The Real-Time Predictive Content of the KOF Economic Barometer

Boriss Siliverstovs[a]

JEL-Classification: C53, C22
Keywords: Leading Indicators, Forecasting, Switzerland

1. Introduction

Decision makers face uncertainty not only about the future development of the economy but also regarding its current state. The uncertainty about the current state of the economy – usually measured by GDP – stems from the fact that quarterly GDP data are only made available with a lag. In case of the United States the delay is about one month after the end of the quarter. European GDP data are usually released with a delay of about six weeks. Moreover, GDP data often undergo substantial revisions as more information becomes available.

Up to date, a significant body of literature has evolved that attempts to reduce the uncertainty about the current and future developments of economy by relying on the leading indicators.[1] The recent economic crisis has increased the interest in this timely source of information.

a ETH Zurich, KOF Swiss Economic Institute, Weinbergstrasse 35, 8092 Zurich, Switzerland, e-mail: boriss.siliverstovs@kof.ethz.ch, +41 44 632 46 24
We are grateful to Niklas Ahlgren, Marc Gronwald, Konstantin A. Kholodilin and to participants at the KOF Brown Bag Seminar (Zurich, Switzerland), the XIV Spring Meeting of Young Economists (Istanbul, Turkey) as well as the DIW End of Year Summit 2010 for constructive comments on an earlier draft of the paper. Two anonymous referees provided very insightful comments and suggestions that greatly helped to shape the manuscript. We also would like to thank Philip Hubbard for the permission to use consensus forecasts data in this paper. The usual disclaimer applies. All computations were performed using Ox version 6.0 (Doornik, 2007).

1 The referee pointed out that decision makers may not be always willing to mechanically and swiftly react to all the news provided by the leading indicators, that often are based on the qualitative survey data, or to the latest GDP releases. Both of these information sources inevitably contain some noise, which can only be sorted out by statistical agencies with a substantial time delay. Nevertheless, at the times decisions are made this is the only up-to-date information that is available to them. Hence in one way or another it helps to shape their opinion on the current stance of the economy.

In Switzerland one of the most closely monitored leading indicators of economic activity is the KOF Economic Barometer. This paper provides an analysis of the predictive content of the KOF Barometer in forecasting the quarterly year-on-year growth rate of real GDP in Switzerland. We compare the forecasting performance of the KOF Barometer to that of the following benchmarks: forecasts from univariate autoregressive models and forecasts published by Consensus Economics Inc. The former allows us to disentangle the predictive content of the KOF Barometer from information contained in the GDP time series itself. Despite its simplicity, it is widely acknowledged that autoregressive models often attain forecast accuracy that it is quite difficult to improve upon. The latter benchmark forecasts are used in order to pit the forecasting performance of the KOF Barometer against predictions of seasoned and skilled professional forecasters that in their judgements may use additional information available or the information provided by the KOF Barometer in a more sophisticated way than we pursue in this paper.

We perform our forecasting exercise in a real-time setting. For this purpose we constructed a real-time data set consisting of all historical *quarterly* vintages released by the State Secretariat of Economic Affairs (SECO) of the quarterly year-on-year growth rate of real GDP and all historical *monthly* vintages of the KOF Economic Barometer starting from April 2006 when the KOF Economic Barometer based on a multisectoral design was introduced (Graff, 2006, 2010). It means that when making our forecasts we only use information that was known to the forecaster. For example, in January 2010 in order to predict the GDP growth rate in the second quarter of 2010, one had to use the latest available vintage of GDP data that was released by SECO in December 2009. Given the publication lag, this vintage contains the latest GDP observation for the third quarter of 2009. For the KOF Barometer it implies that one has to use the corresponding vintage released in January. Since the KOF Barometer has no publication lag, its last available value is for January 2010.

The importance of using real-time instead of latest available data has been emphasized in numerous studies, as it has been shown, for example by Diebold and Rudebusch (1991) and, more recently, by Amato and Swanson (2001) and Croushore (2005) that any favorable conclusion on the forecasting properties of leading indicators obtained using latest available data may be substantially weakened or even reversed when the forecasting exercise is replicated using real-time data. The main contribution of our study is that it is the first paper that assesses leading properties of the KOF Economic Barometer in real time. In addition, our forecast sample includes the recent financial crisis, allowing us to verify the forecasting power of the leading indicator under stress conditions. We demonstrate that it is useful for short-term forecasting of the Swiss GDP providing more accurate forecasts than the benchmark models.

The rest of the paper is structured as follows. Section 2 relates the present paper to previous research. The econometric framework and the results are described in Sections 3 and 4, respectively. The final section concludes.

2. KOF Economic Barometer

In Switzerland the use of business tendency surveys for assessing the economic situation has a long tradition. The first version of the KOF Barometer was developed in 1976. In 1998 it underwent a slight modification. In April 2006 the traditional KOF Barometer was replaced by the new KOF Barometer based on a multi-sectoral design (Graff, 2006, 2010).

Graff (2010) compares the predictive accuracy of the old KOF Barometer with that of a new one for the forecast period from 2003Q1 until 2006Q2. The most interesting feature of Graff (2010) is that a distinction between real-time and latest available data is made in using the barometer in out-of-sample forecasting. However, while coming close to simulating forecasting exercise in real time, Graff (2010) does not take into account that during the forecast period not only the KOF Barometer but also the GDP vintages underwent substantial revisions (Cuche-Curti, Hall and Zanetti, 2008). Hence by utilizing for forecast the latest available figures for the time series as they were known at the end of the forecast period, Graff (2010) is likely to overstate the forecasting accuracy of the leading indicator. Graff (2010) reports a significant improvement in forecast accuracy of the newly designed KOF Barometer over the traditional one. This, however, might be at least partly explained by the fact that the components of the new KOF Barometer were pre-selected using information for the whole forecast period that was clearly not available to forecasters who constructed the old KOF Barometer back in 1976.

Our study distinguishes itself from Graff (2010) in one important aspect. We conduct our exercise in real time; i.e., using real-time vintages both for the KOF Barometer as well as for the GDP growth rate. In doing so, we avoid biasing our results in favour of the leading indicator.

3. Forecasting Framework

We forecast the quarterly year-on-year GDP growth rate in a target quarter $\tau+1$. The release schedule of the GDP and the KOF Barometer is shown in Figure 1. In this figure, $\tau+i$ denotes quarters, with $i=-1,0,1,2$, and the Roman numerals indicate the months of the respective quarters. The Swiss State Secretariat for Economic Affairs (SECO) releases GDP vintages in the beginning of the third month in each quarter. The publication lag comprises one quarter, implying that the first GDP estimate for the quarter $\tau+j$ is released in the quarter $\tau+j+1$, with $j=-2,-1,0,1$. We denote such vintage by $GDP_{(\tau+j,\ \tau+j+1)}$. The KOF Barometer is released at the end of each month and has no publication lag. $KOF_{\tau+k^M}$, with $k=0,1$, stands for vintages of the KOF Barometer that are released in the end of each month $M=$ I,II,III in the respective quarters $\tau+k$.

