Mixed-frequency models for tracking short-term economic developments in Switzerland

Alain Galli, Christian Hepenstrick and Rolf Scheufele

SNB Working Papers

2/2017

SCHWEIZERISCHE NATIONALBANK
BANQUE NATIONALE SUISSE
BANCA NAZIONALE SVIZZERA
BANCA NAZIUNALA SVIZRA
SWISS NATIONAL BANK

Legal Issues

ISSN 1660-7716 (printed version)
ISSN 1660-7724 (online version)

Mixed-frequency models for tracking short-term economic developments in Switzerland*

Alain Galli† Christian Hepenstrick‡ Rolf Scheufele§

January 3, 2017

Abstract

We compare several methods for monitoring short-term economic developments in Switzerland. Based on a large mixed-frequency data set, the following approaches are presented and discussed: factor-based *information combination* approaches (including factor model versions based on the Kalman filter/smoother, a principal component based version and the three-pass regression filter), a *model combination* approach resting on MIDAS regression models and a *model selection* approach using a specific-to-general algorithm. In an out-of-sample GDP forecasting exercise, we show that the considered approaches clearly beat relevant benchmarks such as univariate time-series models and models that work with one or a small number of indicators. This suggests that a large data set is an important ingredient for successful real-time monitoring of the Swiss economy. The models using a large data set particularly outperform others during and after the Great Recession. Forecast pooling of the most-promising methods turns out to be the best option for obtaining a reliable nowcast for the Swiss economy.

JEL classification: C32, C53, E37

Keywords: Mixed frequency, GDP nowcasting, forecasting, Switzerland

*We want to thank Massimiliano Marcellino as well as participants at IAAE conference 2016 in Milano, the ISF 2016 in Santander and the CIRET conference 2016 in Copenhagen for valuable comments. The views, opinions, findings, and conclusions or recommendations expressed in this paper are strictly those of the authors. They do not necessarily reflect the views of the Swiss National Bank. The SNB takes no responsibility for any errors or omissions in, or for the correctness of, the information contained in this paper. In particular, the models described in the paper are not part of the suite of models used for Swiss National Bank's inflation forecast, published on a quarterly basis in its Quarterly Bulletin.

†Swiss National Bank and University of Bern. Swiss National Bank, Economic Analysis, P.O. Box, CH-8022 Zurich, Switzerland, e-mail: alain.galli@snb.ch

‡Swiss National Bank, Economic Analysis, P.O. Box, CH-8022 Zurich, Switzerland, e-mail: christian.hepenstrick@snb.ch

§Swiss National Bank, Economic Analysis, P.O. Box, CH-8022 Zurich, Switzerland, e-mail: rolf.scheufele@snb.ch

1 Introduction

Policy institutions such as central banks depend on the timely assessment of past, current and future economic conditions. However, many official statistics on economic development are released at low frequency and arrive with a substantial time delay. GDP, for example, is released approximately two months after the end of a quarter. This means that between different data releases, it is very difficult to track short-term economic developments.

For Switzerland, the problem of imperfect and incomplete information for recent short-term developments is even more pronounced. Compared to larger countries such as the US, Germany or the UK, it has no official monthly coincident indicators (e.g., industrial production, turnover or new orders). However, despite the shortage of early and frequent official data, other sources of information can be used to assess the current state of the economy, namely, financial indicators or surveys. Moreover, data on foreign and domestic trade are important data sources that are available monthly and prior to the release of GDP data.

In this paper, we assess several methods to extract information for tracking current economic conditions in Switzerland in a timely and efficient manner. This study relates to similar techniques applied for other countries. Evans (2005), Banbura, Giannone, and Reichlin (2011), Banbura and Rünstler (2011) and Kuzin, Marcellino, and Schumacher (2011) are some prominent examples. A common feature of the models we consider is that they are able to handle (a) large data sets and (b) indicators available at different frequencies and, therefore, can be characterized as mixed-frequency methods.

The forecasting methods under consideration can be grouped into three categories: information combination, model combination and model selection. For *information combination*, we rely on factor model techniques.[1] These models assume that a small number of latent variables can be used to describe the fluctuations in the given data set and that those factors are highly correlated with the business cycle and GDP (see, e.g., Stock and Watson 2002b, 2011). Four different factor procedures are compared: first, a standard dynamic factor model (DFM) using the Kalman filter that is estimated in

[1] Large Bayesian VARs would be an alternative, but so far, they have only been used with single frequencies, and adapting them to the use of mixed frequencies would create computational challenges.

a two-step procedure; second, a small-scale factor model that is based on a pre-selected variable set; third, a factor model based on the three-pass regression filter; fourth, an expectation-maximization algorithm for the unbalanced data set, extracting static factors by means of principal component analysis.

An alternative to the factor-based information combination is *model combination*, i.e., combine forecasts from different models. In this case, single indicator models are used to explain GDP within a linear regression framework. For higher-frequency variables (weekly and monthly indicators), MIxed DAta Sampling (MIDAS) methods are used to take into account the temporal aggregation issue.

The final category is *model selection*. In this case, we employ a model selection technique based on a specific-to-general approach that can handle large data sets and mixed-frequency settings.

As benchmark models, we take into account univariate time series models as well as popular Swiss leading indicators such as the KOF economic barometer and the PMI.

To compare the models under investigation, we conduct a pseudo out-of-sample forecast exercise. Using a very large data set of more than 600 variables at weekly, monthly and quarterly frequency and taking into account two different states of information, we generate forecasts from the different models for the period 2005Q1-2015Q2.[2] With all our forecasting procedures, we take into account the problem of missing observations at the end of the sample (ragged-edge problem). We compare the forecasting performances by means of root mean squared forecast errors (RMSEs) and investigate their ranking over time by looking at different subperiods and by looking at rolling windows. Besides RMSEs, we investigate the models' ability to forecast business cycle phases.

Our results show that all approaches using the large data set clearly beat relevant benchmarks such as univariate time series models and indicator models that rely on a very small data set. Factor models based on a large data set are better than small-scale models and are notably more robust after the crisis. This finding is somewhat in contrast to earlier research (before the financial crisis) that points to potential issues of misspecification when including disaggregate information (e.g., Boivin and Ng 2006, Caggiano, Kapetanios, and

[2] The reason for not having a longer evaluation sample is that for the case of Switzerland, many economic indicators start only in 1990 or even later, which automatically limits the estimation sample of the models to some extent.

Labhard 2011). At least for Switzerland, the use of a broad range of indicators leads to relatively good and stable forecasting performance. Model combination techniques also perform well, but do not capture the crisis period very well. After the crisis, the most promising model combination performs similarly well to the best factor models, particularly for the two step ahead forecast. Finally, we show that a combination of our models under consideration produces even better results than the best single model procedure.

2 Forecasting approaches

In this paper, we compare several alternative forecasting approaches to deal with large data sets, irregular publication lags and mixed-frequency data. First, four information combination approaches are presented and discussed. Second, a model combination method based on single indicator models is outlined. Finally, we consider a model selection procedure based on specific-to-general modeling.

2.1 Information combination by factor models

The basic idea of factor models is that a small number of latent variables (factors) can well approximate the fluctuations of many macroeconomic variables. Typically, these factors are found to be highly correlated with the business cycle and GDP (see, e.g., Stock and Watson 2002b, 2011). Thus, a factor model allows us to characterize the interactions of a large number of variables within a relatively parsimonious system framework. Factor models have been extended to cope with ragged edges as well as with mixed-frequency data. Typically, these models perform very well in short-term forecasting (Banbura, Giannone, Modugno, and Reichlin 2013, Kuzin, Marcellino, and Schumacher 2013). Despite their great popularity, there is no final consensus about how to estimate factor models. The proposed methods range from principal component-based estimates to full system estimates (see, e.g., Stock and Watson 2011). In terms of forecasting ability, none of the proposed methods dominate the others in all settings (Kuzin, Marcellino, and Schumacher 2013). Therefore, we consider four different factor models to see which is best suited for the case of Switzerland: first, a standard dynamic factor model (DFM) based on a two-step estimation technique; second, a small-scale factor model that is based on a pre-selected variable set; third, a factor model based on the three-pass regression filter; and fourth, a simple principal component-based approach.

A factor model decomposes a (potentially large) panel of time series x_t into a common component c_t and an idiosyncratic error term u_t

$$x_t = c_t + u_t = \Lambda f_t + u_t. \tag{1}$$

The common component consists of a small number of factors f_t that are linked to the observed time series through the loading coefficients Λ. The idiosyncratic error term has a variance $\text{Var}(u_t) = \Sigma_u = \text{diag}(\sigma_1, ..., \sigma_N)$.

The four factor models we consider differ in the assumptions made to extract the unobserved factors from the data set and how the factors are translated into a GDP forecast. They also differ in how they combine monthly and quarterly indicators and how they deal with missing data at the beginning and end of the sample.

Factor model I: large-scale DFM, estimated in two steps

The first factor model augments (1) with a VAR(p)-law of motion for the factors

$$f_t = \Phi_1 f_{t-1} + ... + \Phi_p f_{t-p} + v_t, \tag{2}$$

where $\text{Var}(v_t) = \Sigma_v$. The two-step estimation of the resulting DFM has become one of the workhorses in factor-based short-term forecasting (see, e.g., Giannone, Reichlin, and Small 2008, Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler 2011).

In our specification, quarterly variables are defined as the average of monthly latent observations: $x^q_{i,t} = 1/3(x^m_{i,t} + x^m_{i,t-1} + x^m_{i,t-2})$ for $t = 3, 6, 9, ...$ To ensure consistency among monthly and quarterly variables, all non-stationary monthly series enter the data set as three month changes. This is consistent with having defined quarterly GDP growth as the three month average of monthly latent observations.[3]

[3]Note that Mariano and Murasawa (2003) and Banbura, Giannone, Modugno, and Reichlin (2013) proposed a slightly different aggregation rule where monthly series are included as monthly changes. However, applying their aggregation rule does lead to inferior forecasting results for Switzerland. Therefore, we apply the aggregation rule proposed by Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011).

The estimation of the parameters and factors follows Giannone, Reichlin, and Small (2008) and Doz, Giannone, and Reichlin (2011) and proceeds in the following way:

1. Principal component analysis is applied to the balanced data set of monthly indicators to estimate the common factors. Based on BIC, the four most important factors, according to their overall explanatory power, are used ($r = 4$) in the subsequent analysis.
2. Loadings and variances for each indicator (monthly and quarterly) are then obtained by OLS regressions on the factors (for quarterly variables, the variances have to be transformed into the higher frequency).
3. A VAR of the r principal components $\widehat{f}_t = \sum_{i=1}^{p} \Phi_i \widehat{f}_{t-i} + \widehat{v}_t$ is estimated to obtain $\widehat{\Phi}_i$ and $\widehat{\Sigma}_v$. The number of lags is set to 1 (which corresponds to the optimal BIC in most cases).
4. The Kalman smoother is used to obtain an updated estimate of $\widehat{f}_t$, and it provides iterative forecasts for the target variable

Factor model II: small-scale DFM

As a second setup, we apply the standard dynamic factor model to a small set of indicators, in the spirit of Aruoba, Diebold, and Scotti (2009) and Camacho and Martinez-Martin (2014). These two approaches use 6 and 11 time series, respectively. In selecting our indicators, we aimed at a comparable number. As discussed by Camacho and Martinez-Martin (2014), indicators should satisfy the following properties: a high statistical correlation with GDP growth, a shorter publication lag than GDP, a sufficiently long time series, and theoretical and empirical relevance. We combined these requirements together with the SNB staffs expert judgment to obtain the following final dataset consisting of 8 indicators (plus GDP): (1) Swiss PMI, which has a very short publication lag and is generally considered to be the best single indicator for the Swiss economy. (2) Real Swiss export growth, which is used in the official GDP estimation as the main indicator for manufacturing value added. (3) Real retail sales growth and (4) real import growth as the two most prominent monthly hard indicators on domestic demand conditions. (5) The percentage change of the number of unemployed as a timely monthly indicator for conditions in the labor market. (6) Production expectations in the manufacturing sector according to the KOF survey representing the second important domestic soft indicator besides the PMI. International demand conditions are captured with (7) Euro area industrial production as an important hard indicator and (8) the German ifo survey as a more timely soft indicator.

