Are Business Tendency Surveys Useful for Short-term Forecasting of GDP? Real-Time Evidence for Switzerland*

Boriss Siliverstovs**

February 13, 2009

Abstract

We investigate whether business tendency surveys collected for Switzerland can be useful for short-term out-of-sample prediction of year-on-year quarterly real GDP growth rates. We find that BTS appear to be useful for prediction of GDP growth rates in Switzerland. Even the earliest forecasts, made six months ahead of first official GDP estimate, allow us to predict GDP growth rates more accurately than forecasts based on univariate autoregressive model. At every subsequent forecast round we observe increase in forecast accuracy as reflected in steadily declining values of RMSFE and MAFE criteria. The corresponding ratios of RMSFE and MAFE criteria observed for the ARDL model to those of the benchmark AR model are 0.72 and 0.78 for the first forecast round and 0.70 and 0.59—for the final fifth forecast round. Based on the values of MAFE criterion, we expect an average forecast error of about 0.5 and 0.3 percentage points for forecasts made from six to four months and for forecasts made three and two months ahead of the first GDP release, respectively.

Keywords: Surveys, forecasting, Bayesian model averaging, GDP

JEL code: C53, C22.

---

*The paper has benefited from comments made at the KOF Research Seminar, Zurich.

**ETH Zurich, KOF Swiss Economic Institute, Weinbergstrasse 35, 8092 Zurich, Switzerland, e-mail: boriss.siliverstovs@kof.ethz.ch

1 Introduction

Various decision-making institutions face a great deal of uncertainty regarding not only the future discourse of the economy but also regarding its current stance. The uncertain knowledge about the current state of economic activity—usually measured by GDP—stems from the fact that quarterly GDP data are only available with a significant delay. In case of the United States such delay is about one month after the end of the reference quarter and in the European countries GDP data are released with delay of about two months. Moreover, as practice shows, the first release of GDP data often undergoes (substantial) revisions made by statistical agencies as more complete information becomes available later.

Up to date, a significant body of literature has evolved that attempts to reduce the uncertainty about current and future developments in economy by relying on the coincident/leading indicators (both quantitative and qualitative) that are readily available to decision makers and whose publication precedes that of quarterly GDP data, or any other data of interest. The quantitative indicators are either macroeconomic or financial variables. A typical example of the quantitative coincident indicators are industrial production, total personal income less transfer payments, total manufacturing and trade sales, and employees on nonagricultural payrolls, available at the monthly frequency, that were used in Stock and Watson (1988) to construct a coincident index model. The qualitative indicators are constructed on basis of business and consumer tendency surveys and they reflect an assessment of the current situation as well as recent and expected developments as perceived by businessmen and consumers, respectively.

In this article, we investigate the usefulness of the business tendency surveys (BTS) collected at the KOF Swiss Economic Institute for short-term forecasting of GDP growth rates in Switzerland. More specifically, we use the multi-sectoral KOF Barometer that is calculated on the basis of 25 various business survey indicators and is regularly published on the monthly basis (Graff, 2006). The principal use of the KOF Barometer is to provide a snapshot of the current economic situation well ahead of the first official release of the quarterly growth rates of real GDP, typically published two months later after the end of a reference quarter. The reference time series is the real GDP observed at the quarterly frequency released by the Swiss State Secretariat for Economic Affairs (Seco). Our aim is to assess predictive value[1] of the KOF Barometer by comparing predictions of GDP growth rates produced with the model that includes the KOF Barometer against those produced with a benchmark univariate autoregressive model. To this end, we compare accuracy of forecasts made starting as early as six months ahead of the first official publication for a reference quarter. Furthermore, we capitalize on the fact that the KOF Barometer is released at the end of each months and, subsequently, produce the sequence of forecasts that precede the first official release by five, four, three, and two months. Such that, the last forecast is made at the very end of a reference quarter. In addition to verifying the presence of the predictive value of the BTS, this sequential approach to forecasting allows us to address questions like, 1) Do earliest forecasts have any predictive value of GDP growth rates? and 2) How quickly improvement in forecast accuracy takes place as additional information is incorporated in forecasting equation?

[1] According to Okun (1962, p. 218), “A variable has predictive value if it makes a positive contribution to the accuracy of forecasting as an addition to other available information”.

Our study contributes to the literature in the following two ways. First, it is worth mentioning that despite of the widespread use of business tendency surveys in forecasting of either GDP or manufacturing/industrial growth rates (e.g., see Abberger, 2007; Hansson et al., 2005; Lemmens et al., 2005; Balke and Petersen, 2002; Lindström, 2000; Kauppi et al., 1996; Öller and Tallbom, 1996; Bergström, 1995; Markku and Timo, 1993; Öller, 1990; Hanssens and Vanden Abeele, 1987; Teräsvirta, 1986; Zarnowitz, 1973, *inter alia*), in most cases, the forecasts are made using the latest-available data. The importance of using real-time instead of latest-available data has been already emphasized in numerous studies as it has been shown, for example, by Diebold and Rudebusch (1991) and, more recently, by Croushore (2005) that the favorable conclusions on forecasting properties of leading indicator indexes obtained using latest-available data may be substantially weakened or even reversed when forecasting exercise is replicated using real-time data sets. Despite of advantages from using real-time data, their use in assessing the forecasting properties of leading indicator models is still limited as collection of such databases is rather a formidable task. In sum, the question on predictive value of BTS is far from being resolved as there is a rather limited number of studies that address this question in real time. Therefore, additional studies further investigating this question are needed. Hence, the main contribution of our study to the forecasting literature is that we provide an additional empirical piece of work that utilizes the real-time approach in assessing predictive value of business tendency surveys for short-term forecasting of GDP growth rates.

