Empirical modelling of survey-based expectations for the design of economic indicators in five European regions

Oscar Claveria[1]*, **Enric Monte[2], Salvador Torra[3]**

[1] AQR-IREA, University of Barcelona
[2] Department of Signal Theory and Communications, Polytechnic University of Catalunya
[3] Riskcenter-IREA, Department of Econometrics and Statistics, University of Barcelona

Abstract

In this study we use agents' expectations about the state of the economy to generate indicators of economic activity in twenty-six European countries grouped in five regions (Western, Eastern, and Southern Europe, and Baltic and Scandinavian countries). We apply a data-driven procedure based on evolutionary computation to transform survey variables in economic growth rates. In a first step, we design five independent experiments to derive a formula using survey variables that best replicates the evolution of economic growth in each region by means of genetic programming, limiting the integration schemes to the main mathematical operations. We then rank survey variables according to their performance in tracking economic activity, finding that agents' "perception about the overall economy compared to last year" is the survey variable with the highest predictive power. In a second step, we assess the out-of-sample forecast accuracy of the evolved indicators. Although we obtain different results across regions, Austria, Slovakia, Portugal, Lithuania and Sweden are the economies of each region that show the best forecast results. We also find evidence that the forecasting performance of the survey-based indicators improves during periods of higher growth.

Keywords

Economic indicators; Qualitative survey data; Expectations; Symbolic regression; Evolutionary algorithms; Genetic programming

JEL Classification: C51; C55; C63; C83; C93

---

* Corresponding Author.

1. Introduction

Agents' expectations about the state of the economy are instrumental for economic modelling. Business and consumer surveys, also known as tendency surveys, are directly addressed to economic agents as a means to measure their expectations. Respondents are asked about the expected direction of change of a wide range of variables (capital expenditures, private consumption, exports, imports, etc.). Accordingly, survey results provide important information about agents' economic expectations, allowing comparisons among different countries' business cycles. On the one hand, sectoral results of the surveys have been often used as partial indicators for the construction of more general aggregate economic indicators and for the estimation of macro magnitudes through their introduction in econometric models (Abberger, 2007; Bruestle and Crain, 2015; Graff, 2010; Hanson et al., 2005; Lehmann and Wohlrabe, 2017; Martinsen et al., 2014; Wilms et al., 2016). On the other hand, survey-based expectations have also been introduced into behavioural equations postulated by economic theory such as the Phillips curve and to evaluate the formation of expectations, as they provide a direct measure of expectations to test the rationality of agents (Altug and Çakmakli, 2016; Bovi, 2013; Jean-Baptiste, 2012; Lee, 1994; Miah et al., 2016; Paloviita, 2006). See Pesaran and Weale (2006) for a review of the uses of survey data for testing and modelling of expectations.

Survey results are available ahead of the publication of quantitative official data, which makes them very useful for monitoring the evolution of the economy. Nevertheless, the fact that survey-based expectations are qualitative in nature has centred research in the development of different approaches to transform survey responses into quantitative measures of agents' expectations. See Driver and Urga (2004), Nardo (2003) and Pesaran (1987) for a review of methods for the quantification of survey results. Recent developments in empirical modelling have allowed to develop conversion approaches based on evolutionary computation. This study extends previous research by Claveria et al. (2016), who proposed an evolutionary-based two-step procedure to generate estimates of economic growth. The authors derived preliminary building blocks defined as simple combinations of survey variables, and then linearly combined the functions to generate estimates of economic growth in Central and Eastern European economies, finding that the forecasting performance of evolved survey-based indicators could be improved by designing ad-hoc quantification procedures for countries with similar characteristics.

These findings have led us to use evolutionary computation to generate indicators of economic growth that combine different survey variables of 26 European countries grouped into five major European regions (Western, Eastern, and Southern Europe, and Baltic and Scandinavian countries). First, we design five independent experiments that link survey expectations to economic growth, limiting the preliminary functions to the main mathematical operations with the aim of facilitating the implementation of the evolved economic indicators. Once we obtain the optimal combination of survey variables that best replicates the evolution of economic activity in each region, we rank the expectations according to the relative weight of each one in the evolved indicators. In a second step, we assess the out-of-sample forecast accuracy of the obtained economic indicators country by country.

Some of the features of empirical modelling are particularly indicated to deal with the problem at hand. First, empirical modelling is especially suitable for finding patterns in large data sets with little or no prior information about the system. Second, empirical modelling allows us to simultaneously evolve both the structure and the parameters of the model without imposing any assumptions regarding agents' expectations. In a recent study, Lahiri and Zhao (2015) found significant improvements in the forecasting performance of quantified expectations when relaxing the assumptions of quantification methods of qualitative survey data.

The empirical modelling approach applied in this research is based on symbolic regression (SR) via genetic programming (GP). While SR is a modelling approach characterised by the search of the space of mathematical expressions that best fit a given dataset, GP is a soft computing search technique for problem-solving (Cramer, 1985). GP is based on the implementation of genetic algorithms (GAs), which are a specific type of evolutionary algorithm (EA). Evolutionary computation can be regarded as a subfield of artificial intelligence, and is being increasingly applied to automated problem solving in economics.

The main aim of this study is twofold. On the one hand, we implement GP to find the optimal combinations of survey expectations to forecast economic growth at a regional level, restricting the integration schemes to the main four mathematical operations so as to obtain easily replicable expressions. This allows us to rank survey variables according to their predictive capacity. On the other hand, we assess the forecasting performance of the evolved economic indicators in each country and compare it to a benchmark model.

The structure of the paper is as follows. The next section reviews the existing literature. In Section 3 we present the methodological approach, describing the data and the experimental set-up. Empirical results are provided in Section 4. Finally, conclusions are drawn in Section 5.

2. Literature review

Economic expectations have been widely studied (Pesaran, 1987; Visco, 1984; Wren-Lewis, 1986). Tendency surveys ask respondents whether they expect a variable to rise, to remain constant, or to fall. The relationship between quantitative data and survey results was first formalised by Anderson (1952), who regressed the actual average percentage change of an aggregate variable on the percentage of respondents expecting a variable to rise and to fall. Carlson and Parkin (1975) developed the theoretical framework for quantifying survey expectations by assuming that respondents report a variable to go up if the mean of their subjective probability distribution lies above a threshold level, also known as indifference interval (Theil, 1952).

This relationship has been also explored by matching individual responses with firm-by-firm realisations, both empirically (Lahiri and Zhao, 2015; Lui et al., 2011a, b; Mitchell et al., 2002, 2005a, b; Mokinski et al., 2015; Müller, 2010), and experimentally via Monte Carlo simulations (Claveria et al., 2006; Nardo, 2003). Common (1985) used experimental expectations to test the rational expectations hypothesis. Muth (1961) assumed that rationality implied that expectations had to be generated by the same stochastic process that generates the variable to be predicted. While Common (1985) rejected the presence of rational agents, Miah et al. (2016) have recently found survey expectations in 18 emerging economies to be mostly unbiased and efficient. Simulation experiments have also been used to assess the forecasting performance of different quantification methods of survey expectations (Claveria, 2010; Löffler, 1999; Nardo and Cabeza-Gutés, 1999; Terai, 2009).

The link between survey expectations and quantitative data at the aggregate level has been further and widely investigated (Abberger, 2007; Balcombe, 1996; Batchelor, 1981, 1982, 1986; Bennett, 1984; Bergström, 1995; Berk, 1999; Białowolski, 2016; Bovi, 2016; Breitung and Schmeling, 2013; Claveria et al., 2007; Franses et al., 2011; Ghonghadze and Lux, 2012; Graff, 2010; Guizzardi and Stacchini, 2015; Jonsson and Österholm,

2011, 2012; Lacová and Král, 2015; Lahiri and Teigland, 1987; Maag, 2009; Martinsen et al., 2014; Mittnik and Zadrozny, 2005; Nolte and Pohlmeier, 2007; Paloviita, 2006; Pesaran, 1985; Robinzonov et al., 2012; Seitz, 1988; Smith and McAleer, 1995; Vermeulen, 2014). These studies use a wide range of econometric techniques, but none of them assesses the relationship between both types of data by means of evolutionary methods.

In this study we fill this gap by linking survey data and economic growth in a SR setting solved by means of evolutionary computation. This approach is based on the implementation of GAs, which adopt Darwinian principles of the theory of natural selection in the context of expensive optimisation (Fogel et al., 1966). GAs are the most common type of EA, and were initially proposed by Holland (1975). GP allows the model structure to vary during the evolution, which makes it particularly indicated for non-linear and empirical modelling. See Banzhaf et al. (2008), Dabhi and Chaudhary (2015) and Poli et al. (2010) for a review of the state of the art in GP.

