A new approach for the quantification of qualitative measures of economic expectations

Oscar Claveria¹,⁴ · Enric Monte² · Salvador Torra³

Abstract In this study a new approach to quantify qualitative survey data about the direction of change is presented. We propose a data-driven procedure based on evolutionary computation that avoids making any assumption about agents’ expectations. The research focuses on experts’ expectations about the state of the economy from the World Economic Survey in twenty eight countries of the Organisation for Economic Co-operation and Development. The proposed method is used to transform qualitative responses into estimates of economic growth. In a first experiment, we combine agents’ expectations about the future to construct a leading indicator of economic activity. In a second experiment, agents’ judgements about the present are combined to generate a coincident indicator. Then, we use index tracking to derive the optimal combination of weights for both indicators that best replicates the evolution of economic activity in each country. Finally, we compute several accuracy measures to assess the performance of these estimates in tracking economic growth. The different results across countries have led us to use multidimensional scaling analysis in order to group all economies in four clusters according to their performance.

Oscar Claveria oclaveria@ub.edu
Enric Monte enric.monte@upc.edu
Salvador Torra storra@ub.edu

¹ AQR-IREA (Institute of Applied Economics Research), University of Barcelona (UB), Diagonal, 690, 08034 Barcelona, Spain
² Department of Signal Theory and Communications, Polytechnic University of Catalunya (UPC), Jordi Girona, 1-3, 08034 Barcelona, Spain
³ Department of Econometrics and Statistics, Riskcenter-IREA, University of Barcelona (UB), Diagonal, 690, 08034 Barcelona, Spain
⁴ Department of Econometrics, Statistics and Applied Economics, Faculty of Economics and Business, University of Barcelona, Diagonal 690, 08034 Barcelona, Spain
1 Introduction

Business and consumer surveys are directly addressed to representatives of firms and agents. Respondents are asked about the expected direction of change of a wide range of variables. Accordingly, these surveys provide important information about agents’ economic expectations and perceptions. Survey results allow comparisons among different countries’ business cycles, and have become an indispensable tool for monitoring the evolution of the economy. The fact that survey results are available ahead of the publication of quantitative official data makes them very useful for prediction.

However, unlike experimental expectations, survey-based expectations are qualitative in nature. With the aim of overcoming this limitation, numerous quantification methods have been proposed in the literature (Nardo 2003; Driver and Urga 2004; Pesaran and Weale 2006; Vermeulen 2014). This line of research centered on the conversion of qualitative responses about the expected direction of change into quantitative measures has evolved in parallel with the application of new econometric techniques.

Recent developments in empirical modelling allow us to generate mathematical models from a given dataset. Empirical modelling has two main advantages over conventional approaches. On the one hand, it is especially suitable for finding patterns in large data sets, where little or no information is known about the system. On the other hand, empirical modelling allows to simultaneously evolve both the structure and the parameters of the model.

In a recent study, Lahiri and Zhao (2015) examined the quality of quantified expectations by comparing them to quantitative realisations at the firm level, obtaining significant improvements when relaxing the assumptions of quantification methods of qualitative survey data, particularly during periods of uncertainty, with high levels of disagreement between respondents.

These findings have led us to look for a data-driven assumption-free approach to transform survey measures of agents’ expectations into quantitative estimates. We aim to break new ground in the quantification of survey responses on the direction of change by presenting a new method based on the implementation of recent developments in evolutionary computation to qualitative survey data.

The CESifo Institute for Economic Research elaborates World Economic Survey (WES), which polls experts in 123 countries about economic trends (Kudymowa et al. 2013). We use 12 survey variables from the WES in 28 countries of the Organisation for Economic Co-operation and Development (OECD) to generate quantitative estimates of economic growth. In the first step, we combined and transformed agents’ expectations about the future state of the economy. We repeated the experiment for agents’ judgements about the present state of the economy. As a result, we derived a leading indicator and a coincident indicator of economic activity. In the second step, we applied index tracking, which is a procedure used for portfolio management, to calculate the optimal relative weights of both indicators that best replicated the evolution of the gross domestic product (GDP) in each country.
With the aim of examining the leading properties of these estimates of economic growth, we computed several accuracy measures to assess their predictive content and to evaluate their cyclical properties in terms of the level of synchronisation with the quantitative variable of reference. Finally, by means of a dimensionality reduction technique, we synthesise all the information provided by these performance measures and the characteristics of the data into two factors that allow us to cluster all economies into four groups.

The structure of the paper is as follows. The next section reviews the existing literature. In Sect. 3 we present the methodological approach and describe the experiment. Empirical results are provided in Sect. 4. Finally, conclusions are presented.

2 Literature review

2.1 Quantification of qualitative survey data

The first attempt to quantify survey expectations was that of Anderson (1951, 1952), who proposed the balance statistic as a measure of the evolution of the quantitative variable it refers to. Aggregating individual replies as percentages of the respondents in each category, and assuming that the expected percentage change in a variable remains constant over time for all agents, the balance statistic is obtained as the subtraction between the percentage of agents reporting an increase and the percentage reporting a decrease. Based on these premises, Pesaran (1984, 1985) developed this framework to allow for an asymmetrical relationship between individual changes and the evolution of the quantitative variable of reference. Using the relationship between actual values and respondents’ perceptions of the past as yardstick for the quantification of expectations about the future, the author proposed the regression approach.

By making positive and negative individual changes dependent on past values of the quantitative variable of reference, Smith and McAleer (1995) proposed a non-linear dynamic regression model to quantify survey responses that can be regarded as an extension of the regression approach. A drawback of the regression approach to quantify survey responses is that there is no empirical evidence that agents judge past values in the same way that they formulate expectations about the future (Nardo 2003). As a result, the regression approach is restricted to expectations of variables over which agents have direct control, be it prices or production. The development of this approach has also been conditioned by the procurement of a rationale for the application, which can only be obtained by means of the analysis of individual data. For an appraisal of individual firm data on expectations see Zimmermann (1997).

Theil (1952) designed a theoretical framework to generate quantitative estimates from the balance statistic proposed by Anderson (1951). Based on the assumption that respondents report a variable to go up (or down) if the mean of their subjective probability distribution lies above (or below) a certain level, the author defined the indifference threshold, also known as the difference limen. This threshold was conceived as an interval around zero within which respondents perceive there are no significant changes in the variable, and respond that the variable remains unchanged. Let $y_{it}$ indicate the percentage change of variable $Y_{it}$ for agent $i$ from time $t - 1$ to time $t$, and $R_t$ and $F_t$, denote the aggregate percentage of respondents at time $t - 1$ expecting a variable to rise or fall at time $t$ respectively. If $y_{it}^e$ is the unobservable expectation that agent $i$ has over the change of
variable $Y_{it}$, the indifference interval can be defined as $(-a_{it}, b_{it})$, where $a_{it}$ and $b_{it}$ are the lower and upper limits of the indifference threshold for agent $i$ regarding time $t$. Assuming that response bounds are symmetrical and fixed both across respondents and over time ($a_{it} = b_{it} = \lambda$, $\forall i,t$), and that agents base their answer on an independent subjective probability distribution that has the same form across respondents, the author generated quantitative estimates of $\hat{y}_i$.

