Journal of Applied Economics. Vol XX, No. 2 (November 2017), 329-349

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

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. 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 for Research on Applied Regional and Public Economics), 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 the projects ECO2016-75805-R and TEC2015-692666-P 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; Claveria et al. 2006; Löffler 1999; Berk 1999; Smith and McAleer 1995; 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; Schmeling and Schrimpf 2011; Franses et al. 2011; Graff 2010; Klein and Özmucur 2010), 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