Forecasting GDP all over the World: Evidence from Comprehensive Survey Data

Garnitz, Johanna and Lehmann, Robert and Wohlrabe, Klaus

ifo Institute, ifo Institute and CESifo, ifo Institute and CESifo

4 October 2017

Forecasting GDP all over the World

Evidence from Comprehensive Survey Data*

Johanna Garnitz Robert Lehmann Klaus Wohlrabe

This version: October 4, 2017

Abstract

Comprehensive and international comparable leading indicators across countries and continents are rare. In this paper, we use a free and fast available source of leading indicators, the World Economic Survey (WES) conducted by the ifo Institute, to forecast growth of Gross Domestic Product (GDP) in 44 countries and three country aggregates separately. We come up with three major results. First, for 35 countries as well as the three aggregates a model containing one of the major WES indicators produces on average lower forecast errors compared to an autoregressive benchmark model. Second, the most important WES indicators are either the economic climate or the expectations on future economic development for the next six months. And last, 70% of all country-specific models contain WES information from at least one of the main trading partners. Thus, by allowing WES indicators from economic important partners to forecast GDP of the country under consideration, increases forecast accuracy.

Keywords: World Economic Survey, Economic Climate, Forecasting GDP
JEL-Classification: E17, E27, E37

*Garnitz: ifo Institute (garnitz@ifo.de). **Lehmann:** ifo Institute and CESifo (lehmann@ifo.de, +49 89/9224-1652); corresponding author. **Wohlrabe:** ifo Institute and CESifo (wohlrabe@ifo.de).

1. Introduction

Macroeconomic projections based on leading indicators is a widely accepted approach when it comes to practical forecasting or by looking at the corresponding scientific literature. Especially survey indicators have often been proved to be very good predictors for the real economy (see, among others, Girardi *et al.*, 2016). Leading indicators, however, crucially differ between countries, which makes a general statement on the usefulness of a specific group of leading indicators between countries nearly impossible. One freely available source of comparable qualitative indicators is the World Economic Survey (WES), conducted by the German ifo Institute. In this paper, we use the main indicators from this survey among economic experts to evaluate their forecasting performance for gross domestic product (GDP) growth in 44 countries and three aggregates.

There are only a few surveys with questionnaires that are comparable across countries. Three examples are the Purchasing Manager Index (PMI) provided by Markit, indicators from the European Commission's Joint Harmonised EU Programme of Business and Consumer Surveys (BCS) and the Composite Leading Indicator (CLI) of the OECD. Whereas the first two are solely business surveys, the CLIs of the OECD are also based on several hard indicators. The PMI covers more than 30 advanced and emerging economies using an identical questionnaire. The BCS ensures harmonized questions across business surveys among almost all European countries. Unfortunately, PMIs are not freely accessible for a large set of countries and the CLIs have a publication lag of two months. The WES, in contrast, is freely available to researchers[1] and covers more than 100 countries. Furthermore, the WES employs comparable questionnaires which allow us to formulate a statement on the WES forecasting performance between countries.

Up to date, a vast literature on country-specific GDP forecasts exists that either focuses on methodological or data issues.[2] A comprehensive study for many countries using identical survey data to forecast national economic activity is, however, missing. One exception is Fichtner *et al.* (2011) who investigate the forecasting properties of the OECD leading indicators for eleven countries. Lehmann (2015) and Lehmann and Weyh (2016) use data from the BCS to forecast export growth or employment growth for various European countries.

Despite the fast and free availability, the WES survey data have only been used by a small number of studies. Henzel and Wollmershäuser (2005) develop a new methodology to elicit inflation expectations from the WES. For 43 countries and two country aggregates, the paper by Kudymowa *et al.* (2013) assesses the in-sample performance of the WES economic climate as a business cycle indicator. They found strong cross-correlations between the WES indicators and country-specific year-on-year growth rates in real GDP. Thus, the

[1] Non-researchers, however, have to pay a small fee to access the data.

[2] See, for example, China: Zhou *et al.* (2013), France: Barhoumi *et al.* (2010), Germany: Drechsel and Scheufele (2012), Greece: Kiriakidis and Kargas (2013), Spain: Pons-Novell (2006), Sweden: Österholm (2014), UK: Barnett *et al.* (2014), US: Banerjee and Marcellino (2006).

climate indicator can be used to assess the state of the economy or even upcoming future economic development. The relevant literature for our purpose, namely the studies that focus on forecasting issues, is also very scarce. For Euro Area real GDP, Hülsewig *et al.* (2008) use three business cycle indicators and ask whether the optimal pooling of nationwide information of these indicators help to increase forecast accuracy of the European aggregate. They find an improvement of their approach over alternative techniques. One of the applied nationwide indicator is the WES economic climate because of its comparability between different countries. Hutson *et al.* (2014) apply the Carlson-Parkin framework and the Pesaran-Timmermann Predictive Failure statistic to several WES indicators for the US economy. As a result, the WES experts provide statistical significant superior directional forecasts for total GDP and sub-components.