Knöbl (1974) and Carlson and Parkin (1975) further developed the probability approach proposed by Theil (1952). As estimates of $\hat{y}_i$ are conditional on a particular value for the imperceptibility parameter $\lambda$, and a specific form for the aggregate density function, Carlson and Parkin (1975) assumed that the individual density functions were normally distributed, and estimated $\lambda$ by assuming that over the sample-period $\hat{y}_i$ is an unbiased estimate of $y_i$. Consequently, the role of $\lambda$ is to scale the aggregate expectations $y^*_i$ such that the average value of $y_i$ equals $y^*_i$. Thus, using the evolution of the observed variable as a yardstick qualitative responses can be transformed into quantitative estimates.

Fishe and Lahiri (1981), Batchelor (1982), Visco (1984), and Foster and Gregory (1977) used alternative distributions. There is inconclusive evidence on the type of probability distribution which aggregate average expectations come from. While Carlson (1975), Batchelor (1981), Batchelor and Dua (1987), Foster and Gregory (1977) and Lahiri and Teigland (1987) reject the hypothesis of normality, Dasgupta and Lahiri (1992), Balcombe (1996), Berk (1999) and Mitchell (2002) find evidence that normal distributions provide expectations that are as accurate as other non-normal distributions.

Another line of research has focused on refining the probability approach by relaxing the assumptions of symmetry and constancy of the indifference bounds. Several strategies have been proposed in the literature in order to introduce dynamic imperceptibility parameters in the probability approach. Bennett (1984), Batchelor (1986), Kariya (1990), and Berk (1999) made the threshold dependent on time-varying quantitative variables. Batchelor and Orr (1988) imposed the unbiasedness condition over predefined subperiods. Mitchell et al. (2007) generalised the Carlson–Parkin procedure to generate cross-sectional and time-varying proxies of the variance.

Using a time-varying parameter model (Cooley and Prescott 1976) together with the Kalman filter (Kalman 1960) for parameter estimation, Seitz (1988) was able to simultaneously introduce asymmetric and time-varying indifference thresholds. The author assumed that the imperceptibility parameters were subject to permanent and temporary shocks. Claveria et al. (2007) extended this framework by using a state-space representation that allowed for asymmetric and dynamic response thresholds generated by a first-order Markov process.

Further improvements of quantification procedures have been developed at the micro level, either by means of experimental expectations generated by Monte Carlo simulations, or by comparing the individual responses with firm-by-firm realisations. Regarding the former option, Common (1985) generated simulated expectations to test the rational expectations hypothesis. Nardo and Cabeza-Gutés (1999) designed a simulation experiment to assess the performance of the different quantification methods, both in terms of the size of the measurement error and of its systematic nature, obtaining the best results with the probabilistic approach with time varying parameters.

By means of simulation-based expectations, Löffler (1999) and Terai (2009) also estimated the measurement error introduced by the probabilistic method. Additionally, Löffler (1999) and Claveria et al. (2006) proposed refinements of the Carlson–Parkin method. Claveria (2010) used computer-generated expectations to assess the forecasting performance of different quantification methods, and presented a variation of the balance
statistic that took into account the proportion of respondents reporting that the variable
remains unchanged.

Using firm-level survey responses, Mitchell et al. (2002) developed a procedure to
quantify individual categorical expectations based on the assumption that responses are
triggered by a latent continuous random variable as it crosses time-varying thresholds, and
found evidence against time invariant thresholds. By introducing the “conditional absolute
null” property, based on the empirical finding that the median of realised quantitative
values corresponding to the “no change” category is zero, Müller (2010) proposed a
variant of the Carlson–Parkin method with asymmetric and time invariant thresholds,
which allows to solve the zero response problem that occurs when all respondents fall into
one of the extreme responses (an increase or a decrease).

The variation of the indifference thresholds across individuals can only be tested by
means of the analysis of individual expectations. Using a matched sample of qualitative
and quantitative individual stock market forecasts, Breitung and Schmeling (2013) cor-
rborated the importance of introducing asymmetric and dynamic indifference parameters,
but found that individual heterogeneity across respondents plays a minor role in forecast
accuracy. On the other hand, Lahiri and Zhao (2015) have recently found strong evidence
against the threshold constancy, symmetry, homogeneity and overall unbiasedness
assumptions of the probability method. The authors have generalised the Carlson–Parkin
framework by means of a hierarchical ordered probit model. Based on a matched sample of
households, they found that when the unbiasedness assumption is replaced by a time-
varying calibration, the resulting quantified series better track the quantitative benchmark.

2.2 Evolutionary computation

Evolutionary computation is a subfield of artificial intelligence that is being increasingly
applied in economics in the context of expensive optimisation. Evolutionary computation
is based on the implementation of algorithms that adopt Darwinian principles of the theory
of natural selection to automated problem solving. These algorithms are known as evo-
lutionary algorithms (EAs). Evolutionary programming was introduced by Fogel et al.
(1966). The most popular type of EA is the genetic algorithm (GA), which was initially
proposed by Holland (1975). Cramer (1985) developed a generalisation of GAs known as
genetic programming (GP). GP is a soft-computing search technique that allows the model
structure to vary during the evolution, which makes it particularly indicated for non-linear
and empirical modelling. See Poli et al. (2010) for a review of the state of the art in GP.

Chen and Kuo (2002) classified the literature on the application of evolutionary com-
putation to economics and finance. Most evolutionary computing in economics has been
implemented in finance (Goldberg 1989). On the one hand, Acosta-González and Fernán-
dez (2014) used a GA to predict the financial failure of firms, and Acosta-González
et al. (2012) also used this method to explain the 2008 financial crisis. Lawrenz and
Westerhoff (2003) modelled exchange rates with a GA. Maschek (2010) evaluated the
performance of the self-adaptation mechanism in GAs for the convergence to the rational
expectations equilibrium. Thinyane and Millin (2011) applied GAs to optimize the signals
generated by technical trading tools. Vasilakis et al. (2013) presented a GP-based tech-
nique to predict returns in the trading of the euro/dollar exchange rate. Wei (2013) used an
adaptive expectation GA to optimise a fuzzy model to forecast stock price trends in
Taiwan. For a review of the applications of GAs for financial forecasting see Drake and
Marks (2002).
On the other hand, Álvarez-Díaz and Álvarez (2005) applied GP to predict exchange rates. Chen et al. (2008) analysed the performance of GP to financial trading. Kaboudan (2000), Larkin and Ryan (2008), and Wilson and Banzhaf (2009) used GP for stock price forecasting. Yu et al. (2004) implemented a GP approach to model short-term capital flows.

Applications of GP in macroeconomics have been very limited. The first GP application is that of Koza (1992), who used GP to reassess the exchange equation relating the price level, gross national product, money supply, and the velocity of money. Chen et al. (2010) applied GP in a vector error correction model for macroeconomic forecasting. Duda and Szydło (2011) developed economic forecasting models by means of gene expression programming (GEP), which can be regarded as a version of GP (Ferreira 2001).

