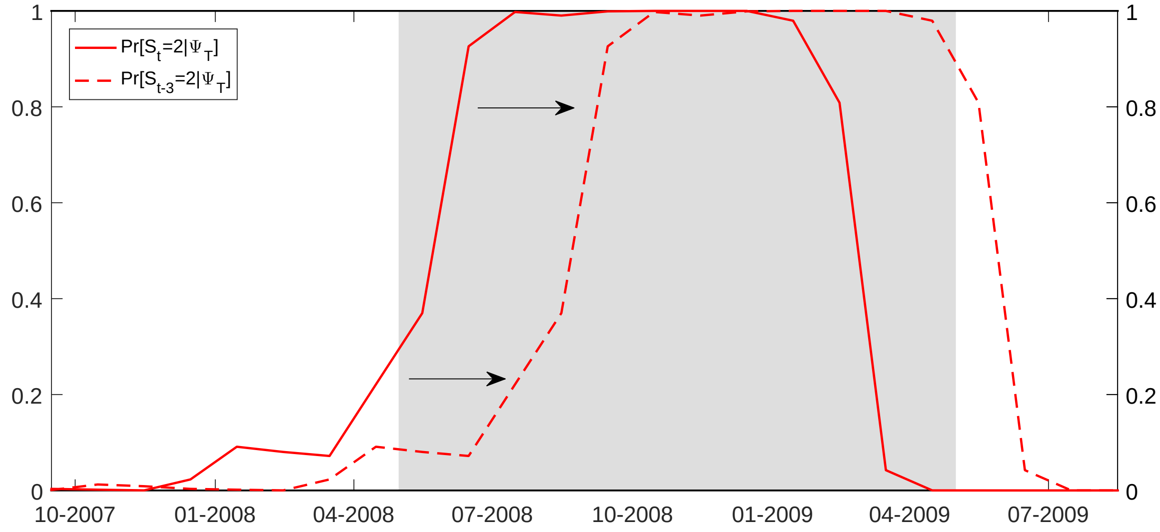
To illustrate the leading properties of the two models, Figure 4 takes a closer look at the Great Recession. In panel (a) the solid line represents the smoothed recession probabilities of the two-state model. Since the factor is designed to lead GDP by one quarter, a recession probability displayed in period t refers to period $t + 3$ if compared to the monthly benchmark series. This implies that the model indicates—at least within sample—a recession three months ahead of time. For the ease of presentation, we thus shift the smoothed recession probabilities three months forward (dashed line). The shifted recession probabilities exceed 0.5 starting in September 2008, the month of the Lehman collapse. While this appears like a sensible result, it is by now conventional wisdom that the Great Recession in Germany started earlier that year[16] while the most severe production declines came a few months later. The root of the problem is again the missing distinction between ordinary and severe recessions. As the two-state model identifies a single "average" recession, it comes late when a recession is mild.

In contrast, the three-state model almost perfectly matches the Great Recession. Panel (b) of Figure 4 displays the smoothed recession probabilities both for the times they are computed (solid lines) and shifted forward by three months (dashed lines). The probability of a normal recession rises above 0.5 in January 2008 indicating that the Great Recession starts three months later in April which compares well with the development of output: the second quarter of 2008 saw the first (small) decline in GDP. The probability of a severe recession exceeds 0.5 during October to December 2008 implying that January to March 2009 are the core recession months.

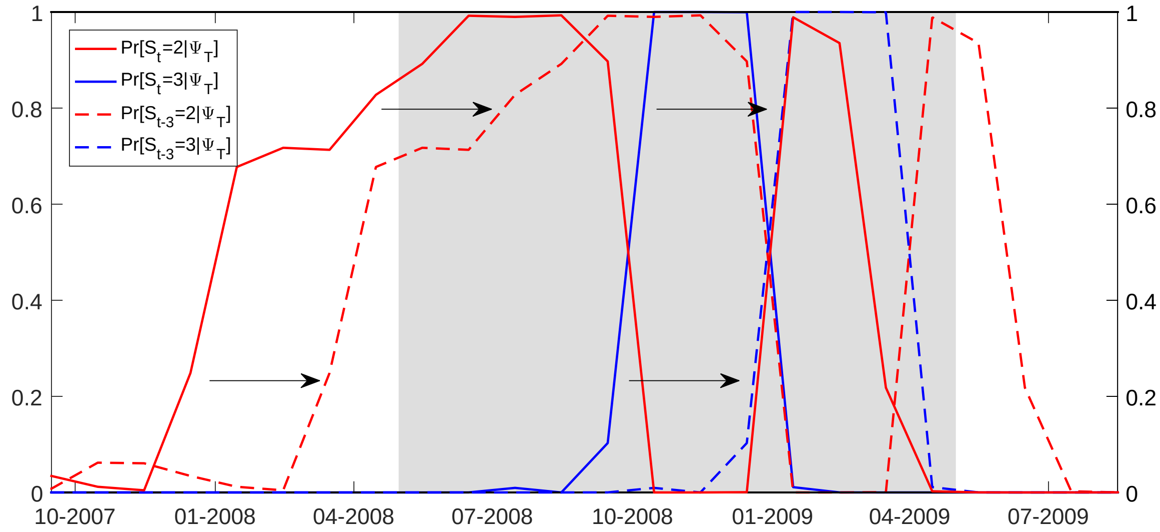
[16]Using a simple rule-of-thumb that defines a recession as at least two consecutive quarters of negative real GDP growth, one would date the start of the recession in the second quarter of 2008. Official business cycle dates from the CEPR Euro Area Business Cycle Dating Committee are only available for the euro area as a whole. According to those the business cycle peak occurred in the first quarter of 2008. Business cycle dates for Germany are released by the Economic Cycle Research Institute (ECRI) which dates the peak of the previous expansion in April 2008.

Figure 4: Recession probabilities of MS(2)-DFM and MS(3)-DFM during the Great Recession

(a) MS(2)-DFM



(b) MS(3)-DFM



Notes: Panel (a) displays smoothed recession probabilities of the MS(2)-DFM during the Great Recession. Panel (b) displays smoothed recession probabilities of the MS(3)-DFM during the Great Recession. The solid lines depict the model-based recession probabilities which lead the business cycle by three months. The dashed lines depict the recession probabilities shifted forward by 3 months so as to fit the business cycle contemporaneously. Shaded areas correspond to the recessions dated by the Bry-Boschan algorithm.

In fact, GDP loss in the first quarter of 2009 was by a large margin the steepest of the Great Recession. Also, industrial production fell maximally in January 2009. Altogether, the three-state model indicates that the Great Recession occurred between April 2008 and May 2009 while the Bry-Boschan algorithm identifies May 2008 to April 2009.

To more formally evaluate the two-state and three-state models against the Bry-Boschan

Table 3: QPS and FPS measures

	QPS				FPS			
k	0	1	2	3	0	1	2	3
MS(2)-DFM	0.1830	0.1702	**0.1661**	0.1725	0.2164	**0.2131**	0.2164	0.2262
MS(3)-DFM	0.1491	0.1240	**0.1089**	0.1121	0.2164	0.1803	0.1574	**0.1541**

Notes: QPS is the quadratic probability measure defined in (11). FPS is the false positives measure defined in (12). $k \in \{0, 1, 2, 3\}$ refers to the lead of of the Markov-switching models compared to the Bry-Boschan benchmark.

benchmark, we employ the quadratic probability score

$$QPS = \frac{1}{T}\sum_{t=1}^{T}[B_{t+k} - \mathrm{P}_t(\text{recession})]^2, \tag{11}$$

where B_{t+k} denotes the binary Bry-Boschan benchmark series with lead equal to $k \in \{0, 1, 2, 3\}$ and P_t(recession) is the smoothed probability to be in a recession (two-state model) or in an ordinary or severe recession (three-state model). QPS takes an optimal value of zero if the smoothed probabilities calculated by a model coincide with the benchmark.

In addition, we compute the false positives measure

$$FPS = \frac{1}{T}\sum_{t=1}^{T}[B_{t+k} - I\{\mathrm{P}_t(\text{recession}) > 0.5\}]^2, \tag{12}$$

where $I\{\mathrm{P}_t(\text{recession}) > 0.5\}$ is an indicator function taking the value of 1 if the smoothed probability of being in a recession is higher than 0.5 and 0 otherwise. Hence, this measure counts the number of false signals, i.e. incorrectly predicted periods, of the model. The lower the FPS is, the better is the model's ability to reliably predict recessions.

Table 3 reports the QPS and FPS measures for the two-state and three-state models. According to both quality measures the three-state approach provides a superior in-sample fit for all measures. This suggests that using an MS(3)-DFM gives a more realistic characterization of the German business cycle than using a more classical two-state model. Since it provides detailed information in terms of regime probabilities we also prefer it over a simple 0-1 classification scheme like the Bry-Boschan algorithm that in addition can only classify downturns that last at least five months as recessions.

Additionally, the QPS and FPS measures corroborate that the Markov-switching models exhibit a lead compared to the Bry-Boschan benchmark. Specifically, the QPS measure is minimal at $k = 2$ suggesting that both models have a lead of two months, while the FPS measure is lowest at $k = 1$ month for the two-state model and $k = 3$ for the three-state model. Taken together, these results indicate that it is possible to achieve a leading property of almost one quarter by carefully selecting a set of leading indicators and integrating them into a Markov-switching dynamic factor model.

4.5 Monthly business cycle chronology for Germany

In some situations it may be valuable to have a dichotomous monthly business cycle chronology (even though recession probabilities are much more informative). Characterizing months with a recession probability greater than 0.5 as recessionary and assuming a lead of three months, we derive such a chronology from our preferred three-state model, see Table 4. We also report the chronologies based on the two-state model and the Bry-Boschan algorithm.[17]

According to our three-state model Germany has experienced eight recessionary phases since January 1991. Particularly pronounced episodes are the post-reunification recession (May 1992 to July 1993), the "dot com" recession (March 2001 to January 2002), the Great Recession (April 2008 to May 2009), and the European sovereign debt crisis which consists of two phases (September 2011 to February 2012 and September to December 2012) summarized in columns 8a and 8b of Table 4.

The two-state model identifies solely those four pronounced recessions. However, the timing is always a little late and the recession lengths appear a bit underestimated. For example, according to the two-state model the Great Recession lasted only nine months and the European debt crisis as little as four months. In contrast, the Bry-Boschan benchmark indicates eight recessionary phases which in most cases coincide well with the three-state model. Exceptions are the two episodes at the sample beginning and the sample end which may be the result of a sample edge problem (in particular, potential data revisions render the 2015 recession tentative), and the episode between August 1998 and February 1999 already discussed in the previous subsection.