Most economic applications of evolutionary computing are in finance (Chen and Kuo, 2002; Fogel, 2006; Goldberg, 1989). GAs have been used to predict the financial failure of firms (Acosta-González and Fernández, 2014), to explain the 2008 financial crisis (Acosta-González et al., 2012), to model exchange rates (Lawrenz and Westerhoff, 2003), to evaluate the convergence to the rational expectations equilibrium (Maschek, 2010), to optimize the signals generated by technical trading tools (Thinyane and Millin, 2011), to forecast stock price trends in Taiwan (Wei, 2013), etc. See Drake and Marks (2002) for a review of the applications of GAs in financial forecasting.

Regarding GP, Vasilakis et al. (2013) proposed a GP-based technique to predict returns in the trading of the euro/dollar exchange rate. GP has also been applied to to model short-term capital flows (Yu et al., 2004), to forecast exchange rates (Álvarez-Díaz and Álvarez, 2005), and for stock price forecasting (Chen et al., 2008; Kaboudan, 2000; Larkin and Ryan, 2008; Wilson and Banzhaf, 2009). Wilson and Banzhaf (2009) compared a developmental co-evolutionary GP approach to standard linear GP for interday stock prices prediction. Alexandridis et al. (2017) have recently compared the forecasting performance of GP in the context of weather derivatives pricing with other state-of-the-art machine learning algorithms and classic linear approaches, finding that non-linear methods outperformed the alternative linear models significantly.

Up until now there have been very few applications of GP in macroeconomics. The first GP application is that of Koza (1992), who used GP to solve a SR problem designed

to reassess the exchange equation, relating the price level, gross national product, money supply, and the velocity of money. More recent macroeconomic applications of GP have been used with forecasting purposes (Chen et al., 2010; Duda and Szydło, 2011). Ferreria (2011) developed a version of GP known as gene expression programming (GEP). Recently, Peng et al. (2014) proposed an improved GEP algorithm especially suitable for dealing with SR problems. Gandomi and Roke (2015) compared the forecasting performance of artificial neural network models to that of GEP techniques.

SR is an empirical modelling technique used to construct regression models. Given a predetermined set of operations and functions, SR searches appropriate models from the space of all possible mathematical expressions that best fit the data. Zelinka et al. (2005) introduced analytical programming in order to synthesise suitable solutions in SR.

Given its versatility, SR has been increasingly used in different areas (Barmpalexis et al., 2011; Ceperic et al., 2014; Sarradj and Geyer, 2014; Vladislavleva et al., 2010; Wu et al., 2008; Yang et al., 2015; Yao and Lin, 2009; Zameer et al., 2017), but there have been very few SR applications in macroeconomics. Claveria et al. (2016) implemented SR via GP to derive a set of building blocks used to estimate economic activity. Kľúčik (2012) used SR to estimate total exports and imports to Slovakia. Kotanchek et al. (2010) implemented SR via GP to predict economic activity. By means of SR, Kronberger et al. (2011) identified interactions between economic indicators in order to estimate the evolution of prices in the US. The authors suggested using SR for the exploration of variable interplay when approaching complex modelling tasks, as it provides a quick overview of the most relevant interactions and can help to identify new unknown links between variables.

In this study we design five independent SR experiments and apply GP in order to find the optimal combinations of survey expectations that best fit the actual evolution of economic activity in each region. We also asses the forecast accuracy of the obtained evolved economic indicators and compare it with several benchmarking models.

3 Data and Methodology

In this study we use SR via GP to formalize the optimal interactions between survey variables that best allow to predict economic growth, restricting them to the main mathematical operations (addition, subtraction, multiplication, and division). In order to

do so, we need to combine two types of information: qualitative survey expectations and quantitative official statistics from 2000:Q2 to 2016:Q3. Regarding the former, we make use of survey data on expectations from the World Economic Survey (WES) carried out quarterly by the Ifo Institute for Economic Research. As a proxy of economic activity we use the year-on-year growth rates of the Gross Domestic Product (GDP) retrieved from the Organisation for Economic Co-operation and Development (OECD) (https://data.oecd.org/gdp/quarterly-gdp.htm#indicator-chart).

The analysis is carried out for 26 European economies grouped in five regions based on the criteria used for statistical processing purposes by the United Nations Statistics Division. As a result, Austria, Belgium, France, Germany, Ireland, the Netherlands and the United Kingdom (UK) are grouped as Western Europe (1); Bulgaria, the Czech Republic, Hungary, Poland, Romania and the Slovak Republic as Eastern Europe (2); Croatia, Greece, Italy, Portugal, Slovenia and Spain as Southern Europe (3); Estonia, Latvia and Lithuania as the Baltic countries (4); Denmark, Finland, Norway and Sweden as the Scandinavian countries (5).

Table 1. Survey variables

Judgements about the *Present economic situation*	
Overall economy	$X1_{it}$
Capital expenditures	$X2_{it}$
Private consumption	$X3_{it}$
Perceptions about the *Economic situation compared to last year*	
Overall economy	$X4_{it}$
Capital expenditures	$X5_{it}$
Private consumption	$X6_{it}$
Expectations about the *Economic situation and foreign trade for the next 6 months*	
Overall economy	$X7_{it}$
Capital expenditures	$X8_{it}$
Private consumption	$X9_{it}$
Volume of exports	$X10_{it}$
Volume of imports	$X11_{it}$
Trade balance	$X12_{it}$

In Table 1 we present the twelve survey variables used in the study, denoted as X_{it}, where *i* refers to each country and *t* to the time period. Survey variables can be divided in judgements, perceptions and expectations, depending on whether they refer to the

expected value in the present, in the present compared to last year, or for the next six months. See Kudymowa et al. (2013), Hutson et al. (2014), and Garnitz et al. (2015) for an appraisal of the WES data.

By means of GP we evolve a symbolic expression for each region combining the different survey variables for each country until a stopping criterion is reached. Regarding this criterion, it can either be a predetermined value of the fitness function or a given number of generations. As there is a trade-off between accuracy and simplicity, we have chosen a maximum number of 50 generations as as stopping criterion. In Table 2 we summarize the steps for implementing the experiment in each of the regions.

Table 2. GP implementation – Steps

1. Creation of an initial population of programs	50,000
2. Evaluation of fitness for each program	Mean absolute error (MAE)
3. Selection of a reproduction strategy	Tournament method (size 3)
4. Application of genetic operators	Mutation probability 0.25
5. Determination of constants	Automatically generated
6. Creation of a new population	Max. generations 150

Genetic operators (crossover and mutation) are applied to the parents selected on the basis of the fitness function. Crossover consists on the recombination of randomly chosen parts of parents, while mutation on randomly altering a part of a parent. Consequently, the fitness of the population increases generation after generation. The output of this process is a set with the best individual functions from all generations for each region. In this study we have used the open source Distributed Evolutionary Algorithms Package (DEAP) framework implemented in Python (Fortin et al., 2012; Gong et al., 2015).

The obtained symbolic expressions are then used to generate out-of-sample forecasts of economic growth in all countries. To evaluate the performance of the evolved economic indicators we compute the accuracy of the forecasts and we compare it to that of the predictions obtained with both naïve and autoregressive (AR) time series models used as a benchmark.

4 Results

In this section we first present the results of the different experiments undertaken for each region (R=1, 2, 3, 4, 5) for the in-sample period (2000:Q2 to 2014:Q1). The output, $\hat{y}_{R,it}$,

is the evolved expression obtained in each region, formed by the optimal combination of survey variables for each set of countries. The following evolved symbolic expressions can be regarded as survey-based indicators of economic activity for each region:

$$\hat{y}_{1,it} = \frac{(X1_{it} + X3_{it} + X4_{it} + X5_{it} - 10)X4_{it}}{6 * X4_{it} - 5} \tag{1}$$

$$\hat{y}_{2,it} = \frac{1}{10}\left[X3_{it}(X10_{it} - 1) - \frac{1}{100}\left[\frac{X3_{it}}{\frac{1}{10}\left[\frac{(X10_{it} + 1) - (X4_{it} - 5)(X3_{it} + 1)}{X10_{it}} - 1 \right] - 1} \right] \right] \tag{2}$$

$$\hat{y}_{3,it} = \frac{X1_{it} - 5}{0.5 * X11_{it}} + \frac{1}{10}(3 * X1_{it} + X9_{it} + X11_{it} - 1) \tag{3}$$

$$\hat{y}_{4,it} = \frac{1}{10}\left[\left(5 + \frac{X2_{it}}{\frac{X6_{it}}{5} - 1 + X4_{it} + X5_{it} - X12_{it}} \right) \frac{X2_{it}}{10} + \frac{X4_{it}}{0.5} \right] \tag{4}$$

$$\hat{y}_{5,it} = \frac{1}{10}\left[X4_{it} + \frac{X3_{it} X4_{it}}{10} - 7 + (X3_{it} + X4_{it})\frac{X11_{it}}{10} + X3_{it} \right] \tag{5}$$

We can observe that variable $X4_{it}$ (perception of the overall economic situation compared to last year) is by far the variable that more frequently appears in the symbolic expressions, being present in all five evolved indicators. The second most frequent variables is $X3_{it}$, which refers to the judgement about the present situation of private consumption. The expectations about the future are the variables with a lower weight, being $X7_{it}$ and $X8_{it}$ (expectations for the next six months about the overall economy and capital expenditures respectively) the only variables that do not appear in any of the regions. Klein and Özmucur (2010) analysed the role of survey expectations in 26 European countries and found that he question related to production expectations was more useful in improving the forecasting performance than the aggregated confidence and sentiment indicators.