Secondly, we employ the Bayesian model averaging framework instead of relying on a single-best model approach based either on minimization of some information criteria or a more sophisticated model selection procedures, like PcGets advocated in Hendry and Krolzig (2001), that is still a rather standard practice while forecasting with leading indicator models, e.g., see a seminal study of Stock and Watson (2002) or a more recent study such as Golinelli and Parigi (2008). Advantages of Bayesian model averaging are well documented in practice (e.g., see Hoeting, Raftery, and Volinsky, 1999). In forecasting context, such an approach allows us to incorporate the following three types of uncertainty in the models forecasts: error term uncertainty, parameter uncertainty, and model selection uncertainty. Observe that predictions based on a single model typically accommodate only the first and, at best, the second sources of uncertainty. At the same time, the third type of uncertainty is typically ignored in a single-best model approach. However, we believe that accounting for model selection uncertainty is especially important when dealing with real-time data vintages that often undergo (substantial) revisions inducing both changes in temporal dependence structure of a time series of interest as well as changes in interdependence structure between the variables.

The rest of the paper is structured as follows. Section 2 relates the present paper to earlier research on forecasting the Swiss GDP using the tendency surveys. Section 3 describes the data used in our predictive exercise. The econometric model utilized in our study is described in Section 4. Section 5 discusses results of out-of-sample predictions. The final section concludes.

2 Literature review

In Switzerland, Business Tendency Surveys are collected at the KOF Swiss Economic Institute at the Swiss Federal Institute of Technology (ETH), Zurich. Consequently, most of the research involving BTS has been done at KOF. An interested reader may consult the following studies: Jacobs and Sturm (2008), Köberl and Lein (2008), Müller and Köberl (2008b), Müller, Wirz, and Sydow (2008), Rupprecht (2008), Schenker (2008), Graff and Etter (2004), and Etter and Graff (2003). However, there are only two studies—Graff (2009) and Müller and Köberl (2008a)—that are directly related to our study as they evaluate predictive value of business tendency surveys for Swiss GDP.

At KOF, assessing of the current economic situation with tendency surveys has a long tradition. The first version of the KOF Barometer was developed in 1976 and its slightly modified version in 1998 has been published until March 2006. Since April 2006, the traditional KOF Barometer has been substituted with the new KOF Barometer based on the multi-sectoral design (Graff, 2006). Graff (2009) compares 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 (2009) is that a distinction between real-time and latest-available data is clearly made in construction and using the constructed barometer in out-of-sample forecasting. However, while coming close to simulating forecasting exercise in real time, Graff (2009) utilizes for forecast comparison the latest-available figures for the reference time series of real GDP as they were known in 2006Q3. This fact may somewhat bias the reported results when compared with those that could have been obtained in a genuine real-time exercise; i.e., when real-time vintages for both time series of a leading indicator and a reference time series are utilized. Graff (2009) reports a significant improvement in forecast accuracy of the new 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 have been pre-selected using the information for the whole forecast period that was not available to a forecaster had he made his predictions in real time.

Müller and Köberl (2008a) suggest a novel approach to using BTS for forecasting of GDP growth rates that is based on semantic cross validation analysis of firms' answers to BTS questionnaires. The main feature of the approach of Müller and Köberl (2008a) is that the constructed indicator is available in real time, undergoes no revisions, and it is based on a single indicator rather than on pooling information from several indicators as done in case of the KOF Barometer. Müller and Köberl (2008a) present the results of an out-of-sample forecasting exercise suggesting that their approach to constructing a leading indicator is useful for out-of-sample forecasting of GDP growth rates, but, again, the latest-available GDP data have been used in evaluating the predictive value of this semantic indicator. Nevertheless, it must be added that the semantic approach to GDP forecasting is an ongoing endeavor and at present real-time forecasts are regularly released every quarter since 2007Q4. Due to the fact that Müller and Köberl (2008a) suggest a rather different way to construct a leading indicator we view their approach to GDP forecasting complementary to ours rather than substitutive. Future research will shed more light on comparative advantages of these two approaches, provided that there will be a sufficient amount of real-time forecasts.

In sum, while we address the same question as in Graff (2009) and Müller and Köberl (2008a) our study distinguishes itself from those two papers at least in two important aspects. First of all, 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 rates. This also means that the composition of the KOF Barometer has not been subject to pre-selection using information for the whole forecast period that was not available in real time. Secondly, Graff (2009) and Müller and Köberl (2008a) utilize a single-best model approach in forecasting of GDP growth rates, whereas we employ a Bayesian model averaging framework allowing us to take into account two additional sources of uncertainty omitted from either of these two studies: parameter estimation as well as, more importantly, model selection uncertainties.

3 Data

The reference time series is the real GDP observed at the quarterly frequency released by the Swiss State Secretariat for Economic Affairs (Seco)[code: TS41808000] being forecast with the KOF Barometer [code: TS12130800]. Both time series were downloaded from the KOF Database. We conduct the exercise in real time. For this purpose, we employ the vintages of the KOF Barometer starting with the earliest vintage released in April 2006. This implies that we can use the KOF Barometer for prediction of GDP growth rates starting with the nowcast for the second quarter of 2006. We end our forecasting exercise in 2008Q3; i.e., the latest quarter for which the data has been officially released to date. Since we aim to predict the GDP growth rates released at the first official publication, we employ the real-time dataset of all GDP releases starting with the second quarter of 2006.

