USING SURVEY DATA TO FORECAST REAL ACTIVITY WITH EVOLUTIONARY ALGORITHMS. A CROSS-COUNTRY ANALYSIS

OSCAR CLAVERIA*
University of Barcelona

ENRIC MONTE
Polytechnic University of Catalunya

SALVADOR TORRA
University of Barcelona

Submitted April 2016; accepted November 2016

Business and consumer surveys are the main source of agents' expectations. In this study we use survey expectations about a wide range of economic variables to forecast real activity. We propose an empirical approach to derive mathematical functional forms that link survey expectations to economic growth. Combining symbolic regression with genetic programming we generate two survey-based indicators: a perceptions index, using agents' assessments about the present, and an expectations index with their expectations about the future. Our examination of the forecast accuracy of both indicators to track the evolution of economic activity in fourteen European countries indicates that the perceptions index always outperforms the expectations index, although the improvements of the perceptions index against the naïve forecasts used as a benchmark are only significant in Austria. When assessing the effect of the 2008 financial crisis on the forecasting performance we find an improvement in accuracy during the crisis, which may be in part caused by a decrease of disagreement among respondents during periods prior to turning points. In order to find the optimal combination of both indexes that best replicates the evolution of economic activity in each country we use a portfolio management procedure known as index tracking. By means of a generalized reduced gradient algorithm we derive the relative weights of both indexes. In most economies, the survey-based predictions generated with the composite indicator outperform the benchmark model for one-quarter ahead forecasts, although these improvements are only significant in Austria, Belgium and Portugal.

***JEL classification codes*:** C51, C55, C63, C83, C93

***Key words*:** business and consumer surveys, forecasting, economic growth, symbolic regression, evolutionary algorithms, genetic programming

* Oscar Claveria (corresponding author): AQR-IREA (Institute of Applied Economics Research), University of Barcelona, Diagonal, 690, 08034 Barcelona, Spain; e-mail oclaveria@ub.edu. Enric Monte: Department of Signal Theory and Communications, Polytechnic University of Catalunya, Jordi Girona, 1-3, 08034 Barcelona, Spain; e-mail enric.monte@upc.edu. Salvador Torra: Riskcenter-IREA, University of Barcelona, Diagonal, 690, 08034 Barcelona, Spain; e-mail storra@ub.edu. This research was supported by by the project ECO2016-75805-R from the Spanish Ministry of Economy and Competitiveness. We would like to thank the Editor and three anonymous reviewers for their useful comments and suggestions. We are indebted to Johanna Garnitz and Klaus Wohlrabe at the Ifo Institute for Economic Research in Munich for providing us with the data.

I. Introduction

Economic expectations are central in macroeconomic time series modelling. Tendency surveys provide detailed information about agents' expectations, but the qualitative nature of agents' responses has led to quantify survey results. Numerous methods to transform responses about the expected direction of change into a quantitative measure of agents' expectations have been proposed in the literature. See Lahiri and Zhao (2015), Vermeulen (2014) and Nardo (2003) for an appraisal of the different quantification methods. The theoretical framework for quantifying survey expectations is based on the assumption 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). Carlson and Parkin (1975) developed this probability approach by using a normal distribution. Mitchell (2002) and Balcombe (1996) found evidence that normal distributions provide expectations as accurate as other stable distributions.

Several refinements of the probabilistic approach have been proposed in order to reduce the measurement error introduced by restrictive assumptions (Breitung and Schmeling, 2013; Mitchell et al., 2007; Löffler, 1999; Berk, 1999; Smith and McAleer, 1995; Seitz, 1988; Batchelor and Orr, 1988; Pesaran, 1987; Batchelor, 1986). By comparing the individual responses with firm-by-firm realizations, Müller (2010) developed a variant of the Carlson-Parkin method with asymmetric and time invariant thresholds. In a recent study, Lahiri and Zhao (2015) linked quantified expectations to quantitative realizations at the firm-level, and obtained a significant improvement in accuracy by allowing for cross-sectional heterogeneity and asymmetric and time-varying thresholds. This improvement was found to be especially relevant during periods of uncertainty with high levels of disagreement between respondents.

This result has led us to evaluate the degree to which survey data on both perceptions and expectations fit the real outcome after the 2008 financial crisis. The relationship between changes in expectations and economic variables has been widely investigated (Martinsen et al., 2014; Ghonghadze and Lux, 2012; Lui et al., 2011a,b; Schmeling and Schrimpf, 2011; Franses et al., 2011; Graff, 2010; Klein and Özmucur, 2010; Claveria et al., 2006, Hansson et al., 2005), but never before by means of symbolic regression (SR). SR can be regarded as an empirical modelling approach, which is particularly indicated to find the most fitting algebraic expression in large data sets, especially when the model structure is unknown or changes over time.

By combining a SR approach with genetic programming (GP), we are able to quantify survey-based expectations in order to generate estimates of economic growth. There are different strategies for finding a solution in SR. Koza (1992) developed GP to implement SR. In spite of its versatility, GP applications in economics are still few (Acosta-González et al., 2012; Álvarez-Díaz and Álvarez, 2005).

In this study we use survey indicators from fourteen different European countries to generate two economic indicators: a perceptions index with agents' assessments about the present economic situation, and an expectations index with their expectations about the future. By linking survey data from the CESifo World Economic Survey (WES) to economic growth in two successive GP experiments we are able to derive an analytical expression for each index. Then we evaluate the forecasting performance of the indexes since the beginning of the crisis. In a second step, we use a generalized reduced gradient algorithm to find the optimal combination of weights for each index that best replicates the evolution of real activity in each country. These weights allow us to design a composite indicator, which we use to forecast economic growth.

We aim to break new ground by presenting a new approach to derive data-driven economic indicators. The proposed methodology is based on evolutionary computation, which through Darwinian competition allows to generate a mathematical functional form that approximates a predefined target variable. The resulting algebraic expressions can be regarded as the fittest empirically-generated combinations of survey variables.

The structure of the paper is as follows. The next section reviews the existing literature on SR via GP. In Section III we present the methodological approach and describe the experiment. Empirical results are provided in Section IV. Finally, conclusions are given in Section V.

II. Methodology

In this study we design two SR experiments that link survey expectations to real activity in order to derive two economic indicators. This data-driven regression approach assumes no model a priori. Using evolutionary algorithms (EAs) that imitate aspects of biological evolution, such as the principle of survival and reproduction of the fittest, an initial population of computer programs are bred through generations to find a set of analytical functions that best fit the data.

As opposed to evolutionary programming (Fogel, 1966), in which the structure of the program to be evolved remains fixed, GP simultaneously evolves the structure and the parameters of the models. Koza (1995) applied GP to assess the non-linear empirical relationship between price level, gross national product, money supply, and the velocity of money. GP is a soft computing search technique for problem-solving. GP's tree-structured programs are evolved by means of genetic operators for model approximation. Dabhi and Chaudhary (2015) have reviewed the main issues related to GP. The versatility of this empirical modelling approach has attracted researchers from different areas (Sarradj and Geyer, 2014; Ceperic et al., 2014; Barmpalexis et al., 2011; Can and Heavey, 2011; Vladislavleva et al., 2010; Yao and Lin, 2009; Wu et al., 2008; Cai et al., 2006). See Chen and Kuo (2002) for a classification of the literature on the application of evolutionary computation to economics and finance.

Acosta-González and Fernández (2014) used a genetic algorithm (GA) to forecast the financial failure of firms. Vasilakis et al. (2013) presented a GP-based technique to predict returns in the trading of the euro/dollar exchange rate based on historical data. Wei (2013) used an adaptive expectation GA to optimize a fuzzy model to forecast stock price trends in Taiwan. Thinyane and Millin (2011) used GAs to optimize the signals generated by technical trading tools. Larkin and Ryan (2008) applied GP to predict stock prices using ordinal sentiment data.

Based upon its performance in eight stock markets and eight foreign exchange markets during three consecutive test periods, Chen et al. (2008) thoroughly analysed the application of GP to financial trading, shedding some light on how GP performance could be connected to the trending and cyclical properties of financial data. Álvarez-Díaz and Álvarez (2005) used GP to forecast exchange rates of the yen and the pound to the US dollar. Yu et al. (2004) implemented a GP approach to model international short-term capital flows. Lawrenz and Westerhoff (2003) modelled exchange rates with a GA. Kaboudan (2000) used GP to forecast stock prices. For a review of the applications of GAs for financial forecasting see Drake and Marks (2002).

There have been few applications of GP in economics. Duda and Szydło (2011) applied an improved version of GP known as gene expression programming (GEP), proposed by Ferreria (2011), to develop a set of economic forecasting models. Chen et al. (2010) introduced GP in a vector error correction model for macroeconomic forecasting. By means of SR via Pareto GP, Kotanchek et al. (2010) provided some insight into Gross Domestic Product (GDP) forecasting. Kronberger et al. (2011) used SR to identify variable interactions between economic indicators in order to estimate the evolution of prices in the US. Kľúčik (2012) used SR via GP in the estimation of total exports and imports to Slovakia. Yang et al. (2015) used SR to predict oil production.