Table 4: Benchmark recession dates for Germany

		Recession no.									
		1	2	3	4	5	6	7	8a	8b	9
MS(3)-DFM	start	–	05.92	04.95	08.98	03.01	10.02	04.08	09.11	09.12	–
	end	–	07.93	09.95	02.99	01.02	04.03	05.09	02.12	12.12	–
MS(2)-DFM	start	–	07.92	–	–	06.01	–	09.08	09.12		–
	end	–	02.93	–	–	01.02	–	05.09	12.12		–
Bry-Boschan	start	01.91	03.92	01.95	–	03.01	09.02	05.08	08.11		08.15
	end	09.91	07.93	10.95	–	11.01	09.03	04.09	01.13		12.15

Notes: Recessions are defined as $PR[S_t = 2|\Psi_T] \geq 0.5$ (MS(2)-DFM) and $PR[S_t = 2|\Psi_T] + PR[S_t = 3|\Psi_T] \geq 0.5$ (MS(3)-DFM), where Ψ_T is the information set available at the sample end. Episodes that last less than 4 months are excluded.

5 Real-time business cycle assessment and forecasting

In this section, we apply the Markov-switching dynamic factor models to nowcast and forecast business cycle turning points, as well as GDP growth rates, in real time. In doing so, we exploit the advantage of these models to indicate turning points instantaneously and thereby circumvent the endpoint problem inherent to the Bry-Boschan algorithm which leads to delayed signals.

[17] In all cases, we exclude episodes of less than four months length.

5.1 Nowcasting German business cycle turning points

To assess the nowcasting ability of the two-state and three-state models, we perform a pseudo nowcasting experiment over the evaluation period January 2001 until June 2016 using real-time data. We choose this evaluation period because it includes five recessions which allows us to judge the results with some confidence, while the initialization sample of ten years (1991M01-2000M12) is still sufficient to estimate a MS-DFM. In addition, we include equally long periods before and after the Lehman bankruptcy which helps us to understand whether adding a third state—which is only hardly identifiable before the Great Recession—would have made a difference in real time.

We construct a real-time dataset consisting of the same pre-selected set of 35 indicators as in the previous section. To this end, we take the series of new orders, employed persons, and inflation from the real-time database of Deutsche Bundesbank,[18] and US industrial production from the real-time database of the OECD. The remaining hard indicators are determined on financial markets and are not revised.[19] The survey indicators are revised only very marginally, hence we neglect these revisions.

In each step of the nowcasting experiment, we go through the selection and estimation stages described in previous sections. To obtain a nowcast for month $\tau \in \{2001M01, \ldots, 2016M06\}$, we first apply the LARS-EN algorithm to the sample available at the end of this month and select a set of three hard and three survey indicators. Subsequently, we feed these indicators into the Markov-switching dynamic factor models with zero lags for the factor and two lags for idiosyncratic component, estimate the parameters, and smooth out the state probabilities. As a result, we not only obtain a series of real-time probabilities but also a time-varying selection of indicators for the period January 2001 until June 2016. Figures 11 and 12 in the Appendix depict the recursive selection of the hard and survey indicators, respectively.

The results of the real-time indicator selection reflect the traditional dependence of the German business cycle on global developments. Of the six indicators, the selection algorithm always picks two hard indicators (foreign orders of capital goods and, with very few exceptions, one of the two commodity prices) and one survey indicator (the Euro-coin indicator until April 2013 and the Ifo export expectations thereafter) that summarize external information while only two survey indicators (the Ifo business expectations and another Ifo indicator) are more closely related to the domestic situation. However, there is also a notable change that may signal an increased relevance of the domestic economy after the Great Recession: the sixth indicator selected by the algorithm is foreign orders of intermediate goods until February 2009 but domestic orders of intermediate goods thereafter.

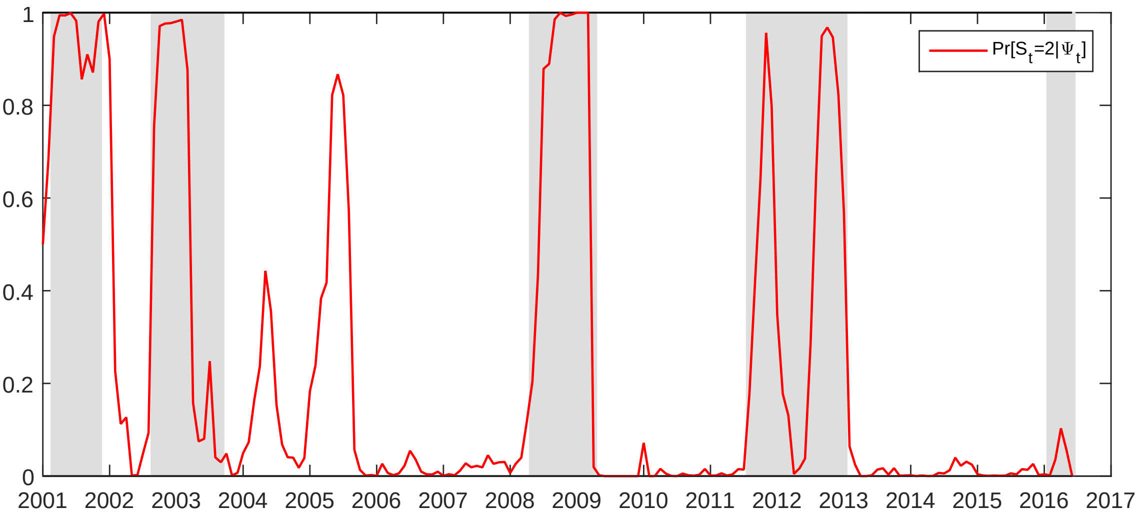
The real-time nowcasts of the recession probabilities are constructed using all available information at a certain point of time. Since we only select indicators that lead the business cycle by at least 3 months, it would be sufficient to include indicators of period $\tau - 3$ and earlier in order to compute filtered probabilities of period τ. However, such an approach would neglect important information as, at the end of period τ, the realizations of, say, new orders for period

[18] Some releases miss some observations at the beginning of the sample. In such cases, we use growth rates from previous releases to fill the gaps by means of backward chaining.

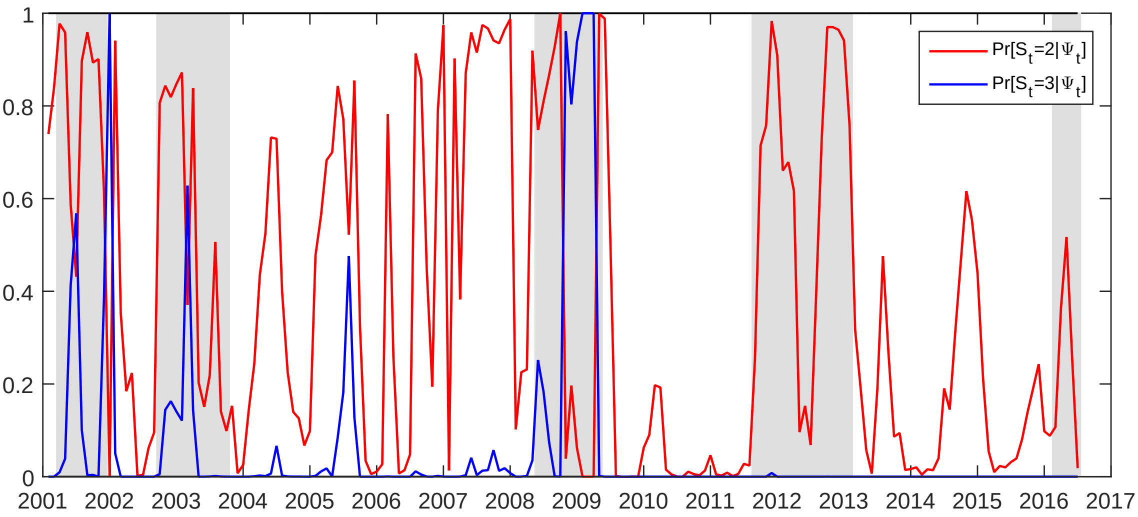
[19] The only exception is the German contribution to EMU M2. However, it is so rarely and slightly revised that we can safely take it as being unrevised.

Figure 5: Real-time nowcasts of recession probability

(a) MS(2)-DFM



(b) MS(3)-DFM



Note: Smoothed recession probabilities of (a) MS(2)-DFM and (b) MS(3)-DFM recursively estimated with real-time data. Ψ_t *denotes the information set as of period t. Shaded areas correspond to the recessions of the benchmark business cycle chronology from Section 4.4.*

$\tau - 2$ and survey indicators for period τ are already known. Therefore, we compute the real-time probabilities by means of backward smoothing taking all observations into account that are known in period τ.[20] Like Chauvet and Hamilton (2006) and Hamilton (2011), we find that these smoothed probabilities are much more stable and reliable than their filtered counterparts.

Figure 5 depicts the smoothed probabilities generated by the MS(2)-DFM and the MS(3)-

[20]Note that this leads to ragged edges in the data structure. We deal with that complication by using the method of Mariano and Murasawa (2003) which is extended to the nonlinear Markov-switching framework by Camacho et al. (2012). It consists of replacing the missing observations at the end of the sample by random numbers distributed independently of the model's parameters. These random numbers are in turn eliminated by an appropriately defined Kalman filter. As shown by Camacho et al. (2012), neither the maximum of the likelihood function nor the estimated filtered probabilities depend on these random numbers.

DFM, respectively. The MS(2)-DFM shows roughly similar real-time recession probabilities as based on the estimates using the full sample in section 4. The main difference is that the model falsely signals a recession in 2005 in real time which in retrospect did not happen. The real-time estimates by the MS(3)-DFM differ substantially from those based on the full sample. Before 2008, the third state is not well identified and the probabilities of being in the second or third state exhibit erratic fluctuations.