Given that $X4_{it}$ and $X3_{it}$ seem to be the most relevant survey variables, we repeated the five experiments using only those two variables, obtaining the following evolved symbolic expressions:

$$\hat{y}_{1,it} = 3.1 + \frac{X3_{it}}{X4_{it}} - \frac{10}{X4_{it}} \tag{6}$$

$$\hat{y}_{2,it} = \frac{X4_{it}}{2} - \frac{3 * X3_{it}}{10} - 1 \tag{7}$$

$$\hat{y}_{3,it} = \frac{X4_{it}}{X3_{it}} - \frac{5}{X3_{it}} \tag{8}$$

$$\hat{y}_{4,it} = \frac{4 * X4_{it}}{10} - 1.1 \tag{9}$$

$$\hat{y}_{5,it} = \frac{X4_{it} X3_{it}}{20} - \frac{X4_{it}}{100} - 1.5 \tag{10}$$

Next, we generate forecasts of GDP growth using both sets of evolved indicators and evaluate their predictive performance in an out-of-sample forecasting comparison for the period 2014:Q2 to 2016:Q3. With this aim we compute several measures of prediction accuracy. First, the the mean absolute error (MAE), the root mean square error (RMSE), and the mean absolute percentage error (MAPE) in order to assess the predictive content in terms of forecast accuracy.

Table 3 summarises the information of the different accuracy measures and the change in precision resulting from the use of the indicators of the second set of experiments. The best forecasting performance is obtained for Austria and the UK in Western Europe, for Slovakia in Eastern Europe, for Portugal in Southern Europe, and for Lithuania in the Baltic countries. If we average the results by region, we obtain the best results for the Baltic countries. These results are in line with those of Claveria et al. (2017), who in a similar experiment obtained the lowest MAE and RMSE values for Austria, Belgium, Bulgaria, Estonia, and Lithuania.

By using only $X4_{it}$ and $X3_{it}$ as the input variables, we find a deterioration of predictive accuracy in most countries, with the exception of Belgium, Germany, Italy, the Netherlands and the UK, where the three accuracy measures (MAE, RMSE and MAPE) decrease with respect to the ones obtained in the first set of experiments.

Finally, we compare the forecast accuracy of the evolved economic indicators obtained by means of all survey variables to the benchmark models. With this aim, we compute the mean absolute scaled error (MASE) and the percentage of periods with lower absolute error (PLAE) to compare the forecasting performance to the benchmarks.

Let us denote y_t as the actual value, and $\hat{y}_t$ as forecast at period t, $t = 1, \ldots, n$. Forecast errors can then be defined as $e_t = y_t - \hat{y}_t$. We have two competing models A and B, where A refers to the forecasting model under evaluation and B stands for benchmark model. Given that there is a delay of more than a quarter between the publication of official quantitative data with regard to survey data, in this study we use two-step ahead naïve forecasts as a baseline. The MASE can then be obtained as the mean of the absolute value of the scaled error q_t:

$$MASE = mean \left| q_t \right| \qquad \text{where } q_t = e_t \Bigg/ \left(\sum_{t=3}^{n} \left| y_t - y_{t-2} \right| \Big/ n-1 \right) \tag{11}$$

The MASE, proposed by Hyndman and Koehler (2006), allows to scale the forecast errors by the mean absolute in-sample errors obtained with a benchmark model. This statistic presents several advantages over other forecast accuracy measures. On the one hand, it is independent of the scale of the data. On the other hand, it is easy to interpret: values less than one indicate that the average prediction computed with the benchmark model is worse than the estimates obtained with the proposed method.

With the aim of finding an easy to interpret measure to compare the forecast accuracy between two models, Claveria et al. (2015) proposed the PLAE statistic, which is also a dimensionless measure. The PLAE is based on the CJ statistic proposed by Cowles and Jones (1937) for testing market efficiency and the 'percent better' measure proposed by Makridakis and Hibon (2000) to compare the forecast accuracy of the models to a random walk. The PLAE consists on a ratio that calculates the proportion of periods in which the model under evaluation obtains a lower absolute forecasting error than the benchmark model:

$$PLAE = \frac{\sum_{t=1}^{n} \lambda_t}{n} \text{ where } \lambda_t = \begin{cases} 1 & \text{if } \left| e_{t,A} \right| < \left| e_{t,B} \right| \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

When comparing the obtained out-of-sample forecasts with the models used as a benchmark (last two columns of Table 3), MASE values show that in the case of Germany, Ireland, Italy and Greece, the forecast accuracy of the evolved indicators does not improve that of the in-sample average prediction of the naive method. The PLAE values obtained in these four countries also highlight that the percentage of out-of-sample periods in which the proposed regional indicator generates lower absolute forecasting errors is very low. Conversely, the high PLAE values obtained for both the Naïve method and the AR model for Croatia, Spain, Lithuania and Sweden are indicative of the good forecasting performance of the generated indicators for these countries. In most cases, the PLAE values obtained for both benchmarks are very similar, with the exception of Bulgaria or Slovakia, where the relative performance of the AR model improves. In Fig. 1 we graphically compare actual and predicted economic growth.

Table 3. Out-of-sample forecast accuracy – SR indicators vs. benchmark models

|	SR indicators with all variables			SR indicators with just $X3_{it}$ and $X4_{it}$			Benchmark models		
							Naïve		AR
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MASE	PLAE (%)	PLAE (%)
Western Europe	**2.554**	**3.140**	**0.761**	**2.482**	**3.017**	**0.753**	**0.972**	**27%**	**21%**
Austria	0.498	0.581	0.663	0.761	0.835	1.059	0.264	50%	40%
Belgium	0.952	1.072	0.672	0.793	0.853	0.562	0.543	10%	10%
France	0.818	0.926	0.918	1.156	1.178	1.366	0.492	20%	20%
Germany	1.900	1.964	1.317	1.256	1.277	0.876	1.439	0%	0%
Ireland	11.800	15.241	0.651	12.206	15.568	0.711	3.032	20%	20%
NL	1.110	1.309	0.743	0.663	0.786	0.454	0.627	50%	50%
UK	0.798	0.886	0.365	0.541	0.623	0.241	0.405	40%	10%
Eastern Europe	**1.058**	**1.241**	**0.395**	**2.481**	**2.625**	**0.759**	**0.379**	**60%**	**25%**
Bulgaria	1.152	1.292	0.497	2.249	2.449	0.721	0.362	90%	10%
Czechia	1.065	1.366	0.395	2.583	2.733	0.763	0.451	70%	60%
Hungary	1.351	1.536	0.681	2.219	2.307	0.823	0.620	60%	10%
Poland	0.972	1.125	0.322	2.594	2.652	0.827	0.249	20%	30%
Romania	1.302	1.549	0.325	2.915	3.195	0.741	0.457	30%	10%
Slovakia	0.505	0.575	0.147	2.325	2.416	0.676	0.132	90%	30%
Southern Europe	**0.860**	**1.011**	**1.354**	**1.250**	**1.370**	**2.430**	**0.817**	**55%**	**53%**
Croatia	0.781	0.893	1.167	1.339	1.442	2.818	0.449	90%	80%
Greece	0.904	1.248	2.904	1.290	1.421	7.455	1.260	50%	50%
Italy	0.669	0.756	2.581	0.342	0.405	2.250	1.446	30%	30%
Portugal	0.646	0.759	0.632	0.804	0.872	0.699	0.753	50%	40%
Slovenia	1.119	1.298	0.478	1.674	1.816	0.656	0.474	10%	20%
Spain	1.039	1.115	0.361	2.050	2.267	0.702	0.517	100%	100%
Baltic countries	**0.828**	**1.043**	**0.483**	**0.990**	**1.236**	**0.515**	**0.192**	**53%**	**73%**
Estonia	0.849	1.123	0.409	1.057	1.399	0.487	0.206	30%	70%
Latvia	0.904	1.058	0.766	0.977	1.199	0.673	0.206	40%	60%
Lithuania	0.730	0.947	0.274	0.935	1.109	0.385	0.163	90%	90%
Scandinavian countries	**0.791**	**0.920**	**0.881**	**1.632**	**1.772**	**1.604**	**0.416**	**58%**	**65%**
Denmark	0.563	0.643	0.699	1.444	1.623	1.143	0.456	50%	70%
Finland	0.594	0.720	1.709	0.969	1.119	2.965	0.269	50%	60%
Norway	0.921	1.060	0.823	1.883	2.035	1.673	0.496	40%	30%
Sweden	1.085	1.255	0.292	2.230	2.309	0.634	0.442	90%	100%

Note: Mean for each group/region in bold. We have used Eurostat GDP data for Bulgaria, Romania and Croatia.