Among recent developments in evolutionary computation, Zelinka et al. (2005) introduced analytical programming (AP), and showed its ability to synthesize suitable solutions in SR. Wilson and Banzhaf (2009) compared a developmental co-evolutionary GP approach to standard linear GP for interday stock prices prediction. Maschek (2010) developed a self-adaptation mechanism and evaluated how it affected an economic application of GAs. 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 ANN models to that of GEP techniques. See Poli et al. (2010) for a review of the state of the art in GP.

III. Econometric design

GP allows finding patterns in large data sets. This feature is particularly suitable where little or no information is known about the system, as in the current study, where there is an arbitrary and unknown functional relationship between the set of survey variables. Therefore we use GP to formalize the interactions between a wide and heterogeneous range of survey-based agents' expectations that best fit the evolution of economic activity. More specifically, by means of SR we link twelve survey-based indicators from the CESifo's WES (Table 1) to year-on-year growth rates of quarterly GDP data from the OECD (https://data.oecd.org/gdp/quarterly-gdp.htm#indicator-chart). The sample period goes from the third quarter of 2000 to the first quarter of 2014.

The WES questions focus on the direction of change of a wide range of economic variables (see Kudymowa et al. 2013, Hutson et al. 2014, and Garnitz et al. 2015 for an appraisal of the WES). The individual replies are combined for each country without weighting, giving a grade of 9 to positive replies, a grade of 5 to indifferent replies and a grade of 1 to negative replies (CESifo World Economic Survey, 2016). As a result, grades within the range of 5 to 9 are indicative of a majority expecting an increasing trend in the variable, revealing a predominant positive perception. The opposite holds true for grades within the range of 1 to 5. In Table 2 we present a descriptive analysis of the twelve survey variables used in the study for fourteen European countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands (NL), Portugal, Spain, Sweden and the United Kingdom (UK).

Table 1. World Economic Survey (WES) – Survey indicators

Perceptions Present	Perceptions Compared to last year	Expectations For the next six months
Economic situation	Economic situation	Economic situation and foreign trade volume
$x1$ overall economy	$x4$ overall economy	$x7$ overall economy
$x2$ capital expenditures	$x5$ capital expenditures	$x8$ capital expenditures
$x3$ private consumption	$x6$ private consumption	$x9$ private consumption
		$x10$ volume of exports
		$x11$ volume of imports
		$x12$ trade balance

Table 2. WES variables – Descriptive analysis

	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$	$x7$	$x8$	$x9$	$x10$	$x11$	$x12$
Austria	5.138	4.552	4.796	5.248	4.957	5.155	5.414	5.277	5.364	5.984	5.921	5.045
	(1.89)	(1.92)	(1.29)	(2.20)	(2.07)	(1.73)	(1.65)	(1.49)	(1.15)	(1.50)	(1.22)	(1.14)
Belgium	4.539	3.941	4.675	4.777	4.677	4.750	5.700	5.554	5.443	5.789	5.934	5.030
	(1.93)	(1.92)	(1.77)	(2.37)	(2.19)	(2.08)	(1.51)	(1.54)	(1.21)	(1.51)	(1.20)	(0.94)
Denmark	5.800	5.270	4.932	5.393	4.952	5.225	5.621	5.275	5.611	6.138	5.991	4.875
	(2.14)	(1.89)	(2.53)	(2.01)	(1.67)	(2.11)	(1.47)	(1.29)	(1.47)	(1.76)	(1.64)	(1.24)
Finland	6.084	4.620	6.714	5.207	4.732	5.350	5.754	5.357	5.152	6.713	6.488	5.720
	(2.03)	(1.92)	(1.71)	(2.30)	(2.14)	(1.83)	(1.57)	(1.52)	(1.24)	(1.73)	(1.61)	(1.30)
France	3.763	3.511	4.757	4.438	4.389	4.334	5.591	5.486	5.023	5.523	5.734	4.463
	(1.77)	(1.67)	(1.71)	(2.31)	(1.98)	(1.76)	(1.34)	(1.28)	(1.17)	(1.20)	(1.05)	(0.88)
Germany	4.943	4.584	4.080	5.545	5.313	5.375	5.993	5.811	6.007	6.332	6.343	5.654
	(2.40)	(2.31)	(2.02)	(2.48)	(2.36)	(1.78)	(1.50)	(1.46)	(1.23)	(1.31)	(1.11)	(1.02)
Greece	4.071	3.721	3.877	3.682	3.538	3.230	5.005	5.021	4.288	5.648	4.782	5.093
	(2.37)	(2.15)	(1.93)	(1.83)	(1.64)	(1.41)	(1.36)	(1.02)	(1.05)	(1.82)	(2.13)	(2.21)
Ireland	5.241	4.582	5.161	4.939	4.427	4.843	5.393	4.991	5.373	6.309	6.014	5.729
	(3.02)	(2.63)	(2.92)	(2.42)	(1.97)	(2.15)	(1.84)	(1.45)	(1.70)	(1.78)	(1.52)	(1.63)
Italy	2.975	2.916	2.884	4.700	4.411	4.316	5.870	5.584	5.604	6.121	5.716	5.214
	(1.43)	(1.35)	(1.13)	(1.97)	(1.75)	(1.68)	(1.10)	(0.98)	(1.05)	(1.14)	(1.11)	(1.03)
NL	4.527	4.089	4.166	4.961	4.680	4.663	6.086	5.920	5.693	6.275	6.189	5.446
	(2.37)	(2.16)	(2.33)	(2.51)	(2.19)	(2.29)	(1.67)	(1.57)	(1.51)	(1.55)	(1.35)	(1.05)
Portugal	2.275	2.264	2.495	3.770	3.600	3.311	5.364	5.302	4.516	6.202	4.439	5.939
	(1.25)	(1.14)	(1.15)	(2.25)	(1.64)	(1.80)	(1.88)	(1.60)	(1.57)	(1.64)	(1.75)	(1.76)
Spain	4.070	3.641	4.205	3.941	3.896	3.734	4.663	4.709	4.143	5.463	5.086	5.080
	(2.47)	(2.05)	(2.54)	(1.77)	(1.50)	(1.61)	(1.25)	(1.01)	(1.17)	(1.40)	(1.31)	(1.33)
Sweden	6.011	5.170	6.396	5.400	5.023	5.654	5.377	5.232	5.286	5.839	6.071	4.882
	(1.98)	(1.90)	(1.45)	(2.59)	(2.32)	(2.23)	(1.59)	(1.53)	(1.45)	(1.60)	(1.48)	(1.04)
UK	4.743	3.691	5.004	4.823	4.409	4.743	5.193	5.107	4.598	6.202	5.923	4.791
	(1.94)	(1.60)	(1.96)	(2.07)	(1.61)	(1.83)	(1.66)	(1.48)	(1.40)	(1.23)	(1.29)	(1.40)

Note: Descriptive statistics – Mean, and Standard deviation in brackets. NL stands for the Netherlands.

We design two independent experiments. Both experiments consist on a SR modelling strategy to find the optimal combination of survey variables to estimate the evolution of economic growth. The first experiment evolves the combination of agents' perceptions about the present economic situation (variables $x1$ to $x6$) that more accurately approximates year-on-year growth rates of quarterly GDP for country i at time t. As a result, we derive a functional expression that will be referred to as the "perception index" $\left(\hat{y}_{1,it}\right)$. In a second experiment we search for the optimal combination of survey variables regarding agents' expectations about the future (variables $x7$ to $x12$) in order to track the evolution of real activity. The evolved symbolic expression will be referred to as the "expectations index" $\left(\hat{y}_{2,it}\right)$. We run both experiments in the fourteen European countries simultaneously.