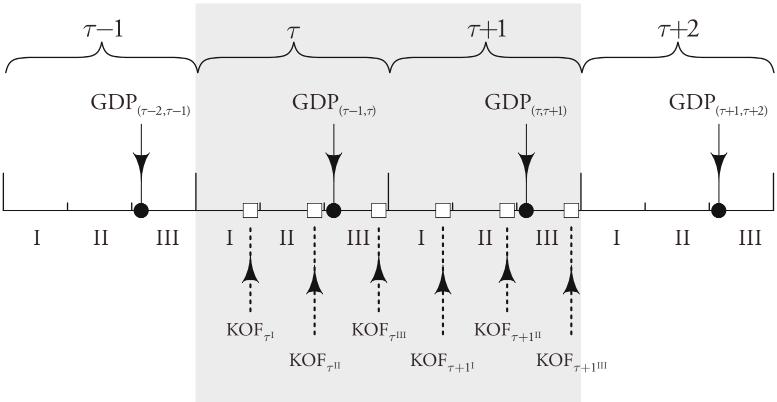
We distinguish between the following six forecasts, corresponding to six monthly releases of the KOF Barometer during quarters τ and $\tau+1$. We denote these forecasts by $r=1,2,\ldots,6$. Let Ω^r_{t+k}, with $k=0,1$, be the information setavailable to a forecaster at each forecast round $r=1,2,\ldots 6$ in respective quarters τ and $\tau+1$. Then, $\Omega^1_\tau=\{GDP_{(\tau-2,\tau-1)};KOF_{\tau^{\mathrm{I}}}\}$, $\Omega^2_\tau=\{GDP_{(\tau-1,\ \tau)};KOF_{r^{\mathrm{II}}}\}$, $\Omega^3_\tau=\{GDP_{(\tau-1,\ \tau)};KOF_{\tau^{\mathrm{III}}}\}$, $\Omega^4_{\tau+1}=\{GDP_{(\tau-1,\tau)};KOF_{\tau+1^{\mathrm{I}}}\}$, $\Omega^5_{\tau+1}=\{GDP_{(\tau,\tau+1)};KOF_{\tau+1^{\mathrm{II}}}\}$, and $\Omega^6_{\tau+1}=\{GDP_{(\tau,\tau+1)};KOF_{\tau+1^{\mathrm{III}}}\}$. The forecasts of quarterly year-on-year GDP growth rate in the quarter $\tau+1$ are compared with the actual values released in a vintage $GDP_{(\tau+j,\tau+2)}$ in the quarter $\tau+2$. Observe that our first forecast is about seven months ahead of the first GDP release by SECO. Our sixth and last forecast precedes an official GDP data release by about two months.

Every forecast is constructed in three steps. First, in cases where the last available values of the KOF Barometer are either for the first or second month of a quarter (this happens for $r=1,2,4,5$) we use a univariate autoregressive model to produce forecasts of the values of the KOF Barometer in the remaining months of the quarter. In these auxiliary regressions, the optimal lag length is automatically selected by the Schwarz Information Criterion (SIC).

Second, since the KOF Barometer is released at a monthly frequency and the GDP growth rate is a quarterly time series, one has to convert the original KOF Barometer to the quarterly frequency. This conversion is achieved by assuming that the value of the KOF Barometer actually recorded (or forecast in the first step) for the last month of a quarter is representative for the whole quarter.[2]

2 An alternative conversion method is to use the average of monthly values of the KOF Barometer as the representative for the whole quarter. This approach yielded slightly worse forecasting performance and for the sake of saving space is not reported.

Figure 1: Data Release Schedule



Notes: $\tau+i$ denotes quarters, with $i=-1,0,1,2$. The Roman numerals indicate the months of the quarters. $GDP_{(\tau+j,\,\tau+j+1)}$ denotes GDP vintages released by SECO in the quarter $\tau+j+1$ that have observations up to the quarter $\tau+j$, with $j=-2,-1,0,1$. $KOF_{\tau+k^M}$ with $k=0,1$ stands for vintages of the KOF Barometer that are released at the end of each month $M=$I,II,III in the respective quarters. Fere Ω^r_{t+1-k} is the information set available to a forecaster at each forecast round $r=1,2,\ldots,6$. Observe that $k=1$ for $r=1,2,3$ and $k=1$ for $r=4,5,6$. Then

$$\Omega^1_\tau=\{GDP_{(\tau-2,\tau-1)};KOF_{\tau^{I}}\},\ \Omega^2_\tau=\{GDP_{(\tau-1,\,\tau)};KOF_{\tau^{II}}\},\ \Omega^3_\tau=\{GDP_{(\tau-1,\,\tau)};KOF_{\tau^{III}}\},$$
$$\Omega^4_{\tau+1}=\{GDP_{(\tau-1,\tau)};KOF_{\tau+1^{I}}\},\ \Omega^5_{\tau+1}=\{GDP_{(\tau,\tau+1)};KOF_{\tau+1^{II}}\},\text{ and }\Omega^6_{\tau+1}=\{GDP_{(\tau,\tau+1)};KOF_{\tau+1^{III}}\}.$$

All the forecasts of the GDP growth rate for the target quarter $\tau+1$ are compared with the first official estimate of GDP growth rate in this quarter released in the vintage $GDP_{(\tau+1,\tau+2)}$. Consensus Economics Inc. releases its forecasts shortly after official publications of GDP. A next-quarter consensus forecast of GDP in $\tau+1$ is available at τ^{III}, shortly after the release of vintage $GDP_{(\tau-1,\tau)}$. A current-quarter consensus forecast of GDP in $\tau+1$ is available at $\tau+1^{III}$, shortly after the release of vintage $GDP_{(\tau,\tau+1)}$.

In the third step the forecasts are computed. In this step we utilize two basic strategies for generating multi-period forecasts: a direct approach and an iterated approach.[3]

3 Previous literature has pointed out that the iterated approach produces more efficient parameter estimates than the direct approach, but the latter method is more robust to model misspecification. Given the ambiguity regarding true data generating process, the choice of either method is an empirical one (Marcellino, Stock and Watson, 2006).

First we describe the direct forecasting approach. Forecasts for a target quarter $\tau+1$ conditional on the information set Ω^{r}_{t+1-k} are generated using the following AutoRegressive Distributed Lag (ARDL) model:

$$\widehat{Y}^{r}_{\tau+1|\Omega^{r}_{\tau+1-k}} = \alpha_0 + \sum_{i=h}^{5} \alpha_i Y^{r}_{\tau+1-i} + \sum_{j=k}^{5} \beta_j \tilde{X}^{r}_{\tau+1-j}, \tag{1}$$

$h=3$ for $r=1$, $h=2$ for $r=2,3,4$, and $h=1$ for $r=5,6$
$k=1$ for $r=1,2,3$, and $k=0$ for $r=4,5,6$,

where Y^{r}_{τ} is the year-on-year quarterly growth rate[4] of real GDP observed in quarter τ for a GDP vintage available in a forecast round r, with $r=1,2,\ldots,6$. $\tilde{X}^{r}_{\tau}$ is the KOF Barometer transformed to quarterly frequency. The model parameters are estimated using data in the window terminating in quarter $\tau+1-h$, with $h=1,2,3$.

The following univariate autoregressive model serves as a benchmark model:

$$\widehat{Y}^{r}_{\tau+1|\Omega^{r}_{\tau+1-k}} = \alpha_0 + \sum_{i=h}^{5} \alpha_i Y^{r}_{\tau+1-i}. \tag{2}$$

Next we describe the iterative forecasting approach. We denote by y^{r}_{τ} the quarterly growth rate[5] of real GDP observed in a particular quarter for a GDP vintage available for a forecast round r, with $r=1,2,\ldots,6$. The corresponding forecasting model is as follows:

$$\hat{y}^{r}_{\tau+1|\Omega^{r}_{\tau+1-k}} = \alpha_0 + \sum_{i=1}^{5} \alpha_i \hat{y}^{r}_{\tau+1-i|\Omega^{r}_{\tau+1-k}} + \sum_{j=k}^{5} \beta_j \tilde{X}^{r}_{\tau+1-j} + \gamma' SEAS_{\tau+1}, \tag{3}$$

where $\hat{y}^{r}_{\tau+1-i|\Omega^{r}_{\tau+1-k}} = y^{r}_{\tau+1-i}$ if $i \geq h$ and $SEAS_{\tau}$ is a vector of seasonal dummies. The parameters are estimated using data in the window terminating in quarter $\tau+1-h$.

4 In order to match the values of the dependent variable with those published by the SECO, we calculate Y^{r}_{τ} by exact formula for obtaining year-to-year growth rates from levels of the reference time series. Denoting the levels of a quarterly time series z_{τ}, the year-to-year growth rates are $(z_{\tau} - z_{\tau-4}) / z_{\tau-4} * 100$.