Following Mariano and Murasawa (2003), we assume that these 8 indicators and GDP are driven by one common factor f_t, so that for the monthly indicators (all but GDP growth), we have

$$x_{i,t} = \lambda_i f_{i,t} + e_{i,t} \tag{3}$$

and for quarterly GDP growth y_t we have

$$y_t = \lambda_y \left(f_t + 2f_{t-1} + 3f_{t-2} + 2f_{t-3} + f_{t-4} \right) + e^q_{y,t}, \tag{4}$$

where $e^q_{y,t} = e_{y,t} + 2e_{y,t-1} + 3e_{y,t-2} + 2e_{y,t-3} + e_{y,t-4}$.

For the state dynamics, it is assumed that the idiosyncratic errors follow an AR(2) process,

$$e_t = \sum_{j=1}^{J=2} \theta_j e_{t-j} + \varepsilon_t, \tag{5}$$

and the common factor follows an AR(1) process[4],

$$f_t = \phi f_t + v_t, \tag{6}$$

with $\varepsilon_{i,t}$ and v_t being i.i.d. We rewrite the model as a state space system that is estimated using Maximum Likelihood (for details, we refer to Mariano and Murasawa 2003, Camacho and Perez-Quiros 2010). Using the estimated state space system, one can filter the data and form forecasts for the factor and for the individual indicators. In our application, we will focus on the resulting forecasts for real GDP growth.

[4]To simplify the analysis, the number of lags for the law of motion of the factor and the idiosyncratic errors are exogenously fixed. The results do not materially change when choosing a different number of lags.

Factor model III: Three-pass regression filter

The three-pass regression filter (3PRF) is used as described in Kelly and Pruitt (2013) and Kelly and Pruitt (2015) and is extended to a mixed-frequency environment by Hepenstrick and Marcellino (2016). Besides frequency-mixing, the latter authors discuss a number of issues that arise in the application of the filter, particularly the presence of indicators with different publication lags and various approaches to generate the GDP forecast. In the present paper, we use one specification of the 3PRF that performs particularly well for Switzerland and that is very easy to implement. For results on alternative specifications, we refer to Hepenstrick and Marcellino (2016).[5] The algorithm can be represented in four steps. Step 0 deals with the ragged edge arising for differing publication lags, steps 1 and 2 ("Pass 1" and "Pass 2" of the 3PRF) extract a factor from the indicator set, and step 3 ("Pass 3") translates the factor into a GDP forecast:

0. Vertical realignment as suggested by Altissimo, Cristadoro, Forni, Lippi, and Veronese (2010): Define T as the last month for which data are observed. For each indicator series, generate a new series of its lags such that the last observation is in T as well. We denote the standardized realigned series as $\tilde{x}_{i,t}$.
1. Pass 1 of 3PRF: for each indicator $i = 1, ..., N$, run a time series regression of the indicator aggregated to the quarterly frequency, $x^q_{i,t}$, on GDP growth, y_t

$$\tilde{x}^q_{i,t} = \phi_i y_t + \epsilon_{i,t}. \tag{7}$$

   Drop all indicators with a p-value of the F-test that is higher than 10%. For the other indicators, retain the slope estimates, $\hat{\phi}_i$.
2. Pass 2 of 3PRF: for each month $t = 1, ..., T$, run a cross-section regression of the remaining indicators $x_{i,t}$ on the slope estimates

$$x_{i,t} = \alpha_t + f_t \hat{\phi}_i + \varepsilon_{i,t}. \tag{8}$$

   The time series of the resulting slopes $\hat{f}_t$ is the 3PRF factor.
3. Pass 3 of 3PRF: Translate the factor estimates $\hat{f}_t$ into a GDP forecast using the

[5] Specific results for Switzerland are available upon request.

U-MIDAS approach of Foroni, Marcellino, and Schumacher (2015). To do this, K quarterly series, $\hat{f}_t^1, ..., \hat{f}_t^K$, are generated from the monthly factor, $\hat{f}_t$. The first series consists of the factor values of the month within the running quarter with the last observation. The second series contains the values of the month before, etc.[6] Next, the new quarterly series, $\hat{f}_t^1, ..., \hat{f}_t^K$, are used as explanatory variables in the forecast for GDP growth in $t + h$

$$y_{t+h} = \beta_{0,h} + \sum_{k=0}^{K} \beta_{1,h,k} \hat{f}_t^k + \eta_{h,t}. \tag{9}$$

We chose $K = 4$ because of the time-aggregation restriction, as proposed by Mariano and Murasawa (2003): if the factor captures (unobserved) monthly GDP growth, quarterly GDP growth is a weighted average of the monthly factor in the quarter of interest and in months two and three of the previous quarter.

Factor model IV: Approximate DFM based on static factors

The last factor approach under investigation is based solely on principal components and a simple linear regression model. Clearly, this method requires fewer assumptions than a fully parametric model. However, it does not fully take into account the time series dimension within the factor estimation (see, Stock and Watson 2002b,a, Bai 2003, for applications and theoretical justifications,). The standard principal component approach is augmented by the expectation-maximization (EM) algorithm to deal with missing observations (for ragged edges and indicators measured at the lower frequency). This method has been successfully applied by Stock and Watson (2002b), Schumacher and Breitung (2008) and Kaufmann and Scheufele (2015) in a mixed-frequency set-up. More specifically, there exists a certain relation between the observed variables $x_i^{\text{obs}} = A_i x_i$, where x_i^{obs} is a subset of x_i and A_i relates the unobserved to the observed values.

This forecasting approach basically consists of two steps:

1. Extract one factor from the entire data set by use of principal components. The

[6] As an example, assume that we have estimates of the monthly factor up to February 2016. The first new quarterly series therefore contains values of February 2016, November 2015, August 2015, etc. The second series contains the values of January 2016, October 2015, etc., and the third series contains the values of December 2015, September 2015, etc. The fourth (fifth/sixth) series is the one-quarter lag of the first (second/third) series, and so on.

EM-algorithm is used to deal with missing observations (see Stock and Watson 2002b, appendix).

2. Quarterly GDP growth is regressed on the monthly factors

$$y_{t+h} = \delta_{0,h} + \sum_{j=0}^{J} \delta_{1,h,j} \hat{f}_t^j + u_{h,t}. \tag{10}$$

using the U-MIDAS approach (Foroni, Marcellino, and Schumacher 2015). Insignificant $\delta_{1,h,j}$'s are dropped.

2.2 Model combination

In the following, we present an alternative to combining information into one or a few variables (or factors) used to explain current and future GDP developments. This procedure is based on simple dynamic linear regression models that relate GDP to leading indicators. We focus on single indicator models where GDP is regressed on one leading indicator with its potential lags. Therefore, we basically follow the work of Kitchen and Monaco (2003) or Stock and Watson (2003) using so-called bridge models. Since leading indicators may be available at a higher frequency than the quarterly target variable, we apply MIxed DAta Sampling (MIDAS) methods to take into account the temporal aggregation issue in an efficient way, (as proposed by Ghysels, Santa-Clara, and Valkanov 2004, Andreou, Ghysels, and Kourtellos 2011). At the end, we present methods that combine the information of the different indicators and methods to arrive at a pooled indicator-based forecast.

Specification of single indicator models using MIDAS

To take into account the complete weekly information flow, we employ the framework proposed by Ghysels, Santa-Clara, and Valkanov (2004), Ghysels, Sinko, and Valkanov (2007) and Andreou, Ghysels, and Kourtellos (2011). This framework has been applied successfully by Clements and Galvão (2009) and Marcellino and Schumacher (2010) to macroeconomic forecasting. We follow their MIxed DAta Sampling (henceforth, MIDAS) regression model procedure, which circumvents problems of quarterly conversion of the higher-frequency indicator. MIDAS models are closely related to distributed lag models (see Judge, Griffiths, Hill, Lütkepohl, and Lee 1985) and use parsimonious polynomials to reflect the dynamic response of a target variable to changes in the explanatory variables. This specification is particularly useful for time series that do not change much from one month to another (which may imply that explanatory variables are nearly linearly dependent). Thus, one does not need to estimate an unrestricted model using all observed monthly data points, which would result in a highly parameterized dynamic model. The main advantage is that only a small number of parameters have to be estimated for the distributed lag specification, although long lags can be captured.

Generally, all our specified models fall into the category of autoregressive distributed lag (ARDL) models. The models using monthly or weekly indicators are MIDAS models. Similar to Andreou, Ghysels, and Kourtellos (2013), we estimate the following regression model:

$$y_{t+h}^{Q} = \mu + \sum_{k=0}^{p_x^i} \omega_k^i(\theta) x_{t+h,N_i-J_i-k}^i + u_{t+h} \tag{11}$$

where i denotes the frequency of the indicator variable x. This can be quarterly, monthly or weekly. p_x^i reflects the lags of the explanatory variable.[7] N_i is the number of observations (months or weeks) per quarter, and J_i denotes the number of missing observations relative to the reference quarter $t + h$. In the case of quarterly indicators, eq. 11 simplifies to a standard ARDL model (see, e.g., Stock and Watson 2003), where $N_i = 0$ and $\omega(\theta)$ is linear with coefficients θ_k for each lag.

In the case of higher-frequency indicators (available monthly or weekly), this paper employs an Almon-Distributed Lag model in order to link indicators to the low-frequency target variable (as applied by Drechsel and Scheufele 2012a). In this case, $\omega_k^i(\theta)$ is specified as

$$\omega_k^i(\theta) = \omega_k^i(\theta_0, \theta_1, ...\theta_q) = \theta_0 + \theta_1 k + \theta_2 k^2 + ... + \theta_q k^q, \tag{12}$$

where q is the polynomial degree, which can be substantially lower than p_x^i. Even with

[7] p_x^i is selected by the Schwartz criterion. Note that we restrict ourselves to models excluding any lagged endogenous terms. In a previous version of the paper, we allowed for an AR(1) term in the specification for monthly and weekly indicators and for lags up to four in the case of quarterly variables. However, this led to inferior forecasting results. Therefore, we opted for the simpler specification.

very small q, many flexible forms can be approximated.[8] The great advantage of using a simple Almon-Distributed Lag model instead of a more complex functional form is that it can be easily estimated by restricted least squares (and no non-linear optimization is required). Subscript k is specified in terms of subperiods (weeks or months) and counts back for quarter $t+h$ into the past. Note that J_i (missing observations for the nowcast) depends on the variable, the horizon and the forecasting round. Within each recursive estimation window, we re-optimize the optimal lag-length and polynomial degree of the models.