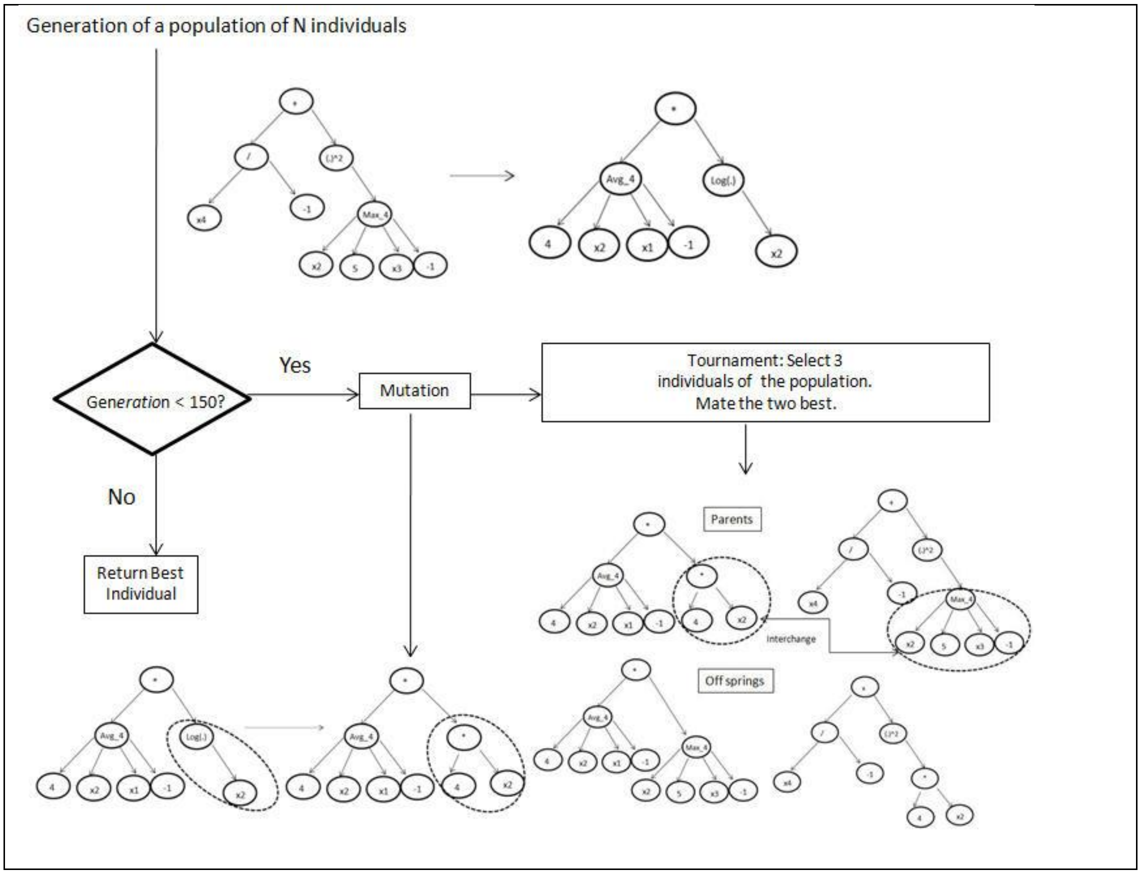
By means of SR we derive the combination of survey variables that best fits economic growth in each experiment. We then assess the forecasting performance of both indicators to track the evolution of GDP, prior, during and after the 2008 financial crisis. In Table 3 we present a detailed description of the parameters of the experiment. To limit the complexity of the resulting expressions, the set of functions is restricted to the mean, the maximum, the minimum, the ratio, and the logarithm. Regarding the termination criterion, we set a maximum number of 150 generations. The determination of the maximum number of generations is done in a heuristic way, with the aim of guaranteeing that the system converges the predetermined minimum error. See Figure 1 for a graphical description of the experiment. We use the Distributed Evolutionary Algorithms Package (DEAP) framework implemented in Python (Fortin et al. 2012; Gong et al. 2015).

Table 3. Description of the experiment

Initial population	3,000,000
Normal population	500,000
Max. generations	150
Selection of individuals	Tournament size =3
Replacement	1-Elitism
Initialization	Select 1000 best of random sample of size 2000
Crossover	Sub-tree-swapping
Mutation prob.	0,1, with a random subtree of depth 2
Tree constraints	Dynamic depth limit (initial limit = 7)
Model selection	Best on validation
Stopping criterion	max. Generations
Fitness function	RMSE
Function set	{+, -, *, /, avg_4, log(.)sign(.), (.)^2, sqrt(.)sign(.), max_4(.), min_4(.)}
Terminal	Set constants={0,5-1,10,5}, variables

Note: RMSE – stands for root mean square error; avg – stands for average; log – stands for logarithm; sqrt – stands for square root; max – stands for maximum; min – stands for minimum.

Figure 1. Design of the SR experiment



IV. Results

In this section, we first present the output of the two SR experiments undertaken. Expression (1) is the evolved perceptions index, which represents the optimal combination of agents' perceptions about the present to track economic activity:

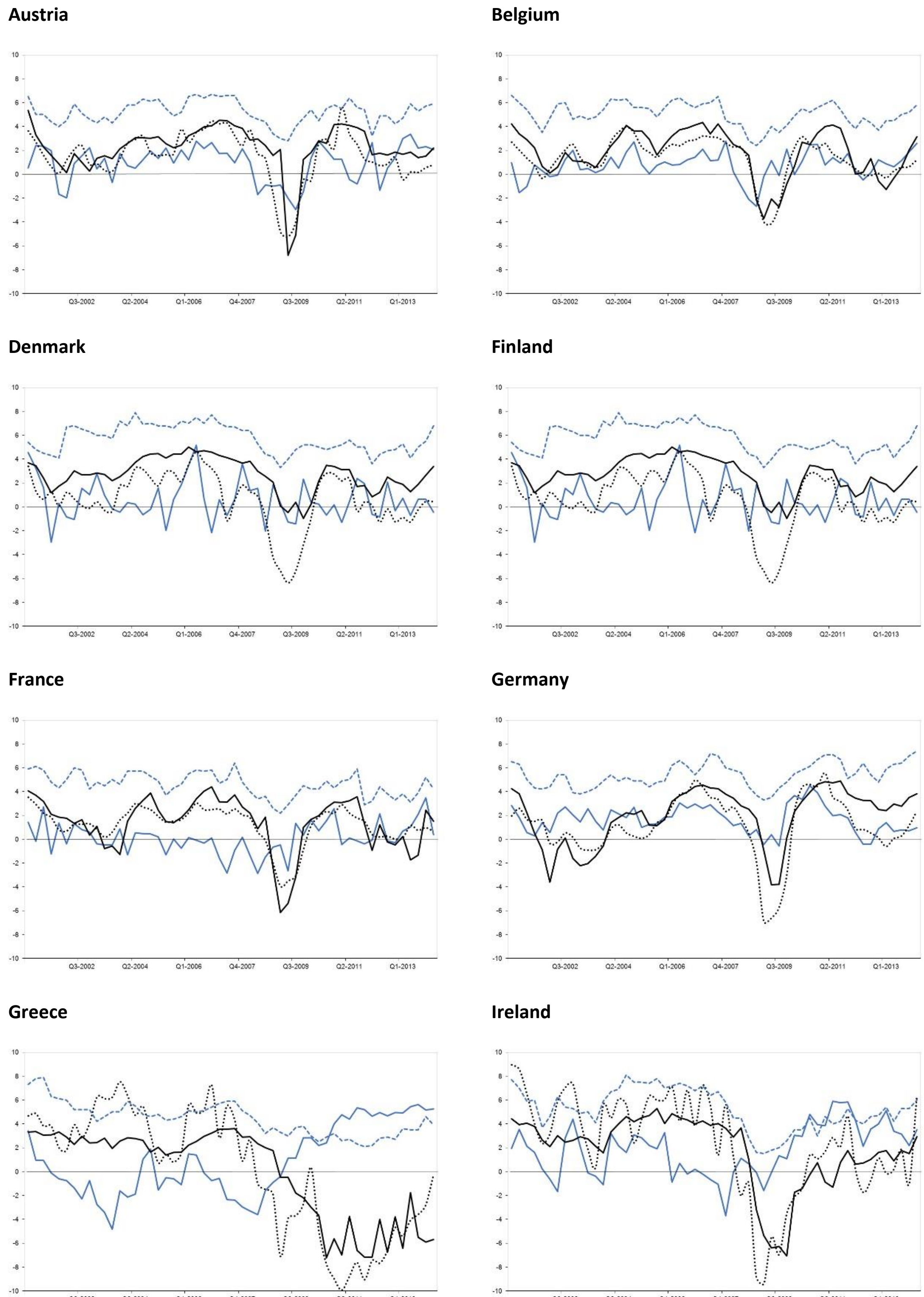
$$\hat{y}_{1,it} = \frac{1 + x4_{it} + x3_{it} + x5_{it} + \dfrac{10\left(\left[\log\left(x1_{it} + x2_{it}\right)\right]^2 - 12\right)}{\sqrt{x3_{it} - 1}\left(x4_{it} - 1\right) + 1}}{4} - 11 \quad , \tag{1}$$

Expression (2) presents the expectations index, which show the evolved optimal combination of survey variables regarding agents' expectations about the future economic situation:

$$\hat{y}_{2,it} = \frac{\log\left[10 * \log\left(\dfrac{x7_{it} + \sqrt{x10_{it}} + x10_{it} + x9_{it}}{4} \Big/ \left(\sqrt{x12_{it} - x8_{it}} + 1\right)\right) - 1\right]^2}{x12_{it}} - 12 \quad , \tag{2}$$