5 We calculate y^{r}_{τ} by exact formula for obtaining quarterly growth rates from levels of the reference time series. Denoting the levelsof a quarterly time series z_{τ}, the quarterly growth rates are $(z_{\tau} - z_{\tau-1}) / z_{\tau-1} * 100$.

Once forecasts of quarterly growth rates have been computed we convert them to year-on-year growth rates as follows:

$$\hat{Y}^{r}_{\tau+1|\Omega^{r}_{\tau+1-k}} = \left(\prod_{i=0}^{3}\left(1+\frac{\hat{y}^{r}_{\tau+1-i|\Omega^{r}_{\tau+1-k}}}{100}\right)-1\right)*100, \tag{4}$$

where $\hat{y}^{r}_{\tau+1-i|\Omega^{r}_{\tau+1-k}} = y^{r}_{\tau+1-i}$ for $i \geq h$ and $k=0,1$.

For the iterative method, we employ another benchmark autoregressive model based on Equation (3) with imposed zero restriction on all β_j coefficients:

$$\hat{y}^{r}_{\tau+1|\Omega^{r}_{\tau+1-k}} = \alpha_0 + \sum_{i=1}^{p}\alpha_i\,\hat{y}^{r}_{\tau+1-i|\Omega^{r}_{\tau+1-k}} + \gamma' SEAS_{\tau+1}. \tag{5}$$

Equations (1), (2), (3), and (5) represent general specifications. When making forecasts, an optimal lag augmentation was selected by minimizing the Schwarz Information Criterion.

We illustrate the presented notation with the following example of predicting GDP growth rate in the target quarter 2010Q2 for $r=1$, implying that $h=3$ and $k=1$.[6] For the direct forecasting method we have:

$$\hat{Y}^{1}_{2010Q2} = f(Y^{1}_{2009Q3}, Y^{1}_{2009Q2}, \ldots, Y^{1}_{2008Q3}; \tilde{X}^{1}_{2010Q1}, \ldots, \tilde{X}^{1}_{2009Q1}). \tag{6}$$

For the iterative forecasting method we chain one-step ahead forecasts as follows:

$$\begin{aligned}
\hat{y}^{1}_{2009Q4} &= f(y^{1}_{2009Q3}, y^{1}_{2009Q2}, \ldots, y^{1}_{2008Q3}; \tilde{X}^{1}_{2009Q3}, \ldots, \tilde{X}^{1}_{2008Q3}),\\
\hat{y}^{1}_{2010Q1} &= f(\hat{y}^{1}_{2009Q4}, y^{1}_{2009Q3}, \ldots, y^{1}_{2008Q4}; \tilde{X}^{1}_{2009Q4}, \ldots, \tilde{X}^{1}_{2008Q4}),\\
\hat{y}^{1}_{2010Q2} &= f(\hat{y}^{1}_{2010Q1}, \hat{y}^{1}_{2009Q4}, \ldots, y^{1}_{2009Q1}; \tilde{X}^{1}_{2010Q1}, \ldots, \tilde{X}^{1}_{2009Q1}).
\end{aligned}$$

$$\hat{Y}^{1}_{2010Q2} = \left(\left(1+\frac{\hat{y}^{1}_{2010Q2}}{100}\right)\left(1+\frac{\hat{y}^{1}_{2010Q1}}{100}\right)\left(1+\frac{\hat{y}^{1}_{2009Q4}}{100}\right)\left(1+\frac{y^{1}_{2009Q3}}{100}\right)-1\right)*100. \tag{7}$$

In addition we compare the forecast accuracy of the ARDL models with that of forecasts provided by Consensus Economics Inc. Consensus Economics Inc.

6 For notational simplicity we omit the statement that forecasts are conditional on the relevant information set Ω^{1}_{2010Q1}.

releases its forecasts shortly after official publications of GDP. A next-quarter consensus forecast of GDP in $\tau+1$ is available at τ^{III}, shortly after the release of vintage $\mathrm{GDP}_{(\tau-1,\tau)}$. A current-quarter consensus forecast of GDP in $\tau+1$ is available at $\tau+1^{\mathrm{III}}$, shortly after the release of vintage $\mathrm{GDP}_{(u,\tau+1)}$. Hence, the next-quarter and current-quarter consensus forecasts are released approximately at the same time as our forecasts for $r=2$ and $r=5$ are computed, respectively.

The forecast sample is 2006Q3–2010Q2. For the estimation of model parameters we use a rolling estimation window of fixed size. Given real-time availability of the data, the initial estimation window used for predicting the growth rate in 2006Q3 is 1993Q4–2005Q4 for $r=1$. For $r=2,3,4$ the initial estimation window is 1993Q4–2006Q1, and for $r=5,6$ is 1993Q4–2006Q2. For the next target quarter 2006Q4 we move the estimation windows one quarter forward; i.e., we use the estimation windows 1994Q1–2006Q1, 1994Q1–2006Q2, and 1994Q1–2006Q3 for $r=1$, $r=2,3,4$, and $r=5,6$, respectively.

4. Results

In this section we report the results of both in-sample and out-of-sample model comparison. First, we consider the in-sample evidence. The latest available vintages of both time series (adjusted to have the same mean and range) are displayed in Figure 2. For the most of the observation period – with the exception of the year 1994 – the KOF Barometer shows itself as a truly leading indicator. It appears to lead most of the turning points in GDP growth and, more importantly, it timely reacts to the current recession.

We follow AMATO and SWANSON (2001) by presenting the results of standard Granger causality tests. The null hypothesis is that the KOF Barometer does not Granger cause GDP growth rate in Switzerland. For every estimation window we performed a sequence of Wald tests of the coefficients $\beta_j=0$ in Equations (1) and (3) for the direct and iterative forecasts, respectively. The corresponding p-values are reported in Table 1 in the upper and lower panels in columns labeled GC. The null hypothesis can be decisively rejected for every estimation window. We also compared competing models in terms of the Schwarz Information Criterion (GRANGER, KING and WHITE, 1995) presented in Table 1. According to the SIC, the ARDL model should be preferred to the univariate autoregressive models in *all* and in *all but a few* cases for direct and iterating forecasts, respectively.

Next we present the results of the out-of-sample forecasting exercise. The ability of a leading indicator to predict GDP growth rate out of sample is a more informative test regarding causation between the two time series than the

Figure 2: Quarterly Year-On-Year Real GDP Growth Rate[a] and the KOF Barometer[b]

a Vintage released on 02/09/2010.
b Released on 27/08/2010; only the values of the last month of each quarter are reported.
Both time series were adjusted to have the same mean and range.

in-sample evidence. As noted in Ashley, Granger and Schmalensee (1980, p. 1149): "In our view the out-of-sample forecasting performance ... provide[s] the best information bearing on hypotheses about causation." Table 2 summarizes our findings for direct (upper panel) and iterative (lower panel) forecasts. We measure forecast accuracy by computing the Root Mean Squared Forecast Error (RMSFE) reported in columns (1), (2), and (3) for the AR, ARDL models as well as for the forecasts from Consensus Economics Inc., respectively. Recall that consensus forecasts are released only once per quarter. Columns (4) and (5) report the relative RMSFE of the ARDL model to the two benchmark forecasts.