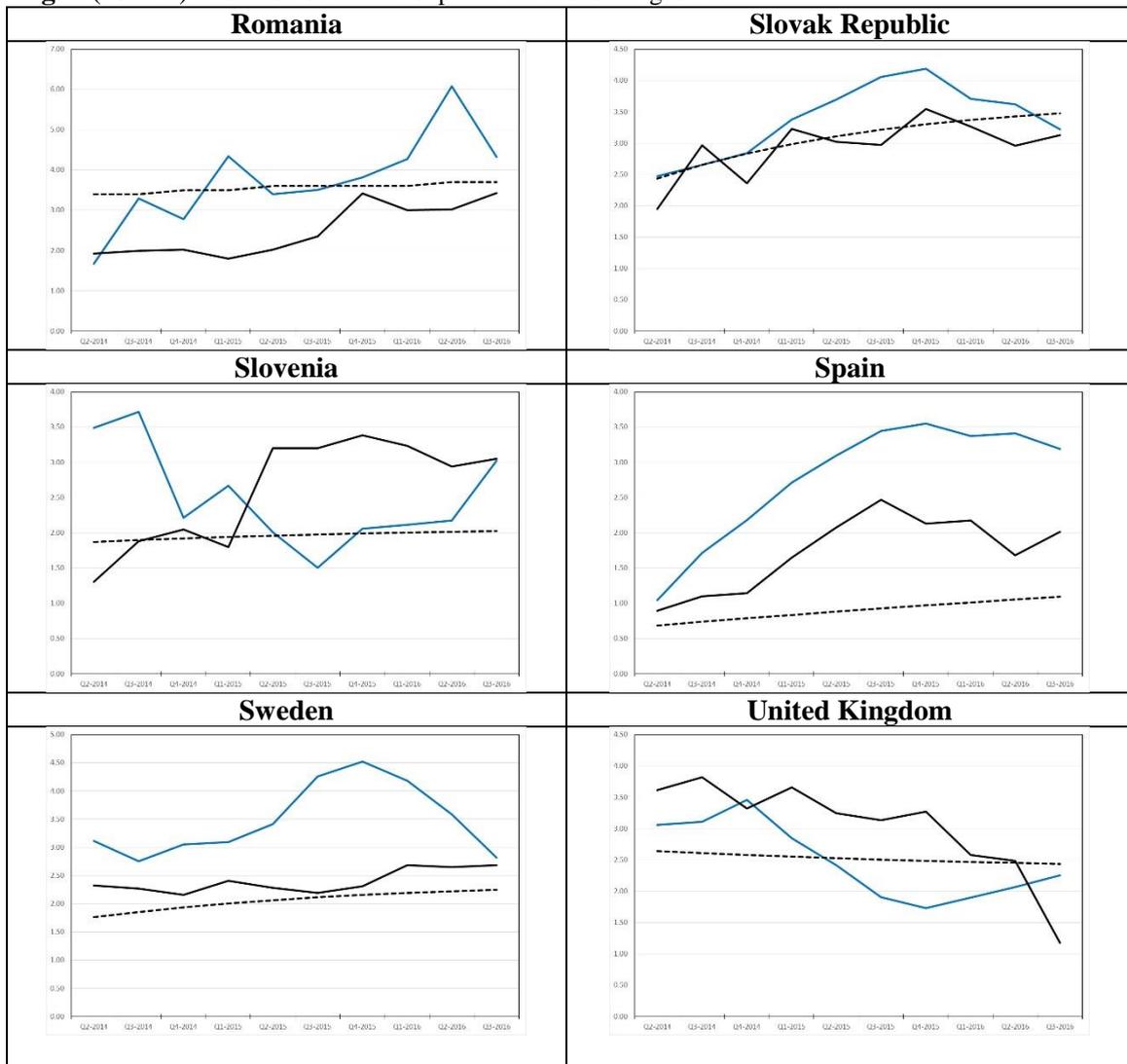**Fig. 1** Evolution of actual vs. predicted economic growth
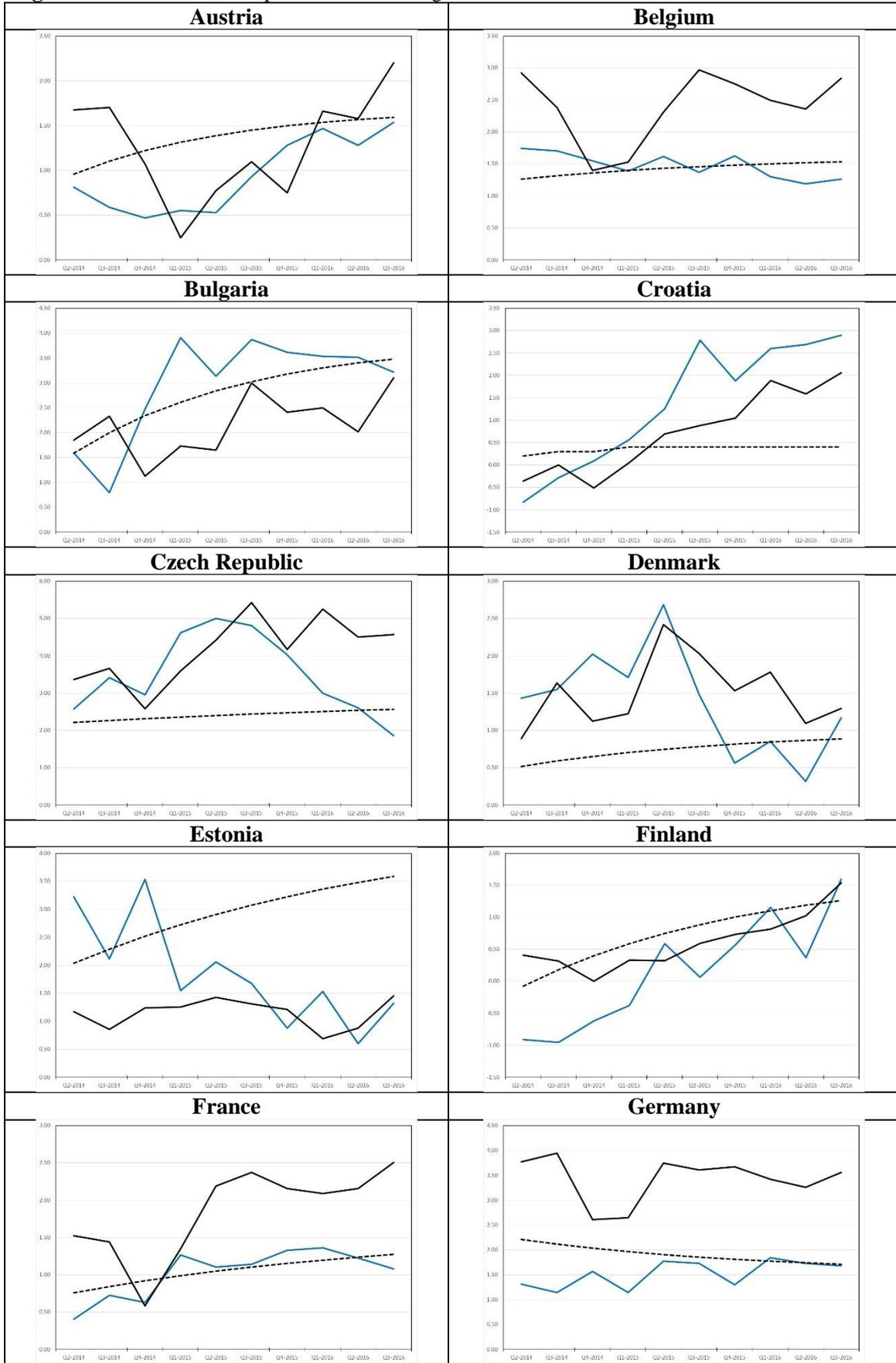
Note: The grey line represents the year-on-year growth rate of GDP in each country. The black line represents the evolution of the proposed evolved economic indicator. The black dotted line shows the evolution of the autoregressive forecasts.