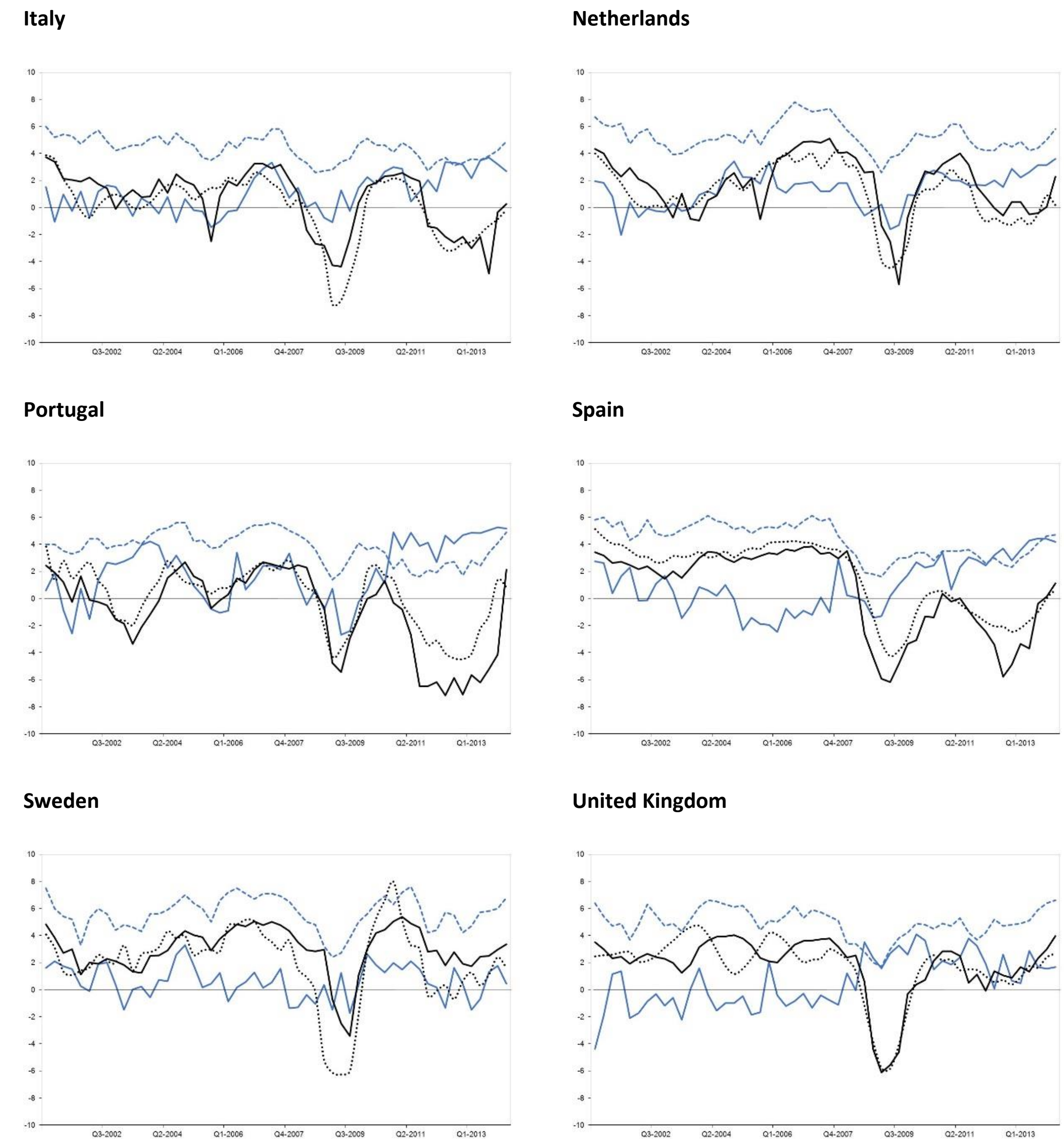
One of the main advantages of using survey data with forecasting purposes is that survey results are available before the GDP release (Klein and Özmucur, 2010). The publication delay of quarterly GDP data and survey data varies widely across countries, but the major European economies publish their quarterly GDP data within about 100 days after the end of a quarter, and survey data within less than 20 days after. Survey results are also used for the design of economic indicators. The Ifo Institute uses WES results to construct the Economic Climate Index (ECI), which is an aggregate indicator obtained as the arithmetic mean of assessments of the general economic situation and the expectations for the economic situation in the next six months. In Figure 2 we graphically compare the evolution of the two SR-generated indicators to that of the ECI and the GDP.

Figure 2. Evolution of year-on-year GDP growth rates vs. survey-based economic indexes ($\hat{y}_{1,it}$ vs. $\hat{y}_{2,it}$)



Note: The black dotted line represents the year-on-year growth rate of GDP in each country. The black line represents the evolution of the proposed perceptions index. The grey line represents the evolution of the proposed expectations index. The grey dotted line represents the evolution of the ECI in each country.

Figure 2 (cont.)**. Evolution of year-on-year GDP growth rates vs. survey-based economic indexes (** $\hat{y}_{1,it}$ **vs.** $\hat{y}_{2,it}$ **)**



Note: See note of Figure 2.

Figure 3. Cross-correlations – Perceptions ($\hat{y}_{1,it}$) and expectations index ($\hat{y}_{2,it}$) vs. GDP

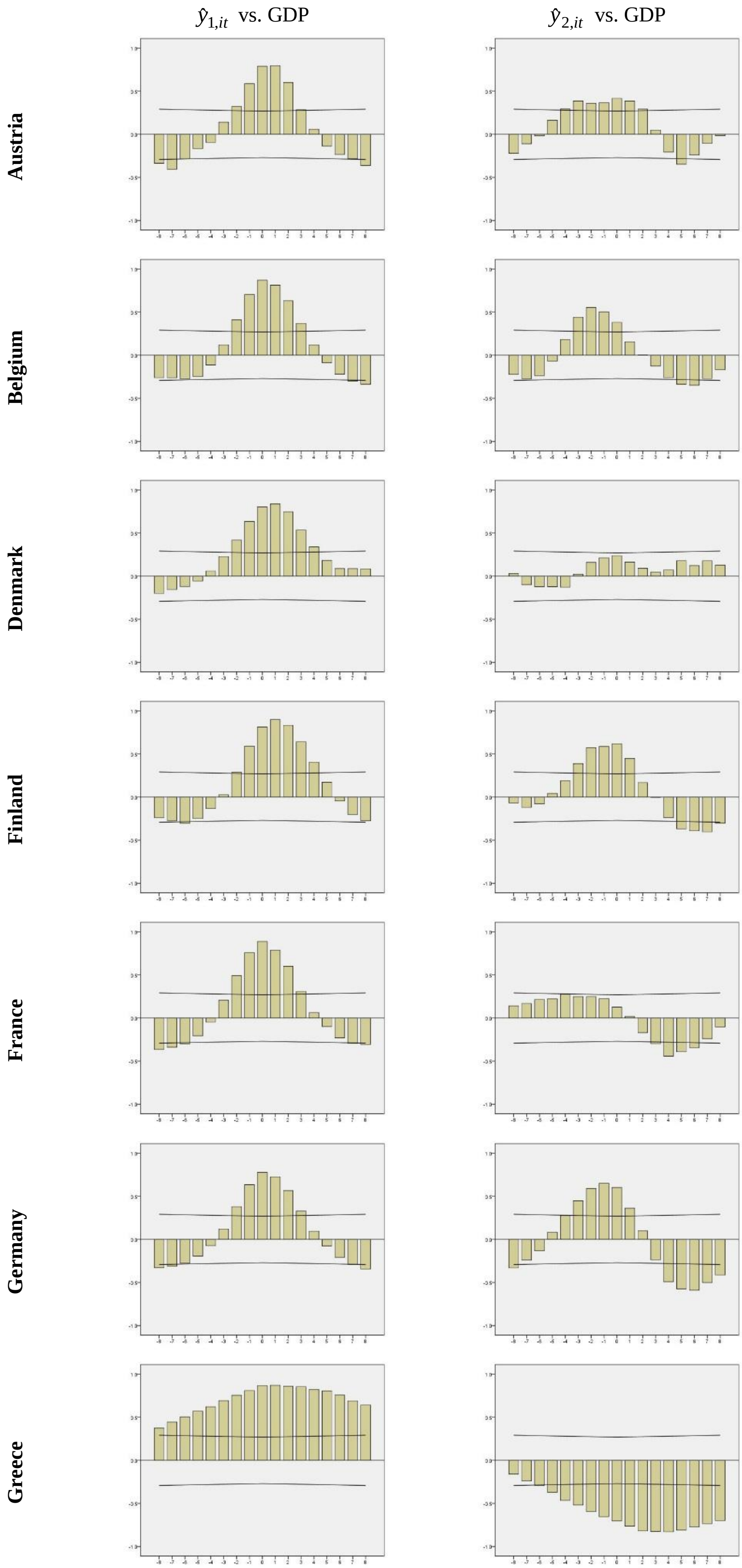
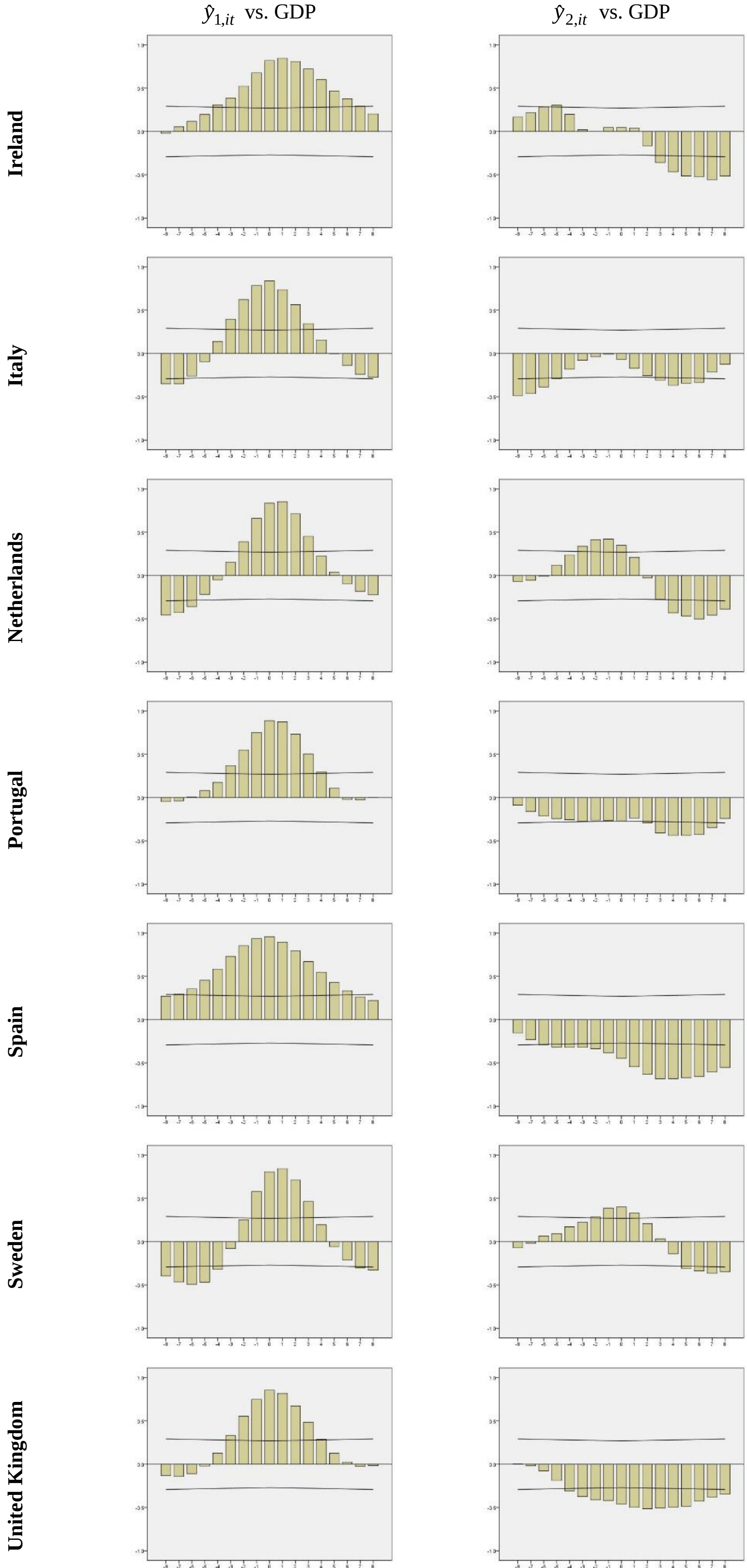