4 Model

Since the Seco releases GDP figures in the beginning of the third month in each quarter; i.e., two months later after the end of the reference quarter, and since the KOF Barometer is published at the end of every month, we have opted for the following forecast timing setup, see Table 1. Table 1 illustrates our sequential approach to making forecasts of GDP growth rates subject to availability of both KOF Barometer and of GDP figures in real time. Our first GDP forecast for the target quarter $\tau+1$ is made in the beginning of the third month of the previous quarter τ when the values of the KOF Barometer are available for the first two months of the current quarter τ and a GDP figure is available for the previous quarter $\tau-1$. The dark-gray color correspondingly illustrates for which months and quarter(s) both the barometer and the GDP values are available at this forecast round. Observe that the third month in the quarter τ and all three months in the next quarter $\tau+1$ are colored in light-gray, indicating the missing end point problem. We solve it by shifting the whole time series of the KOF Barometer four months forward. Similarly, we make the second and the third forecasts when our information set has been increased by the values of the KOF Barometer for the third month of the quarter τ and for the first months of the quarter $\tau+1$, respectively. Observe that the fourth and the final fifth forecasts are made when information set increases not also because of the values of the KOF Barometer for the second and the third months of the quarter $\tau+1$ but also due to newly published GDP figures for the quarter τ. In sum, we produce the sequence of five forecasts for every quarter (subject to data availability at the end points of our sample). This means that our first forecast precedes the first official release of GDP data by six months and our last forecast—by two months. Note that our forecasting setup allows us to address the following questions:

1) Do earliest forecasts have any predictive value of GDP growth rates?

and

2) How quickly improvement in forecast accuracy takes place as additional information is incorporated in forecasting equation?

The model we use is the autoregressive distributed lag model in the following form:

$$Y_\tau = \alpha_0 + \sum_{i=i^\star}^{p} \alpha_i Y_{\tau-p} + \sum_{j=0}^{q} \beta_j X_{\tau-q} + \varepsilon_\tau \tag{1}$$

where Y_τ is the year-to-year quarterly growth rates of real GDP observed in quarter τ. We calculate Y_τ by taking the seasonal difference of the logarithmic transformation of the reference time series. X_τ is an appropriate quarterly aggregation of monthly values of the KOF Barometer $X_{\tau,t}$ for $t = 1, 2, 3$. Specifically, we first solve the missing end point problem by shifting forward the whole monthly time series of the KOF Barometer. Then we aggregate them using a simple average of the respective values observed in each calendar quarter. Observe that the $i^\star$ takes values of two for the first, second, and third forecast rounds and it takes value of one for the fourth and fifth forecast rounds, reflecting the availability of GDP data for the respective forecast rounds. ε_τ is a disturbance term satisfying usual model assumptions.

In general, an ARDL equation allows 2^k combinations of regressors, where k is the number of regressors except the constant term, which is always retained in estimation. Given such a multitude of equation specifications, we chose to conduct our exercise using the Bayesian model averaging (BMA) approach, rather than concentrating on a 'single-best' model approach. The BMA approach allows us to incorporate three following sources of uncertainty while making now- and forecasts: error term uncertainty, parameter uncertainty, and model selection uncertainty. Observe that predictions based on a single-model approach typically accommodate only the first and, at best, the second sources of uncertainty. Assessment of model uncertainty and, henceforth, its incorporation in the prediction process, *per definition*, is ruled out in the latter approach. The equation parameters have been estimated using the Monte Carlo Markov Chain simulation algorithm, which allows us easily to produce the finite-sample predictive densities, rather than those based on the asymptotic approximation. On the basis of these predictive densities, the point- as well as the interval forecasts of GDP growth rates can be readily calculated.

The BMA approach allows us to consider either all possible combinations of the regressors in our predictive exercise or to concentrate out a subset of the most likely models. According to the former approach, for model comparison one has to evaluate posterior probabilities for all the possible combinations of lags of Y and X. This may require a significant computational time. To get around this, we followed Madigan and Raftery (1994) and applied an approach of model selection based on Occam's window. According to this approach we exclude "(a) models that are much less likely than the most likely model-say 20 times less likely, corresponding to a BIC (or BIC') difference of 6; and (optionally) (b) models containing effects for which there is no evidence-that is, models that have more likely submodels nested within them. The models that are left are said to belong to Occam's window, a generalization of the famous Occam's razor, or principle of parsimony in scientific explanation. When both (a) and (b) are used, Occam's window is said to be strict, and when only (a) is used it is said to be symmetric" (Raftery, 1995, p. 146). One can adjust the severity of model selection procedure by changing ratio in (a), and/or apply a strict rather than symmetric Occam's window.

5 Results

In this section we present our results which are summarized in Tables 2 and 3 for ARDL and AR models, respectively, and in Table 4 for both models. Observe that Tables 2 and 3 contain detailed forecast results whereas Table 4 provides a summary of forecast accuracy of these two models. We start discussion of the results that are reported in Tables 2 and 3, followed by comparison of forecasting accuracy of both competing models that is based on Table 4.