Fig. 1 (cont. 1) Evolution of actual vs. predicted economic growth
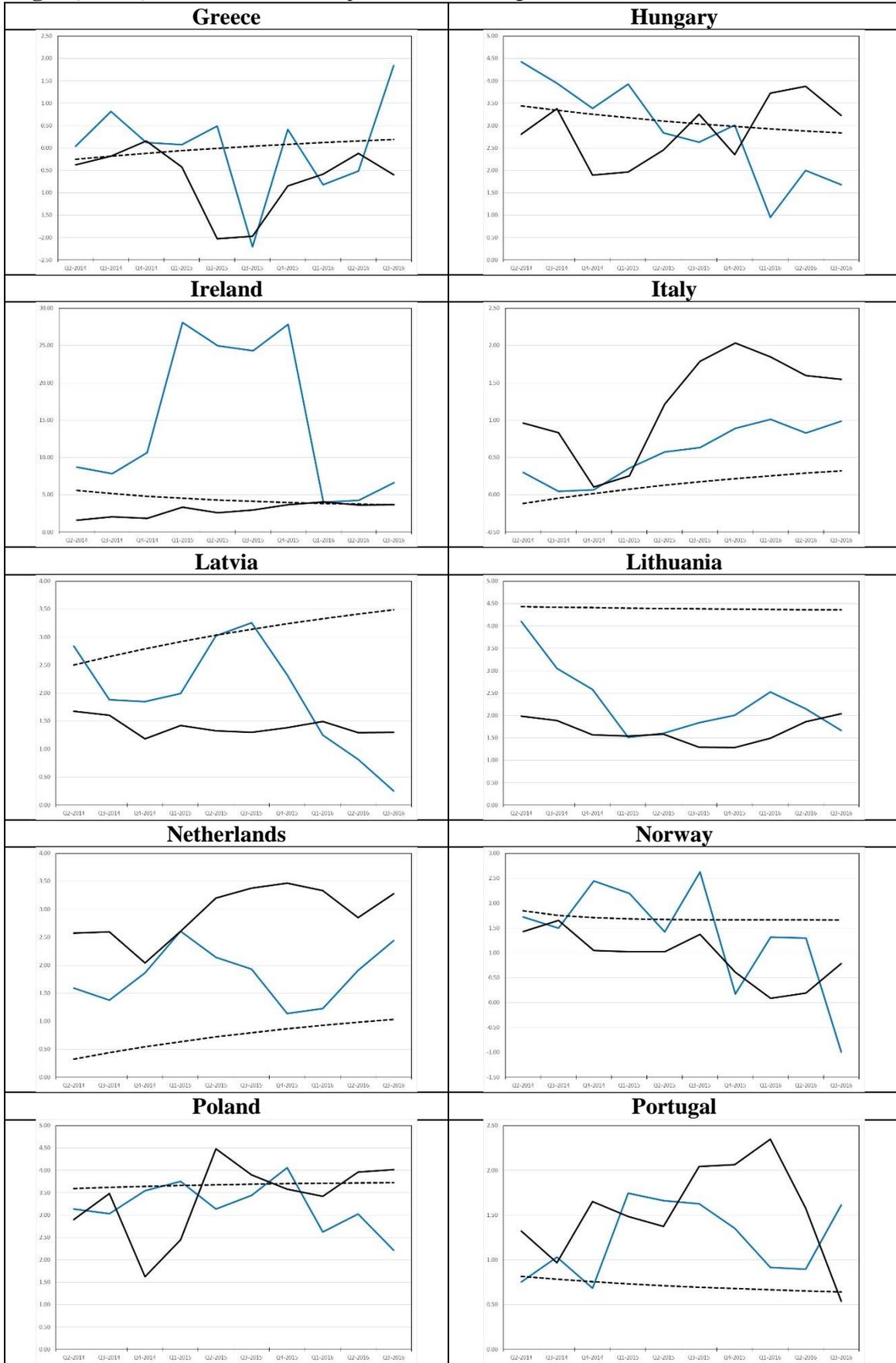
Greece	Hungary

Ireland	Italy

Latvia	Lithuania

Netherlands	Norway

Poland	Portugal

Note: The grey line represents the year-on-year growth rate of GDP in each country. The black line represents the evolution of the proposed evolved economic indicator. The black dotted line shows the evolution of the autoregressive forecasts.

Fig. 1 (cont. 2) Evolution of actual vs. predicted economic growth

Romania	Slovak Republic

Slovenia	Spain

Sweden	United Kingdom

Note: The grey line represents the year-on-year growth rate of GDP in each country. The black line represents the evolution of the proposed evolved economic indicator. The black dotted line shows the evolution of the autoregressive forecasts.

In Fig. 1 we can observe that proposed evolved economic indicators seem to better capture the evolution of GDP growth in most countries. The worst results are obtained in Ireland, which shows a particularly high economic growth during 2015. Therefore, with the aim of graphically assessing whether there are differences in the accuracy of the estimates of economic activity across regions depending on the level of growth, we compute the correlation between actual and predicted economic growth, differentiating between those periods in which economic growth lies out or within the interquartile range (IQR) of the distribution in the European Union. The IQR, also known as midspread, is a measure of statistical dispersion, obtained as the difference between upper and lower quartiles, Q3−Q1.

Fig. 2. Correlations between GDP and expected GDP within and out of IQR
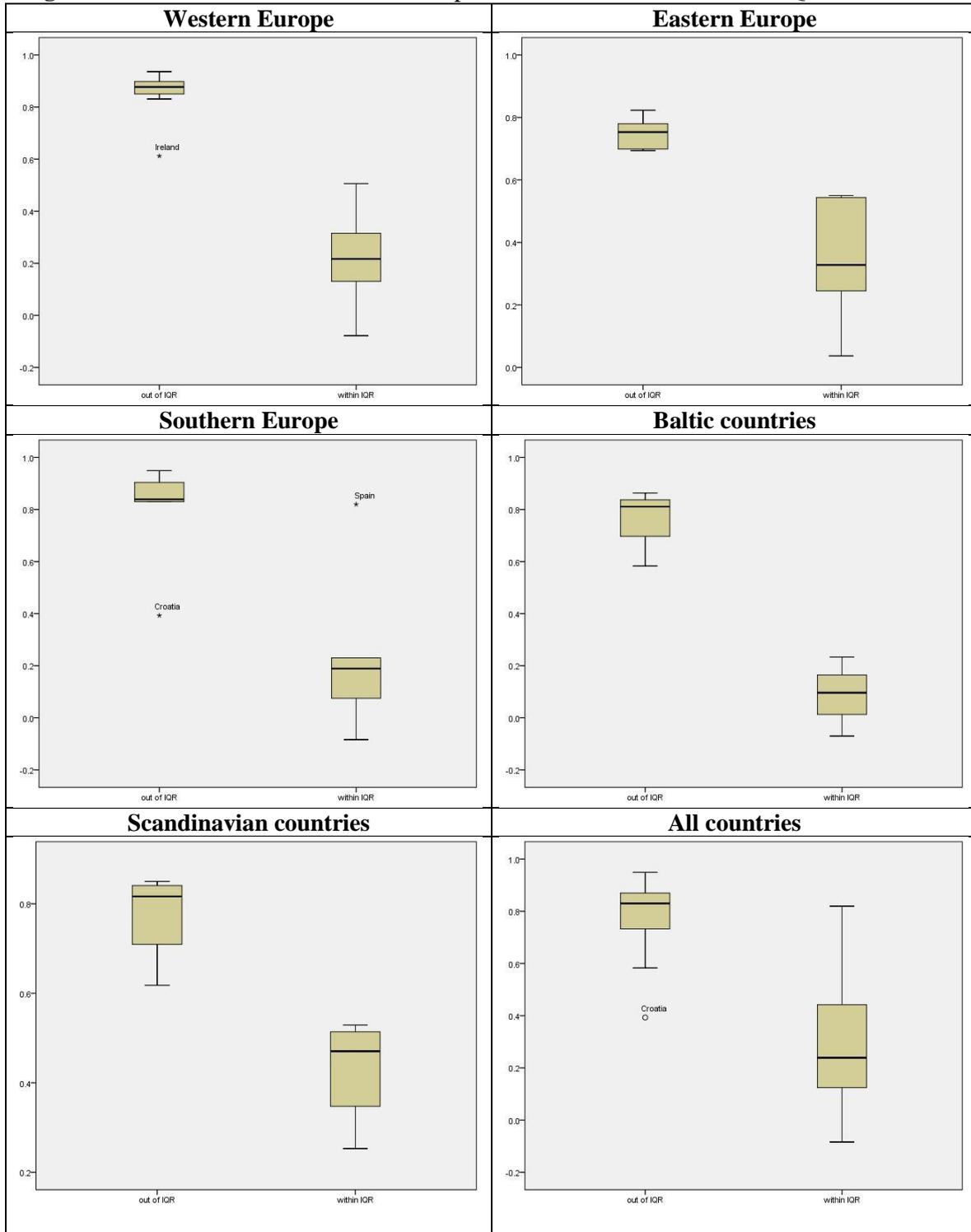
Note: IQR stands for interquartile range.

By discriminating between these two states of growth, we can graphically determine whether there are notable differences in the accuracy of the estimates of economic activity across regions. In Fig. 2 we present the boxplots for each region. We want to note that empirical correlation values in the smaller samples containing the extreme values are likely to be higher than in the subsets containing the remaining larger samples.