Our paper has two major contributions to the literature. First, as there is no comprehensive out-of-sample forecasting study for a large set of countries, this paper evaluates the performance of WES indicators for 44 countries and three country aggregates to forecast national GDP. We use the three major indicators from the WES (the assessment of the current economic situation, the expectations on future economic development for the next six months, and the economic climate) and ask whether one of these indicators has a higher forecast accuracy to a simple autoregressive benchmark. Our second contribution deals with the question whether national GDP forecasts can be improved by additionally using the WES survey results from the country-specific most important trading partners. Since business cycle synchronization between countries rises the higher their trading intensity is (Inklaar *et al.*, 2008; Duval *et al.*, 2016), one can suggest that country-specific forecast accuracy of GDP can be increased by adding WES indicators from economically important countries. Our results show that forecasting models based on WES indicators have a higher forecast accuracy compared to the benchmark for 35 out of our 44 countries as well as the three aggregates. Only for a small number of countries, the WES indicators cannot improve GDP forecasts. Additionally, 70% of the best performing indicator models contain WES information of the main trading partners. Thus, relying on economic signals from economic important countries to the home country leads to a higher forecast accuracy in most of the cases.

The remainder of the paper is organized as follows: Section 2 briefly describes the data set and the WES. The forecasting approach is introduced in Section 3. In Section 4, we present the results. We end by offering some conclusions in Section 5.

2. Data

2.1. Countries and Target Series to Forecast

Forecasting gross domestic product (GDP) all over the world requires a large sample of countries. We build our exercise on 44 single countries and three additional aggregates (the European Union, the Eurozone and the World). This sample comprises emerging countries such as Argentina or Brazil as well as highly developed countries such as Norway or the United States of America. The country selection is driven by both the availability of a long quarterly GDP series and a sufficient number of respondents in the WES. Table 2 in the Appendix lists all countries and aggregates in our sample.

As the target variable, we use GDP as the main indicator to measure economic activity. With the exceptions of China and the World[3], for which we only have nominal figures, we can rely on GDP in real terms. Most of the GDP figures are already provided as seasonally adjusted series; for China, Hong Kong, Thailand, and Uruguay we manually adjusted the series with standard parameters of Census X-12-ARIMA. All GDP series are transformed into quarter-on-quarter growth rates after the seasonal adjustment. Since official statistics have developed differently in various countries, the length of the GDP series differ between the countries in our sample. The earliest starting point in our sample is Q1-1989 (for example, Canada).[4] For Uruguay, we observe the shortest GDP series (first quarterly growth rate for Q1-2005). Unfortunately, we cannot rely on real-time GDP data. To the best of our knowledge, a real-time database for such a large number of countries is not available. We thus decided to be consistent over the whole set of countries by using the latest available GDP figures. Table 2 in the Appendix also shows the starting points for all country GDP figures, along with the source from which we obtained the data.

2.2. ifo World Economic Survey

The ifo World Economic Survey (WES) is one of the standard surveys provided by the ifo Institute in Munich (Becker and Wohlrabe, 2008). Its aim is to detect worldwide economic trends. To this end, the ifo Institute in Munich currently polls over 1,000 economists worldwide from international and national organizations on current economic developments in their respective countries (see Stangl, 2007b; Boumans and Garnitz, 2017). Unlike quantitative information from official statistics, the WES focuses on qualitative information by asking economists to assess main economic indicators for the present and the near-term future. This allows for a rapid, up-to-date assessment of the economic situation around the

[3] In this article, world GDP is the weighted average of advanced countries (Canada, the EU-28, Hong Kong, Japan, Norway, Singapore, South Korea, Switzerland, Taiwan, and the USA) and emerging countries (Argentina, Brazil, Chile, China, Colombia, India, Indonesia, Malaysia, Mexico, the Philippines, Russia, Thailand, Turkey, and Venezuela).

[4] We have to mention that longer GDP series are available. However, as indicated in the next section, our quarterly survey indicator first starts in 1989.

world, and particularly in developing and transition economies that often lack a number of official statistics. The uniform questionnaire, methodology and data processing guarantee comparability between countries and over time as well as the aggregation of country results to various country groups. At present, the survey covers almost 120 countries. The WES was launched via two trial runs in 1981 and conducted three times a year from 1983 to 1988 (Stangl, 2007a). Since 1989 the WES is a quarterly survey, conducted in January, April, July, and October. This is the main reason to start our analysis in 1989 at the earliest possible, because the WES survey results and GDP are both available on a quarterly frequency from that point in time.

The WES is an expert survey that applies a top-down approach, i.e., the surveyed experts assess the present and future economic situation in their country by taking into account all of the aspects that they regard as important. The panel includes representatives of multinational enterprises, academic institutions, foundations, economic research institutes, national and international chambers of industry and trade. Although the panel members are heterogeneous with respect to their professional affiliation, all of the respondents are highly qualified, either being in a leading position or occupied with economic research within their institution. The participation in the survey is absolutely voluntary. As it is common in panel surveys, some economists have left or joined the panel over time and not all participants respond to every survey, thus, the composition of the panel varies with each wave. At present, about 1,100 responses are received each quarter, which leads to a return rate of about 70% of filled questionnaires. Table 2 in the Appendix shows the average number of respondents for the 44 countries and three aggregates for the years 1990 to 2015.