Based on Table 2 several observations can be made. First, the ARDL model produces more accurate forecasts than the univariate AR model and, more interestingly, than the consensus forecasts as the corresponding RMSFE ratios are below one. Second, compared to the AR model forecasts the most sizeable improvement in forecast accuracy takes place for $r = 1$ and $r = 4$, i.e., when a new value of the KOF Barometer is released in the first month of a new quarter. Thirdly, the most

Table 1: In-Sample Model Comparison: Direct and Iterated Forecasts

Direct forecasts

Forecast quarter	GC[a]	SIC[b] ARDL	SIC[b] AR	GC	SIC ARDL	SIC AR	GC	SIC ARDL	SIC AR
Forecast		$r=1$			$r=2$			$r=3$	
2006Q3	[0.000]	0.387	0.797	[0.000]	0.380	0.664	[0.000]	0.380	0.664
2006Q4	[0.000]	0.384	0.825	[0.000]	0.286	0.665	[0.000]	0.287	0.665
2007Q1	[0.000]	0.339	0.840	[0.000]	0.289	0.659	[0.000]	0.290	0.659
2007Q2	[0.000]	0.354	0.821	[0.000]	0.178	0.610	[0.000]	0.180	0.610
2007Q3	[0.000]	0.217	0.810	[0.000]	0.021	0.615	[0.000]	0.021	0.615
2007Q4	[0.000]	–0.025	0.814	[0.000]	0.159	0.698	[0.000]	0.159	0.698
2008Q1	[0.000]	–0.066	0.866	[0.000]	–0.050	0.700	[0.000]	–0.048	0.700
2008Q2	[0.000]	–0.151	0.877	[0.000]	–0.199	0.659	[0.000]	–0.199	0.659
2008Q3	[0.000]	–0.257	0.884	[0.000]	–0.223	0.661	[0.000]	–0.222	0.661
2008Q4	[0.000]	–0.459	0.863	[0.000]	–0.185	0.586	[0.000]	–0.181	0.586
2009Q1	[0.000]	–0.251	0.857	[0.000]	–0.210	0.592	[0.000]	–0.212	0.592
2009Q2	[0.000]	–0.248	0.878	[0.000]	–0.138	0.674	[0.000]	–0.139	0.674
2009Q3	[0.000]	–0.168	0.932	[0.000]	–0.066	0.801	[0.000]	–0.070	0.801
2009Q4	[0.000]	–0.100	1.076	[0.000]	–0.006	0.751	[0.000]	–0.006	0.751
2010Q1	[0.000]	–0.017	1.122	[0.000]	–0.013	0.774	[0.000]	–0.019	0.774
2010Q2	[0.000]	–0.029	1.167	[0.000]	0.027	0.813	[0.000]	0.027	0.813
Forecast		$r=4$			$r=5$			$r=6$	
2006Q3	[0.000]	0.450	0.664	[0.002]	–0.087	–0.041	[0.002]	–0.088	–0.041
2006Q4	[0.000]	0.364	0.665	[0.001]	–0.142	–0.060	[0.001]	–0.142	–0.060
2007Q1	[0.000]	0.360	0.659	[0.001]	–0.145	–0.070	[0.001]	–0.144	–0.070
2007Q2	[0.000]	0.251	0.610	[0.000]	–0.321	–0.099	[0.000]	–0.321	–0.099
2007Q3	[0.000]	0.080	0.615	[0.001]	–0.341	–0.216	[0.001]	–0.343	–0.216
2007Q4	[0.000]	0.157	0.698	[0.000]	–0.419	–0.212	[0.000]	–0.419	–0.212
2008Q1	[0.000]	–0.039	0.700	[0.000]	–0.452	–0.176	[0.000]	–0.452	–0.176
2008Q2	[0.000]	–0.158	0.659	[0.000]	–0.465	–0.162	[0.000]	–0.465	–0.162
2008Q3	[0.000]	–0.186	0.661	[0.000]	–0.440	–0.237	[0.000]	–0.437	–0.237
2008Q4	[0.000]	–0.142	0.586	[0.000]	–0.519	–0.223	[0.000]	–0.519	–0.223
2009Q1	[0.000]	–0.172	0.592	[0.000]	–0.461	–0.136	[0.000]	–0.461	–0.136
2009Q2	[0.000]	–0.153	0.674	[0.000]	–0.452	–0.047	[0.000]	–0.453	–0.047
2009Q3	[0.000]	–0.108	0.801	[0.000]	–0.730	–0.333	[0.000]	–0.729	–0.333
2009Q4	[0.000]	–0.136	0.751	[0.000]	–0.750	–0.326	[0.000]	–0.752	–0.326
2010Q1	[0.000]	–0.141	0.774	[0.000]	–0.706	–0.286	[0.000]	–0.706	–0.286
2010Q2	[0.000]	–0.120	0.813	[0.000]	–0.712	–0.340	[0.000]	–0.712	–0.340

Table 1 continued

Iterated forecasts

Forecast quarter	GC	SIC ARDL	SIC AR	GC	SIC ARDL	SIC AR	GC	SIC ARDL	SIC AR
Forecast		$r=1$			$r=2$			$r=3$	
2006Q3	[0.003]	–0.254	–0.160	[0.003]	–0.268	–0.169	[0.003]	–0.268	–0.169
2006Q4	[0.003]	–0.242	–0.140	[0.004]	–0.209	–0.137	[0.004]	–0.206	–0.137
2007Q1	[0.005]	–0.173	–0.109	[0.007]	–0.181	–0.147	[0.007]	–0.179	–0.147
2007Q2	[0.001]	–0.341	–0.184	[0.001]	–0.376	–0.212	[0.001]	–0.375	–0.212
2007Q3	[0.001]	–0.390	–0.184	[0.000]	–0.424	–0.206	[0.000]	–0.424	–0.206
2007Q4	[0.000]	–0.476	–0.183	[0.001]	–0.181	–0.034	[0.001]	–0.181	–0.034
2008Q1	[0.000]	–0.468	–0.155	[0.000]	–0.504	–0.185	[0.000]	–0.504	–0.185
2008Q2	[0.000]	–0.474	–0.217	[0.000]	–0.445	–0.141	[0.000]	–0.445	–0.141
2008Q3	[0.000]	–0.412	–0.116	[0.000]	–0.442	–0.143	[0.000]	–0.442	–0.143
2008Q4	[0.000]	–0.531	–0.139	[0.000]	–0.455	–0.164	[0.000]	–0.452	–0.164
2009Q1	[0.000]	–0.422	–0.136	[0.000]	–0.424	–0.135	[0.000]	–0.425	–0.135
2009Q2	[0.000]	–0.394	–0.104	[0.000]	–0.323	–0.048	[0.000]	–0.324	–0.048
2009Q3	[0.000]	–0.288	–0.056	[0.000]	–0.316	–0.032	[0.000]	–0.316	–0.032
2009Q4	[0.000]	–0.294	–0.008	[0.003]	–0.428	–0.328	[0.003]	–0.428	–0.328
2010Q1	[0.002]	–0.426	–0.317	[0.001]	–0.512	–0.348	[0.001]	–0.518	–0.348
2010Q2	[0.002]	–0.508	–0.377	[0.001]	–0.482	–0.346	[0.001]	–0.482	–0.346
Forecast		$r=4$			$r=5$			$r=6$	
2006Q3	[0.002]	–0.250	–0.169	[0.005]	–0.197	–0.167	[0.005]	–0.195	–0.167
2006Q4	[0.006]	–0.162	–0.137	[0.008]	–0.174	–0.174	[0.008]	–0.173	–0.174
2007Q1	[0.010]	–0.136	–0.147	[0.008]	–0.174	–0.174	[0.008]	–0.172	–0.174
2007Q2	[0.001]	–0.341	–0.212	[0.001]	–0.377	–0.232	[0.001]	–0.377	–0.232
2007Q3	[0.000]	–0.393	–0.206	[0.001]	–0.176	–0.061	[0.001]	–0.175	–0.061
2007Q4	[0.001]	–0.150	–0.034	[0.001]	–0.186	–0.063	[0.001]	–0.185	–0.063
2008Q1	[0.000]	–0.493	–0.185	[0.000]	–0.471	–0.150	[0.000]	–0.472	–0.150
2008Q2	[0.000]	–0.435	–0.141	[0.000]	–0.471	–0.168	[0.000]	–0.471	–0.168
2008Q3	[0.000]	–0.443	–0.143	[0.000]	–0.403	–0.192	[0.000]	–0.404	–0.192
2008Q4	[0.000]	–0.411	–0.164	[0.000]	–0.424	–0.163	[0.000]	–0.425	–0.163
2009Q1	[0.000]	–0.393	–0.135	[0.000]	–0.366	–0.078	[0.000]	–0.365	–0.078
2009Q2	[0.000]	–0.345	–0.048	[0.000]	–0.380	–0.026	[0.000]	–0.379	–0.026
2009Q3	[0.000]	–0.345	–0.032	[0.000]	–0.586	–0.359	[0.000]	–0.585	–0.359
2009Q4	[0.000]	–0.550	–0.328	[0.000]	–0.609	–0.361	[0.000]	–0.611	–0.361
2010Q1	[0.000]	–0.619	–0.348	[0.000]	–0.614	–0.325	[0.000]	–0.614	–0.325
2010Q2	[0.000]	–0.589	–0.346	[0.000]	–0.620	–0.349	[0.000]	–0.621	–0.349

Notes to Table 1

The table presents the outcome of in-sample evaluation of the competing nested models. For each forecast quarter the six consecutive forecasting models $r=1,2,\ldots,6$ were estimated using appropriate vintages of the KOF Barometer and GDP according to their availability in real time; see the end of Section 3 for an illustrating example.