Table 2 presents detailed forecast results obtained by the ARDL model. The table is arranged in such a way that results are grouped according to forecast rounds. Recall that the first forecast round for quarter $\tau+1$ starts starts in the beginning of the third month of quarter τ when both the value of the KOF Barometer for the second months of this quarter and the first release of GDP figures for the previous quarter $\tau-1$ are available (see Table 1). This means that the first forecast round precedes the first official publication of GDP by six months. The second forecast round is carried out in one months when the value of the KOF Barometer is known. Naturally, the second forecast precedes the first official publication of GDP by five months. Continuing, the final fifth forecast for quarter $\tau+1$ is made when the value of the KOF Barometer for the third month of this quarter is known. This means that our final forecast is made two months in advance of GDP release by Seco. In this way, we can follow the evolution of accuracy of point forecasts as well as of related uncertainty around these point forecasts, captured by 95% predictive interval, as additional information in the form of new values of the KOF Barometer is incorporated in our forecasting equation. Note that in the fourth forecast round, our information set is increased by released values of GDP for a previous quarter. Increase in information set is illustrated by the content of columns *Period* τ, *Period* $\tau^\star$, and *Period* t where we report for each forecasting round a quarter for which forecast is made, a latest quarter for which value of GDP is available, and a latest month for which value of the KOF Barometer is available, respectively.

Based on Tables 2, 3, and 4 a number of interesting observations can be made. First, as seen in Table 2 point forecasts generated by the ARDL model in general are associated with a quite high degree of uncertainty as reflected in rather wide 95% predictive intervals. Thus, for the first three forecast rounds a typical width of 95% predictive intervals is slightly less than four percentage points, whereas for the fourth and the fifth forecast rounds—it is about 2.5 percentage points. A rather sharp decrease in the width of 95% predictive intervals in the fourth forecast round indicates that the largest marginal decrease in forecast uncertainty takes place in this round. This is due to the fact that in this forecast round GDP figures for the previous quarter are incorporated into the information set when forecast is made.

Second, comparing the width of the 95% predictive intervals of the ARDL and the AR models, it is easy to verify that those of the latter model are generally wider than those of the former model. This implies that incorporation of the KOF Barometer to the information set signficantly narrows predictive intervals and therefore reduces reported uncertainty around the point forecasts. We interpret it as the first sign that the KOF Barometer possesses at least some degree of predictive value for GDP growth rates.

Third, the chosen Bayesian Model Averaging framework allows us to analyze how does the model selection uncertainty evolves within a certain forecast round as well as from one to another forecasting round. The relevant information is reported in column $Model$, which contains number of model selected in Occam's window. The columns *Posterior probability* report associated maximum and minimum posterior probabilities of models selected in Occam's window. In Table 2, for the first two forecast rounds we observe only slight, if any, evidence in favor of reduction in the model selection uncertainty as the estimation period increases, as both the number of models selected as well as reported maximum posterior probability take similar values in the end and in the beginning of each of these two forecast rounds. Only in the third forecast round we observe some signs of a decrease in model selection uncertainty—a decrease in number of models selected in Occam's window and an increase in maximum posterior probability—as the size of estimation window increases. Finally, for the last two forecasting rounds an associated decrease in model selection uncertainty is much more pronounced. For example, for the fourth forecast round the number of models selected in Occam's window when making forecast for 2006(2) is 45. This number has decreased to 8 when making forecast for 2008(4); the last observation in our forecast sample. The corresponding maximum posterior probability has gone up from 0.112 till 0.456. A similar conclusion on a substantial reduction in model selection uncertainty within the fifth forecast round in response to expanding estimation window can easily be reached. When we compare reduction of model selection uncertainty from one forecast round to another, we notice again that the largest marginal decrease happens in the fourth forecast round. This results reinforces our earlier conclusion on importance of accounting for newly released GDP data for a previous quarter when making forecast for a current quarter; i.e., so-called nowcast.

Fourth, the earlier conclusion on importance of newly released GDP in the information set is also supported by the results reported in Table 3 for the univariate AR model. Here, again, in the fourth forecast round we notice a substantial reduction both in the width of the 95% predictive interval as well as in the associated model selection uncertainty, measured by the number of models selected in Occam's window as well as by reported maximum posterior probability for models in Occam's window.

Fifth, Table 4 reports two summary measures of forecast accuracy of the competing models: Root Mean Squared Forecast Error (RMSFE) and Mean Absolute Forecast Error (MAFE). The first encouraging observation is that both RMSFE and MAFE steadily decrease for the ARDL model with every forecast round as more information contained in the latest releases of the KOF Barometer (and release of GDP data at the fourth forecast round) are incorporated into forecasting equation. From the first till the fifth forecast round RMSFE and MAFE decrease from 0.620 till 0.347 and from 0.565 till 0.272, respectively.

This corresponds to drop of about 44% and 52% in RMSFE and MAFE criteria, respectively. Observe that the largest drop in RMSFE and MAFE takes place during the fourth forecast round when the first GDP estimate for the previous quarter is released. This finding conforms well with the results discussed above on the largest marginal reduction both in the width of 95% predictive interval as well as in model selection uncertainty. The importance of incorporation of the GDP figures for the previos quarter into forecasting equation is also supported by substantial increase in forecast accuracy of the univariate AR model. As seen, at the forth forecast round the corresponding RMSFE and MAFE criteria drop from 0.860 till 0.497 and from 0.727 till 0.457, or by about 42% and 37%. More importantly, for every forecast round either RMSFE or MAFE reported for the ARDL model are consistently lower than those reported for the benchmark AR model. Even for the earliest forecast round (made six-months ahead of the first official publication) the corresponding ARDL-to-AR RMSFE and MAFE ratios are 0.72 and 0.78. This indicates that even at the earliest forecast round the KOF Barometer substantially contributes to accuracy of GDP forecasts. At the fifth forecast round—made two months ahead of the first official release of GDP data—RMSFE and MAFE values of the ARDL model are smaller than those of the AR model by impressing 30% and 41%, respectively. This finding allows us to conclude that the information contained in the Swiss Business Tendency Surveys has a definite predictive value for GDP growth rates.