In the past 20 years, the number of respondents varies strongly from at least 3 up to 50 experts per country. Generally, the higher a country's economic importance – according to the country's share in world GDP – the more WES experts participate. For our analysis we only consider countries with at least four WES respondents on average as well as a sufficient number of observations.

All tendency questions contained in the WES have, in general, three possible and qualitative answers each: '*good, better, higher*' for a positive assessment or an improvement (+), '*satisfactory, about the same, no change*' for a neutral assessment (=), and '*bad, worse, lower*' for a negative assessment or a deterioration (−). For each quarterly survey, the percentage shares of each tendency category (+), (=), and (−) are calculated from the individual replies. Therefore, no specific weighting of the individual answers per country exist, thus, a simple arithmetic mean is applied. As common in the ifo surveys, a balance statistic is calculated from the percentage shares of positive and negative responses. This results in a statistic ranging from −100 to +100 balance points. If positive and negative shares equal each other, the balance statistic has a value of zero. The GDPs measured in purchasing power parities serve as weights to calculate country groups or regions.

For our forecasting exercise, we use the three main indicators which catch the most attention by the public: the assessment of the present economic situation (SIT), expectations for the economic situation in the next six months (EXP), and the resulting indicator of both questions, the economic climate (CLI). The underlying assessment for the three indicators is as follows: '*This country's general situation regarding the overall economy is:*'. For the judgment of the present economic situation, the respondents can choose either '*good*', '*satisfactory*' or '*bad*'. For the expected situation by the end of the next six months, the answers are '*better*', '*about the same*', and '*worse*'. The economic climate is the geometric mean of the balance statistics for the present situation and the expectation indicator according to the following formula:

$$CLI = \sqrt{(SIT + 200)(EXP + 200)} - 200 \, . \tag{1}$$

This is the usual way of the ifo Institute to calculate its composite indicators such as the most important leading indicator for the German economy, the ifo Business Climate for Industry and Trade (Seiler and Wohlrabe, 2013). Long time series for the ifo World Economic Climate or the ifo Economic Climate for the Euro Area are available free of charge at the ifo homepage.[5] The survey results for other countries are published in the journal *ifo World Economic Survey* or are available upon request.

3. Forecasting Approach

3.1. Univariate One-Indicator Models

As a starting point for our pseudo out-of-sample forecasting exercise, we consider the following very simple indicator model,

$$y^j_{i,t+h} = c + WES^j_{i,t} + \varepsilon_{i,t} \, , \tag{2}$$

where $y^j_{i,t}$ is the quarter-on-quarter growth rate of GDP for a specific country i and a given point in time t. One of the three possible WES indicators (present economic situation SIT, expectations for the next six months EXP or the economic climate (CLI) is denoted by $WES^j_{i,t}$. Each h-step ahead direct forecast is calculated by shifting the specific indicator back in time in the estimation equation. The forecast horizon h is defined in the range of $h \in \{0, 1, 2\}$ quarters, whereas $h = 0$ defines the nowcast and $h = 2$ the maximum forecast of a half year. We assume that the forecast is produced at the end of each quarter t, thus, the GDP growth rate of $t - 1$ as well as the contemporaneous WES indicator are known to the forecaster. Such a simple indicator model has been proved to do a good job in forecasting Euro Area GDP growth (see Girardi *et al.*, 2016). We, however, also experimented with lags

[5] `http://www.cesifo-group.de/ifoHome/facts/Survey-Results/World-Economic-Survey.html` is the exact link to find the described time series.

for both the target series as well as the survey indicators. The results remained qualitatively the same.[6] As the benchmark model we use an AR(1), which proved to be a quite good competitor in the forecasting literature.

We also keep it simple for the calculation of our forecasts. For each country we have a different number of observations (T_i). As this difference prevents us from applying a fix starting point for all countries to forecast GDP, we decided to use the first $T_i/3$ observations as the initial estimation period. First, the model parameters are estimated via ordinary least squares (OLS). Second, based on these estimates, we calculate the forecasts for all three horizons. And last, the estimation window is expanded by one quarter ($T_i/3 + 1$). After this expansion, the model is re-estimated and new forecasts are calculated. This iterative procedure is continued until the end of our observation period.