To further understand what happens inside the two models, Figure 6 takes a closer look at their recursively estimated state-specific means. The two-state model (panel a) exhibits a break at the beginning of the Great Recession. Before, the model is remarkably stable with a first state that has a positive mean and a second state that has a negative mean of similar absolute magnitude. Since the factor is extracted from standardized indicators and thus has a sample mean of approximately zero, the first state can be interpreted as expansion, while the second state represents a recession. During the Great Recession, however, the expansion mean is estimated as approximately zero whereas the recession mean falls dramatically. At that time, a user of this model would have found the model's result unconvincing, both because of its instability and—perhaps more importantly—because of its interpretation: neither an upswing with a growth rate that merely equals the sample average nor an extreme contraction could have been easily reconciled with what was observed as expansions and recessions before the onset of the Great Recession. These findings probably would have been interpreted as a signal that "this time is different" and that a third state is necessary to characterize the German business cycle properly.

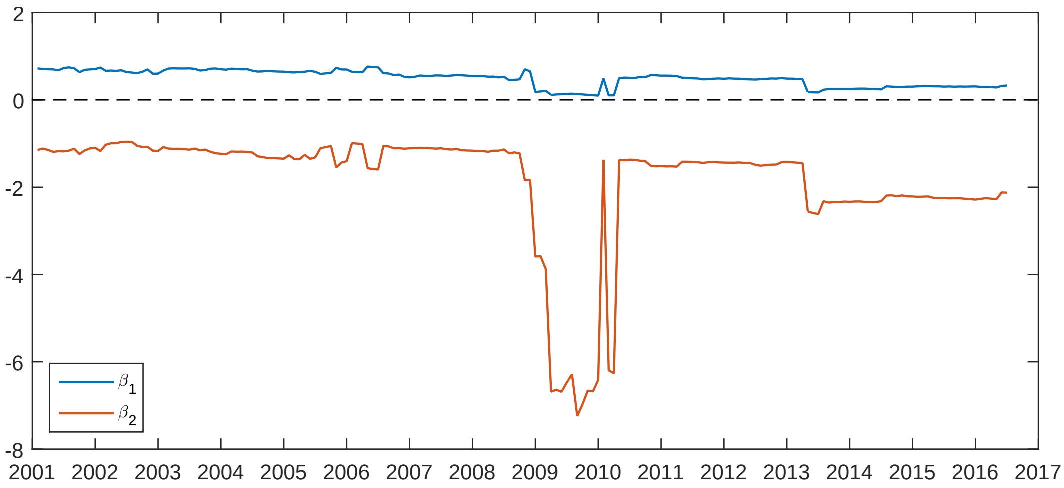
In contrast, the three-state model is instable before the Great Recession because the third state is only weakly identified during this time. Until 2005 the first two states would have been interpreted as expansion and ordinary recession, while the third state having a mean considerably smaller than the second state would have been labeled a severe recession. However, during the boom of 2006 to mid 2008 which preceded the Great Recession, the first state signals a strong boom and the third state a recession of similar absolute magnitude whereas the second state indicates "average times" with mean zero and thus average growth—an interpretation difficult to reconcile with prior experience. This changes again with the beginning of the Great Recession. As more and more bad news come in, the model starts to extract a severe recession regime with a very negative mean that fluctuates—after a few months of undershooting—in a range that is considerably below the pre-crisis level. In addition, the means of the first and second state stabilize at levels that lend to the interpretation of expansion and mild recession, respectively. This stabilization is also visible in Figure 5 where the smoothed real-time recession probabilities largely coincide with those based on the full sample shown in Figure 3.

5.2 Model selection in real time

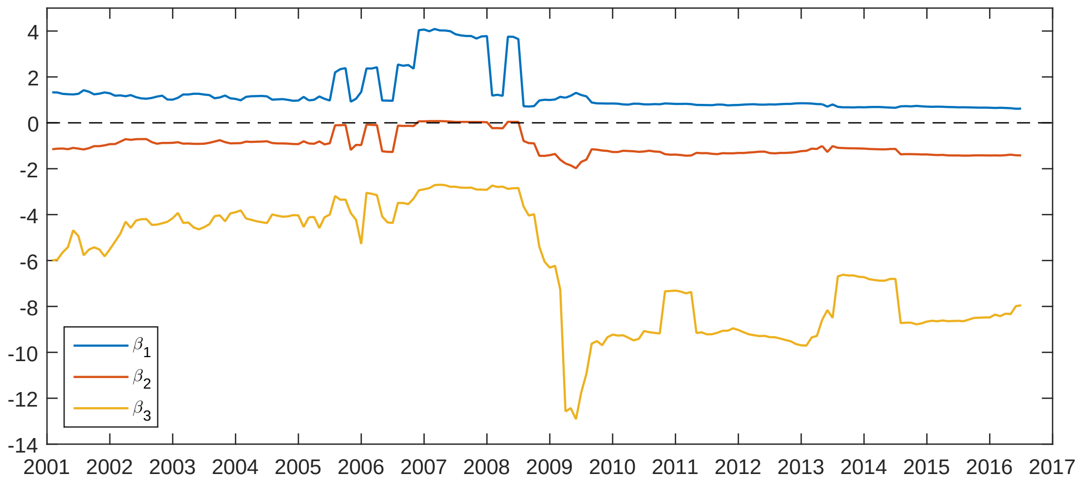
The results of the nowcasting experiment directly raise the issue of model selection in real time. We suggest to use either of the following two criteria to compare the two-state and three-state models. The first criterion is the QPS which measures how closely the Markov-switching models match the business cycle turning points identified by the Bry-Boschan algorithm applied to

Figure 6: Recursively estimated means for MS(2)-DFM and MS(3)-DFM

(a) MS(2)-DFM



(b) MS(3)-DFM



Note: Means of a MS(2)-DFM and MS(3)-DFM, estimated recursively with real-time data. Blue line: first state, red line: second state, yellow line: third state.

German industrial production in real time.[21] We select the model with the better fit. The second criterion is the BIC which may have the advantage over the QPS that it balances fit against parsimonity.[22]

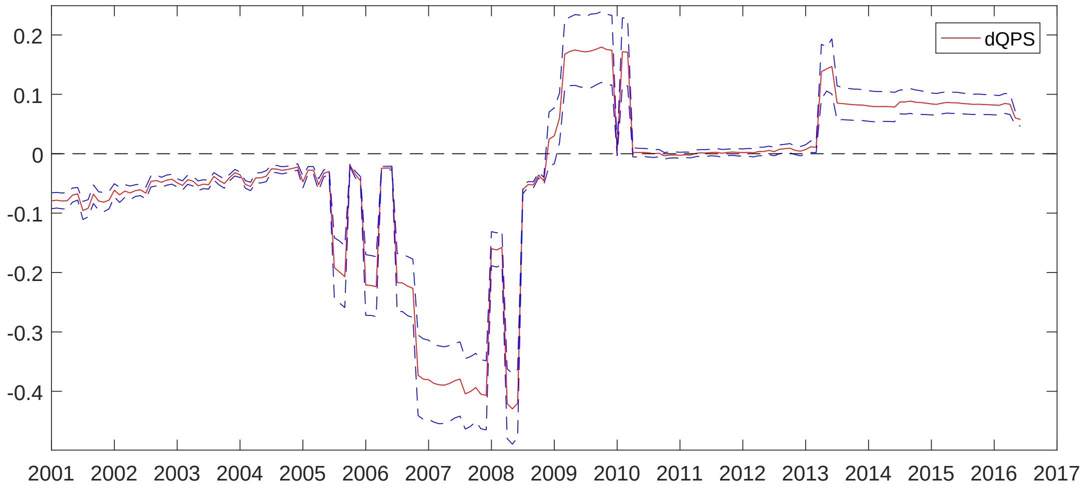
Figure 7 plots the differences between the two-state and three-state models in terms of the QPS, $dQPS = QPS(2) - QPS(3)$, and in terms of the BIC, $dBIC = BIC(2) - BIC(3)$, based on the real-time estimates over the period January 2001 until July 2016. In both cases, a positive value indicates an advantage of the MS(3)-DFM. We find that the two-state model is superior

[21] We take the industrial production data from the real-time database of the Bundesbank.

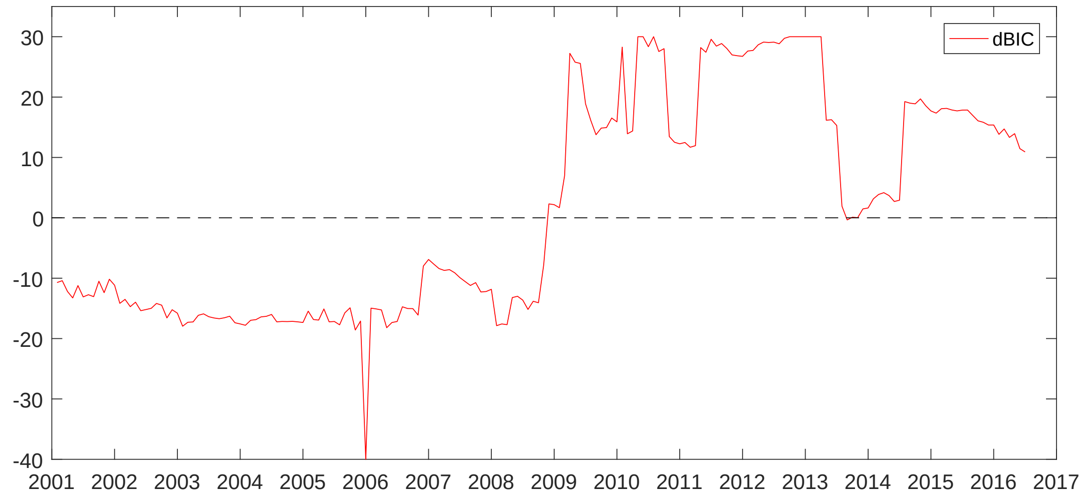
[22] Smith et al. (2006) propose a specific Markov-switching specific criterion. However, it is designed for models in which all parameters switch and thus does not work with our model. They also show that the Akaike information criterion (AIC) always selects the model with more states. Hence, we prefer the BIC over the AIC.