Figure 3 (cont.). **Cross-correlations – Perceptions ($\hat{y}_{1,it}$) and expectations index ($\hat{y}_{2,it}$) vs. GDP**



In Figure 3 we present the cross-correlations between both SR-generated indicators (the perceptions and the expectations indexes) and the evolution of GDP. We can observe that the perceptions index is coincident with GDP in most countries (Austria, Belgium, France, Germany, Greece, Italy, Portugal, Spain and the UK), but it lags one period in Denmark, Finland, Ireland, the Netherlands and Sweden. Regarding the expectations index, it leads one quarter in Germany and the Netherlands, and two quarters in Belgium; it coincides in Austria, Denmark, Finland and Sweden; and lags in Greece, Ireland, Italy, Portugal, Spain and the United Kingdom.

In Table 4 we present the results of several forecast accuracy measures to evaluate the forecasting performance of both SR-generated indicators. Apart from the mean absolute error (MAE) and the root mean square error (RMSE), we complement the forecast accuracy analysis by computing the mean absolute scaled error (MASE) proposed by Hyndman and Koehler (2006). The MASE scales the errors by the mean absolute errors obtained with a random walk. As survey data refer to expectations, and are available ahead of the publication of quantitative official data, we use two-step ahead naïve forecasts as a benchmark.

The MASE statistic presents several advantages over other forecast accuracy measures. First, it is independent of the scale of the data. Second, it does not suffer from some of the problems presented by other relative measures of forecast accuracy (Hyndman and Koehler, 2006). The MASE is also easy to interpret: values larger than one are indicative that the GP-based forecasts are worse than the average prediction computed with the benchmark model. If we denote the forecast error obtained by means of GP as $e_t = Y_t - \hat{Y}_t$, the scale error is defined as:

$$MASE = mean \left| e_t \Big/ \frac{1}{n-2} \sum_{i=3}^{n} \left| Y_i - Y_{i-2} \right| \right| \quad , \tag{3}$$

To test whether the reduction in MAE is statistically significant between the best three models, we also compute the Diebold-Mariano (DM) statistic of predictive accuracy (Diebold and Mariano, 1995). The null hypothesis of the test is that the difference between the two competing series is non-significant. A negative sign of the statistic implies that the second model has bigger forecasting errors.

As it could be expected, the perceptions index shows a better performance than the expectations index in all economies. Nevertheless, in both cases the magnitude of the obtained errors is not negligible. We also find remarkable differences across countries. Belgium is the country with the lowest MAE and RMSE values for both the leading and the coincident indicator. In the case of the perceptions index, France and the UK also obtain the lowest MAE and RMSE values together with Belgium. In the other extreme, Greece is the economy with the least accurate predictions, followed by Ireland, Denmark, and Finland.

When comparing the obtained results with those of the benchmark, we find that the perceptions index yields lower forecasting errors than the benchmark in all countries except Denmark, Greece and Spain, but this difference is only significant in Austria. The opposite is observed in the case of the expectations index, which shows higher forecasting errors in all countries with the exception of Germany. The reduction in MAE is significant in five countries (France, Greece, Ireland, Portugal, Spain and the UK).

These results provide mixed evidence on the usefulness of survey-based expectations for forecasting purposes. While Altug and Çakmakli (2016), Guizzardi and Stacchini (2015), Altavilla et al. (2014), Hutson et al. (2014), Österholm (2014), Martinsen et al. (2014), Girardi (2014), Dees et al. (2013), Jean-Baptiste (2012), Ghonghadze and Lux (2012), Klein and Özmucur (2010), Mitchell et al. (2005) and Hansson et al. (2005) have found that survey-based expectations provide useful information for forecasting purposes, Lehmann (2015), Breitung and Schmeling (2013), Robinzonov et al. (2012), Jonsson and Österholm (2011), Lui et al. (2011a,b), Claveria et al. (2007) and Batchelor and Dua (1992, 1998) have also obtained mixed results.

In order to evaluate the effect of the dispersion of economic growth on forecast results, we compare the different forecast accuracy measures to the standard deviation of the year-on-year growth rates of GDP for each country. In Figure 4 we present the scatterplots for the perceptions index, and in Figure 5 the scatterplots for the expectations index. We can observe a positive relation between the MAE and the standard deviation of GDP growth rates. This is not the case for the MASE, which is a relative measure of forecast accuracy. In the case of the perceptions index, we obtain the highest forecast errors for Finland, Denmark, Greece and Ireland. In the other extreme, Belgium is the economy with the lowest values. Germany is the only country with a MASE lower than one for the expectations index, while Spain the one with the highest MASE value. In Figure 5 we can observe that Greece and Ireland are the countries with the highest forecast errors and GDP dispersion.

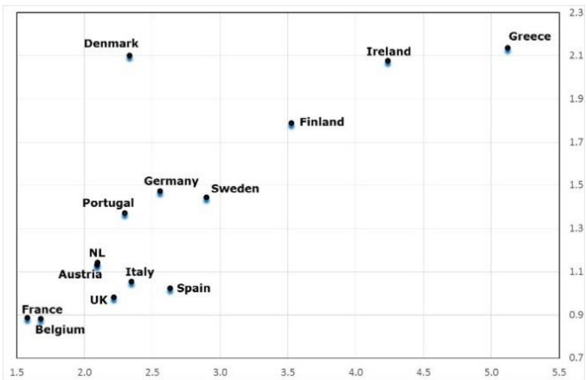
Table 4. Forecast accuracy – MAE, RMSE and MASE

$\hat{y}_{1,it}$	MAE	RMSE	MASE	DM
Austria	1.131	1.527	0.756	**-2.388**
Belgium	0.882	1.100	0.738	-1.364
Denmark	2.100	2.553	1.381	1.905
Finland	1.789	2.688	0.873	-0.755
France	0.885	1.129	0.912	-0.475
Germany	1.473	1.865	0.910	-0.425
Greece	2.136	2.627	1.071	0.554
Ireland	2.076	2.582	0.816	-1.679
Italy	1.053	1.407	0.751	-1.436
Netherlands	1.141	1.377	0.969	-0.094
Portugal	1.370	1.808	0.907	-0.522
Spain	1.023	1.306	1.208	1.099
Sweden	1.444	2.044	0.799	-1.107
UK	0.980	1.206	0.766	-1.029
$\hat{y}_{2,it}$	MAE	RMSE	MASE	DM
Austria	1.723	2.115	1.137	0.499
Belgium	1.384	1.739	1.160	0.999
Denmark	1.964	2.543	1.280	1.705
Finland	2.124	2.844	1.023	0.263
France	1.844	2.214	1.894	**3.035**
Germany	1.550	2.157	0.966	-0.092
Greece	6.477	7.517	3.224	**4.988**
Ireland	4.136	4.705	1.600	**2.971**
Italy	2.236	2.995	1.572	2.017
Netherlands	1.672	2.055	1.406	1.698
Portugal	3.020	4.013	1.977	**2.442**
Spain	3.567	3.920	4.152	**6.166**
Sweden	2.419	2.995	1.338	1.524
UK	3.147	3.781	2.498	**5.037**

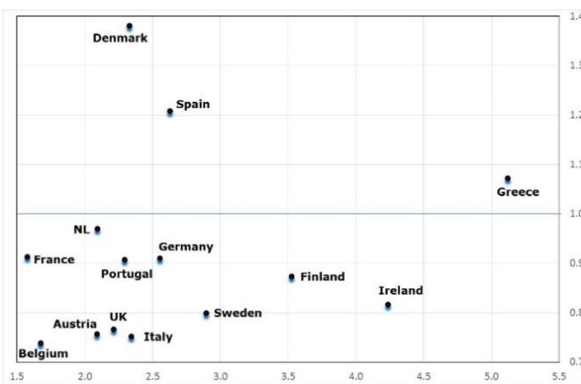
Note: MAE stands for mean absolute error; RMSE stands for root mean square error; MASE stands for mean absolute scaled error; DM stands for Diebold-Mariano test statistic with NW estimator. Null hypothesis of the test: the difference between the two competing series is non-significant. A negative sign of the statistic implies that the second model (naïve) has bigger forecasting errors. The 5% level critical value is 2.028. Significant values in bold.