Koza (1992) developed GP to find the best single computer program to implement symbolic regression (SR). SR can be regarded as a new approach to empirical modelling. 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. Due to its versatility, SR is being increasingly used in different areas: from industrial data analysis (Vladislavleva et al. 2010) and the experimental design of manufacturing systems (Can and Heavey 2011), to signal processing (Yao and Lin 2009) and various other applications (Barmpalexis et al. 2011; Cai et al. 2006; Ceperic et al. 2014; Sarradj and Geyer 2014; Wu et al. 2008).

There have been very few applications in macroeconomics. Claveria et al. (2016) implemented SR via GP to derive a set of building blocks used with forecasting purposes. Klůčik (2012) applied SR to estimate total exports and imports to Slovakia. Kotanchek et al. (2010) used SR via GP for GDP forecasting. By means of SR, Kronberger et al. (2011) identified interactions between economic indicators in order to estimate the evolution of prices in the US. Yang et al. (2015) used SR for production forecasting of crude oil. Recently, Peng et al. (2014) have proposed an improved GEP algorithm especially suitable for dealing with SR problems.

3 Data and methods

3.1 Data

This study matches two sources of information for 28 countries of the OECD: Austria, Belgium, Bulgaria, Croatia, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Latvia, Lithuania, the Netherlands, Norway, Poland, Portugal, Romania, the Slovak Republic, Slovenia, Spain, Sweden, the United Kingdom (UK), and the United States (US). On the one hand, we use quantitative official statistics regarding the evolution of economic activity. Specifically, the 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. On the other hand, we use qualitative survey data reflecting agents’ expectations about the future, and their judgements about the present economic situation.

We focus on the main survey variables from the WES, which assesses worldwide economic trends by polling professionals and experts on current economic developments in their respective countries (Kudymowa et al. 2013). Białowolski (2016) notes that
Professional respondents are characterised by significantly lower biases in responding to survey questions than consumers. Franses et al. (2011) also find evidence in favour of experts’ forecasts when compared with pure model forecasts. See Henzel and Wollmershäuser (2005), Stangl (2007) and Hutson et al. (2014) for an appraisal of the WES.

Respondents are asked about the economic situation in three different forms: their expectation by the end of the next 6 months (variables $X_7$–$X_{12}$), their present judgement (variables $X_1$–$X_3$), and their assessment compared to the same time last year (variables $X_4$–$X_6$). The economic situation is assessed with respect to three items: the overall economy (variables $X_1$, $X_4$ and $X_7$), capital expenditures (variables $X_2$, $X_5$ and $X_8$), and private consumption (variables $X_3$, $X_6$ and $X_9$). Respondents are also asked about their expectations about the volume of exports ($X_{10}$), imports ($X_{11}$) and the balance of trade ($X_{12}$). All twelve variables are presented in Table 1.

In order to present the survey results, the Ifo uses a grading procedure that is conceptually equal to calculating balances: positive replies are assigned a grade of nine; indifferent replies, a grade of five; and negative replies, a grade of one. A country’s results are weighted according to its share of exports and imports in total world trade (CESifo World Economic Survey 2016). The Ifo also constructs 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 6 months: the economic climate indicator (ECI). The ECI tends to correlate closely with the actual business-cycle trend measured in annual growth rates of real GDP (Claveria et al. 2016; Garnitz et al. 2015).

### 3.2 Methods

SR is based on the search for relationships between a given set of variables. The major difference in relation to conventional regression analysis resides in the fact that while the former is based on a certain model specification, SR does not rely on a specific a priori determined model structure. The only assumption made in SR is that the response surface can be described by an algebraic expression.

GP can be regarded as an extension of GAs in which the solutions are expressed in the form of computer programs. The main difference between them is in the representation of the structure: while GP codes potential solutions by means of tree-structured, variable length representations, GAs use fixed length binary string representations. GP’s more general representation scheme allows the model structure to vary during the evolution. This feature is particularly suitable in the current study, in which the functional relationship between the set of survey variables is arbitrary and unknown.

| Present Economic situation | Compared to last year Economic situation | For the next 6 months Economic situation and foreign trade volume |
|---------------------------|------------------------------------------|---------------------------------------------------------------|
| $X_1$ overall economy     | $X_4$ overall economy                   | $X_7$ overall economy                                         |
| $X_2$ capital expenditures| $X_5$ capital expenditures              | $X_8$ capital expenditures                                    |
| $X_3$ private consumption | $X_6$ private consumption               | $X_9$ private consumption                                    |
|                           |                                          | $X_{10}$ volume of exports                                    |
|                           |                                          | $X_{11}$ volume of imports                                    |
|                           |                                          | $X_{12}$ balance of trade                                     |
Consequently, we use GP to solve the SR experiment, and to transform qualitative survey data into quantitative estimates of economic activity, formalising the interactions between a wide range of survey-based indicators. The implementation of GP for SR was based on the following sequence of steps:

First—the creation of an initial population. We determined a population size of 3 million individuals.

Second—determination of a fitness function. In order to evaluate the fitness of each member of the population we use the root mean square error (RMSE).

Third—determination of a strategy for the selection of parents for replacement, which are the programs used to create offspring programs. In order to guarantee the diversity in the population, we use the tournament method of selection of individuals for crossover. This method is based on the selection of the fittest individual in each of the tournaments, which are run among a group of individuals chosen at random from the population. One of the main advantages over other alternative methods is that the selection pressure can be easily adjusted, and it is efficient to code.

Fourth—determination of the probability of a new generation and application of genetic operators to the parents. The main genetic operations are reproduction (copy), crossover (recombination of randomly chosen parts of parents), and mutation (random alteration of a part of a parent). We select a 0.1 mutation probability to prevent trapping into local optima.

Fifth—determination of constants. With the aim of avoiding the search path to deviate from the optimum we include the automatic generation of constants provided by the algorithm, which are optimised after a number of generations according to their correlation relative to the functional form.

Sixth—determination of a stopping criterion. We set a maximum number of 150 generations as the termination criterion. Steps three and four are repeated until a new generation is created. If no individual in the population has a required minimal fitness, or the stopping criterion is fulfilled, everything is repeated using the new generation as the population. As a result, the fitness of the population is ever increasing.

The search process is characterised by a trade-off between accuracy and simplicity. To limit the complexity of the resulting expressions, the set of functions is restricted to some elementary functions. We use the Distributed Evolutionary Algorithms Package (DEAP) framework implemented in Python (Fortin et al. 2012; Gong et al. 2015).

By matching qualitative survey indicators from the WES to quantitative official data in two successive SR experiments, we are able to derive two analytical expressions: one linking agents’ expectations about the future (variables $X_7$–$X_{12}$) to economic growth, and another one combining agents’ judgements about the present (variables $X_1$–$X_6$).