Figure 7: Recursive differences in QPS and BIC between MS(2)-DFM and MS(3)-DFM

(a) Recursive difference in QPS



(b) Recursive difference in BIC



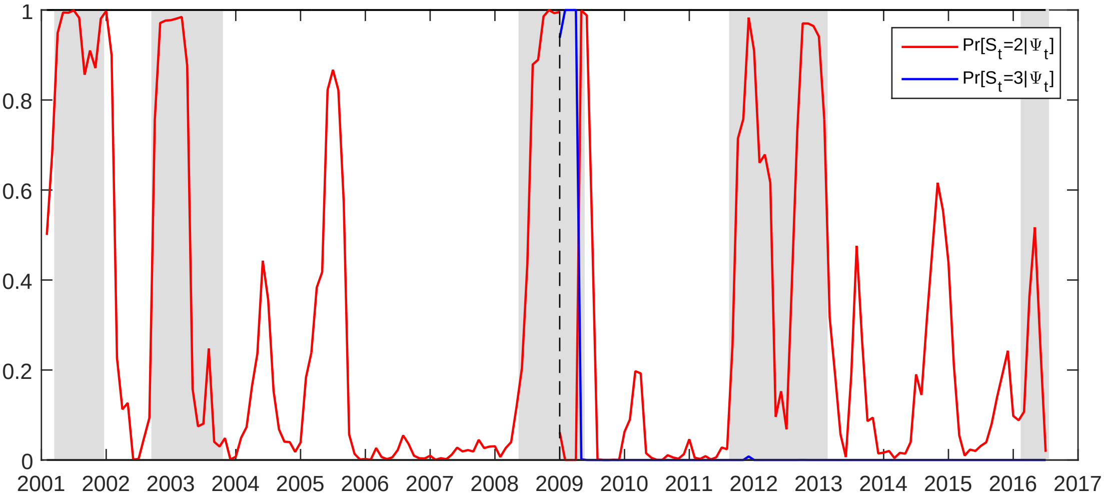
Notes: Panel (a) shows the recursively computed $dQPS = QPS(2) - QPS(3)$ *with* 95% *confidence bands. Panel (b) shows the recursively computed* $dBIC = BIC(2) - BIC(3)$*. In panel (b) we trim the observation of December 2005 (*-72.56*) to* -40 *to make the graph better readable.*

up to the end of 2008, while the the three-state model is favored thereafter. The exact change dates are very similar: the $dBIC$ selects November 2008 as the first month with an advantage of the three-state model, while the $dQPS$ identifies December 2008. Note that this date coincides with the aforementioned break in the recursively estimated state-specific means of the two-state model, see panel (a) of Figure 6. Hence, a user of these models would have noticed by December 2008 that introducing a third state is necessary to obtain a well-specified model.

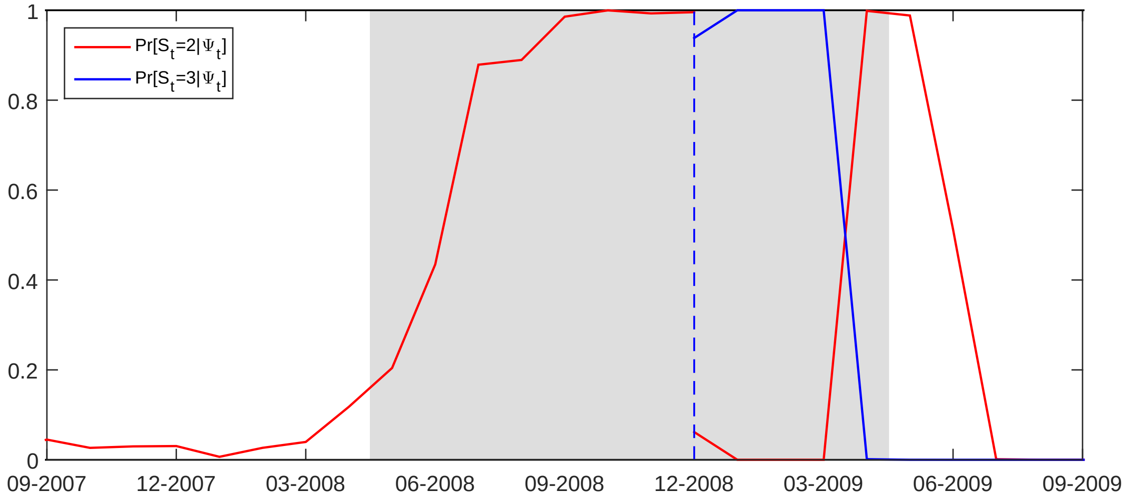
Panel (a) of Figure 8 shows the smoothed nowcast probabilities of a combination of the MS(2)/MS(3)-DFM, with the shift implemented in December 2008, for the whole sample, while panel (b) zooms in on the Great Recession period. The shift occurs when the probability for a

Figure 8: Real-time nowcast of recession probabilities using a MS(2)/MS(3)-DFM

(a) Full evaluation sample 2001M01 to 2016M06



(b) Great Recession sample



Note: Smoothed recession probabilities of a MS(2)/MS(3)-DFM during the Great Recession recursively estimated with real-time data. The dashed line indicates the shift from MS(2)-DFM to MS(3)-DFM. Shaded areas correspond to the recessions from the benchmark business cycle chronology from Section 4.4.

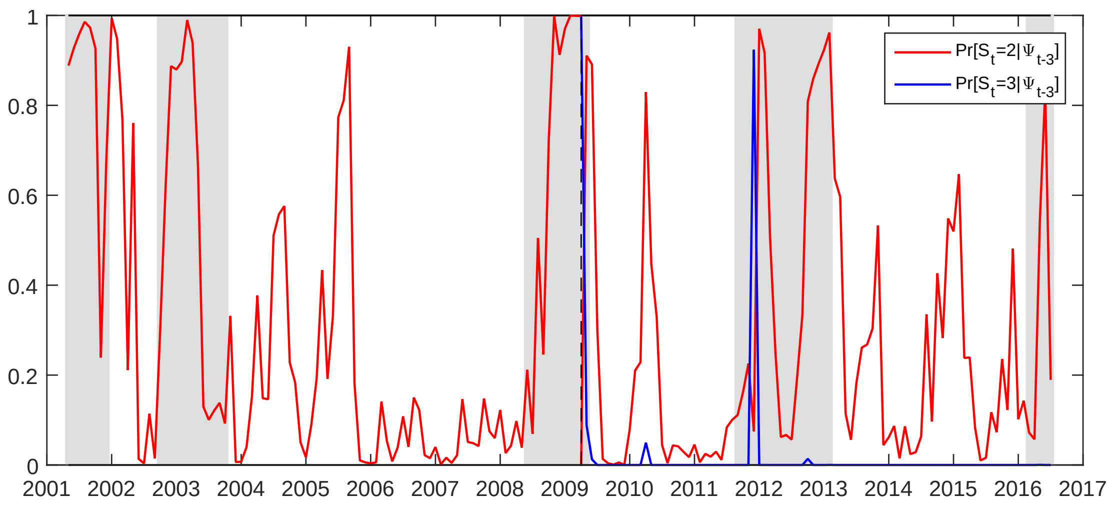
severe recession reaches one in December 2008. The economy gets back to an ordinary recession in April 2009. This information about the magnitude of the recession might have been extremely helpful at this point in time as it perfectly matches the steepest part of the Great Recession: industrial production dropped by -7.2% in January 2009 and GDP dropped by -4.6% in the first quarter of 2009. Further taking the publication lag of two months for industrial production and one quarter for GDP into account, the model could have given timely information about the economic situation at that time and thus provided background for policy-makers to counteract the situation before knowing how deep the recession really was.

5.3 Forecasting German business cycle turning points

Markov-switching models can also be used to forecast future turning points. While nowcasting business cycle turning points in real time is generally difficult enough and accuracy deteriorates quickly with the forecast horizon (Hamilton, 2011), our selection of early indicators that lead GDP by up to three months enables us to directly filter the probabilities $Pr[S_t = i|\Psi_{t-3}]$ from the data. It turns out that the probability forecasts of both the two-state and three-state models are somewhat more volatile than the corresponding nowcasts. This is not surprising because less information is available. Technically, this is reflected in the fact that the nowcasts are smoothed probabilities while the 3-month ahead forecasts are only filtered.

To save space, we solely report the predicted recession probabilities of the combined MS(2)/MS(3)-DFM with the shift taking place in December 2008 as discussed above.[23] Figure 9 shows that they contain very useful information. For example, in July 2008 the model forecasts a recession with almost 100 percent probability for October which is remarkable as most forecasters identified the recession not before November (Heilemann and Schnorr-Bäcker, 2016). It also predicts the recovery very timely. The forecast made in January 2009 already predicts for April 2009 that the severe recession ends and the economy is back in a normal recession. And in March 2009 the model first predicts that the recession ends three months later in June 2009. We believe that this information would have been valuable at that time. For example, the German parliament passed a large stimulus package known as "Konjunkturpaket II" in February 2009, and in April the German public started to discuss another stimulus package because the end of the recession seemed far away.

Figure 9: Real-time forecast recession probabilities of MS(2)/MS(3)-DFM



Note: One-quarter ahead recession probability forecasts $Pr[S_t = i|\Psi_{t-3}]$, $i = 2, 3$, of a MS(2)/MS(3)-DFM recursively estimated with real-time data. The dashed line indicates the shift from MS(2)-DFM to MS(3)-DFM which identified in December 2008 and thus effective for a three-month ahead forecast in March 2009. Shaded areas correspond to the recessions from the benchmark business cycle chronology from Section 4.4.

[23] Results for the single MS(2)-DFM and MS(3)-DFM models are available upon request.

5.4 Point forecasts of German GDP

Chauvet and Potter (2013) compare a large number of GDP-forecasting models including linear univariate and multivariate time series models, DSGE models and Markov-switching models. They find that MS-DFMs are by a large difference the most successful models in predicting GDP during US recessions in real time and even outperform expert forecasts from the Blue Chip Survey. To check whether they are also useful for predicting German GDP, we conduct an out-of-sample forecast experiment using real-time data. Our MS-DFMs do not include GDP and thus do not provide directly a GDP forecast. Therefore, we augment an autoregressive distributed lag (ADL) model for quarterly GDP growth with the estimated factor and the smoothed recession probabilities,

$$\Delta \log(\mathrm{GDP}_{t+h}) = c + \sum_{j=1}^{p} \alpha_j \Delta \log(\mathrm{GDP}_{t-j}) + \sum_{j=0}^{r} \gamma_j f_{t-j} + \sum_{j=0}^{s} \delta_j \Pi_{t-j} + \varepsilon_t, \tag{13}$$

where h denotes the forecast horizon, f_t denotes the quarterly average of the monthly factor, and Π_{t-j} is the quarterly average of the smoothed probability that period $t-j$ experiences a recession. Note that we use a direct rather than an iterative forecasting procedure. We compare the performance of the following forecasting models including a nested benchmark AR-model:

- AR: Our benchmark is a purely autoregressive model with p lags ($\gamma_j = \delta_j = 0$).
- ADL-DFM(1): This is a one-state, i.e. linear, dynamic factor model including p lags of GDP growth and r lags of the factor and $\delta_j = 0$ as there are not switches between states. We consider this model in order to check whether including additional information via a linear factor is already sufficient to improve upon the benchmark AR forecasts or whether a Markov-switching framework is essential.
- ADL-DFM(2) and ADL-DFM(3): These are ADL models which include p lags of GDP growth, r lags of a state-dependent factor and s lags of the recession probabilities generated by the MS-DFM(2)and the MS-DFM(3) model, respectively. For the latter it turned out that distinguishing between mild and severe recessions did not improve forecasting power which is why we only report results based on the joint probability $Pr[S_t = 2|\Psi_T] + Pr[S_t = 3|\Psi_T]$.
- ADL-DFM(2&3): This is an ADL model which includes p lags of GDP growth and r lags of the factor and s lags of the recession probabilities generated by the MS-DFM(2) or MS-DFM(3) depending on which one is preferred by the BIC. The switch from the MS-DFM(2) to the MS-DFM(3) occurs in the fourth quarter of 2008.