In Fig. 2 we can observe that the highest correlations during periods of high growth rates are obtained in Western Europe, with the exception of Ireland. It can also be seen that in all regions the performance of the evolved indicators seems to vary depending on the level of dispersion: during periods of average growth the correlation between estimates and actual values is lower than during periods of high growth rates. These results are in line with those obtained by Łyziak and Mackiewicz-Łyziak (2014), who found that the 2008 financial crisis period had led to a decrease in expectational errors in transition economies. Similarly, Lahiri et al. (2016) found survey-based expectations to have a more pervasive effect on the accuracy of forecasts of all components of aggregate consumption during the recession of 2007–2009.

In the present study we also find evidence regarding the informative value of survey-based expectations. Our results are in line with recent findings by Altug and Çakmakli (2016), Klein and Özmucur (2010), Kłopocka (2017) and Lehmann and Wohlrabe (2017). Altug and Çakmakli (2016) found survey expectations useful to improve inflation forecasts. Klein and Özmucur (2010) found evidence that survey expectations improved the forecasting performance of autoregressive time series models in European countries. Kłopocka (2017) showed the usefulness of survey indicators to forecast household saving and borrowing rates in Poland. Lehmann and Wohlrabe (2017) found that consumers' unemployment expectations and new orders improved predictions of employment growth in Germany.

While there is ample evidence in the literature in favour of the usefulness of expectations to improve the predictive capacity at the macroeconomic level (Batchelor and Dua, 1992, 1998; Batchelor and Orr, 1988; Christiansen et al., 2014; Dees and Brinca, 2013; Girardi, 2014; Hansson et al., 2005; Ivaldi, 1992; Kumar et al., 1995; Leduc and Sill, 2013; Lemmens et al., 2005; Müller, 2009; Qiao et al., 2009; Schmeling and Schrimpf, 2011), several authors have recently proposed refinements in order to enhance the explanatory power of survey expectations in forecasting models. Bruestle and Crain (2015) have showed that controlling for significant versus insignificant changes in consumer confidence improved the accuracy of household expenditure forecasting models. Wilms et al. (2016) have suggested selecting survey indicators from the most predictive industries in order to improve the predictive capacity of survey data. Similarly, Dreger and Kholodilin (2013) have noted that better performing survey-based indicators should be built upon pre-selection methods and data-driven approaches to determine the weights.

In this work we have shown the appropriateness of the SR frame for empirical modelling. SR allows to address complex modelling issues in large data sets where the potential relationships between variables are unknown. In these circumstances, the implementation of SR via evolutionary computation provides researchers with an overview of the most relevant interactions and helps identifying new unknown links between variables. These features make this approach particularly indicated for non-linear modelling.

By means of GP we have simultaneously evolved the structure and the parameters of the models without imposing any a priori assumptions. In this regard, Bruno (2014) has recently noted the importance of avoiding restrictive assumptions about the functional form when modelling using survey indicators. Thus, a SR via GP approach can be of particular interest when it comes to quantify survey expectations, to construct data-driven survey-based indicators or to test economic hypothesis about the formation of agents' expectations.

5 Conclusion

This paper proposes an empirical modelling approach to design survey-based economic indicators at a regional level. By means of SR via GP we find the optimal combination of survey variables that best tracks the evolution of the economic activity in twenty-six European countries grouped in five regions (Western Europe, Eastern Europe, Southern Europe, Baltic countries and Scandinavian countries). This data-driven approach based on evolutionary computation allows us to transform qualitative survey expectations into quantitative estimates of economic activity.

We have used survey variables regarding expectations about the economic situation from the World Economic Survey in order to find the most relevant interactions in each region. This exercise allows us to rank the expectations according to the relative weight of each one in the evolved economic indicators. Although results differ across regions, agents' "perception about the overall economy compared to the same time last year" is the best predictor of economic activity.

In a second step, we assess the out-of-sample forecast accuracy of the evolved survey-based indicators in each region. The best forecasting performance is obtained for Austria and the UK in Western Europe, for Slovakia in Eastern Europe, for Portugal in Southern

Europe, for Lithuania in the Baltic countries, and for Sweden in the Scandinavian countries. At the regional level we obtain the best results for the Baltic and the Scandinavian countries.

Finally, we evaluate if there are differences in the accuracy of the estimates of economic activity across regions depending on the level of growth. We find that during periods of average growth rates the correlation between estimates and actual values is lower in all regions. The highest correlations during periods of high variability are obtained in Western Europe.

In spite of the novelty of the proposed approach, this research is not without limitations. On the one hand, given that we used a data-driven method, the evolved economic indicators are not grounded in any theoretical background. On the other hand, extending the analysis to other survey data would allow us to examine the extent of the similarities in the derived functional forms. Another issue left for further research is testing whether the implementation of alternative algorithms could improve the forecast accuracy of empirically generated quantitative estimates of expectations.

References

Abberger, K. (2007). Qualitative business surveys and the assessment of employment — A case study for Germany. *International Journal of Forecasting* 23(2): 249–258.

Altug, S. and Çakmakli, C. (2016). Forecasting inflation using survey expectations and target inflation: Evidence from Brazil and Turkey. *International Journal of Forecasting* 32(1): 138–153.

Acosta-González, E., and Fernández, F. (2014). Forecasting financial failure of firms via genetic algorithms. *Computational Economics* 43(2): 133–157.

Acosta-González, E., Fernández, F., and Sosvilla, S. (2012). On factors explaining the 2008 financial crisis. *Economics Letters* 115(2): 215–217.

Alexandridis, A. K., Kampouridis, M., and Cramer, S. (2017). A comparison of wavelet networks and genetic programming in the context of temperature derivatives. *International Journal of Forecasting* 33(1): 21–47.

Álvarez-Díaz, M., and Álvarez, A. (2005). Genetic multi-model composite forecast for non-linear prediction of exchange rates. *Empirical Economics* 30(3): 643–663.

Anderson, O. (1952). The Business Test of the IFO-Institute for Economic Research, Munich, and its theoretical model. *Revue de l'Institut International de Statistique* 20: 1–17.

Balcombe, K. (1996). The Carlson-Parkin method applied to NZ price expectations using QSBO survey data. *Economics Letters* 51(1): 51–57.

Banzhaf, W., Nordin, P., Keller, R. E., and Francone, F. D. (2008). *Genetic Programming – An Introduction. On the Automatic Evolution of Computer Programs and its Applications*. San Francisco, CA: Morgan Kaufmann.

Barmpalexis, P., Kachrimanis, K., Tsakonas, A., and Georgarakis, E. (2011). Symbolic regression via genetic programming in the optimization of a controlled release pharmaceutical formulation. *Chemometrics and Intelligent Laboratory Systems* 107(1): 75–82.

Batchelor, R. A. (1981). Aggregate expectations under the stable laws. *Journal of Econometrics* 16(2): 199–210.

Batchelor, R. A. (1982). Expectations, output and inflation: The European experience. *European Economic Review* 17(1): 1–25.

Batchelor, R. A. (1986). Quantitative v. qualitative measures of inflation expectations. *Oxford Bulletin of Economics and Statistics* 48(2): 99–120.

Batchelor, R. and Dua, P. (1992). Survey expectations in the time series consumption function. *The Review of Economics and Statistics* 74(4): 598–606.

Batchelor, R. and Dua, P. (1998). Improving macro-economic forecasts. *International Journal of Forecasting* 14(1): 71–81.

Batchelor, R. and Orr, A. B. (1988). Inflation expectations revisited. *Economica* 55(2019): 317–331.

Bennett, A. (1984). Output expectations of manufacturing industry. *Applied Economics* 16(6): 869–879.

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.

Berk, J. M. (1999). Measuring inflation expectations: A survey data approach. *Applied Economics* 31(11): 1467–1480.

Białowolski, P. (2016). The influence of negative response style on survey-based household inflation expectations. *Quality & Quantity* 50(2): 509–528.

Bovi, M. (2013). Are the representative agent's beliefs based on efficient econometric models?. *Journal of Economic Dynamics & Control* 37(3): 633–648.

Bovi, M. (2016). The tale of two expectations. *Quality & Quantity* 50(6): 2677–2705.

Breitung, J. and Schmeling, M. (2013). Quantifying survey expectations: What's wrong with the probability approach? *International Journal of Forecasting* 29(1): 142–154.

Bruestle, S., and Crain, W. M. (2015). A mean-variance approach to forecasting with the consumer confidence index. *Applied Economics* 47(23): 2430–2444.

Bruno, G. (2014). Consumer confidence and consumption forecast: A non-parametric approach. *Empirica*, 41(1): 37–52.

Carlson, J. A. and Parkin, M. (1975). Inflation expectations. *Economica* 42(166): 123–138.