Last but not least, at the first three earlier forecast rounds the univariate AR model tend to systematically underpredict GDP growth rate, whereas this seems not to be the case for the corresponding forecasts of the ARDL model: compare mean forecast error of 0.499 for the AR model with that for the ARDL model, for example, for the first forecast round—0.167. Also the maximum forecast error of 1.545 observed for the AR model is larger than that observed for the ARDL model during the first three forecast rounds—0.951. During the last two forecast rounds both maximum forecast error and mean forecast error appear to be of a similar magnitude for both competing models.

6 Conclusion

In this paper we investigate whether business tendency surveys collected for Switzerland can be useful for short-term out-of-sample prediction of year-on-year quarterly real GDP growth rates. The forecasts for a reference quarter are continuously made each month starting with the first forecast made about six months ahead of GDP release by the Swiss State Secretariat for Economic Affairs (Seco), followed by the second forecast that precedes GDP release by five months, etc., till the final fifth forecast made about two months ahead of first official GDP estimate. At each forecast round a new BTS information on the current developments of the Swiss economy and its short-term prospects in the form of monthly releases of the KOF Barometer is incorporated into the forecasting equation. Observe that at the fourth forecast round in addition to the update of the KOF Barometer a newly released GDP data for the previous quarter is also incorporated into the forecasting equation.

Our main findings are as follows. First, BTS appear to be useful for prediction of GDP growth rates in Switzerland. Even the earliest forecasts, made six months ahead of first official GDP estimate, allow us to predict GDP growth rates more accurately than forecasts based on univariate autoregressive model.

Second, at every subsequent forecast round we observe increase in forecast accuracy as reflected in steadily declining values of RMSFE and MAFE criteria. The corresponding ratios of RMSFE and MAFE criteria observed for the ARDL model to those of the benchmark AR model are 0.72 and 0.78 for the first forecast round and 0.70 and 0.59—for the final fifth forecast round. It is worth mentioning that the largest incremental increase in forecast accuracy is observed in the fourth forecast round; i.e., when GDP data for previous quarter are released by Seco. At this forecast round we also observe a sharp decline both in the associated uncertainly surrounding point forecasts, measured by the width of a 95% predictive interval, as well as in model selection uncertainty, measured by a number of models selected in Occam's window and by model maximum posterior probability. At the other forecast rounds, associated forecast accuracy increase and model uncertaintly reduction, while still non-negligible, are less pronounced.

Third, based on the values of MAFE criterion, we expect an average forecast error of about 0.5 percentage points for forecasts made from six to four months ahead of the first Seco release and of about 0.3 percentage points for forecasts made three and two months ahead of the first official GDP release.

References

Abberger, K. (2007). Forecasting quarter-on-quarter changes of German GDP with monthly business tendency survey results. Ifo Working paper 40, Ifo Institute for Economic Research at the University of Munich.

Balke, N. S. and D. Petersen (2002). How well does the Beige Book reflect economic activity? Evaluating qualitative information quantitatively. *Journal of Money, Credit and Banking 34*(1), 114–36.

Bergström, R. (1995). The relationship between manufacturing production and different business survey series in Sweden 1968-1992. *International Journal of Forecasting 11*(3), 379–393.

Croushore, D. (2005). Do consumer-confidence indexes help forecast consumer spending in real time? *The North American Journal of Economics and Finance 16*(3), 435 – 450.

Diebold, F. X. and G. D. Rudebusch (1991). Forecasting output with the composite leading index : A real-time analysis. *Journal of the American Statistical Association 86*(415), 603–610.

Etter, R. and M. Graff (2003). Estimating and forecasting production and orders in manufacturing industry from business survey data: Evidence from Switzerland, 1990-2003. *Swiss Journal of Economics and Statistics 139*(4), 507553.

Golinelli, R. and G. Parigi (2008). Real-time squared: A real-time data set for real-time GDP forecasting. *International Journal of Forecasting 24*(3), 368–385.

Graff, M. (2006). Ein multisektoraler Sammelindikator für die Schweizer Konjunktur. *Schweizerische Zeitschrift für Volkswirtschaft und Statistik 142*, 529577.

Graff, M. (2009). Does a multi-sectoral design improve indicator-based forecasts of the GDP growth rate? evidence for Switzerland. *Applied Economics*, forthcoming.

Graff, M. and R. Etter (2004). Coincident and leading indicators of manufacturing industry: Sales, production, orders and inventories in Switzerland. *Journal of Business Cycle Measurement and Analysis 1*(1), 109–113.

Hanssens, D. M. and P. M. Vanden Abeele (1987). A time-series study of the formation and predictive performance of EEC production survey expectations. *Journal of Business & Economic Statistics 5*(4), 507–19.

Hansson, J., P. Jansson, and M. Löf (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting 21*(2), 377–389.

Hendry, D. F. and H.-M. Krolzig (2001). *Automatic Econometric Model Selection Using PcGets*. London: Timberlake Consultants Ltd.

Hoeting, J., D. M. A. Raftery, and C. Volinsky (1999). Bayesian model averaging: A tutorial. *Statistical Science 14*, 382–401.

Jacobs, J. P. and J.-E. Sturm (2008). The information content of KOF indicators on Swiss current account data revisions. Working papers 08-202, KOF Swiss Economic Institute, ETH Zurich.

Kauppi, E., J. Lassila, and T. Terasvirta (1996). Short-term forecasting of industrial production with business survey data: Experience from Finland's great depression 1990-1993. *International Journal of Forecasting 12*(3), 373–381.

Köberl, E. and S. M. Lein (2008). The NAICU and the Phillips Curve - An approach based on micro data. Working paper 08-211, KOF Swiss Economic Institute, ETH Zurich.