- a "GC" stands for Granger Causality. The column entries are marginal significance levels (p-values) for the null hypothesis that the KOF Barometer does not Granger cause real GDP growth rate.
- b The Schwarz Information Criterion (SIC) is reported for the ARDL and AR models shown in Equations (1) and (2) and (3) and (5) for direct and iterative forecasts, respectively.

Notes to Table 2

- a Columns (1)–(3) report the RMSFE for forecasts from AR and ARDL models as well as for forecasts (CF) reported by Consensus Economics Inc. The consensus forecasts are released once per quarter and they refer to the next-quarter forecasts ($r=2$) and the current-quarter forecasts ($r=5$).
- b Columns (4) and (5) report the ratios of RMSFE of the ARDL model to those of the AR model and the consensus forecasts.
- c In column (6), the marginal significance levels (p-values) for the null hypothesis of equal predictive accuracy between the ARDL and the AR models are reported. We use the test of Clark and West (2007) that is appropriate for comparing predictive accuracy between nested models. The one sided alternative hypothesis is used that the ARDL model produces more accurate forecasts than the AR model. The truncation lag for the Newey-West estimator is $h-1$.
- d In column (7), the marginal significance levels (p-values) for the null hypothesis of equal predictive accuracy of between the ARDL and the consensus forecasts are reported. The alternative hypothesis is one sided that the ARDL model forecasts are more accurate than the consensus forecasts. We use the Diebold and Mariano (1995) test statistics modified as suggested in Harvey, Leybourne and Newbold (1997). The loss function is quadratic and the truncation lag for the Newey-West estimator is $h-1$. The critical values are obtained using the Student's t distribution with $P-1$ degrees of freedom, where P is the length of the forecast evaluation period.
- e In columns (8) and (9), the outcome of the forecast encompassing test of Harvey, Leybourne and Newbold (1998) is reported. In column (8), we report the marginal significance levels (p-values) for the null hypotheses that the consensus forecasts encompass those of the ARDL model. In column (9) – that the ARDL model forecasts encompass the consensus forecasts. The truncation lag for the Newey-West estimator is $h-1$. The critical values are obtained using the Student's t distribution with $P-1$ degrees of freedom, where P is the length of the forecast evaluation period.

Table 2: Forecasting Performance Evaluation [2006Q3–2010Q2].

Forecast	Forecast horizon	(1) AR	(2) RMSFE[a] ARDL	(3) CF	(4) RMSFE Ratio[b] ARDL/AR	(5) ARDL/CF	(6) CW[c] ARDL vs AR	(7) MDM[d] ARDL vs CF	(8) HLN[e] CF enc. ARDL	(9) ARDL enc. CF
					Direct forecasts					
$r=1$	$h=3$	1.910	1.095		0.573		[0.030]			
$r=2$	$h=2$	1.370	1.143	1.152	0.834	0.992	[0.029]	[0.470]	[0.105]	[0.162]
$r=3$	$h=2$	1.370	1.018		0.743		[0.021]			
$r=4$	$h=2$	1.370	0.725		0.529		[0.016]			
$r=5$	$h=1$	0.779	0.505	0.807	0.648	0.625	[0.035]	[0.021]	[0.013]	[0.896]
$r=6$	$h=1$	0.779	0.499		0.641		[0.030]			
					Iterated forecasts					
$r=1$	$h=3$	1.731	0.994		0.574		[0.011]			
$r=2$	$h=2$	1.180	0.914	1.152	0.775	0.794	[0.019]	[0.062]	[0.036]	[0.878]
$r=3$	$h=2$	1.180	0.912		0.773		[0.016]			
$r=4$	$h=2$	1.180	0.689		0.584		[0.024]			
$r=5$	$h=1$	0.822	0.630	0.807	0.766	0.781	[0.020]	[0.064]	[0.037]	[0.252]
$r=6$	$h=1$	0.822	0.619		0.752		[0.018]			

accurate forecasts occurs at $h=1$; i.e., when the first estimate of GDP growth rate in the previous quarter is incorporated in the information set.

The superiority of the ARDL forecasts over the AR forecasts is further illustrated in Figures 3 and 4 as well as in Figures 5 and 6 for each r. Figures 3 and 4 present AR and ARDL forecasts as well as actual values of first-released GDP growth rate. In Figures 5 and 6 the difference in absolute forecast errors of the AR and ARDL models is displayed. Bars above the zero line indicate that for a given quarter an ARDL-forecast is more accurate than a corresponding AR forecast. Bars below the zero line indicate the opposite.

In order to assess whether the reported differences in forecast accuracy are statistically significant we employ the test of CLARK and WEST (2007) that allows us comparing forecasting performance of the nested models; i.e., the AR and ARDL models (column (6) in the table). Since one can safely assume that the ARDL model forecasts and the consensus forecasts do not come from nested models, we test for equal predictive ability of this pair of forecasts using the well-known test of DIEBOLD and MARIANO (1995) modified as suggested in HARVEY, LEYBOURNE and NEWBOLD (1997) (column (7)). We also conduct the forecast encompassing test of HARVEY, LEYBOURNE and NEWBOLD (1998) (columns (8) and (9)).

According to the results of the CLARK and WEST (2007) test we can reject the null hypothesis of equal predictive ability of the ARDL and AR models at the 5% significance level for all r both for the direct and iterated forecasts. When comparing the direct forecasts of the ARDL model with the consensus forecasts we can only reject the null hypothesis for the current-quarter forecasts $r=5$ at the 5% level according to the results of the modified Diebol-Mariono test. According to the results of the test of HARVEY, LEYBOURNE and NEWBOLD (1997) the null that the consensus forecasts encompass the direct ARDL forecasts can be rejected at the 5% significance level. The more interesting hypothesis whether the ARDL model forecasts encompass the consensus forecasts. The null of encompassing in this direction can not be rejected at the usual significance levels. Consequently, we conclude that encompassing holds in one direction only, with the ARDL model forecasts encompassing the consensus forecasts. For the next-quarter forecasts at $r=2$ the results of both the Diebold-Mariano and forecast encompassing tests are inconclusive. This is not surprising as the corresponding RMSFE ratio is 0.992. When comparing the iterated ARDL forecasts with the consensus forecasts we can reject the null hypothesis of equal forecast accuracy at the 10% in favor of the former model for $r=5$ and $r=2$. We also establish that encompassing holds in one direction only, with the iterated ARDL forecasts encompassing the consensus forecasts.