Ceperic, V., Bako, N. and Baric, A. (2014). A symbolic regression-based modelling strategy of AC/DC rectifiers for RFID applications. *Expert Systems with Applications* 41(16): 7061–7067.

Chen, X., Pang, Y. and Zheng, G. (2010). Macroeconomic forecasting using GP based vector error correction model. In J. Wang (Ed.), *Business Intelligence in Economic Forecasting: Technologies and Techniques* (pp. 1–15). Hershey, PA: IGI Global.

Chen, S. H. and Kuo, T. W. (2002). Evolutionary computation in economics and finance: A bibliography. In S. H. Chen (Ed.), *Evolutionary Computation in Economics and Finance* (pp. 419–455). Heidelberg: Physica-Verlag.

Chen, S. H., Kuo, T. W. and Hoi, K. M. (2008). Genetic programming and financial trading: How much about "what we know". In C. Zopounidis et al. (Eds.), *Handbook of financial engineering* (pp. 99–154). New York: Springer.

Christiansen, C., Eriksen, J. and Moller, S. (2014). Forecasting US recessions: The role of sentiment. *Journal of Banking & Finance*, 49: 459–468.

Claveria, O. (2010). Qualitative survey data on expectations. Is there an alternative to the balance statistic? In A. T. Molnar (Ed.), *Economic Forecasting* (pp. 181–190). Hauppauge, NY: Nova Science Publishers.

Claveria, O., Monte, E. and Torra, S. (2016). Quantification of survey expectations by means of symbolic regression via genetic programming to estimate economic growth in Central and Eastern European economies. *Eastern European Economics* 54(2): 177–189.

Claveria, O., Monte, E. and Torra, S. (2017). A new approach for the quantification of qualitative measures of economic expectations. *Quality & Quantity* 51(6): 2685–2706.

Claveria, O., Pons, E. and Ramos, R. (2007). Business and consumer expectations and macroeconomic forecasts. *International Journal of Forecasting* 23(1): 47–69.

Claveria, O., Pons, E. and Suriñach, J. (2006). Quantification of expectations. Are they useful for forecasting inflation? *Economic Issues* 11(2): 19–38.

Common, M. (1985). Testing for rational expectations with qualitative survey data. *Manchester School of Economic and Social Statistics* 53(2): 138–148.

Cramer, N. (1985). A representation for the adaptive generation of simple sequential programs. *Proceedings of the International Conference on Genetic Algorithms and their Applications*, 24-26 June. Pittsburgh, PA.

Dabhi, V. K. and Chaudhary, S. (2015). Empirical modeling using genetic programming: A survey of issues and approaches. *Natural Computing* 14(2): 303–330.

Dees, S. and Brinca, P. S. (2013). Consumer confidence as a predictor of consumption spending: Evidence for the United States and the Euro area. International Economics 134: 1–14.

Drake, A. E. and Marks, R. E. (2002). Genetic algorithms in economics and finance: Forecasting stock market prices and foreign exchange – A review. In S. H. Chen (Ed.), *Genetic Algorithms and Genetic Programming in Computational Finance* (pp. 29–54). New York: Springer.

Dreger, C. and Kholodilin, D. (2013). Forecasting private consumption by consumer surveys. *Journal of Forecasting*, 32(1): 10–18.

Driver, C. and Urga, G. (2004). Transforming qualitative survey data: Performance comparisons for the UK. *Oxford Bulletin of Economics and Statistics* 66(1): 71–89.

Duda, J. and Szydło, S. (2011). Collective intelligence of genetic programming for macroeconomic forecasting. In P. Jędrzejowicz et al. (Eds.), *Computational Collective Intelligence. Technologies and Applications* (pp. 445–454). Berlin: Springer.

Ferreira, C. (2001). Gene expression programming: A new adaptive algorithm for solving problems. *Complex Systems* 13(2): 87–129.

Fogel, D. B. (2006). *Evolutionary computation. Toward a new philosophy of machine intelligence* (Third Edition). Hoboken, NJ: John Wiley & Sons.

Fogel, L. J., Owens, A. J. and Walsh, M. J. (1966). *Artificial intelligence through simulated evolution*. New York: John Wiley.

Fortin, F. A., De Rainville, F. M., Gardner, M. A., Parizeau, M. and Gagné, C. (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13(1): 2171–2175.

Franses, P. H., Kranendonk, H. C. and Lanser, D. (2011). One model and various experts: Evaluating Dutch macroeconomic forecasts. *International Journal of Forecasting* 27(2): 482–495.

Gandomi, A. H. and Roke, D. (2015). Assessment of artificial neural network and genetic programming as predictive tools. *Advances in Engineering Software* 88: 63–72.

Garnitz, J., Nerb, G. and Wohlrabe, K. (2015). CESifo World Economic Survey – November 2015. *CESifo World Economic Survey* 14(4): 1–28.

Girardi, A. (2014). Expectations and macroeconomic fluctuations in the euro area. *Economics Letters* 125(2): 315–318.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Boston, MA: Addison-Wesley.

Gong, Y. J., Chen, W. N., Zhan, Z. H., Zhang, J., Li, Y., Zhang, Q. and Li, J. J. (2015). Distributed evolutionary algorithms and their models: A survey of the stat-of-the-art. *Applied Soft Computing* 34: 286–300.

Graff, M. (2010). Does a multi-sectoral design improve indicator-based forecasts of the GDP growth rate? Evidence from Switzerland. *Applied Economics* 42(21): 2759–2781.

Guizzardi, A. and Stacchini, A. (2015). Real-time forecasting regional tourism with business sentiment surveys. *Tourism Management* 47: 213–223.

Hansson, J., Jansson, P. and Löf, M. (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting* 30(1): 65–77.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press (1975).

Hutson, M., Joutz, F. and Stekler, H. (2014). Interpreting and evaluating CESIfo's World Economic Survey directional forecasts. *Economic Modelling*, 38, 6–11.

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.

Ivaldi, M. (1992) Survey evidence on the rationality of expectations. *Journal of Applied Econometrics* 7(3): 225–241.

Jean-Baptiste, F. (2012). Forecasting with the new Keynesian Phillips curve: Evidence from survey data. *Economics Letters*, 117(3) 811–813.

Jonsson, T. and Österholm, P. (2011). The forecasting properties of survey-based wage-growth expectations. *Economics Letters* 113(3): 276–281.

Jonsson, T. and Österholm P. (2012). The properties of survey-based inflation expectations in Sweden. *Empirical Economics* 42(1): 79–94.

Kaboudan, M. A. (2000). Genetic programing prediction of stock prices. *Computational Economics*, 16(3), 207–236.

Klein L. R. and Özmucur, S. (2010). The use of consumer and business surveys in forecasting. *Economic Modelling*, 27(6), 1453–1462.

Klúčik, M. (2012). Estimates of foreign trade using genetic programming. *Proceedings of the 46 the scientific meeting of the Italian Statistical Society*.

Kłopocka, K. (2017). Does consumer confidence forecast household saving and borrowing behavior? Evidence for Poland. Social Indicators Research 133(2): 693–717.

Kotanchek, M. E, Vladislavleva, E. Y. and Smits, G. F. (2010). Symbolic regression via genetic programming as a discovery engine: Insights on outliers and prototypes. In R. Riolo et al., (Eds.), *Genetic Programming Theory and Practice VII, Genetic and Evolutionary Computation* Vol. 8 (pp. 55–72). Springer Science+Business Media, LLC.

Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MA: MIT Press.

Kronberger, G., Fink, S., Kommenda, M. and Affenzeller, M. (2011). Macro-economic time series modeling and interaction networks. In C. Di Chio et al. (Eds.), *EvoApplications, Part II* (pp. 101–110). LNCS 6625.

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.

Kumar, V., Leone, R. and Gaskins, J. (1995). Aggregate and disaggregate sector fore- casting using consumer confidence measures. *International Journal of Forecasting*, 11 (3): 361–377.

Lacová, Ž. and Král, P. (2015). Measurement and characteristics of enterprise inflation expectations in Slovakia. *Procedia Economics and Finance* 30: 505–512.

Lahiri, K. and Teigland, C. (1987). On the normality of probability distributions of inflation and GNP forecasts. *International Journal of Forecasting* 3(2): 269–279.

Lahiri, K. and Zhao, Y. (2015). Quantifying survey expectations: A critical review and generalization of the Carlson-Parkin method. *International Journal of Forecasting* 31(1): 51–62.

Lahiri, K., Monokroussos, G. and Zhao, Y. (2016). Forecasting consumption: The role of consumer confidence in real time with many predictors. *Journal of Applied Econometrics* 31(7): 1254–1275.

Larkin, F. and Ryan, C. (2008). Good news: Using news feeds with genetic programming to predict stock prices. In M. O'Neil et al. (Eds.), *Genetic Programming* (pp. 49–60). Berlin: Springer-Verlag.