Lemmens, A., C. Croux, and M. G. Dekimpe (2005). On the predictive content of production surveys: A pan-European study. *International Journal of Forecasting 21*(2), 363–375.

Lindström, T. (2000). Qualitative survey responses and production over the business cycle. Working Paper Series 116, Sveriges Riksbank (Central Bank of Sweden).

Madigan, D. and A. E. Raftery (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association 89*, 1535–1546.

Markku, R. and T. Timo (1993). Business survey data in forecasting the output of Swedish and Finnish metal and engineering industries: A Kalman filter approach. *Journal of Forecasting 12*(3-4), 255–271.

Müller, C. and E. Köberl (2008a). Business cycle measurement: A semantic identification approach using firm level data. Working paper 08-212, KOF Swiss Economic Institute, ETH Zurich.

Müller, C. and E. Köberl (2008b). The speed of adjustment to demand shocks: A Markov-chain measurement using micro panel data. Working paper 07-170, KOF Swiss Economic Institute, ETH Zurich.

Müller, C., A. Wirz, and N. Sydow (2008). A note on the Carlson-Parkin method of quantifying qualitative data. Working paper 07-168, KOF Swiss Economic Institute, ETH Zurich.

Okun, A. M. (1962). The predictive value of surveys of business intentions. *The American Economic Review 52*(2), 218–225.

Öller, L.-E. (1990). Forecasting the business cycle using survey data. *International Journal of Forecasting 6*(4), 453–461.

Öller, L.-E. and C. Tallbom (1996). Smooth and timely business cycle indicators for noisy Swedish data. *International Journal of Forecasting 12*(3), 389–402.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology 25*, 111–163.

Rupprecht, S. M. (2008). When do firms adjust prices? Evidence from micro panel data. Working paper 07-160, KOF Swiss Economic Institute, ETH Zurich.

Schenker, R. (2008). Comparing quantitative and qualitative survey data. Working paper 07-169, KOF Swiss Economic Institute, ETH Zurich.

Stock, J. H. and M. W. Watson (1988). A probability model of the coincident economic indicators. NBER Working Papers 2772, National Bureau of Economic Research, Inc.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*(2), 147–162.

Teräsvirta, T. (1986). Model selection using business survey data: Forecasting the output of the Finnish metal and engineering industries. *International Journal of Forecasting 2*(2), 191 – 200.

Zarnowitz, V. (1973). A review of cyclical indicators for the United States: Preliminary results. NBER Working Papers 0006, National Bureau of Economic Research, Inc.

Table 1: Forecast timing setup

Forecast round	Quarter / Month	$\tau - 1$ I	$\tau - 1$ II	$\tau - 1$ III	τ I	τ II	τ III	$\tau + 1$ I	$\tau + 1$ II	$\tau + 1$ III
1	KOBA									
	GDP									
2	KOBA									
	GDP									
3	KOBA									
	GDP									
4	KOBA									
	GDP									
5	KOBA									
	GDP									

Notes: Table describes the sequence of forecasts for the target quarter $\tau + 1$. The first forecast is made in the beginning of the third month of the quarter τ, when values of the KOF Barometer are available for the second month of the quarter τ. At this time also the GDP data are released for the quarter $\tau - 1$. The dark-gray color in the table indicates that both the KOF Barometer and the GDP are available up to the corresponding month and the quarter. The next forecast takes place in the beginning of the first month of the quarter $\tau + 1$, when the value of the KOF Barometer is already known for the last month of the quarter τ, etc. The light-gray color indicates the missing end points that are filled with values of the KOF Barometer time series by shifting it forward.