### 4 Results

First, we present the output of the two SR experiments that were undertaken. On the one hand, expression (1), which combines agents’ expectations about the future, and can therefore be regarded as a leading indicator of economic activity ($\hat{y}_{1,it}$). On the other hand, expression (2), which combines agents’ judgements about the present state of the economy, and can be seen as a coincident indicator ($\hat{y}_{2,it}$):

$$\hat{y}_{1,it} = X_{12,it} - 22 - \frac{\sqrt{X_{10,it}}}{\sqrt{X_{12,it}}} + \log \left(\frac{3 \cdot X_{12,it} + X_{7,it} - 10}{4} - \frac{X_{10,it}}{X_{9,it}}\right)$$  \ (1)
\[
\hat{y}_{2,lt} = \sqrt{(X_{2,lt} + X_{5,lt}) \left[ \frac{(X_{6,lt} + X_{2,lt} + \log \sqrt{X_{4,lt}} - 1 + X_{1,lt} - 1)}{4} \right]} - 11
\]  \tag{2}

where the sub index \( i \) refers to each specific country at time \( t \). In Fig. 1 we graphically compare the evolution of the two SR-generated indicators to that of the GDP. We can

---

**Fig. 1** Evolution of year-on-year GDP growth rates versus survey-based economic indicators. The *black dotted line* represents the year-on-year growth rate of GDP in each country. The *grey line* represents the evolution of the proposed leading indicator. The *black line* represents the evolution of the proposed coincident indicator.
Fig. 1 continued
A new approach for the quantification of qualitative...

| Norway               | Poland              |
|----------------------|---------------------|
| ![Norway Graph]      | ![Poland Graph]     |

| Portugal             | Romania            |
|----------------------|---------------------|
| ![Portugal Graph]    | ![Romania Graph]    |

| Slovak Republic      | Slovenia           |
|----------------------|---------------------|
| ![Slovak Republic Graph] | ![Slovenia Graph] |

| Spain                | Sweden             |
|----------------------|---------------------|
| ![Spain Graph]       | ![Sweden Graph]     |

| United Kingdom       | United States      |
|----------------------|---------------------|
| ![United Kingdom Graph] | ![United States Graph] |

Fig. 1 continued
observe that while \( \hat{y}_{2,t} \) is closely correlated to the oscillations of GDP in all countries, \( \hat{y}_{1,t} \) shows a worse performance, especially since the 2008 financial crisis. Łyziak and Mackiewicz-Łyziak (2014) found that the 2008 financial crisis period led to a decrease in expectation errors in transition economies. Claveria et al. (2016) obtained a similar result for ten Eastern European countries.

In a second step, we derive estimates of economic growth by combining the information of both indicators. We use a procedure of constrained optimisation known as index tracking, which is used in finance in order to replicate the performance of stock indexes (Karlow 2012; Kwiatkowski 1992; Rudd 1980). Index tracking consists on the minimisation of a tracking error, which is defined as the expected squared deviation of return from that of the index, with the aim of obtaining the proportion of capital to be invested in each asset. Based on this premise, we use a generalised reduced gradient algorithm to minimise the summation of squared forecast errors. We impose two restrictions on the optimisation process with respect to the value of the weights. First, the sum of both weights must equal one. Second, we impose a non-negativity restriction, so that the weights must be equal to or larger than zero. As a result, we obtain the optimal weights of both the leading and the coincident indicator for each country (Table 2).

While the obtained relative weight of the coincident indicator is higher than that of the leading indicator, we observe numerous differences across countries. In Greece, Hungary, Latvia, Romania, the Slovak Republic, and Slovenia, the algorithm yields a null weight to the leading indicator constructed with agents’ expectations about the future. This result contrasts with that of Lacová and Kráľ (2015), who found that in Slovakia companies are slightly more forward-looking than backward-looking. On the other extreme, in countries such as Norway and Japan, future expectations outweigh judgements about the present. This result brings up the question of whether survey-based indicators should equally weight the information regarding the expectations about the future and the judgements about the present in all countries.

There is no consensus in the literature regarding the information content of survey expectations. Breitung and Schmeling (2013) compared quantified stock market expectations with quantitative forecasts, and found that there was a weak correlation between

| Country        | \( \hat{y}_{1,t} \) | \( \hat{y}_{2,t} \) | \( \hat{y}_{1,t} \) | \( \hat{y}_{2,t} \) |
|----------------|---------------------|---------------------|---------------------|---------------------|
| Norway         | 0.710               | 0.290               | 0.127               | 0.873               |
| Japan          | 0.619               | 0.381               | 0.104               | 0.896               |
| Croatia        | 0.599               | 0.401               | 0.091               | 0.909               |
| Lithuania      | 0.576               | 0.424               | 0.069               | 0.931               |
| Denmark        | 0.569               | 0.431               | 0.039               | 0.961               |
| Germany        | 0.557               | 0.443               | 0.038               | 0.962               |
| Finland        | 0.462               | 0.538               | 0.000               | 1.000               |
| Belgium        | 0.439               | 0.561               | 0.000               | 1.000               |
| Austria        | 0.427               | 0.573               | 0.000               | 1.000               |
| Netherlands    | 0.388               | 0.612               | 0.000               | 1.000               |
| France         | 0.331               | 0.669               | 0.000               | 1.000               |
| Sweden         | 0.329               | 0.671               | 0.000               | 1.000               |
| US             | 0.271               | 0.729               | 0.000               | 1.000               |
| Czech Republic | 0.201               | 0.799               | 0.000               | 1.000               |

Countries are ranked in descending order according to the relative contribution of the leading indicator to the composite indicator.
them. Lacová and Král (2015) found that quantified survey expectations in Slovakia systematically failed to capture changes in the consumer price index. Jonsson and Österholm (2011, 2012), Lui et al. (2011a, b) and Maag (2009) reached similar conclusions. On the other hand, there is ample evidence that survey expectations provide useful information for economic modelling (Altug and Çakmakli 2016; Batchelor and Dua 1992, 1998; Dees and Brinca 2013; Girardi 2014; Hansson et al. 2005; Jean-Baptiste 2012; Klein and Özmucur 2010; Leduc and Sill 2013; Lemmens et al. 2005; Müller 2009; Qiao et al. 2009; Schmeling and Schrimpf 2011).

In order to evaluate the performance of the resulting estimates in monitoring economic activity, we compute several measures of forecast accuracy: the mean absolute error (MAE) and the RMSE to assess the predictive content in terms of forecast accuracy, the mean absolute scaled error (MASE) to compare the forecasting performance to a benchmark model, and the concordance index (CI) proposed by Harding and Pagan (2002) to evaluate the cyclical properties in terms of the level of synchronization with the quantitative benchmark variable.

In Table 3, we present the MAE and the RMSE. We observe differences across countries. Austria, Belgium, Bulgaria, Estonia, and Lithuania present the lowest MAE and RMSE values. In the other extreme, Croatia, Hungary, Ireland, Italy and Poland are the economies where we obtain the least accurate predictions.