Forecast combination

While some single-indicator models may already provide good forecast accuracy, it is generally undesirable to rely on such a limited set of information. Throwing away the majority of information by employing only one single best (in-sample)-fitting model is inefficient in most cases. One way to employ the full set of available information is to pool the results of several indicator models. The literature has shown that combining forecasts often results in improved forecast accuracy compared to univariate benchmark models or to a specific selected model (see Granger and Newbold 1977, Clemen 1989, Stock and Watson 2004, Timmermann 2006). An additional advantage of model averaging is that it guards against instabilities (Hendry and Clements 2004) and often results in a more stable and reliable forecasting performance (see Drechsel and Scheufele 2012b, before and during the financial crisis). In our application we take into account a large set of pooling techniques that combine the forecasts based on MIDAS models (for weekly and monthly indicators) with those based on standard ARDL models (for quarterly indicator variables).

Pooling the individual indicator forecasts $\widehat{y}_{i,t}$, we obtain the total forecast $\tilde{y}_t$ by:

$$\tilde{y}_t = \sum_{i=1}^{n} w_{i,t}\widehat{y}_{i,t} \quad \text{with} \quad \sum_{i=1}^{n} w_{i,t} = 1 \tag{13}$$

where $\omega_{i,t}$ is the weight assigned to each indicator forecast that results from the fit of the i^{th} individual equation. Note that due to subscript t, we allow for time-varying weights.

[8] We set $q^{max} = 3$ and select the q together with p_x^i by BIC following Drechsel and Scheufele (2012a).

n_t is the total number of models retained at time period t.[9]

Although we take into account several forecast combination schemes (see appendix B.1), we will focus on the equal weighting scheme, which is just the average of all candidate models. This implies $w_t = 1/n_t$. Despite its simplicity, this averaging strategy is very difficult to beat in practice (Clemen 1989, Timmermann 2006).

2.3 Model selection using a specific-to-general approach

A well-known alternative to information combination or model combination is model selection. Based on certain criteria, specific indicators from the large panel of indicators are selected and estimated together in a single equation. Typical examples are step-wise procedures that add or eliminate variables from a specification by means of information criteria or statistical tests (see, e.g., Ng 2013). Recent examples include Castle, Doornik, and Hendry (2011) and Chudik, Kapetanios, and Pesaran (2016).

Our setting is specific for two reasons. First, simple general-to-specific methods are infeasible, as $N >> T$ (the number of potential indicators far exceeds the number of observations), and therefore, the most-general model cannot be estimated. Moreover, a search algorithm that compares every combination of indicators would be extremely time consuming. Second, we are confronted with mixed-frequency datapart of our indicator set is available quarterly and the other monthly.

To address these two issues, we propose a simple specific-to-general approach that can handle very large data sets. Monthly indicators are taken into account via *blocking* (see, e.g., McCracken, Owyang, and Sekhposyan 2015). This implies that all monthly indicators are transformed into quarterly frequency, but using all monthly information. More specifically, suppose 2 vectors of variables x_t^m (monthly) and x_t^q (quarterly). Then, our quarterly-based data set W_t would include $W_t = [x_t^m, x_{t-1/3}^m, x_{t-2/3}^m, x_t^q]$, where x_t^m contains series of the third month of a quarter, $x_{t-1/3}^m$, of the second month of a quarter and so forth. As discussed earlier, the indicators are released with different publication lags, and therefore, we have to realign our data set to obtain a balanced data panel.

[9] According to the recursive model selection scheme of each indicator model, only those models survive for the model averaging scheme that obtain a smaller SIC than the AR-model (which exclusively consists of a constant and its own lags). The time varying nature is mostly taken into account by a recursive estimation scheme, which implies that in- and out-of-sample criteria are updated at each point in time.

By employing the specific-to general approach, we basically follow Herwartz (2011a) and Herwartz (2011b). However, we modify his approach by taking into account multiple testing using a p-value correction similar to Chudik, Kapetanios, and Pesaran (2016). W denotes a matrix of potential indicators $x_1, ..., x_N$, and W^{sg} is the selection outcome $W^{sg} = (j, x_1, ..., x_M)$, where j is a vector of ones and M is the number of selected indicators (with $M < N$). The selection strategy works as follows:

1. Initialize $W^{sg} = j$.
2. Project y (the target variable) onto W^{sg} and obtain the estimated residuals $\hat{u}$.
3. Estimate the auxiliary regression of $\hat{u}$ on the variable set $W_k = (W^{sg}, x_k),\ k = 1, ...\tilde{K}$ with $\tilde{K} = N - M$. For each regression, take an LM statistic (Godfrey 1988), $\lambda_k = TR_k^2$, which measures the explanatory content of x_k. R_k^2 is the degree of explanation.
4. The covariate obtaining the highest LM-statistic, λ_{k^*}, is moved from W to W^{sg} if $\lambda_{k^*} > c_{1-\tilde{a}}$, the $(1 - \tilde{a})$ quantile of the $\chi^2(1)$ distribution. $\tilde{a}$ depends on the pre-specified significance level α, $\tilde{a} = f(\alpha)$, and is computed according to the false discovery rate control method (as proposed by Benjamini and Hochberg 1995).
5. Steps 2 to 4 are iterated until $\lambda_{k^*} < c_{1-\tilde{a}}$.

Note that by setting $\tilde{a} \neq \alpha$, we deviate from the original method proposed by Herwartz (2011a) and Herwartz (2011b). In the case of large data sets, without any adjustment for multiple testing, the danger of outfitting is quite evident. However, by applying a standard Bonferroni adjustment, the probability of selecting an indicator would drop to zero when the data set increases dramatically. Therefore, we have to find a compromise between the two issues. Thus, we control the false discovery rate (FDR), which was shown to have much better power properties than standard Bonferroni bounds and still guards against large type I errors.[10]

Generally, the applied model selection approach has several advantages. First, it is very transparent (without path dependency). Second, it can be applied to very large data sets. Third, mixed-frequency data can be handled via blocking. Fourth, it is extremely fast. Finally, it has a well-defined stopping rule. It thus includes additional variables in the specification as long as they offer significant marginal explanatory power.

[10] More specifically, we employ the proposed two step method of Benjamini, Krieger, and Yekutieli (2006).

2.4 Benchmark models

As benchmark models, we consider *(i)* univariate time series models and *(ii)* leading indicator models. As the univariate benchmark time series model, we use an AR model with a maximum of four lags. The lag length is optimized at each forecast round using the BIC. For the leading indicator models, we consider two popular indicators for the Swiss economy: the KOF economic barometer and the PMI in the manufacturing sector. Both indicators are released on a monthly basis, receive some attention in the public and are promising in terms of forecasting performance (see, e.g., Maurer and Zeller 2009, Lahiri and Monokroussos 2013, Abberger, Graff, Siliverstovs, and Sturm 2014). For the two leading indicators, we use the same MIDAS framework as outlined in section 2.2, where the specification is optimized using BIC.

3 Model and forecast evaluation

To evaluate the models, we produced a one- and two-quarter-ahead forecast using a recursive pseudo real-time setup. The different models are compared in terms of root mean squared error (RMSE). The target variable is the annualized quarter on quarter growth of Swiss GDP (real, calendar and seasonally adjusted).

Two specific dates are selected to time the forecast production. The first, “early-quarter” information set, is given by indicator information available on 11th March/June/September/December. At this date, surveys are usually available for one month of the quarter, hard data for one or no month, and financial data for two months. For this information set, we use an underlying data vintage available on 11th September 2015.

The second, “late-quarter” information set is given by data available on 6th May/August/November/February. At this date, surveys are usually available for all three months of the quarter, hard data for two or three months, and financial data for all three months. Theoretically, this additional information should improve the forecasts. For this information set, we use an underlying data vintage available on 6th November 2015.

The evaluation is conducted over the period 2005Q1-2015Q2. It is based on GDP data that came with SECO’s GDP release for the second quarter of 2015. In most cases, RMSEs

Table 1 — Large-scale data set for the Swiss economy

Area	Soft data m	Soft data q	Hard data m	Hard data q	Prominent examples
GDP				27	Total GDP, demand components, value added of sectors.
Labour market	1	4	42	42	OASI statistics, unemployment statistic, job statistics, employment statistics, surveys, hours worked, wage index.
Consumption	4	13	6		Overnight stays of domestic visitors, retail sales, import of fuel, electricity consumption, new car registrations, consumer sentiment.
Investment			3	11	SwissMEM survey, imports of investment goods, industrial production of investment goods.
Foreign trade			19	2	Trade statistics, overnight stays of foreign visitors.
International activity	24	13	39	8	Several indicators covering Germany, euro area, United States, Japan, emerging asia and the CESifo world economic survey.
Financial markets			64		Stock market indices, exchange rates, commodity indices, monetary aggregates, monetary conditions, interest rates, spreads.
Prices			12	3	Consumer prices, real estate prices, import prices, production prices, construction prices.
Construction sector	15	3	6	21	Surveys, production, order books, cement deliveries.
Retail trade sector	6	5			Surveys.
Wholesale trade sector		14			Surveys.
Accomodation sector		18	1		Overnight stays, surveys.
Manufacturing sector	51	41	2	1	Industrial production, surveys, PMI, electricity production, number of working days.
Project engineering sector	8	7			Surveys.
Banking sector	20	22	14		Credit statistics, surveys, balance sheet statistics, illiquidity index.
Insurance sector	16	9			Surveys.
Other	7	2	1		Surveys.

are computed relative to the autoregressive benchmarks.