Lawrenz, C. and Westerhoff, F. (2003). Modeling exchange rate behaviour with a genetic algorithm. *Computational Economics* 21(3): 209–229.

Leduc, S. and Sill, K. (2013). Expectations and economic fluctuations: An analysis using survey data. *The Review of Economic and Statistics* 95(4): 1352–1367.

Lee, K. C. (1994). Formation of price and cost inflation expectations in British manufacturing industries: A multi-sectoral analysis. *The Economic Journal* 104(423): 372–385.

Lehmann, R., and Wohlrabe, K. (2017). Experts, firms, consumers or even hard data? Forecasting employment in Germany. *Applied Economics Letters* 24(4): 279–283.

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

Löffler, G. (1999). Refining the Carlson-Parkin method. *Economics Letters* 64(2): 167–71.

Lui, S., Mitchell, J. and Weale, M. (2011a). The utility of expectational data: firm-level evidence using matched qualitative-quantitative UK surveys. *International Journal of Forecasting* 27(4): 1128–1146.

Lui, S., Mitchell, J. and Weale, M. (2011b). Qualitative business surveys: signal or noise? *Journal of The Royal Statistical Society, Series A (Statistics in Society)* 174(2): 327–348.

Łyziak, T. and Mackiewicz-Łyziak, J. (2014). Do consumers in Europe anticipate future inflation? *Eastern European Economics* 52(3): 5–32.

Maag, T. (2009). On the accuracy of the probability method for quantifying beliefs about inflation. KOF Working Papers, No. 230. Zurich: KOF Swiss Economic Institute.

Martinsen, K., Ravazzolo, F. and Wulfsberg, F. (2014). Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting* 30(1): 65–77.

Maschek, M. K. (2010). Intelligent mutation rate control in an economic application of genetic algorithms. *Computational Economics* 35(1): 25–49.

Martinsen, K., Ravazzolo, F. and Wulfsberg, F. (2014). Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting* 30(1): 65–77.

Miah, F., Rahman, M. S. and Albinali, K. (2016). Rationality of survey based inflation expectations: A study of 18 emerging economies' inflation forecasts. *Research in International Business and Finance* 36: 158–166.

Mitchell, J., Smith, R. and Weale, M. (2002). Quantification of qualitative firm-level survey data. *Economic Journal* 112(478): 117–135.

Mitchell, J., Smith, R. and Weale, M. (2005a). Forecasting manufacturing output growth using firm-level survey data. *The Manchester School* 73(4): 479–499.

Mitchell, J., Smith, R. and Weale, M. (2005b). An indicator of monthly GDP and an early estimate of quarterly GDP growth. *The Economic Journal* 115(501): F108–F129.

Mittnik, S. and Zadrozny, P. (2005). Forecasting quarterly German GDP at monthly intervals using monthly IFO business conditions data. In J. E. Sturm and T. Wollmershäuser (Eds.), *IFO survey data in business cycle analysis and monetary policy analysis* (pp. 19–48). Heidelberg: Physica-Verlag.

Mokinski, F., Sheng, X. and Yang, J. (2015). Measuring disagreement in qualitative expectations. *Journal of Forecasting* 34(5): 405–426.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica* 29(3): 315–335.

Müller, C. (2010). You CAN Carlson-Parkin. *Economics Letters* 108(1): 33–35.

Nardo, M. (2003). The quantification of qualitative data: a critical assessment. *Journal of Economic Surveys* 17(5): 645–668.

Nardo, M. and Cabeza-Gutés, M. (1999). The role of measurement error in rational expectations testing. UAB Working Paper 451. Barcelona: Universitat Autònoma de Barcelona.

Nolte, I. and Pohlmeier, W. (2007). Using forecasts of forecasters to forecast. *International Journal of Forecasting* 23(1): 15–28.

Paloviita, M. (2006). Inflation dynamics in the euro area and the role of expectations. *Empirical Economics* 31: 847–860.

Parigi, G. and Schlitzer, G. (1995). Quarterly forecasts of the Italian business-cycle by means of monthly economic indicators. *Journal of Forecasting* 14(2): 117–141.

Peng, Y., Yuan, C., Qin, X., Huang, J. and Shi, Y. (2014). An improved gene expression programming approach for symbolic regression problems. *Neurocomputing* 137: 293–301.

Pesaran, M. H. (1985). Formation of inflation expectations in British manufacturing industries. *The Economic Journal* 95(380): 948-975.

Pesaran, M. H. (1987). *The limits to rational expectations*. Oxford: Basil Blackwell.

Pesaran, M. H. and Weale, M. (2006). Survey expectations. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 1 (pp. 715–776). Amsterdam: Elsevier North- Holland.

Poli, R., Vanneschi, L., Langdon, W. B. and Mcphee, N. F. (2010). Theoretical results in genetic programming: The next ten years? *Genetic Programming and Evolvable Machines* 11(3): 285–320.

Qiao, Z., McAleer, M. and Wong, W. K. (2009). Linear and nonlinear causality between changes in consumption and consumer attitudes. *Economic Letters* 102(3): 161–164.

Robinzonov, N., Tutz, G., and Hothorn, T. (2012). Boosting techniques for nonlinear time series models. *AStA Advances in Statistical Analysis* 96(1): 99–122.

Sarradj, E. and Geyer, T. (2014). Symbolic regression modeling of noise generation at porous airfoils. *Journal of Sound and Vibration* 333(14): 3189–3202.

Schmeling, M. and Schrimpf, A. (2011). Expected inflation, expected stock returns, and money illusion: What can we learn from survey expectations. *European Economic Review* 55(5): 702–719.

Seitz, H. (1988). The estimation of inflation forecasts from business survey data. *Applied Economics* 20(4): 427–38.

Smith, J. and McAleer, M. (1995). Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing. *Journal of Applied Econometrics* 10(2): 165–185.

Terai, A. (2009). Measurement error in estimating inflation expectations from survey data: an evaluation by Monte Carlo simulations. *Journal of Business Cycle Measurement and Analysis* 8(2): 133–156.

Theil, H. (1952). On the time shape of economic microvariables and the Munich Business Test. *Revue de l'Institut International de Statistique* 20: 105–20.

Thinyane, H. and Millin, J. (2011). An investigation into the use of intelligent systems for currency trading. *Computational Economics* 37(4): 363–374.

Vasilakis, G. A., Theofilatos, K. A., Georgopoulos, E. F., Karathanasopoulos, A., and Likothanassis, S. D. (2013). A genetic programming approach for EUR/USD exchange rate forecasting and trading. *Computational Economics* 42(4): 415–431.

Vladislavleva, E.; Smits, G. and den Hertog, D. (2010). On the importance of data balancing for symbolic regression. *IEEE Transactions in Evolutionary Computation* 14(2): 252–277.

Vermeulen, P. (2014). An evaluation of business survey indices for short-term forecasting: Balance method versus Carlson-Parkin method. *International Journal of Forecasting* 30(4): 882–897.

Visco, I. (1984). *Price expectations in rising inflation*. Amsterdam: North-Holland.

Wei, L. Y. (2013). A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. *Economic Modelling* 33: 893–899.

Wilms, I., Gelper, S. and Croux, C. (2016). The predictive power of the business and bank sentiment of firms: A high-dimensional Granger Causality approach. *European Journal of Operational Research* 254(1): 138–147.

Wilson, G. and Banzhaf, W. (2009). Prediction of interday stock prices using developmental and linear genetic programming. In M. Giacobini et al. (Eds.), *Applications of Evolutionary Computing* (pp. 172–181). Berlin: Springer-Verlag.

Wren-Lewis, S. (1986). An econometric model of U.K. manufacturing employment using survey data on expected output. *Journal of Applied Econometrics* 10(2): 165–185.

Wu, C. H., Chou, H. J. and Su, W. H. (2008). Direct transformation of coordinates for GPS positioning using the techniques of genetic programming and symbolic regression. *Engineering Applications of Artificial Intelligence* 21(8): 1347–1359.

Yang, G., Li, X., Wang, J., Lian, L. and Ma, T. (2015). Modeling oil production based on symbolic regression. *Energy Policy* 82(1): 48–61.

Yao, L. and Lin, C. C. (2009). Identification of nonlinear systems by the genetic programming-based volterra filter. *IET Signal Processing* 3(2): 93–105.

Yu, T., Chen, S. and Kuo, T. W. (2004). A genetic programming approach to model international short-term capital flow. *Applications of Artificial Intelligence in Finance and Economics* 19: 45–70.

Zameer, A., Arshad, J., Khan, A., and Raja, M. A. Z. (2017). Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks. *Energy Conversion and Management* 134: 361–372.

Zelinka, I., Oplatkova, Z., and Nolle, L. (2005). Analytic programming – Symbolic regression by means of arbitrary evolutionary algorithms. *International Journal of Simulation: Systems, Science and Technology* 6(9): 44–56.