Next, we complement the assessment of the estimates by comparing them to those obtained with a benchmark model. With this aim, we compute the MASE proposed by Hyndman and Koehler (2006). The idea behind the MASE is to scale the errors by the mean absolute errors obtained with a benchmark model. The MASE statistic presents several advantages over other forecast accuracy measures, as it is independent of the scale of the data, and it does not suffer from some of the problems presented by other relative measures of forecast accuracy (Hyndman and Koehler 2006). An additional advantage is its easy interpretation, as values less than one indicate that the average prediction computed with the benchmark model is worse than the estimates obtained with the proposed method. If we denote $\hat{e}_t$ as the forecast error, the MASE can be obtained as the mean of the absolute value of the scaled error $q_t$:

| MAE     | RMSE | MAE     | RMSE |
|---------|------|---------|------|
| Austria | 0.926| 1.266   | Japan| 1.506| 2.169 |
| Belgium | 0.978| 1.187   | Latvia| 1.105| 1.384 |
| Bulgaria| 0.889| 1.190   | Lithuania| 0.824| 1.012 |
| Croatia | 3.076| 3.729   | Netherlands| 1.469| 1.753 |
| Czech Republic| 1.816| 2.324| Norway| 1.290| 1.770 |
| Denmark | 2.628| 3.073   | Poland| 4.333| 5.789 |
| Estonia | 0.910| 1.154   | Portugal| 2.284| 2.850 |
| Finland | 2.849| 3.427   | Romania| 2.725| 3.281 |
| France  | 1.499| 1.794   | Slovak Republic| 1.536| 1.934 |
| Germany | 2.365| 2.806   | Slovenia| 2.651| 3.401 |
| Greece  | 1.678| 2.047   | Spain| 1.899| 2.411 |
| Hungary | 5.546| 6.631   | Sweden| 2.727| 3.038 |
| Ireland | 3.075| 4.379   | UK    | 1.090| 1.450 |
| Italy   | 3.233| 3.714   | US    | 2.142| 2.542 |
\[ \text{MASE} = \text{mean}(|q_t|) \quad \text{where} \quad q_t = e_t / \left( \frac{\sum_{i=3}^{n} |Y_i - Y_{i-2}|}{n - 1} \right) \]  

Given that official quantitative data are published with a delay of more than a quarter compared to survey data, we use two-step ahead naïve forecasts as a benchmark. In Table 4 we present the MASE results. In Austria, Belgium, Bulgaria, Estonia, Greece, Japan, Latvia, Lithuania, Norway, the Slovak Republic and the UK, the estimates of GDP obtained with the proposed method outperform those of the benchmark model.

In Fig. 2 we compare the obtained forecast results to the standard deviation of the year-on-year growth rates of GDP. There seems to be no relation between the MASE results and the variability of economic activity. While Estonia and Latvia are the countries that present the highest levels of dispersion, the forecast errors are low, and the opposite holds true for Hungary, Croatia and Poland.

As there is evidence that the trend in the Ifo’s ECI correlates closely with the actual business-cycle trend measured in annual growth rates of real GDP (CESifo WES 2016), next we evaluate the cyclical properties of the proposed SR-generated estimates. We compare it with the ECI in terms of the level of synchronisation. To that end, we use the CI proposed by Harding and Pagan (2002):

\[ CI = \frac{1}{T} \sum_{t=1}^{T} S_x S_y \left( 1 - S_x (1 - S_y) \right) \]

where \(T\) is the number of observations, \(y_t\) refers to the percentage growth rate of GDP and \(x_t\) to the variable under analysis. \(S\) is a binary variable that takes the value of one if the series is in expansion, and zero otherwise. As in the C index developed by Harrell et al. (1996), the CI is expressed as the proportion of sample periods in which the two series are in the same phase of the cycle. Thus, the CI allows us to assess the proposed indicator in terms of regime shifts.

### Table 4 Forecast accuracy—MASE

| Country          | MASE | Country         | MASE |
|------------------|------|-----------------|------|
| Austria          | 0.620| Japan           | 0.752|
| Belgium          | 0.818| Latvia          | 0.375|
| Bulgaria         | 0.533| Lithuania       | 0.613|
| Croatia          | 5.223| Netherlands     | 1.247|
| Czech Republic   | 1.220| Norway          | 0.857|
| Denmark          | 1.728| Poland          | 3.528|
| Estonia          | 0.282| Portugal        | 1.511|
| Finland          | 1.391| Romania         | 1.248|
| France           | 1.544| Slovak Republic | 0.829|
| Germany          | 1.460| Slovenia        | 1.385|
| Greece           | 0.842| Spain           | 2.242|
| Hungary          | 3.808| Sweden          | 1.509|
| Ireland          | 1.208| UK              | 0.851|
| Italy            | 2.306| US              | 1.916|
Results in Table 5 show that, in most cases, there are no major differences in CI values between both proxies. While in countries like Belgium, Finland, Hungary, the Netherlands and Spain, the ECI shows a lower level of synchronisation, the opposite holds true for Japan and Norway.

To synthesise all the information provided by the above performance measures and the characteristics of the data, we finally compute two factors that allow us to cluster all

![Figure 2](image)

**Fig. 2** Forecast accuracy (MASE) versus Standard deviation of GDP growth. The $Y$ axis shows the standard deviation of GDP. The $X$ axis shows the MASE

Table 5 Concordance index—SR estimates versus ECI

| CI          | SR estimates | ECI | SR estimates | ECI |
|-------------|--------------|-----|--------------|-----|
| Austria     | 0.611        | 0.630 | Japan        | 0.685 | 0.759 |
| Belgium     | 0.741        | 0.667 | Latvia       | 0.537 | 0.500 |
| Bulgaria    | 0.500        | 0.537 | Lithuania    | 0.407 | 0.426 |
| Croatia     | 0.444        | 0.556 | Netherlands  | 0.722 | 0.630 |
| Czech Republic | 0.593    | 0.574 | Norway       | 0.556 | 0.667 |
| Denmark     | 0.667        | 0.630 | Poland       | 0.593 | 0.556 |
| Estonia     | 0.630        | 0.611 | Portugal     | 0.685 | 0.648 |
| Finland     | 0.648        | 0.593 | Romania      | 0.611 | 0.574 |
| France      | 0.630        | 0.667 | Slovak Republic | 0.556 | 0.537 |
| Germany     | 0.685        | 0.685 | Slovenia     | 0.667 | 0.667 |
| Greece      | 0.556        | 0.519 | Spain        | 0.722 | 0.593 |
| Hungary     | 0.574        | 0.463 | Sweden       | 0.611 | 0.685 |
| Ireland     | 0.630        | 0.611 | UK           | 0.463 | 0.500 |
| Italy       | 0.630        | 0.611 | US           | 0.667 | 0.630 |

CI stands for concordance index (Harding and Pagan 2002). A one value indicates that the cycles of the variables under comparison are in the same phase one hundred percent of times.
economies into four groups. By transforming the original set of correlated performance measures into a smaller and more understandable set of uncorrelated factors, we aim to summarize the results of the present study. We make use of multidimensional scaling (MDS) analysis to generate a two-dimensional perceptual map in which we position all 28 economies according to their coordinates regarding the two factors. MDS is a multivariate analytical procedure also known as principal coordinates analysis (Torgerson 1952, 1958). For a detailed description of this technique see Hair et al. (2009) and Jolliffe (2002). MDS allows us to visualise the level of similarity of individual cases of a dataset. In our case, the proximity between the different countries in the perceptual map indicates how similar they are in terms of the performance of survey-based measures of economic expectations.