3.1 Data

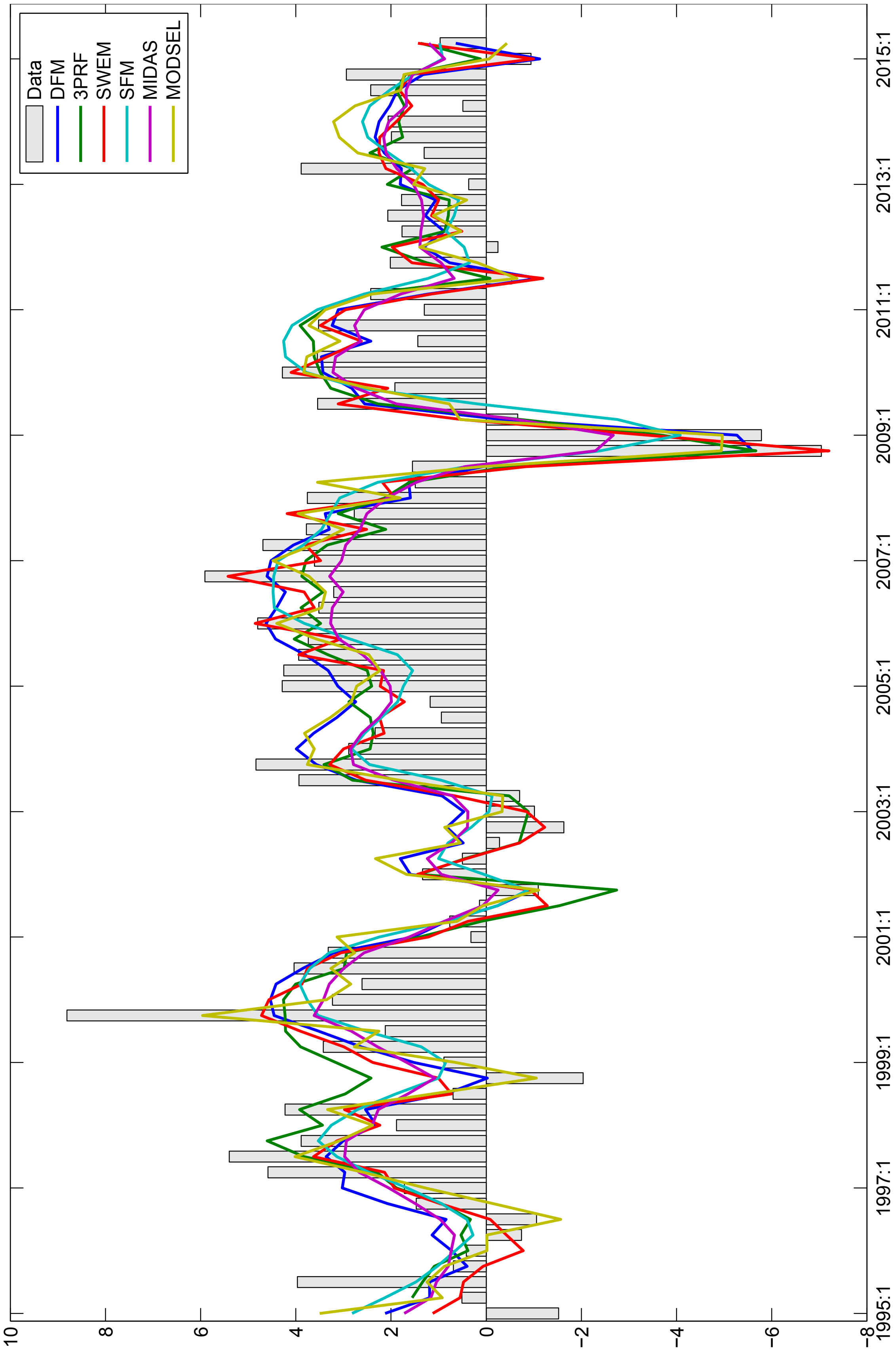
We work with two data sets. For the small-scale dynamic factor model, we use a data set consisting of Swiss quarterly GDP and the following eight selected monthly variables: manufacturing PMI, export of goods, imports of goods, retail sales, number of unemployed, inflows into unemployment, euro area industrial production and the ifo climate index for Germany.

For all other models, we use a novel large-scale data set for the Swiss economy that consists of 620 variables (362 monthly, 285 quarterly) covering 17 areas. All data are used in calendar-, seasonal- and outlier-adjusted terms. Details of the data set are shown in Table 1. In principle, the dataset starts in 1975. However, many indicators start only in 1990 or even later, which automatically limits the estimations sample of models to some extent.

3.2 Insample fit

Figure 1 plots the insample fit of the main models against realized GDP growth. Several observations are in order. *First*, the models general interpretations of the business cycle are similar to phases of weakness around 1996, 2002, and 2009. Also, the models agree that there was a significant deceleration during 2011 amid the strong CHF appreciation and a global slowdown and in 2015 following the end of the EURCHF exchange rate floor. *Second*, the models also agree on a number of quarters, where GDP appears to be driven by idiosyncrasies that are not mirrored in the broad data set from which the model dynamics are constructed. Examples are the strong GDP growth rates of Q1 1995, Q2 2014, and Q3 1999 and the weak GDP growth in Q3 and Q4 2004, Q3 2010, Q1 2011, and Q2 2014. This illustrates the usefulness of a large data set for measuring the underlying business cycle dynamics in a broader sense than what is captured by GDP. *Third*, the smoothness of the insample fit varies considerably, with the model selection approach (MIDAS) and small-scale factor model (SFM) being particularly smooth (standard deviations of 1.3 and 1.7, respectively). The EM-based static factor model (SWEM) is the least smooth, with a standard deviation of 2.0 percentage points. The other models – the three-pass regression filter (3PRF), the large-scale dynamic factor model (DFM) and the model

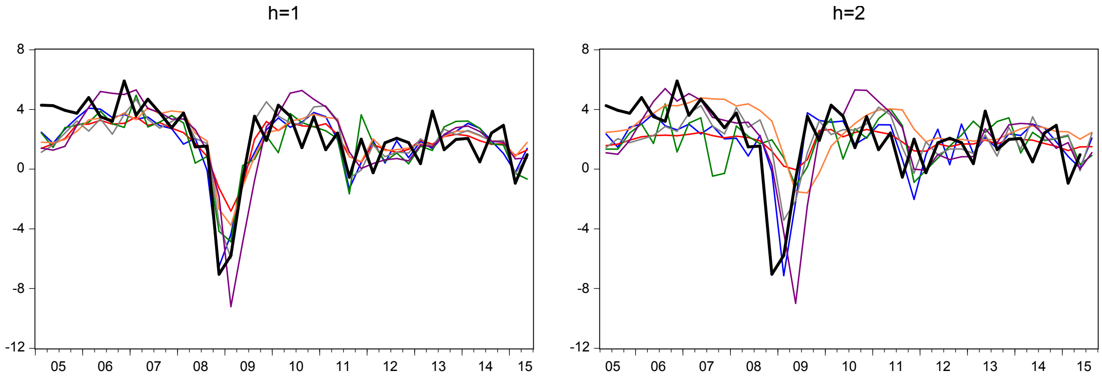
Figure 1 — Insample fit of the main models



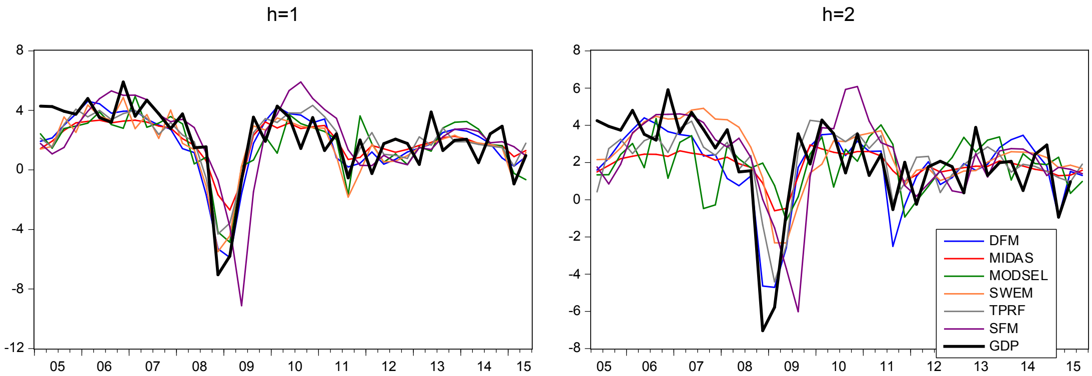
selection approach (MODSEL) – lie in between, with standard deviations of approximately 1.9. For comparison, the standard deviation of GDP is 2.4 percentage points. *Fourth*, and related to the previous point, the tightness of the insample fit varies across models. The smoother a model is, the lower the insample fit tends to be. The two smoothest models, MIDAS and SFM, have insample R^2 of 0.57 and 0.54, respectively. The other models R^2 are 0.62, 0.68, 0.71, and 0.72 for the 3PRF, MODSEL, DFM, and SWEM, respectively. The ability of SWEM and DFM to closely fit GDP is nicely illustrated by how they nearly exactly match the negative growth rates in Q3 2011 and Q1 2015. Whether these good insample fits translate into a good forecasting performance is the question of the next section.

Figure 2 — Model forecasts and realizations

(a) Early-quarter information set:



(b) Late-quarter information set:



3.3 Out-of-sample evaluation results

Figure 2 plots the forecasts of the main models over time. The main model features are directly visible in these graphs. For the nowcast ($h = 1$), all of the investigated models

Table 2 — Relative RMSEs of different forecasting methods

	Total sample 2005Q1-2015Q2		Post crisis sample 2010Q1-2015Q2	
	h=1	h=2	h=1	h=2
AR benchmark	2.1689	2.5081	1.6334	1.2686
I. EARLY-QUARTER INFORMATION				
a. Information combination				
dfm	0.620*	0.725*	0.829*	1.158
3prf	0.683*	0.845	0.812*	1.252
swem	0.703**	1.054	0.820***	1.386
sfm	0.928	1.020	0.990	1.438
b. Model combination				
midas	0.768**	0.870**	0.760***	0.927
c. Model selection				
modsel	0.726*	1.056	0.936	1.470
d. Leading indicator models				
pmi	0.829	0.871*	1.001	1.055
kofbaro	0.850	0.844	0.923	0.958
e. Combination of different procedures				
combo1	0.611*	0.748*	0.805**	1.145
combo2	0.640**	0.758**	0.767***	1.028
combo3	0.645**	0.792**	0.772***	1.060
II. LATE-QUARTER INFORMATION				
a. Information combination				
dfm	0.645**	0.562*	0.818**	1.043
3prf	0.660**	0.662*	0.832*	0.995
swem	0.614**	0.819*	0.753**	1.135
sfm	1.110	1.022	1.069	1.434
b. Model combination				
midas	0.719**	0.818**	0.704***	0.922
c. Model selection				
modsel	0.726*	1.059	0.937	1.486
d. Leading indicator models				
pmi	0.733*	0.845	0.899	1.359
kofbaro	0.938	0.834	1.029	1.057
e. Combination of different procedures				
combo1	0.627**	0.574**	0.792**	0.934
combo2	0.636**	0.623**	0.749***	0.874
combo3	0.614**	0.643**	0.729***	0.912

Notes: The table shows root-mean-squared errors (RMSE) for $h = 1$ (nowcast) and $h = 2$ (one additional quarter ahead) using two different states of information. Besides the benchmark AR model, all numbers are defined relative to the benchmark. Combination of different procedures include the following models using equal weights. combo1: dfm + 3prf, combo2: dfm + 3prf + swem, combo3: dfm + 3prf + swem + midas. ***, ** and * indicates whether a model's predictive ability (using the DM test) is significantly better than the benchmark (at the 1%, 5% and 10% level).

more or less follow the tendency of GDP growth. Clearly, there are periods where the connection is stronger (e.g., before, during and after the financial crisis) and periods where GDP growth deviates persistently from the predictions of the models (in 2005 and 2013), which broadly corresponds to the in-sample results. Generally, forecasts of the different models are highly correlated (particularly for $h = 1$). Additionally, forecasts using the early-quarter information set are not much different from the late-quarter information set. As expected, the forecasts are more accurate for the nowcast ($h = 1$) than for the following quarter ($h = 2$).

Looking at the different models, we see that three factor models (DFM, 3PRF, SWEM) capture the crisis period 2008/2009 very well, whereas the SFM first misses the downturn and then predicts very negative rates when the economy has already recovered. The model combination strategy (MIDAS) is not as successful during the crisis as the factor models are, but it at least got the tendency right. It generally delivers less-volatile forecasts and is therefore less flexible in situations of rapid change. At the same time, the model selection approach also performs quite well for the nowcast ($h = 2$) during the crisis period. However, this forecast method is very volatile for $h = 2$ and seems to be less correlated with GDP growth than most of the other forecasting models.