Figure 3: Quarterly Year-On-Year Real GDP Growth Rate: Actual (First Release) and AR Model Forecasts

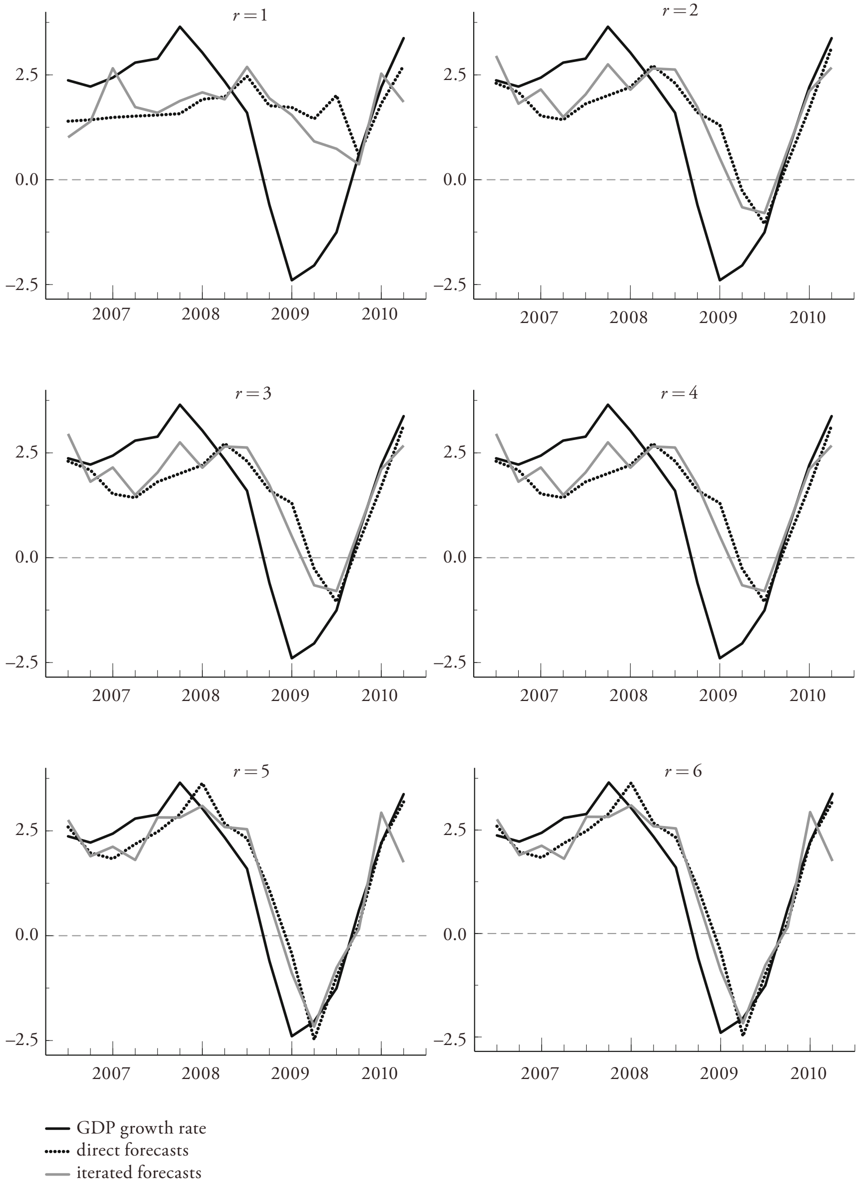


Figure 4: Quarterly Year-On-Year Real GDP Growth Rates: Actual (First Release) and ARDL Model Forecasts

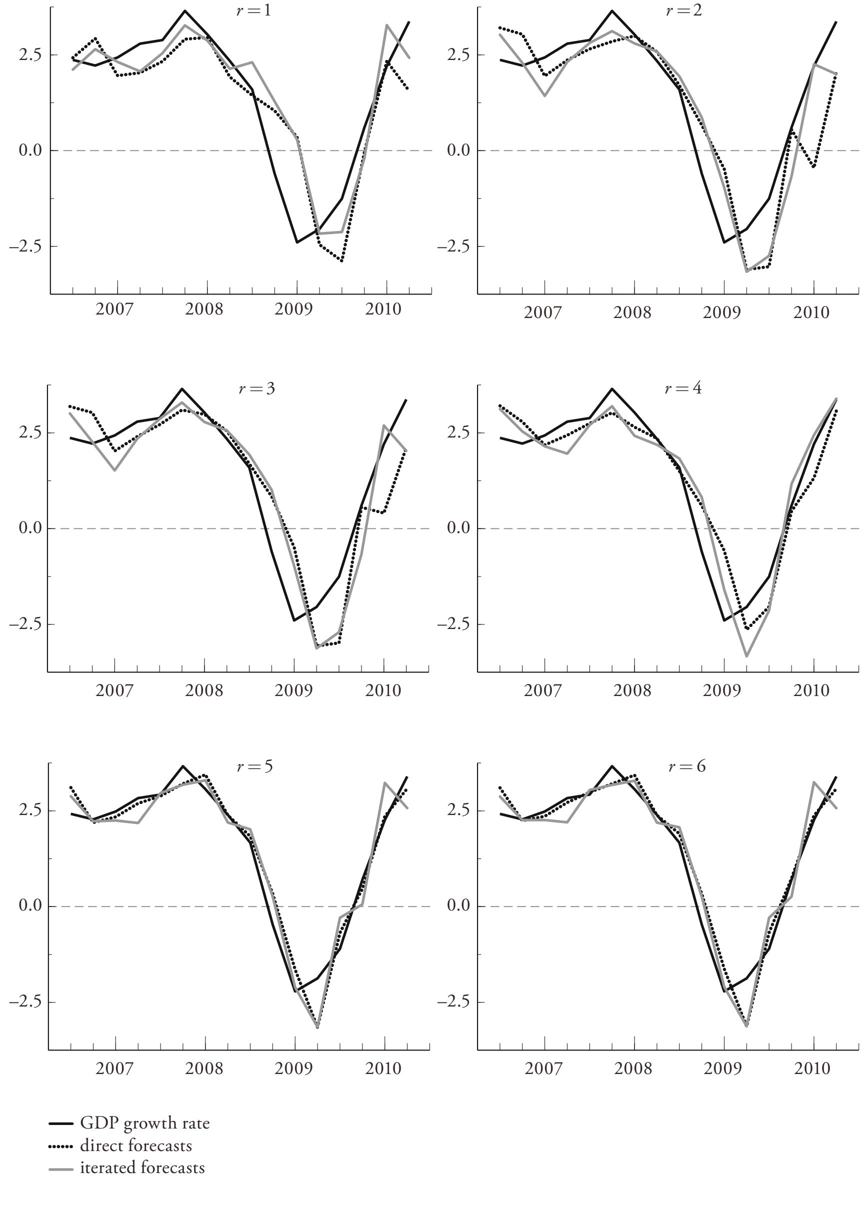
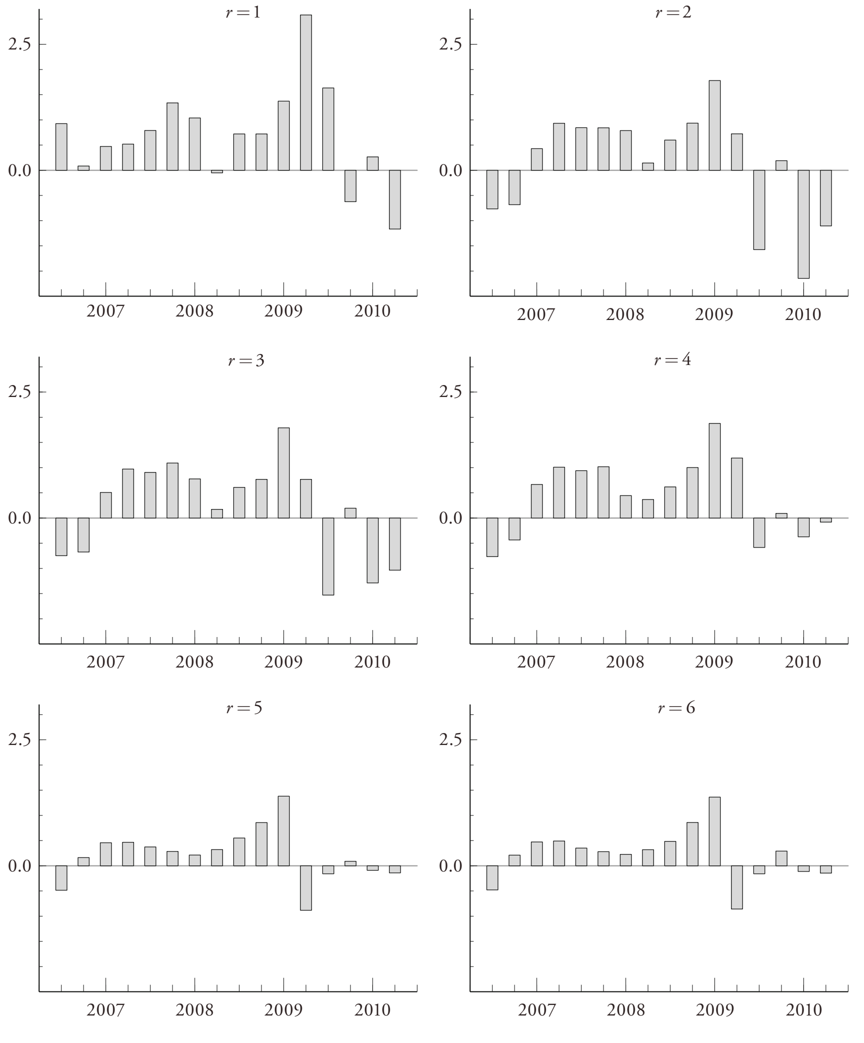
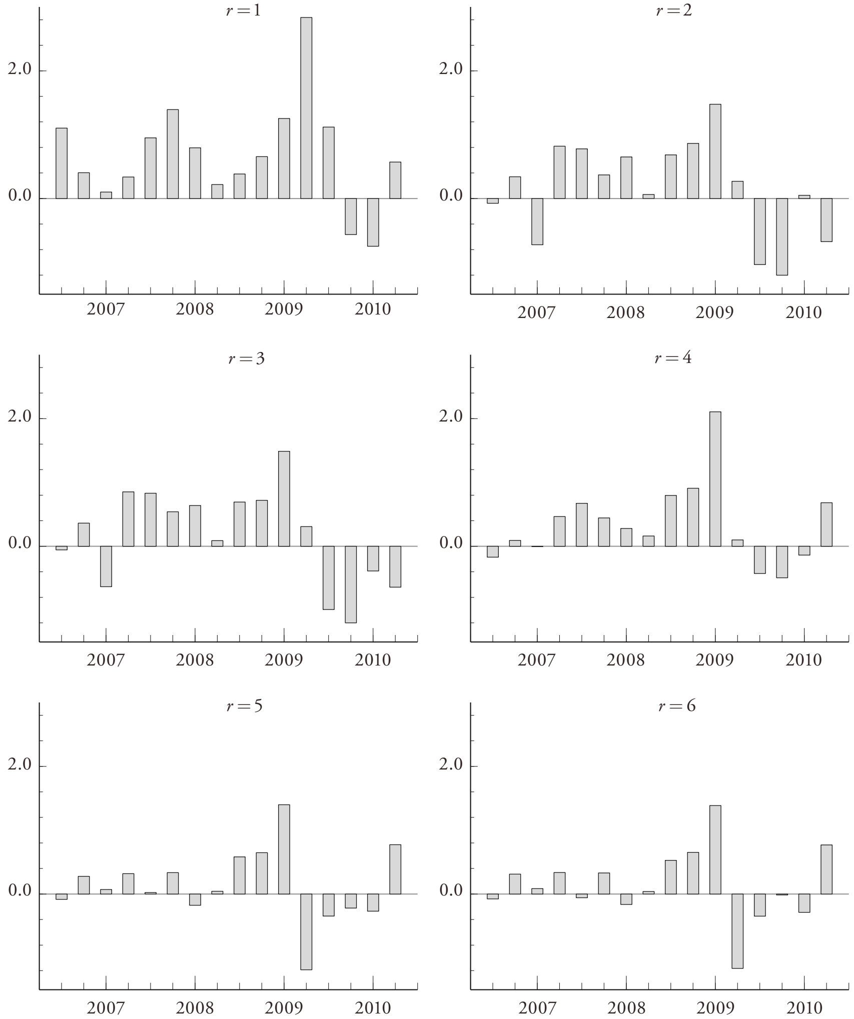