We recursively construct real-time nowcasts ($h = 0$) and h-step forecasts for $h = 1, \ldots, 4$ quarters based on an expanding window of vintage data.[24] Since we apply direct-step forecasting,

[24] A real-time nowcast of, say, 2001Q1 uses the data vintage available at the end of this quarter which includes—due to its one-quarter publication lag—GDP until 2000Q4. Correspondingly, an h-step ahead forecast is based on the data vintage available at the end of 2001Q1$-h$ which includes GDP until 2000Q4$-h$. Note that some indicators also have publication lags but the dynamic factor model has not because it filters out the last observation

for each model we consider one lag order specification per forecast horizon h. We proceed as follows. It is a well-known feature of German GDP growth that it has almost no autocorrelation (see, e.g., Pirschel and Wolters, 2014, for a comparison of autocorrelation functions of German and US GDP). Therefore, we include only one lag of GDP ($p = 1$) in all specifications. The recession probability Π_t is a first-order Markov process and includes by construction all relevant information which is why, at least theoretically, it is not necessary to include distributed lags. Since, in addition, Π_t leads GDP by one quarter, we include solely its first lag in the nowcast specifications ($s = 1$, $\gamma_0 = 0$) and its contemporaneous value in the forecast specifications ($s = 0$).[25] The factor f_t also leads GDP by one quarter. Therefore, we again exclude its current value from the nowcast specifications ($\delta_0 = 0$), but include it in the forecast specifications. At each recursion of our out-of-sample forecasting experiment we then choose the maximum lag order r as the one that minimizes the BIC.

We evaluate nowcasts and forecasts over the sample 2001Q1 to 2016Q2. Since GDP figures are subject to data revisions, we compare each forecast with the realisation published two quarters later. For example, a nowcast of 2001Q1 is compared with the value released by the end of 2001Q3. Exceptions are the major revisions of the German national account in 2005, 2011 and 2014. Here we use the last release before the revision to ensure that we take into account early data revisions but abstract from benchmark revisions which are difficult to forecast.

Before evaluating nowcasts and forecasts systematically based on RMSEs we graphically inspect the main characteristics of the different forecasting models. Figure 10 shows nowcasts and one-step ahead forecasts for the AR, ADL-DFM(1), ADL-DFM(2) and the ADL-DFM(3) model. The simple AR model captures the mean of GDP growth well, but mostly misses expansions and recessions. Both nowcast and forecast are basically flat until the Great Recession when the AR model reacts too late and too moderate. Afterwards, the nowcast becomes somewhat more accurate, while the forecast remains always close to the sample mean. Given the weak autocorrelation of German GDP, the result is not surprising.

The ADL-DFM(1) includes additional information via the factor estimated from a one-state, and thus linear, dynamic factor model. It turns out that this information and in particular the leading property of the factor is extremely valuable in generating accurate predictions. The model detects most turning points and misses only some episodes like the strong expansion in 2006 or the spike of GDP growth in 2010. It gets the timing of the largest drop in GDP during the Great Recession right, even one quarter in advance, though by far not its actual depth. After the Great Recession the model considerably overestimates the strength of the recovery.

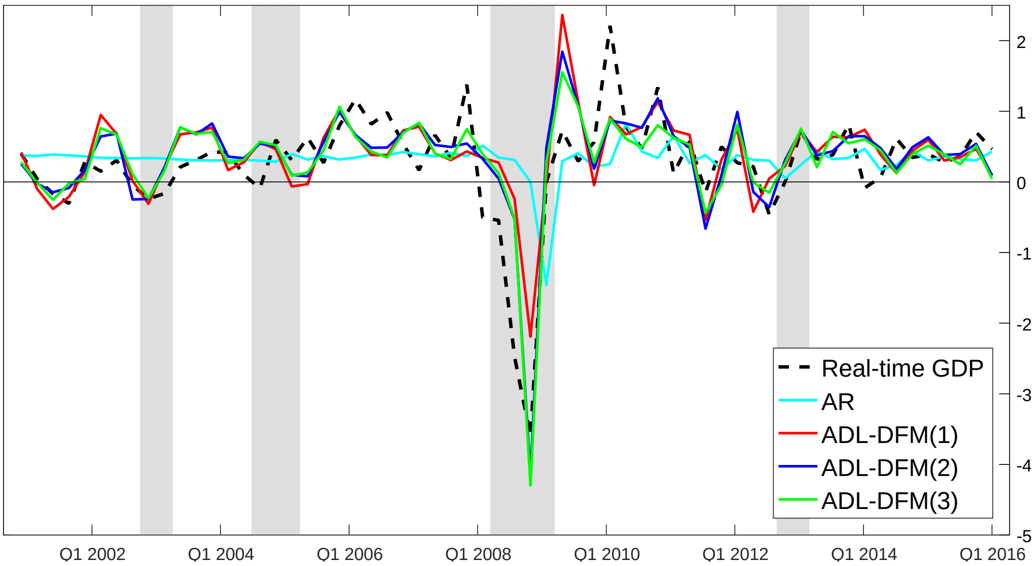
Turning to the Markov-switching models, both the ADL-DFM(2) and ADL-DFM(3) models improve during expansions and normal recessions only slightly upon the ADL-DFM(1) model. This changes, however, during the Great Recession and the subsequent recovery when their nowcasts for 2009Q1 are almost exactly correct and their one-quarter ahead forecasts for 2009Q1 outperform the ADL-DFM(1) model by a noticeable amount, even though they still underpredict the actual depth of the recession. During the recovery from the Great Recession, they again make

of the vintage based on the information contained in surveys and commodity prices that are published without delay.

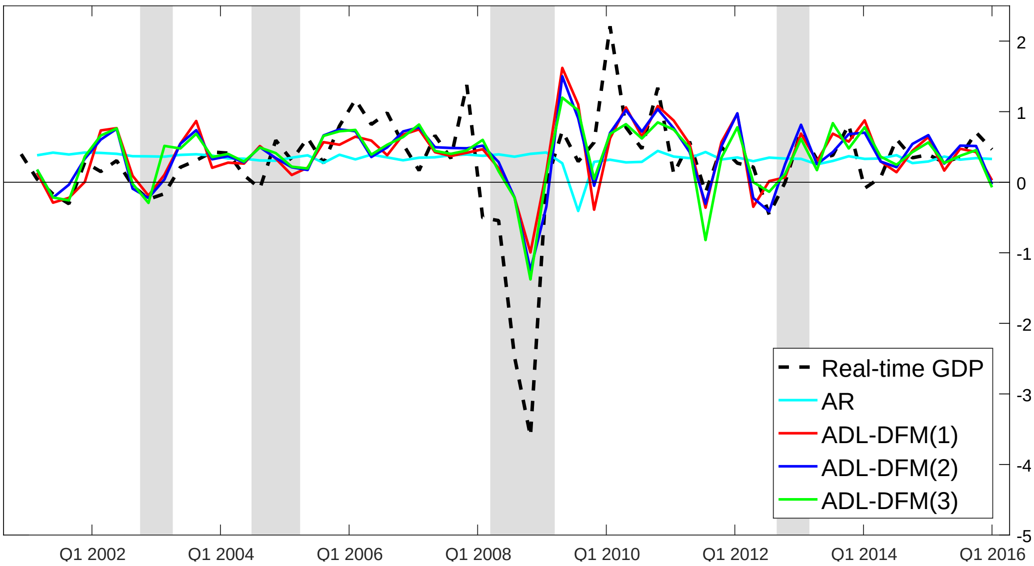
[25] We checked specifications that allowed higher choices for s but got worse RMSEs which supports our argument.

Figure 10: Real-time nowcasts and one-quarter ahead forecasts of GDP growth

(a) Nowcasts ($h = 0$)



(b) Forecasts ($h = 1$)



Note: Nowcasts and forecasts are based on real-time GDP and quarterly averages of monthly recession probabilities and dynamic factors. Shaded areas correspond to recessions according to the Bry-Boschan algorithm.

more accurate predictions than the ADL-DFM(1) model. It is further noticeable that except for the Great Recession the differences between the nowcasts and the 1-quarter ahead forecast are surprisingly small for all versions of the ADL-DFM framework. The 1-quarter forecasts are almost as accurate as the nowcast.

Based on the graphical analysis we conclude that in normal times it is sufficient to use the leading information extracted by a linear factor model. In contrast, during highly volatile times like the Great Recession and the subsequent recovery, predictions improve substantially when applying the MS-DFM(2) and MS-DFM(3) to account for the potential nonlinearity induced by those extraordinary business cycle movements.

As to the question whether to specify two or three states, we find that the predictions of the ADL-DFM(2) and ADL-DFM(3) are extremely close to each other, both before and after the Great Recession. Hence, using the information provided by the three-state Markov-switching model throughout the entire sample does not worsen GDP forecast accuracy despite the erratic switches between states before the Great Recession documented, inter alia, in Figure 5. Consequently, it does not make a difference here when the real-time model selection approach discussed above is applied. Since the shift from two to three states is detected in 2008Q4, the predictions of the ADL-DFM(2&3) model, which uses the information provided by the combined MS(2)/MS(3)-DFM, equal the predictions of the ADL-DFM(2) until the 2008Q4 and the ADL-DFM(3) thereafter. This is why we do not include these predictions in Figure 10.