Average performance of different models

Table 2 reports the relative root mean squared errors of the models against the AR benchmark. Although the forecast errors for the post-crisis sample – which covers 2010Q1-2015Q2 – are, on average, smaller (see, e.g., the AR benchmark), the relative performance of the models under investigation has deteriorated somewhat (relative to the AR benchmark). A similar result is well documented for the pre-crisis period relative to the crisis period (Drechsel and Scheufele 2012b, Kuzin, Marcellino, and Schumacher 2013).

Overall, the three factor approaches DFM, 3PRF and SWEM perform very similarly (see table 3 for pairwise model comparisons). For the total sample, the gains in terms of RMSE are more than 30%. In the post-crisis sample, the gains decline to approximately 20%. For both sample periods, each of these models significantly outperforms the benchmark for the nowcast period ($h = 1$). However, none of the three factor models is able to systematically outperform one of the two other approaches. The improvements

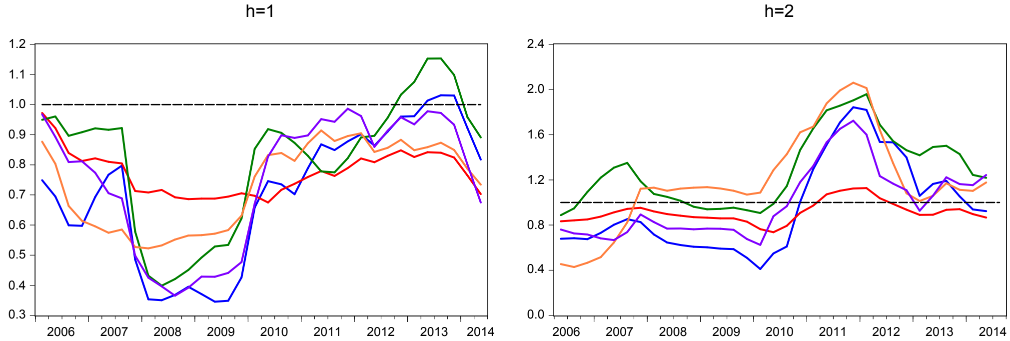
Table 3 — Pairwise forecast comparisons

	h=1						h=2					
Bench	dfm	3prf	sfm	swem	midas	modsel	dfm	3prf	sfm	swem	midas	modsel
I. EARLY-QUARTER INFORMATION												
3prf	1.10						1.17					
sfm	1.50*	1.36					1.41	1.21				
swem	1.13	1.03	0.76**				1.45	1.25	1.03			
midas	1.24	1.12	0.83	1.09			1.20	1.03	0.85	0.83		
modsel	1.17**	1.06	0.78	1.03	0.94		1.46**	1.25**	1.03	1.00	1.21***	
Best model (in % of time)	16.7	16.7	16.7	7.1	16.7	26.2	23.8	9.5	16.7	4.8	19.0	26.2
II. LATE-QUARTER INFORMATION												
3prf	1.02						1.18					
sfm	1.72	1.68					1.82	1.54				
swem	0.95	0.93	0.55				1.46	1.24	0.80			
midas	1.11	1.09	0.65	1.17			1.46	1.24	0.80	1.00		
modsel	1.12	1.10	0.65	1.18	1.01		1.88**	1.60**	1.04	1.29**	1.29***	
Best model (in % of time)	19.0	7.1	16.7	23.8	14.3	19.0	21.4	4.8	21.4	2.4	21.4	28.6

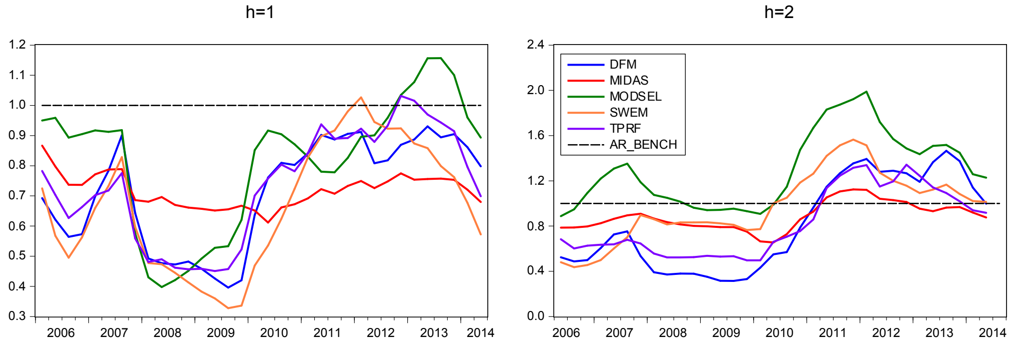
Notes: Results of pairwise comparisons are displayed using relative RMSEs and the test for equal predictive ability of DM. The benchmark is given by the top line (and is used as the denominator of the relative RMSE) . ***, ** and * indicates whether the DM test is significant at the 1%, 5% and 10% level. Additionally, we show the percentage of times when models provides the best forecast.

Figure 3 — Rolling relative RMFEs (2 year windows)

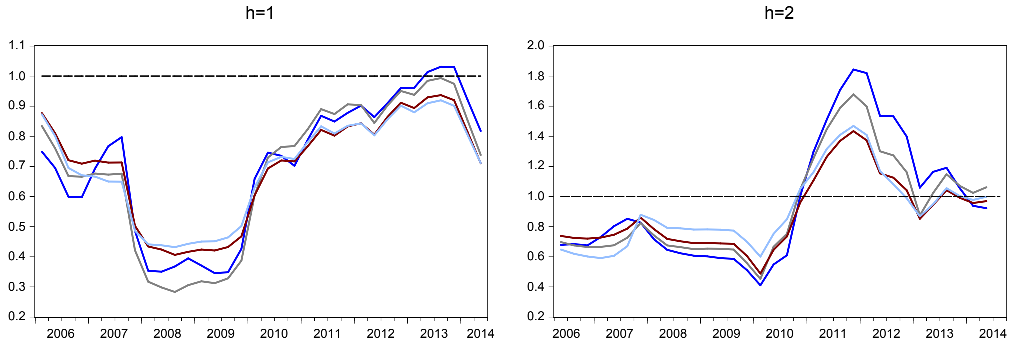
(a) Individual models, early-quarter information set:



(b) Individual models, late-quarter information set:



(c) DFM and model combinations, early-quarter information set:



(d) DFM and model combinations, late-quarter information set:

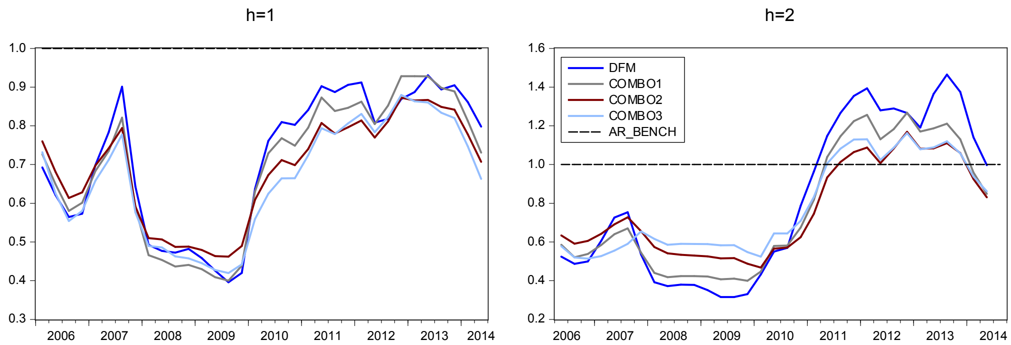


Table 4 — Contingency tables for business cycle phase predictions

I. EARLY-QUARTER INFORMATION SET

	a. dfm Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	b. 3prf Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	c. swem Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$
Prediction									
$(-\infty,0.5]$	11.9	4.8	2.4	11.9	0.0	2.4	7.1	0.0	0.0
$(0.5,2]$	4.8	4.8	19.0	2.4	9.5	14.3	7.1	9.5	21.4
$(2,+\infty)$	2.4	14.3	35.7	4.8	14.3	40.5	4.8	14.3	35.7
% of correct predictions			52.38			61.90			52.38
Phi coefficient			0.58			0.68			0.58
Pearson χ^2			14.22			19.60			14.37
p value			0.007			0.001			0.006

	d. midas Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	e. modsel Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	f. combo3 Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$
Prediction									
$(-\infty,0.5]$	7.1	0.0	2.4	11.9	7.1	0.0	11.9	0.0	2.4
$(0.5,2]$	11.9	11.9	14.3	4.8	4.8	16.7	4.8	9.5	16.7
$(2,+\infty)$	0.0	11.9	40.5	2.4	11.9	40.5	2.4	14.3	38.1
% of correct predictions			59.52			57.14			59.52
Phi coefficient			0.63			0.64			0.69
Pearson χ^2			16.71			17.15			19.82
p value			0.002			0.002			0.001

II. LATE-QUARTER INFORMATION SET

	a. dfm Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	b. 3prf Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	c. swem Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$
Prediction									
$(-\infty,0.5]$	11.9	4.8	2.4	11.9	2.4	0.0	11.9	2.4	2.4
$(0.5,2]$	4.8	7.1	19.0	2.4	11.9	16.7	7.1	9.5	21.4
$(2,+\infty)$	2.4	11.9	35.7	4.8	9.5	40.5	0.0	11.9	33.3
% of correct predictions			54.76			64.29			54.76
Phi coefficient			0.57			0.72			0.64
Pearson χ^2			13.86			21.48			17.19
p value			0.008			0.000			0.002

	d. midas Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	e. modsel Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$	f. combo3 Realisation $(-\infty,0.5]$	$(0.5,2]$	$(2,+\infty)$
Prediction									
$(-\infty,0.5]$	7.1	0.0	0.0	11.9	7.1	0.0	11.9	2.4	0.0
$(0.5,2]$	11.9	11.9	16.7	4.8	4.8	16.7	7.1	9.5	21.4
$(2,+\infty)$	0.0	11.9	40.5	2.4	11.9	40.5	0.0	11.9	35.7
% of correct predictions			59.52			57.14			57.14
Phi coefficient			0.70			0.64			0.72
Pearson χ^2			20.45			17.15			21.53
p value			0.000			0.002			0.000

Notes: Results of pairwise comparisons about three different business cycle phases. The numbers reflect frequencies (in percentage points) how predictions match with realizations. 1: low/negative growth (below 0.5%), 2: moderate growth ($0.5-2$%) and 3: high growth (above 2%). Additionally, statistics about the degree of association are displayed (as well as a test of independence).

in forecasting accuracy from the early-quarter to the late-quarter information set are, on average, very small for these methods. For the total sample, the gain is even slightly negative for the DFM but more moderately positive for the SWEM. Because the SWEM is only based on static factors, it does not take into account the dynamics of the factors in the estimation stem, and thus, it may benefit more from the increasing data set. This is also consistent with the finding that this model does not perform very well for $h = 2$. As expected from the graphical analysis, the SFM performs relatively poorly.