Figure 5: Direct Forecasts: Difference in Absolute Forecast Errors of AR and ARDL Models



Notes: Each of the panels contains a plot of the difference between the absolute forecast errors from the AR model (i.e., without the KOF Barometer) and the absolute forecast errors from the ARDL model (i.e., with the KOF Barometer). Bars above the zero line indicate quarters for which the AR model produced larger forecast errors than the ARDL model. Bars below the line indicate the opposite. All forecasts are constructed using real-time data sets, and the forecast comparison is based on first-released GDP data.

Figure 6: Iterated Forecasts: Difference in Absolute Forecast Errors of AR and ARDL Models



Notes: Each of the panels contains a plot of the difference between the absolute forecast errors from the AR model (i.e., without the KOF Barometer) and the absolute forecast errors from the ARDL model (i.e., with the KOF Barometer). Bars above the zero line indicate quarters for which the AR model produced larger forecast errors than the ARDL model. Bars below the line indicate the opposite. All forecasts are constructed using real-time data sets, and the forecast comparison is based on first-released GDP data.

5. Conclusions

In this paper we investigate whether the KOF Economic Barometer is useful for short-term forecasting of the quarterly year-on-year real GDP growth rate in Switzerland. We employ both direct and iterative methods to produce forecasts. We perform our analysis using real-time vintages of both time series. The forecast sample is 2006Q3–2010Q2. For each forecast quarter we produce a sequence of six forecasts taking into account the flow of conjunctural information in the form of new releases of the KOF Barometer and GDP data. Our first forecast is about seven months ahead of the first GDP release by SECO. Our sixth and last forecast precedes an official GDP data release by about two months. We compare the forecast accuracy of models with the leading indicator against univariate autoregressive model forecasts as well as forecasts published by Consensus Economics Inc.

Our main findings are the following. The model with the KOF Barometer displays superior performance over the univariate autoregressive models both in-sample and out-of-sample. For all estimation windows, we decisively reject the null hypothesis that the KOF Barometer does not Granger cause the GDP growth rate. In the out-of-sample forecast comparisons the model with the KOF Barometer provides a substantial improvement in forecast accuracy over the benchmark model. The largest improvement in forecast accuracy takes place at the first and fourth forecast round for which we observe reductions in the measures of forecast accuracy of about 40%. The largest improvement in forecast accuracy of the model with the leading indicator is achieved at the fifth (and sixth) forecast round for the direct forecasting approach.

Application of formal statistical tests for equal predictive ability suggests that the reduction in the forecast accuracy measures is significant in case of the AR and ARDL forecasts, both direct and iterative. We also establish that the ARDL forecasts are more accurate than the consensus forecasts except when comparing next-quarter consensus forecasts with direct forecasts. In this case the forecast accuracy of both approaches is similar. The forecast encompassing holds in one direction only, with the ARDL forecasts encompassing the consensus forecasts.

Appendix

DIEBOLD and MARIANO (1995) Test

DIEBOLD and MARIANO (1995) suggest a test for comparing the out-of-sample predictive ability of two models. They consider two sequences of forecasts $\{\hat{y}_{1t}\}_{t=1}^{P}$ and $\{\hat{y}_{2t}\}_{t=1}^{P}$ of a time series $\{y\}_{t=1}^{P}$. Assuming a quadratic loss function, the loss at time t is $e_{it}^2 = (y_t - \hat{y}_{it})^2$, for $i = 1,2$. The loss differential is $d_t = e_{1t}^2 - e_{2t}^2$. The null hypothesis of equal predictive ability is $E(d_t) = 0$.