Table 2: Forecast results: ARDL model

Forecast round	Period[a] τ	Period[b] $\tau^\star$	Period[c] t	Y_τ	$\widehat{Y}_\tau$	$Y_\tau - \widehat{Y}_\tau$	95% Predictive interval	Models[d]	Posterior probability max	min
1	2006(2)	2005⟨4⟩	2006{ 2}	3.107	na	na	na	na	na	na
	2006(3)	2006⟨1⟩	2006{ 5}	2.343	2.967	-0.625	[0.773,5.067]	63	0.093	0.005
	2006(4)	2006⟨2⟩	2006{ 8}	2.196	2.872	-0.675	[0.849,4.922]	56	0.121	0.006
	2007(1)	2006⟨3⟩	2006{11}	2.404	1.696	0.708	[-0.418,3.846]	42	0.156	0.008
	2007(2)	2006⟨4⟩	2007{ 2}	2.754	1.888	0.866	[-0.123,3.907]	42	0.150	0.008
	2007(3)	2007⟨1⟩	2007{ 5}	2.846	2.332	0.514	[0.289,4.337]	35	0.176	0.010
	2007(4)	2007⟨2⟩	2007{ 8}	3.583	2.632	0.951	[0.571,4.656]	56	0.104	0.005
	2008(1)	2007⟨3⟩	2007{11}	2.987	2.734	0.253	[0.671,4.768]	53	0.107	0.005
	2008(2)	2007⟨4⟩	2008{ 2}	2.314	2.618	-0.304	[0.646,4.565]	49	0.102	0.005
	2008(3)	2008⟨1⟩	2008{ 5}	1.582	1.767	-0.185	[-0.134,3.667]	47	0.109	0.006
	2008(4)	2008⟨2⟩	2008{ 8}	na	1.163	na	[-0.834,3.178]	45	0.140	0.007
	2009(1)	2008⟨3⟩	2008{11}	na	0.482	na	[-1.557,2.522]	43	0.134	0.007
2	2006(2)	2005⟨4⟩	2006{ 3}	3.107	na	na	na	na	na	na
	2006(3)	2006⟨1⟩	2006{ 6}	2.343	3.093	-0.751	[0.987,5.181]	60	0.096	0.005
	2006(4)	2006⟨2⟩	2006{ 9}	2.196	2.824	-0.627	[0.875,4.778]	56	0.108	0.006
	2007(1)	2006⟨3⟩	2006{12}	2.404	1.849	0.555	[-0.145,3.856]	44	0.137	0.007
	2007(2)	2006⟨4⟩	2007{ 3}	2.754	2.114	0.640	[0.170,4.095]	47	0.128	0.006
	2007(3)	2007⟨1⟩	2007{ 6}	2.846	2.354	0.492	[0.435,4.265]	45	0.138	0.007
	2007(4)	2007⟨2⟩	2007{ 9}	3.583	2.750	0.833	[0.756,4.746]	67	0.067	0.004
	2008(1)	2007⟨3⟩	2007{12}	2.987	2.885	0.102	[0.815,4.932]	67	0.072	0.004
	2008(2)	2007⟨4⟩	2008{ 3}	2.314	2.596	-0.282	[0.661,4.506]	59	0.079	0.004
	2008(3)	2008⟨1⟩	2008{ 6}	1.582	1.756	-0.175	[-0.108,3.627]	55	0.084	0.005
	2008(4)	2008⟨2⟩	2008{ 9}	na	1.094	na	[-0.835,3.053]	49	0.110	0.006
	2009(1)	2008⟨3⟩	2008{12}	na	0.136	na	[-1.941,2.177]	54	0.099	0.005
3	2006(2)	2005⟨4⟩	2006{ 4}	3.107	2.271	0.836	[0.402,4.119]	38	0.119	0.006
	2006(3)	2006⟨1⟩	2006{ 7}	2.343	3.140	-0.797	[1.219,5.057]	45	0.127	0.006
	2006(4)	2006⟨2⟩	2006{10}	2.196	2.680	-0.483	[0.782,4.571]	43	0.125	0.006
	2007(1)	2006⟨3⟩	2007{ 1}	2.404	2.046	0.358	[0.134,3.940]	36	0.149	0.008
	2007(2)	2006⟨4⟩	2007{ 4}	2.754	2.145	0.609	[0.266,4.021]	37	0.144	0.008
	2007(3)	2007⟨1⟩	2007{ 7}	2.846	2.521	0.324	[0.673,4.384]	33	0.162	0.008
	2007(4)	2007⟨2⟩	2007{10}	3.583	2.727	0.856	[0.753,4.680]	42	0.138	0.007
	2008(1)	2007⟨3⟩	2008{ 1}	2.987	2.581	0.405	[0.616,4.525]	42	0.138	0.007
	2008(2)	2007⟨4⟩	2008{ 4}	2.314	2.338	-0.024	[0.476,4.174]	33	0.184	0.010
	2008(3)	2008⟨1⟩	2008{ 7}	1.582	1.769	-0.187	[-0.071,3.603]	32	0.196	0.010
	2008(4)	2008⟨2⟩	2008{10}	na	0.878	na	[-1.011,2.794]	22	0.305	0.016
	2009(1)	2008⟨3⟩	2009{ 1}	na	-0.122	na	[-2.137,1.846]	23	0.294	0.015
4	2006(2)	2006⟨1⟩	2006{ 5}	3.107	3.234	-0.127	[1.756,4.697]	45	0.122	0.006
	2006(3)	2006⟨2⟩	2006{ 8}	2.343	3.127	-0.785	[1.691,4.571]	50	0.104	0.005
	2006(4)	2006⟨3⟩	2006{11}	2.196	2.142	0.054	[0.688,3.601]	42	0.135	0.007
	2007(1)	2006⟨4⟩	2007{ 2}	2.404	2.041	0.363	[0.617,3.470]	42	0.134	0.007
	2007(2)	2007⟨1⟩	2007{ 5}	2.754	2.538	0.216	[1.108,3.961]	40	0.145	0.008
	2007(3)	2007⟨2⟩	2007{ 8}	2.846	2.851	-0.006	[1.458,4.244]	19	0.258	0.016
	2007(4)	2007⟨3⟩	2007{11}	3.583	3.107	0.476	[1.727,4.498]	19	0.256	0.016
	2008(1)	2007⟨4⟩	2008{ 2}	2.987	3.437	-0.450	[2.109,4.772]	16	0.298	0.016
	2008(2)	2008⟨1⟩	2008{ 5}	2.314	2.349	-0.035	[1.031,3.674]	16	0.295	0.017
	2008(3)	2008⟨2⟩	2008{ 8}	1.582	1.977	-0.395	[0.601,3.348]	8	0.453	0.060
	2008(4)	2008⟨3⟩	2008{11}	na	0.629	na	[-0.753,1.997]	8	0.456	0.061
5	2006(2)	2006⟨1⟩	2006{ 6}	3.107	3.269	-0.162	[1.847,4.670]	21	0.282	0.014
	2006(3)	2006⟨2⟩	2006{ 9}	2.343	3.076	-0.733	[1.668,4.494]	26	0.259	0.014
	2006(4)	2006⟨3⟩	2006{12}	2.196	2.166	0.031	[0.761,3.556]	19	0.308	0.016
	2007(1)	2006⟨4⟩	2007{ 3}	2.404	2.144	0.260	[0.765,3.522]	17	0.323	0.016
	2007(2)	2007⟨1⟩	2007{ 6}	2.754	2.573	0.180	[1.203,3.935]	15	0.345	0.018
	2007(3)	2007⟨2⟩	2007{ 9}	2.846	2.907	-0.061	[1.545,4.279]	14	0.341	0.017
	2007(4)	2007⟨3⟩	2007{12}	3.583	3.140	0.443	[1.784,4.504]	14	0.340	0.017
	2008(1)	2007⟨4⟩	2008{ 3}	2.987	3.437	-0.451	[2.125,4.756]	8	0.463	0.063
	2008(2)	2008⟨1⟩	2008{ 6}	2.314	2.341	-0.026	[1.043,3.635]	8	0.463	0.063
	2008(3)	2008⟨2⟩	2008{ 9}	1.582	1.950	-0.368	[0.613,3.290]	8	0.497	0.065
	2008(4)	2008⟨3⟩	2008{12}	na	0.456	na	[-0.920,1.815]	8	0.503	0.065