The fact that the forecasting performance of factor models using a very large data set is very promising for Switzerland contrasts somewhat with previous findings in the literature on dynamic factor models. Boivin and Ng (2006) for the US and Caggiano, Kapetanios, and Labhard (2011) for selected euro area countries show that using all sectorial information in a large data set may be suboptimal. The reason is that including sectorial information could lead to model mis-specification in small samples since it increases the idiosyncratic cross-correlation. This finding is supported by Banbura and Rünstler (2011), Banbura, Giannone, and Reichlin (2011) and Banbura and Modugno (2014), who conclude that disaggregate information does not improve the accuracy of euro area GDP forecasts. More recently, however, Alvarez, Camacho, and Quiros (2016) shed new light on the issue of large-scale versus small-scale factor models. Based on simulations and the Stock and Watson data set, they concluded that the inclusion of disaggregated data in a set of specified predictors could be desirable in terms of forecasting precision. One major reason why using a small data set might not be sufficient for Switzerland is that there are hardly any reliable hard indicators, such as monthly industrial production or new orders available, that show a very strong correlation with GDP growth.[11]

The forecasting performance of the model combination based on MIDAS models is slightly less accurate compared to the most-promising factor models for the total sample. The forecasting gains in terms of RMSEs are approximately 23% (early-quarter information) and 28% (late-quarter information). This is broadly in line with previous research that compares similar methods (Banbura, Giannone, Modugno, and Reichlin 2013, Kuzin, Marcellino, and Schumacher 2013, Foroni and Marcellino 2014). However, for the post-crisis sample and $h = 1$, the model combination based on MIDAS models offers the largest and most significant improvements relative to the benchmark model.

Interestingly, the specific-to-general model selection (MODSEL) approach performs reasonably well for the nowcast period (and is able to outperform the benchmark for the total sample). For one quarter ahead ($h = 2$), this approach is not very reliable, and for the post-crisis sample, there are no significant improvements relative to the benchmark.

It is also instructive to investigate how the models using a large indicator set compare to single leading indicator models. For the total sample, the performance of the PMI

[11] For the DFM, we experimented with a smaller data set including 70 pre-selected time series and with selection criteria in the spirit of Boivin and Ng (2006), but this did not lead to any improvements in forecast accuracy.

and KOF economic barometer is less accurate than our three factor models and slightly inferior to the model combination and model selection approaches. Only the PMI offers some significant improvements relative to the benchmark. The performances of the two leading indicator models have clearly deteriorated in the post-crisis period, and the models contain only a little information. Most interestingly, the KOF economic barometer – which is itself calculated from a large-scale factor model – does not outperform the much simpler calculated survey from the purchasing managers.

Time-varying forecasting performance of different models

Figure 3 gives a more complete picture of the time-varying performance of the models. First, it confirms that the good performance of the factor models comes mainly from the crisis period 2008/2009. Second, it shows how the relative performance of the models has deteriorated in 2010-2013. Most recently, the models' performance gained again, most likely due to the exchange rate shock in 2015. Interestingly, the performance of the model combination approach based on MIDAS models remains very stable over time and is less sensitive.

Pooling different forecasting models

Table 2 and figure 3 show that one may obtain additional gains in average performance and in terms of reliability (significance) when combining different forecast methods. For the total sample and the nowcast period, just combining the DFM and 3PRF (COMBO1) results in the best overall performance. Generally, the different combinations perform nearly identically, and it is difficult to judge which is the clear winner. However, the results confirm Kuzin, Marcellino, and Schumacher (2013), who find that a pooling of factor model forecasts is stable and reliable in terms of forecast accuracy. Our results indicate that for the most recent period, model combination seems to add some additional information. This is suggested by the finding that the best performance during the sample 2010q1 2015q2 is obtained by pooling the best three factor models with the model combination approach (COMBO3). Additionally, figure 3 shows that since 2013, this forecasting approach is difficult to beat in terms of forecasting performance.

Forecasting the tendency of output growth

When looking at other measures of forecasting performance, our general results basically hold. However, the details may be slightly different. By looking at whether the models are able to accurately capture the general tendency of output growth, we go from a continuous measure based on an average distance measure (such as RMSE) to a nominal measure. In practice, the forecaster may not be judged by the specific distance to a target but by whether he got the general tendency right. By using a trichotomous measure, we take into account three different phases: low, moderate and strong growth states. We can then investigate how our forecasts match those categories when compared to the realizations. Our definition of the three different phases is motivated by two different considerations: first, by the empirical distribution of GDP growth since the 1990s, and second, by policy considerations. Thus, a policy maker may be interested in whether the economy expands robustly, which we define by growth that is markedly greater than the long-term average of 1.6% – so we opt for the 2% threshold. Additionally, a policy maker might become worried when growth is only marginally positive or even negative. This motivates our second threshold of 0.5%.

Table 4 shows the main results for the nowcast period. Generally, the main findings are compatible with previous results. First, the differences among the different approaches tend to be small. Second, all models under investigation are informative. Thus, the null hypothesis of independence between forecasts and realizations can be rejected in all cases. On average, the predictions match the corresponding realized categories in approximately 52% - 64% cases. Third, there seems to be a small improvement in terms of accuracy from the early to late information set. Fourth, model averaging performs relatively well. The combination of the three factor models plus MIDAS provides a relatively high categorical association (Phi coefficient). Interestingly, the 3PRF among the factor models performs relatively well in predicting the business cycle phase.

4 Conclusions

Monitoring economic developments in real time is one of the most important but also most challenging tasks that the applied economist working on Switzerland faces. In this paper, we set up a large database containing hundreds of potentially relevant variables. We then considered different approaches to condensing the information of the data set into a GDP forecast. The traditional approaches often used in practice select one or a few indicators based on expert knowledge and derive the forecast using OLS regressions or a small-scale dynamic factor model. Alternatively, one may use all indicators without an expert's pre-selection. We presented three approaches to doing so.

Factor-based *information combination* extracts a small number of common factors from the database and forms GDP forecasts based on these factors. Three different procedures were considered: a Kalman filter-based DFM approach, a principal components-based approach and the three-pass regression filter. While the first two methods try to extract the factors such that they explain as much of the variation in the data set as possible, the latter optimizes the factor(s) towards a particular target (in our case, GDP growth). Our results indicate that the performances of the three procedures are similar. Each of the three methods provides very good results for nowcasting GDP, and they beat relevant benchmarks such as univariate time series models, prominent leading indicator models and a small-scale factor model.

The second approach, *model combination*, performs estimation based on MIDAS equations for each indicator and then combines the resulting forecasts to form a final GDP forecast. This makes communicating a particular indicator's contribution to the forecast very easy. Also, the variation in the distribution of the forecasts over time can be used as a measure of uncertainty. Moreover, it is quite easy to implement. This model performs slightly worse than the large-scale factor models within the total evaluation sample, mainly because it is not fully able to match the large drop in output during the crisis. For the post-crisis sample, however, its forecasting performance is even slightly better than the best factor models.

Our third approach to extract relevant information from the large data set is *model selection*. A specific-to-general approach that can handle very large data sets and mixed frequencies is used for this purpose. Although this approach delivers forecasts that are less accurate than the large factor models and the model combination under investigation, it still offers some improvements against the benchmark in certain cases.

Additionally, we propose to construct a pooled forecast using the most promising methods. Forecast pooling of two or three large-scale factor models with the combined single indicator approach delivers very reliable forecasts. Overall, this suggests that it is best to employ several short-term forecast methods on a large data set and to pool the results of the most-promising methods. This is much better than relying on one method that uses only one indicator or a small number of indicators and is particularly beneficial after the financial crisis. Given this finding, we strongly recommend using a large data set and several forecasting approaches when monitoring the Swiss economy.

References

Abberger, K., M. Graff, B. Siliverstovs, and J.-E. Sturm (2014): “The KOF Economic Barometer, Version 2014: A Composite Leading Indicator for the Swiss Business Cycle,” KOF Working Papers 353, KOF Swiss Economic Institute.

Aiolfi, M., and A. Timmermann (2006): “Persistence in forecasting performance and conditional combination strategies,” *Journal of Econometrics*, 135(1-2), 31–53.

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

Alvarez, R., M. Camacho, and G. P. Quiros (2016): “Aggregate versus disaggregate information in dynamic factor models,” *International Journal Forecasting*, forthcoming.

Andreou, E., E. Ghysels, and A. Kourtellos (2011): “Forecasting with mixed-frequency data,” in *Oxford Handbook of Economic Forecasting*, ed. by M. P. Clements, and D. F. Hendry, chap. 8. Oxford University Press.

——— (2013): “Should Macroeconomic Forecasters Use Daily Financial Data and How?,” *Journal of Business & Economic Statistics*, 31(2), 240–251.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011): “Short-term forecasts of euro area GDP growth,” *Econometrics Journal*, 14(1), C25–C44.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009): “Real-Time Measurement of Business Conditions,” *Journal of Business & Economic Statistics*, 27(4), 417–427.

Bai, J. (2003): “Inferential Theory for Factor Models of Large Dimensions,” *Econometrica*, 71, 135–172.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013): “Now-Casting and the Real-Time Data Flow,” in *Handbook of Economic Forecasting*, ed. by G. Elliot, and A. Timmermann, vol. 2A, chap. 4, pp. 195–237. Elsevier.

Banbura, M., D. Giannone, and L. Reichlin (2011): "Nowcasting," in *The Oxford Handbook of Economic Forecasting*, ed. by M. P. Clements, and D. F. Hendry, chap. 7, pp. 193–224. Oxford University Press.

Banbura, M., and M. Modugno (2014): "Maximum Likelihood Estimation Of Factor Models On Datasets With Arbitrary Pattern Of Missing Data," *Journal of Applied Econometrics*, 29(1), 133–160.

Banbura, M., and G. Rünstler (2011): "A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP," *International Journal of Forecasting*, 27(2), 333–346.

Benjamini, Y., and Y. Hochberg (1995): "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B*, 57, 289–300.

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006): "Adaptive Linear Step-up False Discovery Rate controlling procedures," *Biometrika*, 93, 491–507.

Boivin, J., and S. Ng (2006): "Are more data always better for factor analysis?," *Journal of Econometrics*, 127(1), 169–194.

Caggiano, G., G. Kapetanios, and V. Labhard (2011): "Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK," *Journal of Forecasting*, 30(8), 736–752.

Camacho, M., and J. Martinez-Martin (2014): "Real-time forecasting US GDP from small-scale factor models," *Empirical Economics*, 47(1), 347–364.

Camacho, M., and G. Perez-Quiros (2010): "Introducing the Euro-Sting: Short-term Indicator of Euro Area Growth," *Journal of Applied Econometrics*, 25, 663–694.

Castle, J. L., J. A. Doornik, and D. F. Hendry (2011): "Evaluating automatic model selection," *Journal of Time Series Econometrics*, 3(1), 1–33.

Chudik, A., G. Kapetanios, and M. H. Pesaran (2016): "Big Data Analytics: A New Perspective," Working paper.

Clemen, R. T. (1989): "Combining forecasts: A review and annotated bibliography," *International Journal of Forecasting*, 5(4), 559–583.

Clements, M., and A. Galvão (2009): "Forecasting US output growth using Leading Indicators: An appraisal using MIDAS models," *Journal of Applied Econometrics*, 24(7), 1187–1206.