3.2. Univariate Multi-Indicator Models

In times of a globalized world, we may gain some forecasting improvements for national GDP by adding survey indicators of the most important trading partners. The literature on international linkages has found that a higher trade intensity between countries leads to a more intensive business cycle synchronization between those (see, among others, Inklaar *et al.*, 2008; Duval *et al.*, 2016). We thus sequentially add the WES results of the three most important trading partners to Equation (2), ending up in the following multivariate models,

$$y^j_{i,t+h} = c + WES^j_{i,t} + WES^j_{TP1,t} + \varepsilon_{i,t} \ , \tag{3}$$

$$= c + WES^j_{i,t} + WES^j_{TP1,t} + WES^j_{TP2,t} + \varepsilon_{i,t} \ , \tag{4}$$

$$= c + WES^j_{i,t} + WES^j_{TP1,t} + WES^j_{TP2,t} + WES^j_{TP3,t} + \varepsilon_{i,t} \ . \tag{5}$$

First, we add the same WES indicator j from the most important trading partner (TP1) and repeat the forecasting experiment from the previous section. Second, we also add indicator j from the second most important trading partner (TP2) of country i. Finally, the largest model comprises the survey indicators of all three most important trading partners (TP3). Taking Germany as the example, its three most important trading partners are the USA, France, and UK. If we set up a model with the WES Economic Climate for Germany, we sequentially add the WES Economic Climate of (i) the USA, (ii) France, and (iii) UK. We refrain from allowing a mix of indicators, thus, we have 12 forecasting models per country (3 one-indicator and 9 multi-indicator models). All other steps of the forecasting exercise are as equal as for the univariate one-indicator models. The choice of the most important trading partners is also limited to the availability of WES information. In cases where we

[6] Automatic model selections either by the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) suggested very parsimonious models in the majority of cases. We take this finding as evidence for the application of our simple indicator model.

do not have survey indicators from the WES for a main trading partner, we replace it with information from the next most important trading partner. The last three columns of Table 2 in the Appendix list the three main trading partners per country.

3.3. Forecast Evaluation

We apply the standard root mean squared forecast error (RMSFE) as the measure of forecast accuracy. Let $FE_{i,t+h}^{j} = y_{i,t+h} - \hat{y}_{i,t+h}^{j}$ denote the h-step ahead forecast error resulting from one of the three WES indicators j, then the $RMSFE_{i,h}^{j}$ is defined as

$$RMSFE_{i,h}^{j} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(FE_{i,t+h}^{j,n}\right)^2}\,, \tag{6}$$

with N as the total number of forecasts that were calculated. The respective RMSFE for the benchmark model is: $RMSFE_{i,h}^{AR(1)}$. In order to decide whether the WES indicator model delivers smaller forecast errors on average, we calculate the relative root mean squared forecast error (rRMSFE):

$$rRMSFE_{i,h}^{j} = \frac{RMSFE_{i,h}^{j}}{RMSFE_{i,h}^{AR(1)}}\,. \tag{7}$$

A ratio smaller of one means that the specific WES indicator model j has, on average, a higher forecast accuracy compared to the autoregressive benchmark. The opposite is indicated by ratios larger than one.

The standard way to discriminate between the forecasting performances of two competing models in a statistical way is to apply the forecast accuracy test by Diebold and Mariano (1995) (DM test). This pairwise test evaluates whether the average loss differential between two models is statistically different from zero. Under the null hypothesis,

$$E\left[d_{i,t+h}^{j}\right] = E\left[\mathcal{L}_{i,t+h}^{AR(1)} - \mathcal{L}_{i,t+h}^{j}\right] = 0\,, \tag{8}$$

the DM test examines in a statistical sense whether two models produce equal quadratic losses. In our case, $\mathcal{L}_{i,t+h}^{j}$ is the quadratic loss from one specific indicator model and $\mathcal{L}_{i,t+h}^{AR(1)}$ the quadratic loss of the benchmark.

For many countries, we have to rely on rather small samples, thus, we need to correct for a possible small sample bias in the test. This is done with the modification proposed by Harvey *et al.* (1997). The resulting modified Diebold-Mariano test (MDM) has the following test statistic:

$$\mathrm{MDM}_h = \left(\frac{N+1-2h+N^{-1}h(h-1)}{N}\right)^{1/2}\hat{V}\left(\overline{d}_{i,h}^{j}\right)^{-1/2}\overline{d}_{i,h}^{j}\,. \tag{9}$$

The estimated long-run variance of the loss differential $\overline{d}_{i,h}^{j}$ is denoted as $\hat{V}\left(\overline{d}_{i,h}^{j}\right)$. Critical values to decide on the statistical significance are taken from Student's t-distribution with $N-1$ degrees of freedom.

4. Results

Table 1 shows the relative root mean squared forecast errors for each forecast horizon and country separately. The results for the three aggregates, European Union, Eurozone, and the World, can be found at the bottom of the table. For each forecast horizon, the table contains three columns: (i) the lowest rRMSFE out of the 12 available indicator models, (ii) the corresponding indicator model, and (iii) a statement on the statistical significance between forecast errors based on the modified Diebold-Mariano test (MDM). The best indicator model is always abbreviated as a combination of the specific indicator and the number of additional survey results from the main trading partners. For example, EXP–1 for Argentina is a model with WES economic expectations for the next six months of Argentina and Brazil. The best model for Chile in the nowcast situation is CLI–0, thus, a model with the WES economic climate for Chile and no additional trading partners. An asterisk indicates that the corresponding indicator model significantly produces lower forecast errors than the benchmark model at least to the 10% significance level.