Figure 4. Forecast accuracy vs. Standard deviation of GDP – Perceptions index

MAE



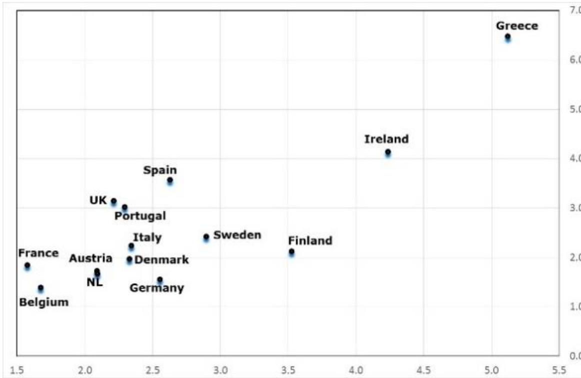
MASE



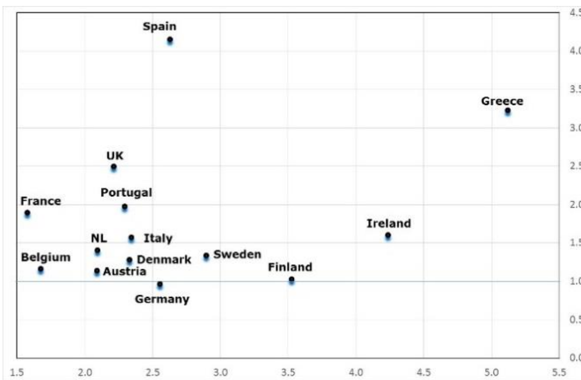
Note: The X axis shows shows the standard deviation of GDP growth. The Y axis shows the forecast accuracy measures. MAE stands for mean absolute error; MASE stands for mean absolute scaled error.

Figure 5. Forecast accuracy vs Standard deviation of GDP – Expectations index

MAE



MASE



Note: See note of Figure 4.

Łyziak and Mackiewicz-Łyziak (2014) found that the 2008 financial crisis period led to a decrease in expectational errors in transition economies. Claveria et al. (2016) obtained a similar result for ten Eastern European countries. To analyze whether the 2008 financial crisis has had an influence on the forecast accuracy of survey-based measures of economic expectations, in Table 5 we evaluate the forecasting performance of the SR-generated indicators to that of the benchmark, differentiating between the pre-crisis sub-period (2000-2007), the crisis (2007-2010), and the post-crisis sub-period.

Table 5. Forecast accuracy – MASE and DM loss-differential test statistic

$\hat{y}_{1,it}$	Pre-crisis	Crisis	Post-crisis	$\hat{y}_{2,it}$	Pre-crisis	Crisis	Post-crisis
Austria	0.705	0.698	0.999		1.234	0.815	1.728
	(**-2.308**)	(**-2.017**)	(0.000)		(0.207)	(-0.684)	(**2.527**)
Belgium	0.760	0.389	2.100		1.181	1.089	1.403
	(-1.590)	(**-2.237**)	(**2.292**)		(0.831)	(0.318)	(0.970)
Denmark	1.419	1.123	2.302		1.722	0.947	1.543
	(1.553)	(0.426)	(**2.766**)		(**2.852**)	(-0.222)	(1.415)
Finland	0.709	0.831	1.225		1.174	0.915	1.164
	(**-2.577**)	(-0.643)	(0.742)		(0.399)	(-0.264)	(0.425)
France	0.969	0.588	1.817		2.263	1.490	2.254
	(-0.240)	(-1.609)	(**2.107**)		(**3.633**)	(1.028)	(**2.250**)
Germany	1.005	0.577	1.762		1.350	0.792	0.921
	(0.006)	(-1.704)	(**2.465**)		(1.446)	(-0.7329	(-0.317)
Greece	1.229	0.748	1.607		3.025	1.746	7.772
	(0.968)	(-1.341)	(1.971)		(**3.493**)	(**3.010**)	(**9.140**)
Ireland	0.892	0.663	0.965		1.667	1.523	1.618
	(-0.756)	(**-2.180**)	(-0.130)		(1.867)	(**2.155**)	(**2.944**)
Italy	1.026	0.606	0.719		1.570	0.939	3.192
	(0.294)	(-1.646)	(-1.060)		(1.264)	(-0.220)	(**3.443**)
NL	1.170	0.744	1.223		1.355	1.029	2.519
	(0.998)	(-1.176)	(0.843)		(0.954)	(0.103)	(**2.440**)
Portugal	0.849	0.480	1.536		2.000	0.604	3.740
	(-0.746)	(**-2.752**)	(**2.486**)		(**2.497**)	(-1.826)	(**5.345**)
Spain	2.290	0.801	1.262		9.462	1.834	5.189
	(**5.355**)	(-1.963)	(0.474)		(**6.251**)	(1.828)	(**6.430**)
Sweden	0.624	0.815	1.007		2.214	1.156	0.800
	(**-2.676**)	(-0.634)	0.023		(**2.493**)	(0.469)	(-0.689)
UK	1.169	0.364	1.238		3.924	1.696	1.630
	(1.083)	(**-2.435**)	(1.420)		(**9.371**)	(**2.096**)	(1.614)

Note: See notes of Table 4. DM between brackets. NL stands for the Netherlands.

In Table 5 we can observe that a common feature between both indexes is the improvement in relative forecast accuracy during the crisis. This result is partly due to the fact that surveys are based on the report of subjective evaluations of participants prior to the publication of official economic data, which are able to incorporate information about the current state of the economy as well as forward-looking information (Altavilla et al., 2014). The decline of forecast errors during the crisis is also related to the decrease of disagreement among experts responding to the survey during periods prior to turning points. Dovern (2015) found that variations of overall disagreement among professional forecasters are driven by economic uncertainty. Mokinski et al. (2015) provided a detailed review and assessment of the measurement of disagreement in qualitative survey data.

These findings are in line with the results obtained by Kauppi et al. (1996), Klein and Özmucur (2010), Dees et al. (2013), Łyziak and Mackiewicz-Łyziak (2014) and Claveria et al. (2016). Kauppi et al. (1996) found that the importance of business survey information increased during recession periods, obtaining a significant improvement in prediction during Finland's great depression. Dees et al. (2013) showed that the contribution of survey indicators in explaining consumption expenditures increased during periods presenting huge changes. Klein and Özmucur (2010) and Claveria et al. (2016) also found that the significance of survey results in forecasting increased at times of greater uncertainty. Łyziak and Mackiewicz-Łyziak (2014) quantified inflation expectations without imposing their unbiasedness and then evaluated the forecasting accuracy of quantified measures of expectations. All these authors found that survey data became more relevant after the beginning of the 2008 financial crisis.

Nevertheless, we find differences across countries regarding the different patterns over the three sub-periods. Spain and the UK are the only countries in which there is an improvement in the relative forecast accuracy of both indexes in the post-crisis period with respect to the pre-crisis period. In the rest of economies, the forecast accuracy of SR-based predictions with respect to the benchmark deteriorates after the crisis, especially for the perceptions index. Austria and Ireland are the only countries where agents' perceptions are more accurate than the predictions obtained with the benchmark model for all three sub-periods. Conversely, Denmark is the only country in which the benchmark model is not outperformed regardless of the sub-period.

With the objective of combining the information of both indexes we use a procedure of constrained optimization to find the optimal weights of both indicators. This procedure is used for portfolio management to replicate the performance of a stock index, and is known as index tracking. See Kwiatkowski (1992) and Rudd (1980) for a discussion. For a detailed description of new techniques applied to index tracking see Karlow (2012). The aim of index tracking is to minimise a tracking error, understood as the expected squared deviation of return from that of the index, in order to obtain the proportion of capital to be invested in each company. Based on this premise, we use a generalized reduced gradient algorithm to minimize the summation of squared forecast errors subject to two constraints: the weights must be equal or larger than zero (non-negativity restriction), and the sum of both weights equals one. This procedure allows us to obtain the relative weights of each index for each country (Table 6).