Doz, C., D. Giannone, and L. Reichlin (2011): "A two-step estimator for large approximate dynamic factor models based on Kalman filtering," *Journal of Econometrics*, 164(1), 188–205.

Drechsel, K., and R. Scheufele (2012a): "Bottom-up or Direct? Forecasting German GDP in a Data-rich Environment," Working Papers 2012-16, Swiss National Bank.

——— (2012b): "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.

Evans, M. D. D. (2005): "Where Are We Now? Real-Time Estimates of the Macroeconomy," *International Journal of Central Banking*, 1(2).

Foroni, C., and M. Marcellino (2014): "A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates," *International Journal of Forecasting*, 30(3), 554 – 568.

Foroni, C., M. Marcellino, and C. Schumacher (2015): "U-MIDAS: MIDAS regressions with unrestricted lag polynomials," *Journal of the Royal Statistical Society - Series A*, 178, 57–82.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004): "The MIDAS touch: Mixed DAta Sampling regression models," mimeo, Chapel Hill, N.C.

Ghysels, E., A. Sinko, and R. Valkanov (2007): "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26(1), 53–90.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The real-time informational content of macroeconomic data releases," *Journal of Monetary Economics*, 55, 665–676.

Godfrey, L. G. (1988): *Misspecification Tests in Econometrics.* Cambridge University Press, Cambridge.

Granger, C. W., and P. Newbold (1977): *Forecasting Economic Time Series.* Academic Press, New York.

Hansen, B. E. (2007): "Least Squares Model Averaging," *Econometrica*, 75(4), 1175–1189.

——— (2008): "Least-squares forecast averaging," *Journal of Econometrics*, 146(2), 342–350.

Hendry, D. F., and M. P. Clements (2004): "Pooling of forecasts," *Econometrics Journal*, 7(1), 1–31.

Hepenstrick, C., and M. Marcellino (2016): "Forecasting with Large Unbalanced Datasets: The Mixed Frequency Three-Pass Regression Filter," *mimeo.*

Herwartz, H. (2011a): "Forecast accuracy and uncertainty in applied econometrics: a recommendation of specific-to-general predictor selection," *Empirical Economics*, 41(2), 487–510.

——— (2011b): "Specific-to-general predictor selection in approximate autoregressions - Monte Carlo evidence and a large-scale performance assessment with real data," *AStA Advances in Statistical Analysis*, 95, 147–168.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T.-C. Lee (1985): *The Theory and Practice of Econometrics.* John Wiley and Sons, 2nd edition edn.

Kaufmann, D., and R. Scheufele (2015): "Business Tendency Surveys and Macroeconomic Fluctuations," KOF Working papers 15-378, KOF Swiss Economic Institute, ETH Zurich.

Kelly, B., and S. Pruitt (2013): "Market Expectations in the Cross-Section of Present Values," *The Journal of Finance*, 68(5), 1721–1756.

——— (2015): "The three-pass regression filter: A new approach to forecasting using many predictors," *Journal of Econometrics*, 186(2), 294 – 316, High Dimensional Problems in Econometrics.

Kitchen, J., and R. Monaco (2003): "Real-Time Forecasting in Practice," *Business Economics*, 38(4), 10–19.

Kuzin, V., M. Marcellino, and C. Schumacher (2011): "MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the Euro Area," *International Journal of Forecasting*, 27(2), 529–542.

——— (2013): "Pooling Versus Model Selection For Nowcasting Gdp With Many Predictors: Empirical Evidence For Six Industrialized Countries," *Journal of Applied Econometrics*, 28(3), 392–411.

Lahiri, K., and G. Monokroussos (2013): "Nowcasting US GDP: The Role of ISM Business Surveys," *International Journal of Forecasting*, 29(4), 644–658.

Lütkepohl, H. (2005): *New introduction to multiple time series analysis*, vol. 26. Springer.

Marcellino, M., and C. Schumacher (2010): "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP," *Oxford Bulletin of Economics and Statistics*, 72(4), 518–550.

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

Maurer, C., and M. Zeller (2009): "PMI - aktuelles Konjunkturbarometer der Schweiz," *Die Volkswirtschaft*, 3-2009, 51–52.

McCracken, M. W., M. T. Owyang, and T. Sekhposyan (2015): "Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR," Working Paper 2015-030A.

Ng, S. (2013): "Variable Selection in Predictive Regressions," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, chap. 14, pp. 753–789. Elsevier.

Schumacher, C., and J. Breitung (2008): "Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data," *International Journal of Forecasting*, 24(3), 386–398.

Stock, J. H., and M. W. Watson (2002a): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

——— (2002b): "Macreconomic forecasting using diffusion indexes," *Journal of Business and Economic Statistics*, 20(2), 147–162.

——— (2003): "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature*, 47(3), 788–829.

——— (2004): "Combination forecasts of output growth in a seven-country data set," *Journal of Forecasting*, 23(6), 405–430.

——— (2011): "Dynamic factor models," in *The Oxford Handbook of Economic Forecasting*, ed. by M. P. Clements, and D. F. Hendry, chap. 7, pp. 35–59. Oxford University Press.

Timmermann, A. (2006): "Forecast Combinations," in *Handbook of Forecasting*, ed. by G. Elliot, C. W. Granger, and A. Timmermann, vol. 1, chap. 4, pp. 135–196. Elsevier.

A Two-step estimation of the large dynamic factor model

In this setup, we apply the standard dynamic factor approach to a large-scale data set for Switzerland (see section 3.1). To select the number of factors (r), we regress quarterly GDP growth on the quarterly aggregation of the factors obtained from a principal component analysis over a balanced sample of the monthly variables and choose the model with the lowest BIC,

$$BIC(r) = ln(\hat{\sigma}_{\epsilon}^2) + \frac{ln(T)}{T} r. \tag{14}$$

An alternative would be to choose the number of factors using an out-of-sample criterion such the mean squared forecasting error. In our case, both criteria yield $r = 4$.

Given the number of factors r, we choose the number of lags p based on the VAR(p) of the factors that is associated with the lowest BIC,

$$BIC(p) = -2l(r,p) + rpln(T), \tag{15}$$

where $\ell(r,p) = -\frac{Tr}{2} log(2\pi) - \frac{T}{2} log|\Omega| - \frac{1}{2}\sum_1^T e_t'\Omega^{-1}e_t$.

Alternatively, one could follow Lütkepohl (2005), pp. 150 and 140, and write the BIC as

$$BIC(p) = ln|\hat{\Sigma}_e(p)| + \frac{lnT}{T} pr^2, \tag{16}$$

where $\hat{\Sigma}_e(p) = \frac{1}{T}(f_t - \hat{\Phi}_1 f_{t-1} ... - \hat{\Phi}_p f_{t-p})(f_t - \hat{\Phi}_1 f_{t-1} - ... - \Phi_p f_{t-p})'$. Applying the BIC criteria yields $p = 1$.

We estimate the model using a two-step approach. In the first step, we estimate the parameters of the model via principal components and linear regressions using the procedure described in the appendix. In the second step, we estimate the factors $\hat{f}_t$ using the Kalman filter and smoother. Our monthly fitted values are then given by $\hat{x}_t = \hat{\Lambda}\hat{f}_t$, and the monthly interpolated values for quarterly variables are given by $\hat{x}_t^* = \hat{\Lambda}\hat{f}_t + \hat{u}_t$. The forecast for quarterly GDP in period t is then given by

$$\hat{y}_t^Q = G^f(L)\hat{y}_t. \tag{17}$$

We define our data vector at a monthly frequency as

$$x_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}, \tag{18}$$

where y_t denotes real GDP growth and x_t denotes our vector of indicator variables. x_t contains both variables measured at a monthly frequency and variables measured at a quarterly frequency. For non-stationary indicators, we use growth rates or first differences. The corresponding data vector at a quarterly frequency is defined as

$$x_t^Q = \begin{bmatrix} y_t^Q \\ x_t^Q \end{bmatrix}. \tag{19}$$

The monthly correspondence for an observation in the quarterly period t, x_t^Q, is the observations in the monthly period t, x_t. This means that, e.g., the observation for the first quarter of a given year corresponds to the month of March.

For quarterly variables, the time aggregation of the monthly observations to quarterly values is given by[12].

$$x_{i,t}^Q = (\frac{1}{3} + \frac{1}{3}L + \frac{1}{3}L^2)x_{i,t} = G(L)x_{i,t}. \tag{20}$$

If the first block of variables in x corresponds to quarterly flow variables and the second to monthly variables, the generalized link between x_t^Q and x_t is given by

$$x_t^Q = (G_0 + G_1 L + G_2 L^2)x_t \tag{21}$$

with $G_0 = \begin{bmatrix} \frac{1}{3}I_{n^Q} & \mathbf{O} \\ \mathbf{O} & I_{n^M} \end{bmatrix}$, $G_1 = \begin{bmatrix} \frac{1}{3}I_{n^Q} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$, $G_2 = \begin{bmatrix} \frac{1}{3}I_{n^Q} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$,
where n^Q denotes the number of quarterly variables and n^M denotes the number of monthly variables.

Combining the dynamic factor model and the time aggregation equations, we obtain the following relationship among the observed variables x_t, the monthly factors f_t and the

[12] Since non-stationary variables enter the data as 3 month growth rates, the time aggregation rule is the same for both stock and flow variables

monthly errors u_t:

$$x_t^Q = \mathbf{G}_t(L)x_t = \mathbf{G}_t(L)(\Lambda f_t + u_t) = \mathbf{G}_t(L)\Lambda f_t + \mathbf{G}_t(L)u_t, \tag{22}$$

where $\mathbf{G}_t(L) = (G_0 + G_1 L + G_2 L^2)$.