The WES indicator models produce lower forecast errors compared to the autoregressive benchmark for 35 countries or aggregates in our sample. There are, however, some few exceptions for which the AR(1) cannot be beaten. These are: Belgium, Canada, China, Indonesia, Mexico, Russia, Spain, Switzerland, and the United Kingdom. From these nine countries, only three show rRMSFE that are larger than one for all forecast horizons (China, Switzerland, and the United Kingdom). Among the best country-specific models, approximately 70% contain WES indicators from the main trading partners. Thus, taking economic signals of main trading partners into account can improve the GDP forecast of the specific domestic economy. This improvement is especially present for the two shorter forecast horizons, since only 50% of the best models for $h = 2$ contain WES indicators from the main trading partners. For $h = 0$ and $h = 1$, this share increases to more than three-fourths.

Sticking to the best indicators, we find that most of the best models either contain the WES economic climate (CLI) or the economic expectations for the next six months (EXP). The models containing the present economic situation (SIT) also produce forecast errors that are smaller than those from the benchmark model. However, SIT does not contain as much information as the other two indicators in terms of forecasting GDP. Nevertheless, we have countries in the sample, such as Finland or New Zealand, for which a model including the WES economic situation produces the smallest rRMSFE (Finland: $rRMSFE_{h=2} = 0.856$, New Zealand: $rRMSFE_{h=0} = 0.981$).

Now we take a closer look at the countries in the sample. The largest improvement for $h = 0$ can be found for the EU ($rRMSFE_{h=0} = 0.695$), followed by Bulgaria ($rRMSFE_{h=0} =$

Table 1: Best Models for each Country

	$h=0$			$h=1$			$h=2$		
Country	rRMSFE	Model	MDM	rRMSFE	Model	MDM	rRMSFE	Model	MDM
Argentina	0.849	EXP–1	*	0.806	EXP–1	*	0.930	EXP–1	
Australia	0.983	EXP–1		1.012	EXP–1		0.978	SIT–1	
Austria	0.898	CLI–2	*	0.935	CLI–2	*	0.928	EXP–3	
Belgium	1.081	CLI–0		1.073	EXP–1		0.951	EXP–1	
Brazil	0.835	CLI–1	*	0.926	CLI–0	*	0.961	CLI–0	*
Bulgaria	0.698	CLI–2		0.691	CLI–2		0.793	CLI–2	
Canada	1.029	CLI–0		1.069	CLI–1		0.977	CLI–0	
Chile	0.965	CLI–0		0.890	CLI–0	*	0.897	CLI–0	*
China	1.322	CLI–2		1.342	SIT–0		1.088	SIT–0	
Czech R.	0.932	EXP–3		1.002	EXP–2		0.932	EXP–3	
Denmark	0.823	CLI–3	*	0.845	EXP–1	*	0.957	EXP–2	*
Estonia	0.813	EXP–3	*	0.845	EXP–3	*	1.102	EXP–0	*
Finland	0.872	CLI–3	*	0.846	EXP–3	*	0.856	SIT–3	*
France	0.918	CLI–3		1.044	CLI–3		0.976	SIT–3	
Germany	0.893	CLI–1	*	0.949	EXP–3		0.974	EXP–0	*
Hong Kong	0.869	EXP–3	*	0.967	EXP–1	*	0.990	EXP–0	
Hungary	0.933	EXP–2		0.895	EXP–1		0.955	EXP–1	
India	0.940	EXP–1		0.942	CLI–0	*	0.980	CLI–0	
Indonesia	1.003	EXP–0		0.950	EXP–0		1.066	EXP–3	
Ireland	0.860	CLI–0	*	0.926	CLI–1		0.962	CLI–0	
Italy	0.949	CLI–3	*	0.915	EXP–3		0.845	EXP–1	*
Japan	0.867	CLI–3	*	0.946	EXP–0	*	0.992	EXP–0	
Latvia	0.851	CLI–2	*	0.773	CLI–2	*	0.859	EXP–1	*
Mexico	1.023	CLI–1		1.058	EXP–0		0.957	SIT–0	*
Netherlands	0.847	CLI–0		0.904	CLI–2		0.910	EXP–1	
New Zealand	0.981	SIT–0		0.997	SIT–0		0.994	EXP–3	
Norway	0.933	EXP–0	*	0.970	SIT–1		0.965	EXP–0	*
Philippines	0.914	EXP–3	*	0.976	EXP–1		0.977	SIT–0	*
Poland	0.911	CLI–0	*	0.922	CLI–0	*	0.951	CLI–0	
Portugal	0.795	CLI–1	*	0.863	CLI–2	*	0.992	SIT–0	
Russia	0.995	CLI–2		1.151	CLI–2		1.024	CLI–0	
Slovakia	0.926	CLI–0		0.958	CLI–0		0.992	EXP–0	
Slovenia	0.867	CLI–3	*	0.873	CLI–3	*	0.972	EXP–1	
South Africa	0.973	SIT–2		1.047	CLI–1		0.955	CLI–0	
South Korea	0.999	CLI–0		1.041	CLI–1		0.999	CLI–1	
Spain	1.045	CLI–0	*	1.098	CLI–3		0.864	CLI–3	*
Sweden	0.811	EXP–1	*	0.840	EXP–1	*	0.932	EXP–1	*
Switzerland	1.004	CLI–2		1.108	CLI–2		1.025	CLI–2	
Taiwan	0.918	EXP–2	*	0.915	EXP–2	*	0.980	SIT–0	*
Thailand	0.907	EXP–1		0.998	SIT–0		1.015	SIT–0	
Turkey	0.941	EXP–0	*	0.943	EXP–1	*	0.988	SIT–1	
UK	1.081	CLI–1		1.082	CLI–3		1.006	CLI–3	
USA	0.963	CLI–0		0.976	EXP–2		1.008	EXP–1	
Uruguay	0.951	EXP–1		1.051	SIT–0		1.017	SIT–0	
EU	0.695	CLI–2		0.835	EXP–1		0.556	EXP–0	
Eurozone	0.806	CLI–2	*	0.972	CLI–3	*	0.905	SIT–1	
World	0.839	EXP–3		1.007	EXP–1		0.875	EXP–0	