First, we rank all 28 countries in decreasing order according to their performance experienced over the sample period for each of the following measures: the weight of the leading indicator \(\omega_{x_{1},\mu}\), the summation of squared forecast errors (SSE), the MAE, the RMSE, the MASE, the CI, and the standard deviation of GDP growth \(\sigma_{yt}\) in each country. Second, we assign a numerical value to each country corresponding to its position. We used the Kaiser–Guttman method (Guttman 1954; Kaiser 1960; Yeomans and Golder 1982) in order to determine the number of factors to retain. According to this criterion, only the factors that have eigenvalues greater than one are retained for interpretation. In our case, only the first two factors have eigenvalues larger than the unity, indicating that the optimal number of dimensions is two.

After deciding on the number of components, we reduce the information of all rankings into two dimensions, which can be regarded as two synthetic indicators that maintain the original ordinal structures. We obtain a Kruskal stress value of 0.012, which indicates the amount of distortion in distances to tolerate. Stress values range from 0 to 1, where zero denotes a perfect representation of the input data in two dimensions. The two-dimensional scatterplot that represents the coordinates of the first two dimensions for each country is presented in Fig. 3.

The perceptual map is divided into four quadrants. In the top right quadrant, we find the countries with the highest scores in the two dimensions: Belgium, Norway, Austria,
Lithuania, Japan and the UK. In this group of economies, the evolution of GDP displays a stable pattern, and expectations show a good forecasting performance. In the other extreme, in the lower left quadrant, we find the economies with the lowest scores in both dimensions: Ireland, Hungary, Slovenia, Poland, Italy, Finland, Sweden and Denmark, which is very close to Spain, Portugal and Germany in top left quadrant. Croatia is grouped apart, obtaining the lowest score in the first dimension, as opposed to Bulgaria, Estonia and Latvia, which are the economies with the highest scores in the first dimension.

According to Lee (1994), the differences between the actual values of a variable and the quantified expectations may arise from three different sources: the measurement or conversion error due to the use of quantification methods to approximate unobservable expectations; the expectational error due to the agents’ limited ability to predict the movements of the actual variable; and the sampling errors. The groupings in Fig. 3 are indicative of different values regarding these three sources of error.

5 Conclusion

This paper proposes an empirical approach to transform qualitative survey responses on the direction of change into quantitative estimates of economic activity by means of SR via GP. We used survey-based agents’ expectations about the economic situation from the World Economic Survey in 28 countries of the OECD to derive a leading indicator, which consists of an optimal combination of survey variables that best tracks the evolution of the economic activity. We repeated the experiment using agents’ perceptions about the present economic situation to obtain a coincident indicator.

We then combined the information from the leading and the coincident indicator by means of index tracking, which is a procedure to find the optimal relative weights of both indicators. By doing so, we generated quantitative proxies of economic activity. To assess the forecasting performance of the generated estimates, we computed several stylised facts and compared them to a benchmark model and to the ECI constructed by the Ifo. The heterogeneity of the results across countries led us to synthesise the information provided by all the forecast accuracy measures by means of a dimensionality reduction technique that allows us to cluster all economies according to their performance.

We obtained significant differences between mean values of each cluster, which indicates that the forecasting performance of survey-based expectations could be improved by designing ad-hoc quantification procedures for countries with similar characteristics.

Due to the novelty of the proposed approach, there are still several limitations to be addressed. Given that we used a data-driven method, the obtained quantitative estimates lack any theoretical background. Extending the analysis to other questionnaires would allow us to examine to what level the obtained functional forms are extensive to different survey data. Another issue left for further research is to assess the effect of GDP updates on the results. Finally, there is the question of whether the implementation of alternative EAs could improve the forecast accuracy of empirically generated quantitative estimates of expectations.

Acknowledgments This paper has been partially financed by the project SpeechTech4All (TEC2012-38939-C03-02). We would like to thank the Editor and an anonymous referee for their useful comments and suggestions. This paper has been partially supported by the Spanish Ministry of Economy and Competitiveness (TEC2015-69266-P). We also wish to thank Johanna Garnitz and Klaus Wohlrabe at the Ifo Institute for Economic Research in Munich for providing us the data used in the study.
References

Altug, S., Çakmakli, C.: Forecasting inflation using survey expectations and target inflation: evidence from Brazil and Turkey. Int. J. Forecast. 32(1), 138–153 (2016)

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

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

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

Anderson, O.: Konjunkturtest und Statistik. Allg. Stat. Arch. 35, 209–220 (1951)

Anderson, O.: 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 (1952)

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

Barmpalexis, P., Kachrimanis, K., Tsakonas, A., Georgarakis, E.: Symbolic regression via genetic programming in the optimization of a controlled release pharmaceutical formulation. Chemom. Intell. Lab. Syst. 107(1), 75–82 (2011)

Batchelor, R.A.: Aggregate expectations under the stable laws. J. Econ. 16(2), 199–210 (1981)

Batchelor, R.A.: Expectations, output and inflation: the European experience. Eur. Econ. Rev. 17(1), 1–25 (1982)

Batchelor, R.A.: Quantitative v. qualitative measures of inflation expectations. Oxf. Bull. Econ. Stat. 48(2), 99–120 (1986)

Batchelor, R., Dua, P.: The accuracy and rationality of UK inflation expectations. Appl. Econ. 19(6), 819–828 (1987)

Batchelor, R., Dua, P.: Survey expectations in the time series consumption function. Rev. Econ. Stat. 74(4), 598–606 (1992)

Batchelor, R., Dua, P.: Improving macro-economic forecasts. Int. J. Forecast. 14(1), 71–81 (1998)

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

Bennett, A.: Output expectations of manufacturing industry. Appl. Econ. 16(6), 869–879 (1984)

Berk, J.M.: Measuring inflation expectations: a survey data approach. Appl. Econ. 31(11), 1467–1480 (1999)

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

Breitung, J., Schmeling, M.: Quantifying survey expectations: what’s wrong with the probability approach? Int. J. Forecast. 29(1), 142–154 (2013)

Cai, W., Pacheco-Vega, A., Sen, M., Yang, K.T.: Heat transfer correlations by symbolic regression. Int. J. Heat Mass Transf. 49(23–24), 4352–4359 (2006)

Can, B., Heavey, C.: Comparison of experimental designs for simulation-based symbolic regression of manufacturing systems. Comput. Ind. Eng. 61(3), 447–462 (2011)

Carlson, J.A.: Are price expectations normally distributed? J. Am. Stat. Assoc. 70(352), 749–754 (1975)