Table 6. Relative weights of perceptions ($\hat{y}_{1,it}$) and expectations index ($\hat{y}_{2,it}$)

	$\hat{y}_{1,it}$	$\hat{y}_{2,it}$
Austria	0.677	0.323
Belgium	0.701	0.299
Denmark	1.000	0.000
Finland	0.587	0.413
France	0.734	0.266
Germany	0.619	0.381
Greece	1.000	0.000
Ireland	1.000	0.000
Italy	0.939	0.061
Netherlands	0.738	0.262
Portugal	0.735	0.265
Spain	0.887	0.113
Sweden	0.795	0.205
UK	0.919	0.081

While the obtained relative weight of the perceptions index is always higher than that of the expectations index, we observe numerous differences across countries. In Denmark, Greece and Ireland, the algorithm yields a null weight to the expectations index, while in countries such as Austria, Finland and Germany we obtain weights close or superior to a third. This result brings up the question of whether survey-based indicators shall equally weight the information regarding the expectations about the future and the perceptions about the present.

Finally, we compute a composite economic indicator by combining both the perceptions and the expectations indexes according to the weights in Table 6. In Table 7 we present the MASE results obtained with the composite indicator. While in most countries there is an improvement with respect to the benchmark, this improvement is only significant in Austria, Belgium and Portugal.

Table 7. Forecast accuracy by country – Composite indicator

	MASE	DM		MASE	DM
Austria	0.639	**-2.903**	Ireland	0.841	-1.663
Belgium	0.567	**-2.785**	Italy	0.871	-0.812
Denmark	1.029	0.280	Netherlands	0.824	-1.094
Finland	0.831	-0.924	Portugal	0.676	**-2.427**
France	0.715	-1.567	Spain	1.221	1.099
Germany	0.895	-0.425	Sweden	0.745	-1.562
Greece	1.052	0.554	UK	0.734	-1.189

Note: See notes of Table 4.

V. Concluding remarks

This paper proposes an empirical approach to derive indicators of economic growth from qualitative survey responses about the state of the economy in fourteen European countries by means of SR via GP. We use survey-based agents' assessments about the present economic situation from the WES to derive a perceptions index which consists of the optimal combination of variables that best tracks the evolution of the economic activity. We repeat the experiment using agents' expectations about the future economic situation to obtain an expectations index.

We analyze the forecasting performance of both indexes, and we find that the perceptions index yields more accurate estimates of the evolution of GDP than the expectations index, although these improvements are not significant. With the aim of analyzing the impact of the 2008 financial crisis on agents' expectations, we assess the capacity of SR-generated expectations to anticipate future economic growth, prior, during, and after the financial crisis. We find an improvement in relative forecast accuracy during the crisis and a subsequent deterioration in the forecasting performance of agents' expectations in all countries after the crisis.

In order to combine the information from both indicators we use a constrained optimization procedure known as index tracking to find the optimal relative weights of both the perceptions and the expectations indexes. By doing so, we generate a composite indicator of economic activity that yields more accurate forecasts than the ones obtained separately with both indexes in all countries except Spain.

Due to the novelty of this approach, there are still several limitations to be addressed. As we use a data-driven method, the obtained indicator lacks any theoretical background. By extending the analysis to other questionnaires and countries, we could examine to what extent the obtained functional forms are sensitive to different survey data. Another question to be considered in further research is whether the implementation of alternative evolutionary algorithms may improve the forecast accuracy of empirically-generated indicators.

References

Altavilla, Carlo, Raffaella Giacomini, and Giuseppe Ragusa (2014). Anchoring the yield curve using survey expectations. European Central Bank Working Paper Series, No. 1632.

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

Acosta-González, Eduardo, Fernando Fernández-Rodríguez, and Simón Sosvilla-Rivero (2012). On factors explaining the 2008 financial crisis. *Economics Letters* 115: 215–217.

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

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

Barmpalexis, Panagiotis, Kyriakos Kachrimanis, Athanasios Tsakonas, and Emmanouel Georgarakis (2011). Symbolic regression via genetic programming in the optimization of a controlled release pharmaceutical formulation. *Chemometrics and Intelligent Laboratory Systems* 107: 75–82.

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

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

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

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

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

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

Cai, Weihua, Arturo Pacheco-Vega, Mihir Sen, and Kwang-Tzu Yang (2006). Heat transfer correlations by symbolic regression. *International Journal of Heat and Mass Transfer* 49: 4352–4359.

Can, Birkan, and Cathal Heavey (2011). Comparison of experimental designs for simulation-based symbolic regression of manufacturing systems. *Computers & Industrial Engineering* 61: 447–462.

Carlson, John A., and Michael Parkin (1975). Inflation expectations. *Economica* 42: 123–138.

CESifo World Economic Survey (2016). Volume 15, no. 3, August 2016.

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

Chen, Xi, Ye Pang, and Guihuan Zheng (2010). Macroeconomic forecasting using GP based vector error correction model. In J. Wang, editor, *Business intelligence in economic forecasting: Technologies and techniques* (pp. 1–15). Hershey, PA: IGI Global.

Chen, Shu-Heng, and Tzu-Wen Kuo (2002). Evolutionary computation in economics and finance: A bibliography. In S. H. Chen, editor, *Evolutionary computation in economics and finance* (pp. 419–455). Heidelberg: Physica-Verlag.

Chen, Shu-Heng, Tzu-Wen Kuo, and Kong-Mui Hoi (2008). Genetic programming and financial trading: How much about "what we know". In C. Zopounidis et al., editors, *Handbook of financial engineering* (pp. 99–154). New York: Springer.

Claveria, Oscar, Enric Monte, and Salvador Torra (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: 177–189.

Claveria, Oscar, Ernest Pons, and Raul Ramos (2007). Business and consumer expectations and macroeconomic forecasts. *International Journal of Forecasting* 23: 47–69.

Claveria, Oscar, Ernest Pons, and Jordi Suriñach (2006). Quantification of expectations. Are they useful for forecasting inflation? *Economic Issues* 11: 19–38.
Dabhi, Vipul K., and Sanjay Chaudhary (2015). Empirical modeling using genetic programming: A survey of issues and approaches. *Natural Computing* 14: 303–330.
Dees, Stephane, and Pedro Soares Brinca (2013). Consumer confidence as a predictor of consumption spending: Evidence for the United States and the Euro area. *International Economics* 134: 1–14.
Diebold, Francis X., and Roberto S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13: 253–263.
Dovern, Jonas (2015). A multivariate analysis of forecast disagreement: Confronting models of disagreement with survey data. *European Economic Review* 80: 16–35.
Drake, Adrian E., and Robert E. Marks (2002). Genetic algorithms in economics and finance: Forecasting stock market prices and foreign exchange –A review. In S. H. Chen, editor, *Genetic algorithms and genetic programming in computational finance* (pp. 29–54). New York: Springer.
Duda, Jerzy, and Stanisław Szydło (2011). Collective intelligence of genetic programming for macroeconomic forecasting. In P. Jędrzejowicz et al., editors, *Computational collective intelligence. Technologies and applications* (pp. 445–454). Berlin: Springer.
Ferreira, Cândida (2001). Gene expression programming: A new adaptive algorithm for solving problems. *Complex Systems* 13: 87–129.
Fogel, Lawrence Jerome, Alvin J. Owens, and Michael John Walsh (1966). *Artificial intelligence through simulated evolution*. New York: John Wiley.
Fortin, Félix-Antoine, François-Michel De Rainville, Marc-André Gardner, Marc Parizeau, and Christian Gagné (2012). DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13: 2171–2175.
Franses, Philip Hans, Henk C. Kranendonk, and Debby Lanser (2011). One model and various experts: Evaluating Dutch macroeconomic forecasts. *International Journal of Forecasting* 27: 482–495.
Gandomi, Amir H., and David Roke (2015). Assessment of artificial neural network and genetic programming as predictive tools. *Advances in Engineering Software* 88: 63–72.
Garnitz, Johanna, Gernot Nerb, and Klaus Wohlrabe (2015). CESifo World Economic Survey – November 2015. *CESifo World Economic Survey* 14: 1–28.
Ghonghadze, Jaba, and Thomas Lux (2012). Modelling the dynamics of EU economic sentiment indicators: An interaction-based approach. *Applied Economics* 44: 3065–3088.
Girardi, Alessandro (2014). Expectations and macroeconomic fluctuations in the euro area. *Economics Letters* 125: 315–318.
Gong, Yue-Jiao, Wei-Neng Chen, Zhi-Hui Zhan, Jun Zhang, Yun Li, Y., Qingfu Zhang, and Jing-Jing Li (2015). Distributed evolutionary algorithms and their models: A survey of the stat-of-the-art. *Applied Soft Computing* 34: 286–300.
Graff, Michael (2010). Does a multi-sectoral design improve indicator-based forecasts of the GDP growth rate? Evidence from Switzerland. *Applied Economics* 42: 2759–2781.
Guizzardi, Andrea, and Annalisa Stacchini (2015). Real-time forecasting regional tourism with business sentiment surveys. *Tourism Management* 47: 213–223.
Hansson, Jesper, Per Jansson, and Mårten Löf (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting* 21: 377–389.
Hutson, Mark, Fred Joutz, and Herman Stekler (2014). Interpreting and evaluating CESIfo's World Economic Survey directional forecasts. *Economic Modelling* 38: 6–11.
Hyndman, Rob J., and Anne B. Koehler (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22: 679–688.
Jean-Baptiste, Frédo (2012). Forecasting with the new Keynesian Phillips curve: Evidence from survey data. *Economics Letters* 117: 811–813.
Jonsson, Thomas, and Pär Österholm (2011). The forecasting properties of survey-based wage-growth expectations. *Economics Letters* 113: 276–281.
Kaboudan, Mak A. (2000). Genetic programing prediction of stock prices. *Computational Economics* 16: 207–236.
Karlow, Denis (2012). Comparison and development of methods for index tracking. Dissertation. Frankfurt: Frankfurt School of Finance & Management.
Kauppi, Eija, Jukka Lassila, and Timo Teräsvirta (1996). Short-term forecasting of industrial production with business survey data: Experience from Finland's great depression 1990-1993. *International Journal of Forecasting* 12: 373–381.
Klein, Lawrence R., and Süleyman Özmucur (2010). The use of consumer and business surveys in forecasting. *Economic Modelling* 27: 1453–1462.
Kľúčik, Miroslav (2012). Estimates of foreign trade using genetic programming. Proceedings of the 46 the scientific meeting of the Italian Statistical Society.
Kotanchek, Mark E, Ekaterina Y. Vladislavleva, and Guido F. Smits (2010). Symbolic regression via genetic programming as a discovery engine: Insights on outliers and prototypes. In R. Riolo et al., editors, *Genetic programming theory and practice VII, Genetic and evolutionary computation Vol. 8* (pp. 55–72). Springer Science+Business Media, LLC.
Koza, John R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MA: MIT Press.
Koza, John R. (1995). Genetic Programming for econometric modeling. In S. Goonatilaje and P. Treleaven, editors, *Intelligent systems for finance and business* (pp. 251–269). London: John Wiley and Sons.
Kronberger, Gabriel, Stefan Fink, Michael Kommenda, and Michael Affenzeller (2011). Macro-economic time series modeling and interaction networks. In C. Di Chio et al., editors, *EvoApplications, Part II* (pp. 101–110). LNCS 6625.
Kudymowa, Evgenia, Johanna Plenk, and Klaus Wohlrabe (2013). Ifo World Economic Survey and the business cycle in selected countries, *CESifo Forum* 14: 51–57.
Kwiatkowski, Jan Witold (1992). Algorithms for index tracking. *IMA Journal of Management Mathematics* 4: 279–299.
Lahiri, Kajal, and Yongchen Zhao (2015). Quantifying survey expectations: A critical review and generalization of the Carlson-Parkin method. *International Journal of Forecasting* 31: 51–62.
Larkin, Fiacc, and Conor Ryan (2008). Good news: Using news feeds with genetic programming to predict stock prices. In M. O'Neil et al., editors, *Genetic programming* (pp. 49–60). Berlin: Springer-Verlag.
Lawrenz, Claudia, and Frank Westerhoff (2003). Modeling exchange rate behaviour with a genetic algorithm. *Computational Economics* 21: 209–229.
Lehmann, Robert (2015). Survey-based indicators vs. hard data: What improves export forecasts in Europe? Ifo Working Paper, No. 196.
Löffler, Gunter (1999). Refining the Carlson-Parkin method. *Economics Letters* 64: 167–71.