Note: For each forecast horizon and country or aggregate, the table reports the smallest rRMSFE of the 12 possible indicator models; the columns 'Model' show the abbreviation of this best model. The indicators are abbreviated as: SIT ... WES present economic situation, EXP ... WES expectations for the next six months and CLI ... WES economic climate. Numbers in the model's name indicate either an one-indicator model (–0) or a multi-indicator approach with WES indicators of one (–1), two (–2) or three (–3) main trading partners. The benchmark is always the AR(1). A * in column 'MDM' indicates a significant improvement in forecast accuracy due to the modified Diebold-Mariano test at least to the 10% significance level.

0.697) and Portugal ($rRMSFE_{h=0} = 0.795$). For $h = 1$, the top 3 improvements are observable for Bulgaria, Latvia and Argentina ($rRMSFE_{h=1}$: $0.691, 0.773,$ and 0.806). We again find the EU and Bulgaria ($rRMSFE_{h=2}$: $0.556, 0.793$), in addition to Italy ($rRMSFE_{h=2} = 0.845$), among the smallest relative forecast errors for the largest forecast horizon. By grouping the countries into advanced and emerging economies, the correlation between being an emerging economy and the rRMSFE is negative (≈ -0.3), thus, the relative forecast errors are on average smaller for advanced economies. This holds true for all three forecast horizons.

The finding that the relative errors are smaller for advanced countries on average raises the question whether the performance of the WES indicators depends on the number of interviewed experts. There seems to be a slight statistical relationship between the relative forecast errors and the number of experts for the specific country. Furthermore, this correlation is negative, indicating that the rRMSFEs are on average smaller the more experts are interviewed. A composition effect of the pool of experts on the relative forecast performance is also imaginable. However, the corresponding affiliation of the expert is only captured in the data set since 2015. For all countries together, approximately 50% of the experts are either affiliated with a research institution (institute or university) or a financial institution (central bank, commercial bank or other financial organization). The composition of experts may deliver more insights into the heterogeneity of forecast accuracy between countries. We, however, have to leave such a question for future research activities.

5. Conclusion

A comprehensive international study on forecasting GDP in which the accuracy for countries is comparable, requires the same set of indicators. Since official data varies between countries, such a comparability is hard to reach. In this paper, we use fast and free available indicators that are, on top, international comparable: the survey results from the World Economic Survey (WES). By applying the three main indicators from the WES (the assessment of the current economic situation, the expectations on future economic development for the next sixth months, and the economic climate), our paper studies the forecasting performance of these indicators for 44 countries and three country aggregates separately. Additionally, we ask whether the national-specific forecast accuracy for GDP can be improved by adding WES indicators of the three main trading partners by country. For 35 countries in the sample as well as the three country aggregates (European Union, Eurozone, and the World), a model containing WES information produces lower forecast errors than a simple autoregressive benchmark up to two quarters ahead. Only for three countries (China, Switzerland, and the United Kingdom), the indicator model cannot beat the benchmark at all. We also find that the root mean squared forecast errors relative to the benchmark model are on average smaller for advanced economies compared to emerging economies. The most important indicators are the economic climate and the expectations on future economic development for the next six months. The assessment of the current economic situation plays only a minor role in forecasting GDP. Sticking to our second contribution, 70% of all indicator models contain at least one indicator of one main trading partner. Thus, using survey information for economic important partners to the specific country improves national GDP forecasts.