Table 5 reports RMSEs relative to the AR model for forecasts up to $h=4$. To test whether the forecast are significantly different from the benchmark AR-model, we employ the test proposed by Clark and West (2007). We find that all factor models provide significantly better predictions than the AR benchmark up to forecast horizon $h = 2$, with decreasing margin as the forecast horizon h increases. For a horizon of $h = 3$, only the ADL-DFM(1) outperforms the benchmark, and for $h = 4$ the AR model dominates even if not significantly so. These results are not surprising as by construction the factor leads GDP by only one quarter. Hence, for higher forecast horizons, the information provided by the factor models is much less relevant while the additional parameter estimation uncertainty remains unchanged. However, using the information from the MS-DFM is beneficial for forecasting GDP growth in spite of the low persistence of German GDP. In line with the graphical inspection, we also find that differences between the ADL-DFM(2), ADL-DFM(3), and ADL-DFM(2&3) models are rather small, especially for forecast horizons of up to two quarters.

Table 5: Relative RMSEs

Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
ADL-DFM(1)	0.7669***	0.8815***	0.9205***	0.9431***	1.2943
ADL-DFM(2)	0.6565***	0.8215***	0.8839***	1.3114	1.4523
ADL-DFM(3)	0.6472***	0.8616***	0.8983***	1.2471	1.5145
ADL-DFM(2&3)	0.6426***	0.8602***	0.9072***	1.2653	1.5201

Notes: Root mean squared errors relative to an AR-benchmark model. *, ** and *** denotes significance on the 15% level, 10% level and 5% level, respectively, according to the Clark-West test with Newey-West standard errors.

The graphical analysis showed that the Markov-switching models perform particularly well during recessions. Hence, it is of interest to analyze differences in forecast precision between recessions and expansions systematically. To this end, we employ the quarterly version of the Bry-Boschan algorithm (Harding and Pagan, 2002) to GDP. The recession subsample includes 11 quarters (2002Q4-2003Q1, 2004Q3-2005Q1, 2008Q2-2009Q1, and 2012Q4-2013Q1), while the expansion subsample covers the remaining 55 quarters. Table 6 reports the corresponding RMSEs relative to the AR-benchmark model. These results confirm the finding by Chauvet and Potter (2013) that the advantage of Markov-switching models is largest during recessions. Interestingly, these models also improve upon the linear factor model during expansions, albeit to a smaller extent.

Table 6: Relative RMSEs for recessions

Model	$h=0$	$h=1$	$h=2$	$h=3$	$h=4$
			Recessions		
ADL-DFM(1)	0.6186**	0.8154**	0.8567**	0.9100**	1.3229
ADL-DFM(2)	0.4910**	0.7274***	0.7634**	0.9923	1.5005
ADL-DFM(3)	0.5111**	0.7920**	0.8611**	1.1746	1.4960
ADL-DFM(2&3)	0.5019**	0.7831**	0.8643**	1.1715	1.5006
			Expansions		
ADL-DFM(1)	0.9812***	1.0186***	1.0665***	1.0184	1.2160
ADL-DFM(2)	0.8816***	1.0066***	1.1353***	1.8687	1.3172
ADL-DFM(3)	0.8403***	1.0048***	0.9871**	1.4066	1.5621
ADL-DFM(2&3)	0.8406***	1.0171***	1.0087***	1.4669	1.5705

Notes: Relative root mean squared errors during recessions and expansions. *, ** and *** denotes significance on the 15% level, 10% level and 5% level, respectively, according to the Clark-West test with Newey-West standard errors.

6 Conclusion

We demonstrate that Markov-switching dynamic factor models together with a flexible variable pre-selection algorithm are an appropriate device to predict and date business cycle turning points for the German economy. It turns out that a three-state model is more sensitive than a two-state model and provides a better ex-post characterization of the German business cycle, especially because it identifies the Great Recession as a severe recession. Using real-time data we show that nowcasts and one-quarter ahead forecasts capture business cycle dynamics in Germany well even though German GDP growth is characterized by very low persistence.

During the Great Recession the model predicts the timing of events one quarter in advance starting with the initially mild downturn, the severe recessionary phase afterwards, and finally the recovery. Further, a comparison of the two- and three-state model clearly signals that the three-state model would have been preferable in December 2008 right before the biggest downturn of the German economy. Hence, for professional forecasters using this framework during the Great Recession would have been valuable to predict events systematically based on leading indicators. Moreover, the framework would have been highly useful for policymakers in order to plan the timing of policies to mitigate the crisis without the danger of stimulating the economy when the recovery was already on the way.

References

Aastveit, K. A., Jore, A. S., and Ravazzolo, F. (2016). Identification and real-time forecasting of Norwegian business cycles. International Journal of Forecasting, 32(2):283–292.

Abberger, K. and Nierhaus, W. (2010). Markov-Switching and the Ifo Business Climate: the Ifo Business Cycle Traffic Lights. OECD Journal: Journal of Business Cycle Measurement & Analysis, 2010(2).

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., and Veronese, G. (2010). New Eurocoin: tracking economic growth in real time. The Review of Economics and Statistics, 92(4):1024–1034.

Anas, J., Billio, M., Ferrara, L., and Mazzi, G. L. (2008). A system for dating and detecting turning points in the euro area. Manchester School, 76(5).

Artis, M., Krolzig, H.-M., and Toro, J. (2004). The European business cycle. Oxford Economic Papers, 56(1):1–44.

Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. Journal of Econometrics, 146(2):304–317.

Bandholz, H. and Funke, M. (2003). In search of leading indicators of economic activity in Germany. Journal of Forecasting, 22(4):277–297.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? Journal of Econometrics, 132:169–194.

Boldin, M. D. (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 & Econometrics, 1(1).

Camacho, M. and Perez-Quiros, G. (2007). Jump-and-rest effect of US business cycles. Studies in Nonlinear Dynamics & Econometrics, 11(4).

Camacho, M., Perez-Quiros, G., and Poncela, P. (2012). Markov-switching dynamic factor models in real time. CEPR Discussion Papers 8866, C.E.P.R.,Centre for Economic Policy Research.

Camacho, M., Perez-Quiros, G., and Poncela, P. (2014). Green shoots and double dips in the euro area: A real time measure. International Journal of Forecasting, 30(3):520–535.

Camacho, M., Perez-Quiros, G., and Poncela, P. (2015). Extracting nonlinear signals from several economic indicators. Journal of Applied Econometrics, 30(7):1073–1089.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. International Economic Review, 39(4):969–996.

Chauvet, M. and Hamilton, J. D. (2006). Dating business cycle turning points. Contributions to Economic Analysis, 276:1–54.

Chauvet, M. and Potter, S. (2013). Forecasting output. Handbook of Economic Forecasting, 2(Part A):141–194.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics, 138(1):291–311.

Diebold, F. X. and Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. The Review of Economics and Statistics, 78(1):67–77.

Doz, C. and Petronevich, A. (2016). Dating Business Cycle Turning Points for the French Economy: An MS-DFM approach. Dynamic Factor Models (Advances in Econometrics, Volume 35) Emerald Group Publishing Limited, 35:481–538.

Drechsel, K. and Scheufele, R. (2012). The performance of short-term forecasts of the German economy before and during the 2008/2009 recession. International Journal of Forecasting, 28(2):428–445.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. The Annals of Statistics, 32(2):407–499.

Ferrara, L. (2003). A three-regime real-time indicator for the US economy. Economics Letters, 81(3):373–378.

Fritsche, U. and Stephan, S. (2002). Leading Indicators of German Business Cycles. An Assessment of Properties/Frühindikatoren der deutschen Konjunktur. Eine Beurteilung ihrer Eigenschaften. Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics, 222(3):289–315.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57(2):357–384.

Hamilton, J. D. (2005). What's real about the business cycle? Federal Reserve Bank of St. Louis Review, (Jul):435–452.

Hamilton, J. D. (2011). Calling recessions in real time. International Journal of Forecasting, 27(4):1006–1026.

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

Heilemann, U. and Schnorr-Bäcker, S. (2016). Could the start of the German Recession 2008-2009 have been foreseen? Evidence from real-time data. George Washington University Research Program on Forecasting Working Paper No. 2016-003.

Ho, W.-Y. A. and Yetman, J. (2012). Does US GDP stall? BIS Working Papers 387, Bank for International Settlements.

Ivanova, D., Lahiri, K., and Seitz, F. (2000). Interest rate spreads as predictors of german inflation and business cycles. International Journal of Forecasting, 16:39–58.

Kholodilin, K. A. (2005). Forecasting the German Cyclical Turning Points: Dynamic Bi-Factor Model with Markov Switching. Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics, 225(6):653–674.

Kholodilin, K. A. and Siliverstovs, B. (2006). On the forecasting properties of the alternative leading indicators for the German GDP: Recent evidence. Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics, 226(3):234–259.

Kim, C.-J. (1994). Dynamic linear models with Markov-switching. Journal of Econometrics, 60(1-2):1–22.

Kim, C.-J. and Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. Review of Economics and Statistics, 80(2):188–201.

Kim, M.-J. and Yoo, J.-S. (1995). New index of coincident indicators: A multivariate Markov switching factor model approach. Journal of Monetary Economics, 36(3):607–630.

Krolzig, H.-M. and Toro, J. (2001). A new approach to the analysis of business cycle transitions in a model of output and employment. Oxford University Department of Economics Discussion Paper 59.

Layton, A. P. and Smith, D. (2000). A further note on the three phases of the US business cycle. Applied Economics, 32(9):1133–1143.

Lehmann, R. and Wohlrabe, K. (2016). Looking into the black box of boosting: the case of Germany. Applied Economics Letters, 23(17):1229–1233.

Mariano, R. S. and Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. Journal of Applied Econometrics, 18(4):427–443.

McAdam, P. (2003). US, Japan and the euro area: comparing business cycle features. ECB Working Paper 283.

Morley, J. and Piger, J. (2012). The asymmetric business cycle. Review of Economics and Statistics, 94(1):208–221.

Nalewaik, J. J. (2011). Forecasting recessions using stall speeds. In Finance and Economics Discussion Series 2011-24, Federal Reserve Board. Citeseer.

Pirschel, I. and Wolters, M. (2014). Forecasting German key macroeconomic variables using large dataset methods. Kiel Working Papers 1925, Kiel Institute for the World Economy.

Schirwitz, B. (2009). A comprehensive German business cycle chronology. Empirical Economics, 37(2):287–301.