Lui, Silvia, James Mitchell, and Martin Weale (2011a). The utility of expectational data: firm-level evidence using matched qualitative-quantitative UK surveys. *International Journal of Forecasting* 27: 1128–1146.
Lui, Silvia, James Mitchell, and Martin Weale (2011b). Qualitative business surveys: signal or noise? *Journal of The Royal Statistical Society, Series A (Statistics in Society)* 174: 327–348.
Łyziak, Tomasz, and Joanna Mackiewicz-Łyziak (2014). Do consumers in Europe anticipate future inflation? *Eastern European Economics* 52: 5–32.
Martinsen, Kjetil, Francesco Ravazzolo, and Fredrik Wulfsberg (2014). Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting* 30: 65–77.
Maschek, Michael Kurtis (2010). Intelligent mutation rate control in an economic application of genetic algorithms. *Computational Economics* 35: 25–49.
Mitchell, James (2002). The use of non-normal distributions in quantifying qualitative survey data on expectations. *Economics Letters* 76: 101–107.
Mitchell, James, Kostas Mouratidis, and Martin Weale (2007). Uncertainty of UK manufacturing: Evidence from qualitative survey data. *Economic Letters* 94: 245–252.
Mitchell, James, Richard J. Smith, Martin R. Weale, Stephen Wright, and Eduardo L. Salazar (2005). An indicator of monthly GDP and an early estimate of quarterly GDP growth. *The Economic Journal* 115: F108–F129.
Mokinski, Frieder, Xuguang (Simon) Sheng, and Jingyun Yang (2015). Measuring disagreement in qualitative expectations. *Journal of Forecasting* 34: 405–426.
Müller, Christian (2010). You CAN Carlson-Parkin. *Economics Letters* 108: 33–35.
Nardo, Michela (2003). The quantification of qualitative data: a critical assessment. *Journal of Economic Surveys* 17: 645–668.
Österholm, Pär (2014). Survey data and short-term forecasts of Swedish GDP growth. *Applied Economics Letters* 21: 135–139.
Peng, YuZhong, ChangAn Yuan, Xiao Qin, JiangTao Huang, and YaBing Shi (2014). An improved gene expression programming approach for symbolic regression problems. *Neurocomputing* 137: 293–301.
Pesaran, M. Hashem (1987). *The limits to rational expectations*. Oxford: Basil Blackwell.
Poli, Riccardo, Leonardo Vanneschi, William B. Langdon, and Nicholas Freitag Mcphee (2010). Theoretical results in genetic programming: The next ten years? *Genetic Programming and Evolvable Machines* 11: 285–320.
Robinzonov, Nikolay, Gerhard Tutz, and Torsten Hothorn (2012). Boosting techniques for nonlinear time series models. *AStA Advances in Statistical Analysis* 96: 99–122.
Rudd, Andrew (1980). Optimal selection of passive portfolios. *Financial Management* 9: 57–66.
Sarradj, Ennes, and Thomas Geyer (2014). Symbolic regression modeling of noise generation at porous airfoils. *Journal of Sound and Vibration* 333: 3189–3202.
Schmeling, Maik, and Andreas Schrimpf (2011). Expected inflation, expected stock returns, and money illusion: What can we learn from survey expectations. *European Economic Review* 55: 702–719.
Seitz, Helmut (1988). The estimation of inflation forecasts from business survey data. *Applied Economics* 20: 427–38.
Smith, Jeremy, and Michael McAleer (1995). Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing. *Journal of Applied Econometrics* 10: 165–185.
Theil, Henri (1952). On the time shape of economic microvariables and the Munich Business Test. *Revue de l'Institut International de Statistique* 20: 105–20.
Thinyane, Hannah, and Jonathan Millin (2011). An investigation into the use of intelligent systems for currency trading. *Computational Economics* 37: 363–374.
Vasilakis, Georgios A., Konstantinos A. Theofilatos, Efstratios F. Georgopoulos, Andreas Karathanasopoulos, and Spiros D. Likothanassis (2013). A genetic programming approach for EUR/USD exchange rate forecasting and trading. *Computational Economics* 42: 415–431.
Vladislavleva, Ekaterina, Guido Smits, Dick den Hertog (2010). On the importance of data balancing for symbolic regression. *IEEE Transactions in Evolutionary Computation* 14: 252–277.
Vermeulen, Philip (2014). An evaluation of business survey indices for short-term forecasting: Balance method versus Carlson-Parkin method. *International Journal of Forecasting* 30: 882–897.
Wei, Liang-Ying (2013). A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. *Economic Modelling* 33: 893–899.
Wu, Chih-Hung, Hung-Yu Chou, and Wei-Han Su (2008). Direct transformation of coordinates for GPS positioning using the techniques of genetic programming and symbolic regression. *Engineering Applications of Artificial Intelligence* 21: 1347–1359.
Yang, Guangfei, Xianneng Li, Jianliang Wang, Lian Lian, and Tieju Ma (2015). Modeling oil production based on symbolic regression. *Energy Policy* 82: 48–61.
Yao, Liangzhu, and Cheng-Chun Lin (2009). Identification of nonlinear systems by the genetic programming-based volterra filter. *IET Signal Processing* 3: 93–105.
Wilson, Garnett C., and Wolfgang Banzhaf (2009). Prediction of interday stock prices using developmental and linear genetic programming. In M. Giacobini et al., editors, *Applications of evolutionary computing* (pp. 172–181). Berlin: Springer-Verlag.
Yu, Tina, Shu-Heng Chen, and Tzu-Wen Kuo (2004). A genetic programming approach to model international short-term capital flow. *Applications of Artificial Intelligence in Finance and Economics* 19: 45–70.
Zelinka, Ivan, Zuzana Oplatkova, and Lars Nolle (2005). Analytic programming – Symbolic regression by means of arbitrary evolutionary algorithms. *International Journal of Simulation: Systems, Science and Technology* 6: 44–56.