The parameters of the large dynamic factor model are estimated using a two-step approach. In the first step, the parameters of the model are estimated via principal components and OLS using the following procedure:

- Perform a principal component analysis with monthly variables only and select r (for example by regressing quarterly GDP on the factors).
- Obtain the estimate for Λ by regressing all variables on the first r principal components using the following regression setup:
  - For quarterly variables: $x_{i,t}^Q = \Lambda_i G(L) f_t + u_{i,t}^Q$, where $u_{i,t}^Q = G(L)u_{i,t}$
  - For monthly variables: $x_{i,t} = \Lambda_i f_t + u_{i,t}$
- Obtain the estimate for Σ_u by using the resulting residual from the regressions above in the following way:
  - For quarterly variables, we have that $u_{t,i}^Q = G(L)u_{i,t}$, so that $V[u_t^Q] = V[G(L)u_t] = \frac{1}{9}V(u_t) + \frac{1}{9}V(u_{t-1}) + \frac{1}{9}V(u_{t-2}) = \frac{3}{9}V(u_t)$ and $V(u_t) = \frac{9}{3}V(u_t^Q)$.
  - For monthly variables, we have that $V(u_t) = V(u_t)$.
- Obtain the estimates for Φ and Σ_v by estimating a simple VAR(p) of the first r principal components

In the second step, we estimate the factors $\hat{f}_t$ using the Kalman filter and smoother. Our monthly fitted values are then given by $\hat{x}_t = \hat{\lambda}\hat{f}_t$, and the monthly interpolated values for quarterly variables are given by $\hat{x}_t^* = \hat{f}_t + \hat{u}_t$. The underlying state space system is given by the measurement equation

$$z_t = H_t \underbrace{\begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \end{bmatrix}}_{\substack{\xi_t \\ 3r\times 1}} + \underbrace{\begin{bmatrix} G(L)u_t^Q \\ u_t^M \end{bmatrix}}_{\epsilon_t}, \tag{23}$$

$$\text{where } H_t = \begin{cases} \begin{bmatrix} G_0\Lambda & G_1\Lambda & G_2\Lambda \end{bmatrix} & \text{for t=3,6,9,...} \\ \\ \begin{bmatrix} G\Lambda & \mathbf{O}_{n\times 2r} \end{bmatrix} & \text{otherwise} \end{cases} \tag{24}$$

$$\text{and } \epsilon_t \sim N(\mathbf{0}_n, R), \text{ where } R = \begin{cases} \begin{bmatrix} \frac{3}{9}\Sigma_u^Q & \mathbf{O}_{n^Q\times n^M} \\ \mathbf{O}_{n^M\times n^Q} & \Sigma_u^M \end{bmatrix} & \text{for t=3,6,9,...} \\ \\ \begin{bmatrix} I_{n^Q} & \mathbf{O}_{n^Q\times n^M} \\ \mathbf{O}_{n^M\times n^Q} & \Sigma_u^M \end{bmatrix} & \text{otherwise} \end{cases} \tag{25}$$

For $t = 3, 6, 9, ..$ and for all monthly variables, we always have that $z_{i,t} = x_{i,t}$. For $t \neq 3, 6, 9, ..,$ we have that $z_{i,t} = 0$ when i is a quarterly variable. The zeros (missing monthly values for quarterly variables) are modeled as draws from a N(0,1) distribution.

The transition equation, representing the VAR of the factors, is given by:

$$\begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \end{bmatrix} = \begin{bmatrix} \overset{r\times rp}{\Phi_1 \cdots \Phi_p} & \mathbf{O}_{r\times(5-p)r} \\ I_{2r} & \mathbf{O}_{2r\times r} \end{bmatrix} \begin{bmatrix} f_{t-1} \\ f_{t-2} \\ f_{t-3} \end{bmatrix} \tag{26}$$

$$\text{where } e_t = \begin{bmatrix} v_t \\ \mathbf{0}_{2r} \end{bmatrix} \sim N\left(\mathbf{0}_{3r}, Q\right), \text{ where } Q = \begin{bmatrix} \Sigma_v & \mathbf{O}_{r\times 2r} \\ \mathbf{O}_{2r\times r} & \mathbf{O}_{2r\times 2r} \end{bmatrix} \tag{27}$$

B Robustness checks

B.1 Alternative forecast combination schemes

Besides rank-based weights, we additionally take into account other forecast combination strategies, namely:

1. Approximate a Bayesian weighting scheme that depends on the model's Bayesian (or Schwarz) information criterion. This scheme implies that $w_i^{SIC} = \frac{\exp\left(-0.5\cdot\Delta_i^{SIC}\right)}{\sum_{i=1}^{n}\exp\left(-0.5\cdot\Delta_i^{SIC}\right)}$ with $\Delta_i^{SIC} = SIC_i - SIC_{\min}$.

2. As a third model averaging scheme employing the full in-sample covariance information, we consider the Mallows Model Averaging (*MMA*) criterion proposed by Hansen (2007) and Hansen (2008). This measure is based on Mallows' criterion for model selection. The goal of this measure is to minimize the MSE over a set of feasible forecast combinations. This implies minimizing the function $C = (y - Fw)'(y - Fw) + w'Ks^2$, where K is a vector including the number of coefficients of each model and $s^2 = \frac{\hat{e}(M)'\hat{e}(M)}{T-k(M)}$ is an estimate σ^2 from the model with the smallest estimated error variance. We apply the constraints $0 \leq w \leq 1$ and $\sum_{i=1}^{n} \omega_i = 1$. Note that MMA explicitly takes into account the number of estimated parameters of the model.

3. Weights that are inversely proportional to the models' past mean square forecast errors (*mse*). Weights based on mean square forecast errors (MSFEs) for a specific training sample of size p. This combination scheme has been applied quite successfully by Stock and Watson (2004) and Drechsel and Scheufele (2012b) for output predictions based on leading indicators. This implies that S_{it} is the recursively computed mean square forecast error of model i over a training sample of size p, which is computed as $S_{it} = \sum_{s=t-h-p}^{t-h} (\hat{e}_{i,s})^2$. The mean square forecast error weights are based on

$$w_{i,t} = \frac{S_{it}^{-1}}{\sum_{j=1}^{n} S_{jt}^{-1}} \tag{28}$$

4. The rank-based weighting scheme. This scheme is based on the models past forecast performance, whereas the weights are computed according to ranks of the specific models (Aiolfi and Timmermann 2006, Drechsel and Scheufele 2012a). It is thus very closely related to combination schemes that use the past MSFEs to calculate the weights (as proposed by Stock and Watson 2004). Instead of relying directly on the MSFEs, the weights are assigned according to the model ranks. Let S_{it} be the computed mean square forecast error of model i within a training sample, with $S_{it} = \sum_{s=t-h-p}^{t-h} \frac{1}{p} (\hat{e}_{i,s})^2$. For each model, i, the rank for a h-step ahead forecast up to time t is then assessed by $\mathcal{R}_{t,t-h,i} = f(S_{1t}, ..., S_{nt})$. The model with the best MSFE forecasting performance obtains a rank of 1, the second-best a rank of 2 and so on. The individual weights are then calculated as $w_{i,t} = \frac{\mathcal{R}_{i,t-h}^{-1}}{\sum_{j=1}^{N} \mathcal{R}_{j,t-h}^{-1}}$. One

advantage of ranks compared to direct MSFE-weights is that they are less sensitive to outliers and, thus, should be more robust. In practice, the weighting scheme based on ranks places very high weight on the group of best models and nearly zero weight on models with less-accurate past performance.

B.2 Additional results

TABLE 5 — ALTERNATIVE FORECAST COMBINATION APPROACHES

	Total sample 2005Q1-2015Q2		Post crisis sample 2010Q1-2015Q2	
	h=1	h=2	h=1	h=2
I. EARLY-QUARTER INFORMATION				
Benchmark (mean)	1.6660	2.1831	1.2411	1.1755
sic	0.999	0.999	1.005*	0.999
mma	1.140*	1.072	1.283**	1.134
mse	0.995	1.007	1.037	1.010
rank	0.982	0.990	1.158*	1.027
II. LATE-QUARTER INFORMATION				
Benchmark (mean)	1.5604	2.0528	1.1498	1.1701
sic	0.997	1.000	1.005	1.000
mma	1.074	1.068	1.153	1.095
mse	0.979	1.006*	1.028	1.011
rank	0.969	0.965	1.115	0.995

Notes: The table shows root-mean-squared errors (RMSEs) for $h = 1$ (nowcast) and $h = 2$ (one additional quarter ahead) using two different states of information. Besides the benchmark based on equal weights, all numbers are defined relative to the benchmark. sic: weights based on information criterion, mma: Mallows model averaging, mse: inverse mse of previous performance, rank: inverse of the rank of the previous mean squared performance. ***, ** and * indicate whether a model's predictive ability (using the DM test) is significantly different from the benchmark (at the 1%, 5% and 10% levels, respectively).

Table 6 — Modifications of the specific-to-general approach

	Total sample 2005Q1-2015Q2		Post crisis sample 2010Q1-2015Q2	
	h=1	h=2	h=1	h=2
I. EARLY-QUARTER INFORMATION				
Benchmark (fdr)	1.5736	2.6475	1.5286	1.8642
Bonferroni	1.004	0.995	0.930	0.876
No correction	1.182**	0.976	1.007	1.255
II. LATE-QUARTER INFORMATION				
Benchmark (fdr)	1.5739	2.6653	1.5303	1.8846
Bonferroni	1.004	0.992	0.931	0.866
No correction	1.263**	0.978	1.064	1.253

Notes: The table shows root-mean-squared errors (RMSE) for $h = 1$ (nowcast) and $h = 2$ (one additional quarter ahead) using two different states of information. Besides the benchmark procedure, which uses the control of false discovery rate (FDR) of the significance level, all numbers are defined relative to the benchmark. Bonferroni: Bonferroni adjustment, No correction: no correction for multiple testing. ***, ** and * indicate whether a model's predictive ability (using the DM test) is significantly different from the benchmark (at the 1%, 5% and 10% levels, respectively).

Recent SNB Working Papers

2017-2 Alain Galli, Christian Hepenstrick and Rolf Scheufele: Mixed-frequency models for tracking short-term economic developments in Switzerland

2017-1 Matthias Gubler and Christoph Sax: The Balassa-Samuelson Effect Reversed: New Evidence from OECD Countries

2016-19 Jens H. E. Christensen and Signe Krogstrup: A Portfolio Model of Quantitative Easing.

2016-18 Jonas Meuli, Thomas Nellen and Thomas Nitschka: Securitisation, loan growth and bank funding: the Swiss experience since 1932.

2016-17 Rita Fleer, Barbara Rudolf and Mathias Zurlinden: Price change dispersion and time-varying pass-through to consumer prices.

2016-16 Gregor Bäurle, Daniel Kaufmann, Sylvia Kaufmann and Rodney W. Strachan: Changing dynamics at the zero lower bound.

2016-15 Adrian Jäggi, Martin Schlegel and Attilio Zanetti: Macroeconomic surprises, market environment and safe - haven currencies.

2016-14 Alain Galli: Sticky consumption and wealth effects in Switzerland.

2016-13 Gregor Bäurle and Rolf Scheufele: Credit cycles and real activity – the Swiss case.

2016-12 Silvio Schumacher: Networks and lending conditions: Empirical evidence from the Swiss franc money markets.

2016-11 Lucas Marc Fuhrer, Benjamin Müller and Luzian Steiner: The Liquidity Coverage Ratio and Security Prices.

2016-10 Nikola Mirkov, Igor Pozdeev and Paul Söderlind: Toward Removal of the Swiss Franc Cap: Market Expectations and Verbal Interventions.

2016-9 Severin Bernhard and Till Ebner: Cross-border Spill over Effects of Unconventional Monetary Policieson Swiss Asset Prices.

2016-8 Pinar Yeşin: Capital Flows and the Swiss Franc.

2016-7 Signe Krogstrup and Cédric Tille: On the roles of different foreign currencies in European bank lending.

2016-6 Pascal Towbin and Sebastian Weber: Price expectations and the US housing boom.

2016-5 Raphael A. Auer and Cédric Tille: The banking sector and the Swiss financial account during the financial and European debt crises.

2016-4 Christian Hepenstrick and Massimiliano Marcellino: Forecasting with Large Unbalanced Datasets: The Mixed-Frequency Three-Pass Regression Filter.

2016-3 Alain Galli: How reliable are cointegration-based estimates for wealth effects on consumption? Evidence from Switzerland.

2016-2 Pinar Yeşin: Exchange Rate Predictability and State-of-the-Art Models.

SCHWEIZERISCHE NATIONALBANK
BANQUE NATIONALE SUISSE
BANCA NAZIONALE SVIZZERA
BANCA NAZIUNALA SVIZRA
SWISS NATIONAL BANK