The fast and free availability of the WES makes it a powerful tool to forecast GDP all over the world. Since the ifo Institute plans to conduct the WES on a monthly basis, the indicators will be available on a even higher frequency, making the WES interesting to forecast other important economic variables such as the inflation rate or industrial production. However, such considerations have to be left for long-term research activities for which the time series are long enough. Follow up studies to ours can also go into more detail of the WES forecast accuracy. As indicated in the results section, the composition of the pool of experts and thus the cross-section variance may explain country differences in relative forecasting performance. Other studies may also investigate the performance of the WES compared to other, very prominent leading indicators such as the Purchasing Manager Index or the Composite Leading Indicator of the OECD.

References

Banerjee, A. and Marcellino, M. (2006). Are there any reliable leading indicators for US inflation and GDP growth? *International Journal of Forecasting*, **22** (1), 137–151.

Barhoumi, K., Darné, O. and Ferrara, L. (2010). Are disaggregate data useful for factor analysis in forecasting French GDP? *Journal of Forecasting*, **29** (1-2), 132–144.

Barnett, A., Mumtaz, H. and Theodoridis, K. (2014). Forecasting UK GDP growth and inflation under structural change. A comparison of models with time-varying parameters. *International Journal of Forecasting*, **30** (1), 129–143.

Becker, S. O. and Wohlrabe, K. (2008). Micro Data at the Ifo Institute for Economic Research – The "Ifo Business Survey" Usage and Access. *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, **128** (2), 307–319.

Boumans, D. and Garnitz, J. (2017). Ifo World Economic Survey Database – An International Economic Expert Survey. *Journal of Economics and Statistics*, **237** (1), 71–80.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, **13** (3), 253–263.

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

Duval, R., Li, N., Saraf, R. and Seneviratne, D. (2016). Value-Added Trade and Business Cycle Synchronization. *Journal of International Economics*, **99**, 251–262.

FICHTNER, F., RÜFFER, R. and SCHNATZ, B. (2011). The Forecasting Performance of Composite Leading Indicators: Does Globalisation Matter? *OECD Journal: Journal of Business Cycle Measurement and Analysis*, **2011** (1), 55–72.

GIRARDI, A., GAYER, C. and REUTER, A. (2016). The Role of Survey Data in Nowcasting Euro Area GDP Growth. *Journal of Forecasting*, **35** (5), 400–418.

HARVEY, D. I., LEYBOURNE, S. J. and NEWBOLD, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, **13** (2), 281–291.

HENZEL, S. and WOLLMERSHÄUSER, T. (2005). Quantifying Inflation Expectations with the Carlson-Parkin Method: A Survey-based Determination of the Just Noticeable Difference. *Journal of Business Cycle Measurement and Analysis*, **2** (3), 321–352.

HÜLSEWIG, O., MAYR, J. and WOLLMERSHÄUSER, T. (2008). *Forecasting Euro Area Real GDP: Optimal Pooling of Information*. CESifo Working Paper No. 2371.

HUTSON, M., JOUTZ, F. and STEKLER, H. O. (2014). Interpreting and evaluating CESIfo's World Economic Survey directional forecasts. *Economic Modelling*, **38** (February), 6–11.

INKLAAR, R., JONG-A-PIN, R. and DE HAAN, J. (2008). Trade and Business Cycle Synchronization in OECD Countries – A Re-Eexamination. *European Economic Review*, **52** (4), 646–666.

KIRIAKIDIS, M. and KARGAS, A. (2013). Greek GDP forecast estimates. *Applied Economics Letters*, **20** (8), 767–772.

KUDYMOWA, E., PLENK, J. and WOHLRABE, K. (2013). Ifo World Economic Survey and the Business Cycle in Selected Countries. *CESifo Forum*, **14** (4), 51–57.

LEHMANN, R. (2015). *Survey-based indicators vs. hard data: What improves export forecasts in Europe?* ifo Working Paper No. 196.

— and WEYH, A. (2016). Forecasting Employment in Europe: Are Survey Results Helpful? *Journal of Business Cycle Research*, **12** (1), 81–117.

ÖSTERHOLM, P. (2014). Survey data and short-term forecasts of Swedish GDP growth. *Applied Economics Letters*, **21** (2), 135–139.

PONS-NOVELL, J. (2006). An analysis of a panel of Spanish GDP forecasts. *Applied Economics*, **38** (11), 1287–1292.

SEILER, C. and WOHLRABE, K. (2013). Das ifo Geschäftsklima und die deutsche Konjunktur. *ifo Schnelldienst*, **66** (18), 17–21.

Stangl, A. (2007a). Ifo World Economic Survey Micro Data. *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, **127** (3), 487–496.

— (2007b). World Economic Survey. In G. Goldrian (ed.), *Handbook of Survey-Based Business Cycle Analysis*, Elgar, Cheltenham, pp. 57–65.

Zhou, R., Wang, X. and Tong, G. (2013). Forecasting macroeconomy based on the term structure of credit spreads: evidence from China. *Applied Economics Letters*, **20** (15), 1363–1367.