Notes:
[a] Period for which forecast is made.
[b] Latest period for which GDP data is available.
[c] Latest period for which value of the KOF Barometer is available.
[d] Indicates number of models selected in Occam's window.

Table 3: Forecast results: AR model

Forecast round	Period[a] τ	Period[b] $\tau^\star$	Y_τ	$\widehat{Y}_\tau$	$Y_\tau - \widehat{Y}_\tau$	95% Predictive interval	Models[c]	Posterior probability max	Posterior probability min
1, 2, 3	2006(2)	2005⟨4⟩	3.107	1.969	1.138	[-0.519,4.469]	8	0.296	0.019
	2006(3)	2006⟨1⟩	2.343	2.301	0.042	[-0.219,4.799]	8	0.320	0.017
	2006(4)	2006⟨2⟩	2.196	2.060	0.137	[-0.408,4.540]	7	0.355	0.033
	2007(1)	2006⟨3⟩	2.404	1.574	0.830	[-0.830,4.000]	7	0.358	0.033
	2007(2)	2006⟨4⟩	2.754	1.589	1.165	[-0.817,4.012]	7	0.359	0.033
	2007(3)	2007⟨1⟩	2.846	1.863	0.983	[-0.523,4.217]	7	0.343	0.032
	2007(4)	2007⟨2⟩	3.583	2.038	1.545	[-0.386,4.475]	6	0.519	0.043
	2008(1)	2007⟨3⟩	2.987	2.184	0.803	[-0.224,4.629]	6	0.487	0.045
	2008(2)	2007⟨4⟩	2.314	2.669	-0.355	[0.273,5.087]	8	0.382	0.020
	2008(3)	2008⟨1⟩	1.582	2.261	-0.679	[-0.121,4.657]	8	0.372	0.020
	2008(4)	2008⟨2⟩	na	1.760	na	[-0.682,4.229]	8	0.300	0.024
	2009(1)	2008⟨3⟩	na	1.458	na	[-0.977,3.885]	8	0.306	0.023
4, 5	2006(2)	2006⟨1⟩	3.107	2.951	0.157	[1.167,4.740]	13	0.279	0.016
	2006(3)	2006⟨2⟩	2.343	2.583	-0.240	[0.847,4.350]	13	0.253	0.017
	2006(4)	2006⟨3⟩	2.196	1.914	0.282	[0.203,3.611]	13	0.261	0.017
	2007(1)	2006⟨4⟩	2.404	1.850	0.555	[0.148,3.554]	13	0.269	0.016
	2007(2)	2007⟨1⟩	2.754	2.174	0.580	[0.490,3.867]	13	0.265	0.015
	2007(3)	2007⟨2⟩	2.846	2.447	0.398	[0.795,4.104]	4	0.671	0.073
	2007(4)	2007⟨3⟩	3.583	2.820	0.763	[1.181,4.429]	4	0.670	0.074
	2008(1)	2007⟨4⟩	2.987	3.542	-0.555	[1.949,5.106]	3	0.760	0.103
	2008(2)	2008⟨1⟩	2.314	2.636	-0.322	[1.071,4.203]	3	0.764	0.103
	2008(3)	2008⟨2⟩	1.582	2.298	-0.716	[0.685,3.923]	3	0.750	0.101
	2008(4)	2008⟨3⟩	na	1.138	na	[-0.493,2.745]	3	0.745	0.096

Notes:
[a] Period for which forecast is made.
[b] Latest period for which GDP data is available.
[c] Indicates number of models selected in Occam's window.

Table 4: Forecast results: Summary

Forecast round	ARDL RMSFE	ARDL MAFE	ARDL |max|[a]	ARDL mean[b]	AR RMSFE	AR MAFE	AR |max|	AR mean	Obs.	Ratio[c] RMSFE	Ratio[c] MAFE
1	0.620	0.565	0.951	0.167	0.860	0.727	1.545	0.499	9	0.72	0.78
2	0.551	0.495	0.833	0.087	0.860	0.727	1.545	0.499	9	0.64	0.68
3[d]	0.517	0.449	0.856	0.118	0.860	0.727	1.545	0.499	9	0.60	0.62
4	0.374	0.291	0.785	-0.069	0.497	0.457	0.763	0.090	10	0.75	0.64
5	0.347	0.272	0.733	-0.089	0.497	0.457	0.763	0.090	10	0.70	0.59

Notes:
[a] Denotes absolute value of maximum forecast error in the corresponding forecast round.
[b] Denotes mean forecast error.
[c] Denotes ratio of RMSFE and MAFE of the ARDL model to those of the benchmark AR model.
[d] For this forecast round table entries correspond to the forecast sample period 2006(3)—2008(3) in order to make results comparable with those reported for the first and the second forecast rounds.