Schumacher, C. (2010). Factor forecasting using international targeted predictors: The case of German GDP. Economics Letters, 107(2):95–98.

Sichel, D. E. (1994). Inventories and the three phases of the business cycle. Journal of Business & Economic Statistics, 12(3):269–277.

Smith, A., Naik, P. A., and Tsai, C.-L. (2006). Markov-switching model selection using Kullback–Leibler divergence. Journal of Econometrics, 134(2):553–577.

Stock, J. H. and Watson, M. W. (2005). Understanding Changes In International Business Cycle Dynamics. Journal of the European Economic Association, 3(5):968–1006.

Vanhaelen, J.-J., Dresse, L., and De Mulder, J. (2000). The Belgian industrial confidence indicator: leading indicator of economic activity in the euro area? National Bank of Belgium Working Paper 12.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2):301–320.

7 Appendix

7.1 Construction of the state space form

We start defining the $(12 + nq)$-dimensional state vector

$$a_t = [f_t, \ldots, f_{t-11}, z_t', \ldots, z_{t-q+1}']'.$$

Now the measurement equations (1) and (2) can be jointly written as

$$y_t = Ba_t,$$

where

$$B = \begin{bmatrix} \gamma^{(h)} & 0_{n_h \times 1} & \cdots & 0_{n_h \times 1} & I_{n_h} & 0_{n_h \times n_s} & 0_{n_h \times (nq-q)} \\ \gamma^{(s)} & \gamma^{(s)} & \cdots & \gamma^{(s)} & 0_{n_s \times n_h} & I_{n_s} & 0_{n_s \times (nq-q)} \end{bmatrix}$$

and $\gamma^{(h)} = [\gamma_1^{(h)}, \ldots, \gamma_{n_h}^{(h)}]'$ and $\gamma^{(s)} = [\gamma_1^{(s)}, \ldots, \gamma_{n_s}^{(s)}]'$.

The transition equation can be written as

$$a_t = \mu_{S_t} + Fa_{t-1} + R\omega_t$$

using the following definitions. The system matrix is

$$F = \begin{bmatrix} F_{11} & 0_{12 \times nq} \\ 0_{nq \times 12} & F_{22} \end{bmatrix}$$

where F_{11} is the (12×12)-dimensional companion matrix of an AR(12) process with lag coefficients ϕ_1 to ϕ_{12} of which coefficients 3 to 12 restricted to zero because we only allow a maximum lag order of $p = 2$ for f_t, and F_{22} is the $(nq \times nq)$-dimensional companion matrix of an n-dimensional VAR process with q lags and coefficient matrices ψ_1 to ψ_q. The intercept vector is nonzero only for f_t and thus is

$$\mu_{S_t} = \left[\beta_{S_t}, 0_{1 \times (11+nq)}\right]'.$$

The vector of iid shocks, $\omega_t = [\eta_t, \varepsilon_t']'$, is iid normally distributed with mean zero and diagonal covariance matrix

$$Q \equiv \mathrm{E}(\omega_t \omega_t') = \begin{bmatrix} 1 & 0_{1 \times n} \\ 0_{n \times 1} & \Sigma_z \end{bmatrix}.$$

Finally, we define the coefficient matrix

$$R = \begin{bmatrix} R_{11} & 0_{12 \times n} \\ 0_{nq \times 1} & R_{22} \end{bmatrix},$$

where $R_{11} = [1, 0_{1 \times 11}]'$ and $R_{22} = [I_n, 0_{n \times (nq-n)}]'$.

7.2 Estimation of the MS-DFM

We employ the filter proposed by Kim (1994) to estimate the MS-DFM. Based on the initialization $a_{0|0} = (I - F)^{-1}\mu_{S_t}$ and $P_{0|0} = (I - F \otimes F)^{-1}\text{vec}(Q)$, the recursion consists of the usual prediction and updating steps. To this end, let us define $P_{t|t-1}^{(j,i)}$ as the variance of z_t conditional on Ψ_{t-1}, the information available in $t-1$, and on $S_t = j$ and $S_{t-1} = i$, $P_{t|t}^{(i)}$ as the variance of z_{t-1} conditional on Ψ_t and $S_{t-1} = i$, and equivalently $a_{t|t-1}^{(j,i)}$ and $a_{t-1|t-1}^{(i)}$. Then the prediction step is

$$a_{t|t-1}^{(j,i)} = Fa_{t-1|t-1}^{(i,k)} + \mu_{S_t}^{(j)}, \tag{14}$$

$$P_{t|t-1}^{(j,i)} = FP_{t-1|t-1}^{(i,k)}F' + RQR', \tag{15}$$

and the updating step is

$$a_{t|t}^{(j,i)} = a_{t|t-1}^{(j,i)} + K_t^{(j,i)}(y_t - Ba_{t|t-1}^{(j,i)}), \tag{16}$$

$$P_{t|t}^{(j,i)} = (I_{2n+p} - K_t^{(j,i)}B)P_{t|t-1}^{(j,i)}, \tag{17}$$

where the Kalman gain is defined by $K_t^{(j,i)} = P_{t|t-1}^{(j,i)}B'(BP_{t|t-1}^{(j,i)}B')^{-1}$. However, each recursion generates an m-fold increase in the number of states to be considered. Therefore, we apply the approximation by Kim (1994),

$$a_{t|t}^{(j)} = \frac{\sum_{i=1}^{m} Pr[S_{t-1} = i, S_t = j|\Psi_t]a_{t|t}^{(j,i)}}{Pr[S_t = j|\Psi_t]} \tag{18}$$

$$P_{t|t}^{(j)} = \frac{\sum_{i=1}^{m} Pr[S_{t-1} = i, S_t = j|\Psi_t](P_{t|t}^{(j,i)} + (a_{t|t}^{(j)} - a_{t|t}^{(j,i)})(a_{t|t}^{(j)} - a_{t|t}^{(j,i)})')}{Pr[S_t = j|\Psi_t]}, \tag{19}$$

which reduces the number of possible states of $a_{t|t}$ and $P_{t|t}$ to m per period by taking weighted averages over the states and feeding them into the prediction steps (14) and(15).

The corresponding log likelihood function is obtained by Hamilton (1989):

$$ln\ L = \sum_{t=1}^{T} ln\left(\sum_{S_t=1}^{m}\sum_{S_{t-1}=1}^{m} f(y_t|S_t, S_{t-1}, \Psi_{t-1})Pr(S_t{=}j, S_{t-1}{=}i|\Psi_{t-1})\right). \tag{20}$$

Evaluating it requires calculating the weights $Pr(S_t{=}j, S_{t-1}{=}i|\Psi_{t-1})$, which can be expressed as the product of the probability of being in a certain regime at period $t{-}1$ and the corresponding transition probability:

$$Pr(S_t{=}j, S_{t-1}{=}i|\Psi_{t-1}) = p_{ij}Pr(S_{t-1} = i|\Psi_{t-1}). \tag{21}$$

Updating this probability with information up to period t yields the filtered probabilities:

$$\begin{aligned} Pr(S_t=j, S_{t-1}=i|\Psi_t) &= \frac{f(S_t=j, S_{t-1}=i, y_t|\Psi_{t-1})}{f(y_t|I_{t-1})} \\ &= \frac{f(y_t|S_t=j, S_{t-1}=i, \Psi_{t-1})Pr(S_t=j, S_{t-1}=i|\Psi_{t-1})}{\sum_{S_t=1}^{m}\sum_{S_{t-1}=1}^{m} f(y_t|S_t=j, S_{t-1}=i, \Psi_{t-1})Pr(S_t=j, S_{t-1}=i|\Psi_{t-1})} \end{aligned}$$

and

$$Pr(S_t=j|\Psi_t) = \sum_{i=1}^{m} Pr[S_{t-1}=i, S_t=j|\Psi_t].$$

Based on an initialization—we employ the unconditional probabilities as derived by Hamilton (1989)— the steps can be iterated forward over the sample to obtain the filtered probabilities for each period. Along with the filter recursions, this yields all the information we need to estimate the latent dynamic regime dependent factor as well as the Markov-switching probabilities.

7.3 Data: indicators, sources, and real-time selection

The majority of the series is downloaded from Thomson Reuters Datastream, while the remaining indicators are directly obtained from the German Bundesbank, the ECB and the OECD. Tables 7 and 8 list the hard and survey indicators, respectively, together with their sources and the transformations we applied. For the hard indicators we report the sources for both our ex-post analysis and our real-time analysis. The survey indicators are stationary by construction and thus left untransformed. They are published without (noticeable) revisions, hence the use of a specific real-time data set is not necessary.

The hard and survey indicators selected by the LARS algorithm in each step of our real-time analysis are reported in Figures 11 and 12. Note that we exclude from the Figures all variables that are never selected.