The Diebold-Mariano (DM) test statistic is

$$DM = \frac{\overline{d}}{\sqrt{V(\overline{d})}}, \tag{8}$$

where $\overline{d}$ is a sample mean of the time series $\{d_t\}_{t=1}^{P}$ and

$$V(\overline{d}) = \frac{1}{P}\left(\gamma(0) + 2\sum_{\tau=1}^{h-1}\gamma(\tau)\right). \tag{9}$$

$\gamma(\tau)$ stands for the τ-th autocovariance of d_t, and h denotes the forecast horizon. The limiting distribution of the DM test statistic is a standard Gaussian in case one compares forecasts from non-nested models.

HARVEY, LEYBOURNE and NEWBOLD (1997) argue that in small samples the t-distributed modified Diebold-Mariano (MDM) test statistic should be prefered:

$$MDM = \frac{DM}{\sqrt{\frac{P+1-2h+\frac{h(h-1)}{P}}{P}}}. \tag{10}$$

HARVEY, LEYBOURNE and NEWBOLD (1998) Test

HARVEY, LEYBOURNE and NEWBOLD (1998) suggest a test for forecast encompassing. Forecast $\hat{y}_{1t}$ encompasses $\hat{y}_{2t}$ if the latter forecast adds no predictive power to the former forecast. We denote the composite forecast by $\hat{y}_{ct} = (1-\lambda)\hat{y}_{1t} + \lambda\hat{y}_{2t}$, for $\lambda > 0$. The associated forecast error of the composite forecast is $e_{ct} = y_t - \hat{y}_{ct}$ which transforms to $e_{1t} = \lambda(e_{1t} - e_{2t}) + e_{ct}$. The null hypothesis of forecast encompassing corresponds to $\lambda = 0$, which can be tested in the regression of e_{1t} on $(e_{1t} - e_{2t})$. Harvey, Leybourne and Newbold (1998) suggest to cast the test for forecast encompassing in a framework similar to that of the DM test. To this end, they define $f_t = e_{1t}(e_{1t} - e_{2t})$. Then the null hypothesis of forecast encompassing is $E(f_t) = 0$.

The forecast encompassing test statistic (HLN) is

$$HLN = \frac{\bar{f}}{\sqrt{V(\bar{f})}}, \tag{11}$$

where $\bar{f}$ is a sample mean of the time series $\{f_t\}_{t=1}^{P}$ and

$$V(\bar{f}) = \frac{1}{P}\left(\psi(0) + 2\sum_{\tau=1}^{h-1}\psi(\tau)\right). \tag{12}$$

$\psi(\tau)$ stands for the t-th autocovariance of f_t, and h denotes the forecast horizon. The limiting distribution of the HLN test statistic is a standard Gaussian in case one compares forecasts from non-nested models. A similar small sample correction of Harvey, Leybourne and Newbold (1997) can also be applied in this case.

Clark and West (2007) Test

In case of nested models the limiting distribution of the DM test is non-Gaussian. Nevertheless, Clark and West (2007) suggest the following modification that allows to test the null hypothesis of equal predictive ability using Gaussian critical values also when comparing forecasts from nested models. Assume that $\{\hat{y}_{1t}\}_{t=1}^{P}$ is the sequence of forecasts from the parsimonious model. Then $\{\hat{y}_{2t}\}_{t=1}^{P}$ represent the forecasts from the larger model that nests the smaller model. Assuming a quadratic loss function, the loss at time t is $e_{it}^2 = (y_t - \hat{y}_{it})^2$, for $i = 1,2$. Recall that Diebold and Mariano (1995) define the loss differential by $d_t = e_{1t}^2 - e_{2t}^2$. In contrast, Clark and West (2007) define the loss differential by $d_t^{adj} = e_{1t}^2 - e_{2t}^2 + (\hat{y}_{1t} - \hat{y}_{2t})^2$. As before, the null hypothesis of equal predictive ability is $E(d_t^{adj}) = 0$.

The Clark-West (CW) test statistic is

$$CW = \frac{\bar{d}^{adj}}{\sqrt{V(\bar{d}^{adj})}}, \tag{13}$$

where $\bar{d}^{adj}$ is a sample mean of the time series $\{d_t^{adj}\}_{t=1}^{P}$ and

$$V(\bar{d}^{adj}) = \frac{1}{P}\left(\phi(0) + 2\sum_{\tau=1}^{h-1}\phi(\tau)\right). \tag{14}$$

$\phi(\tau)$ stands for the τ-th autocovariance of d_t^{adj}, and h denotes the forecast horizon. CLARK and WEST (2007) suggest to use one-sided alternative hypothesis that the larger model produces more accurate forecasts than the smaller model.
Observe that the term $(\hat{y}_{1t} - \hat{y}_{2t})^2$ can be rewritten as $(e_{2t} - e_{1t})^2$. Then the adjusted loss differential is $d_t^{adj} = 2e_{1t}(e_{1t} - e_{2t})$, resembling f_t in the forecast encompassing test presented above. Hence the test of CLARK and WEST (2007) could be interpreted also as forecasting encompassing test.

References

AMATO, JEFFREY D., and NORMAN R. SWANSON (2001), "The Real-Time Predictive Content of Money for Output", *Journal of Monetary Economics*, 48, pp. 3–24.

ASHLEY, RICHARD, CLIVE W. J. GRANGER, and RICHARD SCHMALENSEE (1980), "Advertising and Aggregate Consumption: An Analysis of Causality", *Econometrica*, 48, pp. 1149–1167.

CLARK, TODD E., and KENNETH D. WEST (2007), "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models", *Journal of Econometrics*, 138, pp. 291–311.

CROUSHORE, DEAN (2005), "Do Consumer-Confidence Indexes Help Forecast Consumer Spending in Real Time?", *The North American Journal of Economics and Finance*, 16, pp. 435–450.

CUCHE-CURTI, NICOLAS, PAMELA HALL, and ATTILIO ZANETTI (2008), "Swiss GDP Revisions: A Monetary Policy Perspective", *Journal of Business Cycle Measurement and Analysis*, 4, pp. 183–214.

DIEBOLD, FRANCIS X. and ROBERT S. MARIANO (1995), "Comparing Predictive Accuracy", *Journal of Business & Economic Statistics*, 13, pp. 253–63.

DIEBOLD, FRANCIS X. and GLENN D. RUDEBUSCH (1991), "Forecasting Output with the Composite Leading Index : A Real-Time Analysis", *Journal of the American Statistical Association*, 86, pp. 603–610.

DOORNIK, JURGEN A. (2007), *Object-Oriented Matrix Programming Using Ox*. London.

Graff, Michael (2006), „Ein multisektoraler Sammelindikator für die Schweizer Konjunktur“, *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 142, pp. 529–577.

Graff, Michael (2010), “Does a Multi-Sectoral Design Improve Indicator-Based Forecasts of the GDP Growth Rate? Evidence for Switzerland”, *Applied Economics*, 42, pp. 2579–2781.

Granger, Clive W. J., Maxwell L. King, and Hall White (1995), “Comments on Testing Economic Theories and the Use of Model Selection Criteria”, *Journal of Econometrics*, 67, pp. 173–187.

Harvey, David I., Stephen J. Leybourne, and Paul Newbold (1997), “Testing the Equality of Prediction Mean Squared Errors”, *International Journal of Forecasting*, 13, pp. 281–291.

Harvey, David I., Stephen J. Leybourne, and Paul Newbold (1998), “Tests for Forecast Encompassing”, *Journal of Business & Economic Statistics*, 16, pp. 254–259.

Marcellino, Massimiliano, James H. Stock, and Mark W. Watson (2006), “A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series”, *Journal of Econometrics*, 135, pp. 499–526.

SUMMARY

We investigate whether the KOF Economic Barometer – a leading indicator released by the KOF Swiss Economic Institute – is useful for short-term prediction of quarterly year-on-year real GDP growth in Switzerland. Using a real-time data set consisting of historical vintages of GDP data and the leading indicator we find that the model augemented with the KOF Barometer produces more accurate forecasts of the Swiss GDP than purely autoregressive models and consensus forecasts.