A. Data Set Description

Table 2: Countries, Data Sources and Main Trading Partners

					Main Trading Partners			
Country	GDP	Source	Start	WES	First	Second	Third	Source
Argentina	R, SA	OECD	Q1-2004	9	Brazil	China	USA	World Bank
Australia	R, SA	OECD	Q1-1989	11	China	Japan	South Korea	World Bank
Austria	R, SA	Eurostat	Q1-1996	13	Germany	USA	Italy	Eurostat
Belgium	R, SA	Eurostat	Q1-1995	15	Germany	France	Netherlands	Eurostat
Brazil	R, SA	OECD	Q1-1996	21	China	USA	Argentina	World Bank
Bulgaria	R, SA	Eurostat	Q1-2000	14	Germany	Italy	Turkey	Eurostat
Canada	R, SA	OECD	Q1-1989	11	US	China	UK	World Bank
Chile	R, SA	OECD	Q1-1995	9	China	USA	Japan	World Bank
China	N, mSA	National	Q1-1989	43	USA	Hong Kong	Japan	World Bank
Czech Republic	R, SA	Eurostat	Q1-1996	10	Germany	Slovakia	Poland	Eurostat
Denmark	R, SA	Eurostat	Q1-1995	7	Germany	Sweden	USA	Eurostat
Estonia	R, SA	Eurostat	Q1-1995	20	Sweden	Finland	Latvia	Eurostat
Finland	R, SA	Eurostat	Q1-1990	17	Germany	Sweden	USA	Eurostat
France	R, SA	Eurostat	Q1-1989	17	Germany	Spain	USA	Eurostat
Germany	R, SA	Eurostat	Q1-1991	48	USA	France	UK	Eurostat
Hong Kong	R, mSA	National	Q1-1989	8	China	USA	Japan	World Bank
Hungary	R, SA	Eurostat	Q1-1995	11	Germany	Slovakia	Austria	World Bank
India	R, SA	OECD	Q2-1996	13	USA	Hong Kong	China	World Bank
Indonesia	R, SA	OECD	Q1-1990	7	Japan	China	USA	World Bank
Ireland	R, SA	Eurostat	Q1-1997	7	USA	UK	Belgium	Eurostat
Italy	R, SA	Eurostat	Q1-1995	21	Germany	France	USA	Eurostat
Japan	R, SA	OECD	Q1-1989	29	USA	China	South Korea	World Bank
Latvia	R, SA	Eurostat	Q1-1995	6	Russia	Estonia	Germany	Eurostat
Mexico	R, SA	OECD	Q1-1989	12	USA	Canada	China	World Bank
Netherlands	R, SA	Eurostat	Q1-1996	15	Germany	Belgium	UK	Eurostat
New Zealand	R, SA	OECD	Q1-1989	10	China	Australia	USA	World Bank
Norway	R, SA	Eurostat	Q1-1989	6	UK	Germany	Netherlands	World Bank
Philippines	R, SA	National	Q1-1998	6	Japan	US	China	World Bank
Poland	R, SA	Eurostat	Q1-2002	16	Germany	UK	Czech R.	Eurostat
Portugal	R, SA	Eurostat	Q1-1995	11	Spain	France	Germany	Eurostat
Russia	R, SA	OECD	Q1-1995	19	Netherlands	China	Italy	World Bank
Slovakia	R, SA	Eurostat	Q1-1997	10	Germany	Czech R.	Poland	Eurostat
Slovenia	R, SA	Eurostat	Q1-1995	7	Germany	Italy	Austria	Eurostat
South Africa	R, SA	OECD	Q1-1989	20	China	USA	Germany	World Bank
South Korea	R, SA	OECD	Q1-1989	9	China	USA	Hong Kong	World Bank
Spain	R, SA	Eurostat	Q1-1995	24	France	Germany	UK	Eurostat
Sweden	R, SA	Eurostat	Q1-1993	13	Germany	USA	UK	Eurostat
Switzerland	R, SA	Eurostat	Q1-1989	14	Germany	USA	Hong Kong	World Bank
Taiwan	R, SA	National	Q1-1989	10	China	Hong Kong	USA	WTO
Thailand	R, mSA	National	Q1-1993	8	USA	China	Japan	World Bank
Turkey	R, SA	OECD	Q1-1998	11	Germany	UK	Italy	World Bank
United Kingdom	R, SA	Eurostat	Q1-1989	18	USA	Germany	Switzerland	Eurostat
USA	R, SA	OECD	Q1-1989	27	Canada	Mexico	China	World Bank
Uruguay	R, mSA	National	Q1-2005	5	Brazil	China	USA	World Bank
EU	R, SA	Eurostat	Q1-1995	292	USA	China	Switzerland	Eurostat
Eurozone	R, SA	Eurostat	Q1-1995	252	USA	China	Switzerland	Eurostat
World	N, SA	–	Q2-1994	809	USA	China	Germany	World Bank

Note: For each country or aggregate, the table reports the characteristics of the GDP series, its corresponding data source as well as starting point and the average sample size of the WES between 1990 and 2015. The last four columns show the three main trading partners of each country or aggregate and again the data source from which we obtained the trade data. *Abbreviations:* SA...seasonally adjusted, mSA...manual seasonal adjustment, R...real-terms, N...nominal-terms.