Table 7: Hard Indicators

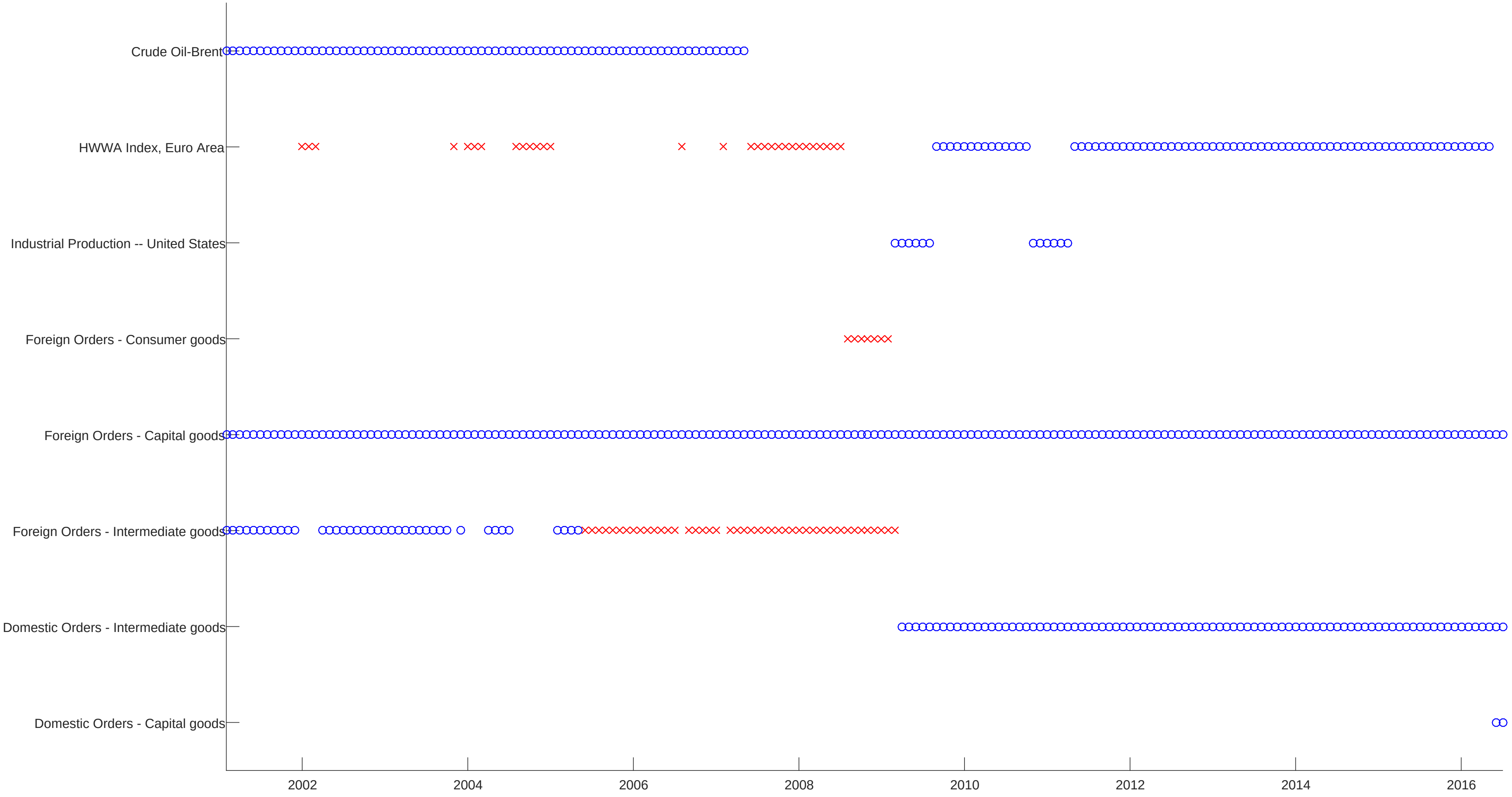
Name of Series	**Datastream code** (ex post data set)	**Real-time code** (real-time data set)	**Transformation** diff	log
New Orders				
Domestic Orders – Capital goods, Germany	BDDCPORDG	BBKRT.M.DE.Y.I.IO2.ACM03.C.I	yes	yes
Domestic Orders – Consumer goods, Germany	BDDCNORDG	BBKRT.M.DE.Y.I.IO2.ACM04.C.I	yes	yes
Domestic Orders – Intermediate goods, Germany	BDDBPORDG	BBKRT.M.DE.Y.I.IO2.ACM02.C.I	yes	yes
Foreign Orders – Consumer goods, Germany	BDOCNORDG	BBKRT.M.DE.Y.I.IO3.ACM04.C.I	yes	yes
Foreign Orders – Capital goods, Germany	BDOCPORDG	BBKRT.M.DE.Y.I.IO3.ACM03.C.I	yes	yes
Foreign Orders – Intermediate goods, Germany	BDOBPORDG	BBKRT.M.DE.Y.I.IO3.ACM02.C.I	yes	yes
Interest rates				
Yield on German Federal securities, residual maturity 9 to 10 years	BDT0557	no revisions	yes	no
Fibor - 3 Month (Monthly Average)	BDINTER3	no revisions	yes	no
Term Spread on German Federal securities - (10y-3m)	BDT0557-BDINTER3	no revisions	yes	no
Commodity prices				
Crude oil – Brent	OILBREN	no revisions	yes	yes
HWWA Index, Euro	BDHWWAINF	no revisions	yes	yes
General economic indicators				
Dax performance index	DAXINDX	no revisions	yes	yes
German Contibution to EMU M2	BDTXI302A	no revisions	yes	yes
Employed persons - Overall economy, Germany	last vintage	BBKRT.M.DE.S.L.BE1.CA010.P.A	yes	yes
Consumer Prices - All categories, Germany	last vintage	BBKRT.M.DE.Y.P.PC1.PC100.R.I	yes	yes
Foreign markets				
US Industrial Production	last vintage	from OECD.Stat	yes	yes

Notes: “no revisions” indicates that the series is assumed to be published without revisions, “last vintage” indicates that we use the last vintage of the real-time data set as ex post data.

Table 8: Survey Indicators

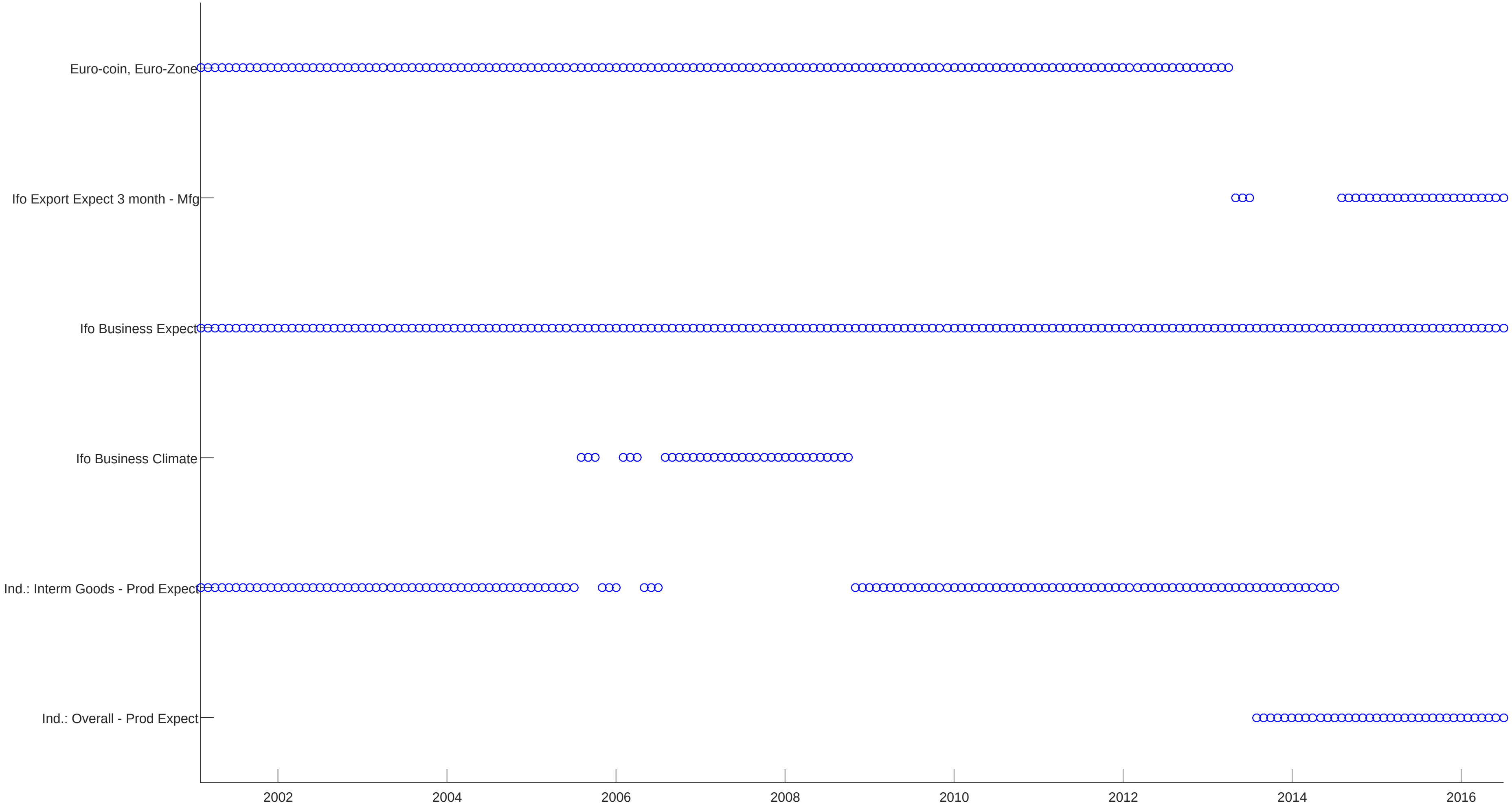
Name of Series	**Datastream-Code**	**Source**
Industry overall production expectation, Germany	BDTTA5BSQ	European commission
Industry intermediate goods production expectation, Germany	BDITM5.BQ	European commission
Industry investment goods production expectation, Germany	BDIVE5.BQ	European commission
Industry overall employment expectations, Germany	BDTTA7BSQ	European commission
Industry overall Order Books, Germany	BDTTA2BSQ	European commission
Consumer confidence indicator, Germany	BDCNFCONQ	European commission
Consumer Survey: Economic Situation next 12 mth., Germany	BDEUSCEYQ	European commission
Economic sentiment indicator, Germany	BDEUSESIG	European commission
Economic sentiment indicator, Euro Zone	EKEUSESIG	European commission
Ifo business climate index, Germany	BDCNFBUSQ	Ifo institute
Ifo business climate manufacturing, Germany	BDIFDMTLQ	Ifo institute
Ifo business climate manufacturing capital goods, Germany	BDIFDMPLQ	Ifo institute
Ifo business expectations index, Germany	BDCYLEADQ	Ifo institute
Ifo business expectation (6 month) manufacturing, Germany	BDIFDMTKQ	Ifo institute
Ifo business expectation (6 month) capital good producers, Germany	BDIFDMPKQ	Ifo institute
Ifo export expectation (3 month), Germany	BDIFDMTJQ	Ifo institute
ISM purchasing managers index, USA	USCNFBUSQ	ISM institute
Belgium business indicator, Belgium	BGCNFBUSQ	National Bank of Belgium
Euro-coin, Euro-Zone	EMECOIN.Q	CEPR

Figure 11: Real-time variable selection—hard indicators



Note: Recursive real time variable selection with LARS-EN. Variables selected at ○ = *lag* 1, × = *lag* 2, ∗ = *lag* 3. *Variables which do not appear are never selected.*

Figure 12: Real-time variable selection—survey indicators



Note: Recursive real time variable selection with LARS-EN. Variables selected at ○ = *lag* 1, × = *lag* 2, ∗ = *lag* 3. *Variables which do not appear are never selected.*