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

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

CESifo World Economic Survey (2016). Vol. 15(2), May 2016

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

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

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

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

Claveria, O., Monte, E., Torra, S.: Quantification of survey expectations by means of symbolic regression via genetic programming to estimate economic growth in Central and Eastern European economies. East. Eur. Econ. 54(2), 177–189 (2016)
Claveria, O., Pons, E., Ramos, R.: Business and consumer expectations and macroeconomic forecasts. Int. J. Forecast. 23(1), 47–69 (2007)
Claveria, O., Pons, E., Suriñach, J.: Quantification of expectations. Are they useful for forecasting inflation? Econ. Issue 11(2), 19–38 (2006)
Common, M.: Testing for rational expectations with qualitative survey data. Manch. School Econ. Social Stat. 53(2), 138–148 (1985)
Cooley, T.F., Prescott, E.C.: Estimation in the presence of stochastic parameter variation. Econometrica 44(1), 167–184 (1976)
Cramer, N: A representation for the adaptive generation of simple sequential programs. In: Proceedings of the International Conference on Genetic Algorithms and their Applications, Pittsburgh, PA, 24–26 June 1985
Dasgupta, S., Lahiri, K.: A comparative study of alternative methods of quantifying qualitative survey responses using NAPM data. J. Bus. Econ. Stat. 10(4), 391–400 (1992)
Dees, S., Brinca, P.S.: Consumer confidence as a predictor of consumption spending: evidence for the United States and the Euro area. Int. Econ. 134, 1–14 (2013)
Drake, A.E., Marks, R.E.: Genetic algorithms in economics and finance: Forecasting stock market prices and foreign exchange: a review. In: Chen, S.H. (ed.) Genetic Algorithms and Genetic Programming in Computational Finance, pp. 29–54. Springer, New York (2002)
Driver, C., Urga, G.: Transforming qualitative survey data: performance comparisons for the UK. Oxf. Bull. Econ. Stat. 66(1), 71–89 (2004)
Duda, J., Szydło, S.: Collective intelligence of genetic programming for macroeconomic forecasting. In: Jędrzejowicz, P., et al. (eds.) Computational Collective Intelligence. Technologies and Applications, pp. 445–454. Springer, Berlin (2011)
Ferreira, C.: Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst. 13(2), 87–129 (2001)
Fishe, R.P.H., Lahiri, K.: On the estimation of inflationary expectations from qualitative responses. J. Econom. 16(1), 89–102 (1981)
Fogel, L.J., Owens, A.J., Walsh, M.J.: Artificial Intelligence Through Simulated Evolution. Wiley, New York (1966)
Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., Gagné, C.: DEAP: evolutionary algorithms made easy. J. Mach. Learn. Res. 13(1), 2171–2175 (2012)
Foster, J., Gregory, M.: Inflation expectations: the use of qualitative survey data. Appl. Econ. 9(4), 319–329 (1977)
Frenses, P.H., Kranendonk, H.C., Lanser, D.: One model and various experts: evaluating Dutch macroeconomic forecasts. Int. J. Forecast. 27(2), 482–495 (2011)
Garnitz, J., Nerb, G., Wohlrabe, K.: CESifo World Economic Survey: November 2015. CESifo World Econ. Surv. 14(4), 1–28 (2015)
Girardi, A.: Expectations and macroeconomic fluctuations in the euro area. Econ. Lett. 125(2), 315–318 (2014)
Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading (1989)
Gong, Y.J., Chen, W.N., Zhan, Z.H., Zhang, J., Li, Y., Zhang, Q., Li, J.J.: Distributed evolutionary algorithms and their models: a survey of the stat-of-the-art. Appl. Soft Comput. 34, 286–300 (2015)
Guttman, L.: Some necessary and sufficient conditions for common factor analysis. Psychometrika 19(2), 149–161 (1954)
Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis, 7th edn. Prentice Hall, Upper Saddle River (2009)
Hansson, J., Jansson, P., Löf, M.: Business survey data: do they help in forecasting GDP growth? Int. J. Forecast. 30(1), 65–77 (2005)
Harding, D., Pagan, A.: Dissecting the cycle: a methodological investigation. J. Monet. Econ. 49(2), 365–381 (2002)
Harrell, F.E., Lee, K.L., Mark, D.B.: Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat. Med. 15(4), 361–387 (1996)
Henzel, S., Wollmershäuser, T.: An alternative to the Carlson-Parkin method for the quantification of qualitative inflation expectations: evidence from the Ifo World Economic Survey. J. Bus. Cycle Meas. Anal. 2(3), 321–352 (2005)
Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, MI (1975)
Hutson, M., Joutz, F., Stekler, H.: Interpreting and evaluating CESifo’s World Economic Survey directional forecasts. Econ. Model 38, 6–11 (2014)

Hyndman, R.J., Koehler, A.B.: Another look at measures of forecast accuracy. Int. J. Forecast. 22(4), 679–688 (2006)

Jean-Baptiste, F.: Forecasting with the new Keynesian Phillips curve: evidence from survey data. Econ. Lett. 117(3), 811–813 (2012)

Jolliffe, I.T.: Principal Component Analysis, vol. 2. Springer, New York (2002)

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

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

Kaboudan, M.A.: Genetic programing prediction of stock prices. Comput. Econ. 16(3), 207–236 (2000)

Kaiser, H.F.: The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20, 141–151 (1960)

Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic Eng. 82(1), 35–45 (1960)

Karlow, D. (2012). Comparison and development of methods for index tracking. Dissertation. Frankfurt: Frankfurt School of Finance & Management

Kariya, T.: A generalization of the Carlson-Parkin method for the estimation of expected inflation rate. Econ. Stud. Q. 41(2), 155–165 (1990)

Klein, L.R., Ozmuçur, S.: The use of consumer and business surveys in forecasting. Econ. Model. 27(6), 1453–1462 (2010)

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

Knöbl, A.: Price expectations and actual price behaviour in Germany. IMF Staff Papers 21 (1974)

Kotanchev, M.E., Vladislavleva, E.Y., Smits, G.F.: Symbolic regression via genetic programming as a discovery engine: insights on outliers and prototypes. In: Riolo, R., et al. (eds.) Genetic Programming Theory and Practice VII, Genetic and Evolutionary Computation, vol. 8, pp. 55–72. Springer, Ann Arbor (2010)

Koza, J.R.: Genetic Programming: On the Programming of Computers by means of Natural Selection. MIT Press, Cambridge (1992)

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

Kudymowa, E., Plenk, J., Wohlrabe, K.: Ifo World Economic Survey and the business cycle in selected countries. CESifo Forum 14(4), 51–57 (2013)

Kwiatkowski, J.W.: Algorithms for index tracking. IMA J. Manag. Math. 4(3), 279–299 (1992)

Lacová, Ž., Král, P.: Measurement and characteristics of enterprise inflation expectations in Slovakia. Proc. Econ. Financ. 30, 505–512 (2015)

Lahiri, K., Teigland, C.: On the normality of probability distributions of inflation and GNP forecasts. Int. J. Forecast. 3(2), 269–279 (1987)

Lahiri, K., Zhao, Y.: Quantifying survey expectations: a critical review and generalization of the Carlson-Parkin method. Int. J. Forecast. 31(1), 51–62 (2015)

Larkin, F., Ryan, C.: Good news: using news feeds with genetic programming to predict stock prices. In: O’Neil, M., et al. (eds.) Genetic Programming, pp. 49–60. Springer, Berlin (2008)

Lawrence, C., Westerhoff, F.: Modeling exchange rate behaviour with a genetic algorithm. Comput. Econ. 21(3), 209–229 (2003)

Lee, K.C.: Formation of price and cost inflation expectations in British manufacturing industries: a multi-sectoral analysis. Econ. J. 104(423), 372–385 (1994)

Leduc, S., Sill, K.: Expectations and economic fluctuations: an analysis using survey data. Rev. Econ. Stat. 95(4), 1352–1367 (2013)

Lemmens, A., Croux, C., Dekimpe, M.G.: On the predictive content of production surveys: a pan-European study. Int. J. Forecast. 21(2), 363–375 (2005)

Löffler, G.: Refining the Carlson-Parkin method. Econ. Lett. 64(2), 167–171 (1999)

Lui, S., Mitchell, J., Weale, M.: The utility of expectation data: firm-level evidence using matched qualitative-quantitative UK surveys. Int. J. Forecast. 27(4), 1128–1146 (2011a)

Lui, S., Mitchell, J., Weale, M.: Qualitative business surveys: signal or noise? J. R. Stat. Soc. Ser. A 174(2), 327–348 (2011b)

Łyziak, T., Mackiewicz-Łyziak, J.: Do consumers in Europe anticipate future inflation? East. Eur. Econ. 52(3), 5–32 (2014)
Maag, T.: On the accuracy of the probability method for quantifying beliefs about inflation. KOF working paper. 2030. KOF Swiss Economic Institute, Zurich (2009)
Maschek, M.K.: Intelligent mutation rate control in an economic application of genetic algorithms. Comput. Econ. 35(1), 25–49 (2010)
Mitchell, J.: The use of non-normal distributions in quantifying qualitative survey data on expectations. Econ. Lett. 76(1), 101–107 (2002)
Mitchell, J., Mouratidis, K., Weale, M.: Uncertainty of UK manufacturing: evidence from qualitative survey data. Econ. Lett. 94(2), 245–252 (2007)
Mitchell, J., Smith, R., Weale, M.: Quantification of qualitative firm-level survey data. Econ. J. 112(478), 117–135 (2002)
Müller, C.: The information content of qualitative survey data. J. Bus. Cycle Meas. Anal. 2(1), 1–12 (2009)
Müller, C.: You CAN Carlson-Parkin. Econ. Lett. 108(1), 33–35 (2010)
Nardo, M.: The quantification of qualitative data: a critical assessment. J. Econ. Surv. 17(5), 645–668 (2003)
Nardo, M., Cabeza-Gutiés, M.: The role of measurement error in rational expectation testing. UAB working paper, 451. Universitat Autònoma de Barcelona, Barcelona (1999)
Peng, Y., Yuan, C., Qin, X., Huang, J., Shi, Y.: An improved gene expression programming approach for symbolic regression problems. Neurocomputing 137, 293–301 (2014)
Pesaran, M.H.: Expectation formation and macroeconomic modelling. In: Malgrange, P., Muet, P.A. (eds.) Contemporary Macroeconomic Modelling, pp. 27–55. Basil Blackwell, Oxford (1984)
Pesaran, M.H.: Formation of inflation expectations in British manufacturing industries. Econ. J. 95(380), 948–975 (1985)
Pesaran, M.H., Weale, M.: Survey Expectations. In: Elliott, G., Granger, C.W.J., Timmermann, A. (eds.) Handbook of Economic Forecasting, vol. 1, pp. 715–776. Elsevier, Amsterdam (2006)
Poli, R., Vanneschi, L., Langdon, W.B., Mcphee, N.F.: Theoretical results in genetic programming: the next ten years? Genet. Program Evol. Mach. 11(3), 285–320 (2010)
Qiao, Z., McAleer, M., Wong, W.K.: Linear and nonlinear causality between changes in consumption and consumer attitudes. Econ. Lett. 102(3), 161–164 (2009)
Rudd, A.: Optimal selection of passive portfolios. Financ. Manag. 9(1), 57–66 (1980)
Sarradj, E., Geyer, T.: Symbolic regression modeling of noise generation at porous airfoils. J. Sound Vib. 333(14), 3189–3202 (2014)
Schmeling, M., Schrimpf, A.: Expected inflation, expected stock returns, and money illusion: what can we learn from survey expectations. Eur. Econ. Rev. 55(5), 702–719 (2011)
Seitz, H.: The estimation of inflation forecasts from business survey data. Appl. Econ. 20(4), 427–438 (1988)
Smith, J., McAleer, M.: Alternative procedures for converting qualitative response data to quantitative expectations: an application to Australian manufacturing. J. Appl. Econ. 10(2), 165–185 (1995)
Stangl, A.: Ifo World Economic Survey micro data. J. Appl. Social Sci. Stud. 127(3), 487–496 (2007)
Terai, A.: Measurement error in estimating inflation expectations from survey data: an evaluation by Monte Carlo simulations. J. Bus. Cycle Meas. Anal. 8(2), 133–156 (2009)
Theil, H.: On the time shape of economic microvariables and the Munich Business Test. Revue de l’Institut International de Statistique 20, 105–120 (1952)
Thinyane, H., Millin, J.: An investigation into the use of intelligent systems for currency trading. Comput. Econ. 37(4), 363–374 (2011)
Torgerson, W.S.: Multidimensional scaling: I. Theory and method. Psychometrika 17(4), 401–419 (1952)
Torgerson, W.S.: Theory & Methods of Scaling. Wiley, Oxford (1958)
Vasilakis, G.A., Theofilatos, K.A., Georgopoulos, E.F., Karathanasopoulos, A., Likothanassis, S.D.: A genetic programming approach for EUR/USD exchange rate forecasting and trading. Comput. Econ. 42(4), 415–431 (2013)
Vladislavleva, E., Smits, G., den Hertog, D.: On the importance of data balancing for symbolic regression. IEEE Trans. Evol. Comput. 14(2), 252–277 (2010)
Vermeulen, P.: An evaluation of business survey indices for short-term forecasting: balance method versus Carlson-Parkin method. Int. J. Forecast. 30(4), 882–897 (2014)
Visco, L.Y.: Price Expectations in Rising Inflation. North-Holland, Amsterdam (1984)
Wei, L.Y.: A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIX. Econ. Model. 33, 893–899 (2013)
Wu, C.H., Chou, H.J., Su, W.H.: Direct transformation of coordinates for GPS positioning using the techniques of genetic programming and symbolic regression. Eng. Appl. Artif. Intell. 21(8), 1347–1359 (2008)
Yang, G., Li, X., Wang, J., Lian, L., Ma, T.: Modeling oil production based on symbolic regression. Energy Policy 82(1), 48–61 (2015)
Yao, L., Lin, C.C.: Identification of nonlinear systems by the genetic programming-based volterra filter. IET Signal Proc. 3(2), 93–105 (2009)

Yeomans, K.A., Golder, P.A.: The Guttman-Kaiser criterion as a predictor of common factors. J. R. Stat. Soc. Ser. D 31(3), 221–229 (1982)

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

Yu, T., Chen, S., Kuo, T.W.: A genetic programming approach to model international short-term capital flow. Appl. Artif. Intell. Financ. Econ. 19, 45–70 (2004)

Zelinka, I., Oplatkova, Z., Nolle, L.: Analytic programming—symbolic regression by means of arbitrary evolutionary algorithms. Int. J. Simul. Syst. Sci. Technol. 6(9), 44–56 (2005)

Zimmermann, K.F.: Analysis of business surveys. In: Pesaran, M.H., Schmidt, P. (eds.) Handbook of Applied Econometrics. Microeconomics, vol. II, pp. 407–441. Blackwell Publishers, Oxford (1997)