

Unemployment expectations: A socio-demographic analysis of the effect of news

Petar Sorić^{1*}, Ivana Lolić¹, Oscar Claveria², Enric Monte³, Salvador Torra⁴

¹ University of Zagreb, Faculty of Economics and Business

² AQR-IREA, University of Barcelona

³ Department of Signal Theory and Communications, Polytechnic University of Catalunya

⁴ Riskcenter-IREA, Department of Econometrics and Statistics, University of Barcelona

Abstract. In this study, we evaluate the effect of news on consumer unemployment expectations for sixteen socio-demographic groups. To this end, we construct an unemployment sentiment indicator and extract news about several economic variables. By means of genetic programming we estimate symbolic regressions that link unemployment rates in the Euro Area to qualitative expectations about a wide range of economic variables. We then use the evolved expressions to compute unemployment expectations for each consumer group. We first assess the out-of-sample forecast accuracy of the evolved indicators, obtaining better forecasts for the leading unemployment sentiment indicator than for the coincident one. Results are similar across the different socio-demographic groups. The best forecast results are obtained for respondents between 30 and 49 years. The group where we observe the bigger differences among categories is the occupation, where the lowest forecast errors are obtained for the unemployed respondents. Next, we link news about inflation, industrial production, and stock markets to unemployment expectations. With this aim we match positive and negative news with consumers' unemployment sentiment using a distributed lag regression model for each news item. We find asymmetries in the responses of consumers' unemployment expectations to economic news: they tend to be stronger in the case of negative news, especially in the case of inflation.

Keywords: unemployment; expectations; news; consumer behaviour; forecasting; genetic programming

JEL Classification Codes: C53; C55; D12; E24; E27; J10

Acknowledgements

This research was supported by the project HRZZ IP-2018-01-4189 from the Croatian Science Foundation, and the projects ECO2016-75805-R and TEC2015-69266-P from the Spanish Ministry of Economy and Competitiveness. We would like to thank the Editor and the referees for their useful comments and suggestions.

* Corresponding Author.

1. Introduction

The global financial crisis, although it emerged in the US subprime mortgage market, has induced harsh worldwide spillover effects to the real economy. One of the short-run consequences of such an abrupt downturn was a 19.1% increase in global unemployment in 2009 with respect to 2007 (International Labour Office, 2010). In the same period, the Euro Area (EA) unemployment raised by only 2 percentage points (Pissarides, 2013). However, the major EA economic turmoil had yet to begin. The outbreak of the sovereign debt crisis has forced peripheral EA countries (especially Greece, Ireland, Spain, and Portugal) to request financial assistance packages from the troika. These heavily indebted countries had to implement severe austerity measures to correct for their fiscal imbalances, which in turn triggered high and persistent unemployment (Dias, 2017; Lane, 2012; Lucarelli, 2011). In March 2013, the EA unemployment rate reached 12.1%, its highest level since EA inception. This led the European institutions to identify market labour disequilibrium as an obstacle to efficient monetary policy transmission mechanism and a driver of high adjustment costs due to idiosyncratic shocks such as Great Recession or the sovereign debt crisis. In a recent paper, Krištić et al. (2018) found evidence of a certain degree of unemployment hysteresis in the EA countries.

In the crisis aftermath, most economic indicators started to pick up and the EA headed towards a slow and steady expansion. However, since the beginning of 2015, the labour market has experienced imbalances. Although official unemployment figures have not significantly changed, immigration pressure has influenced agents' unemployment expectations (European Commission, 2018). This background has led us to analyse the effect of news on consumer unemployment expectations.

We add to the literature in four specific aspects. First, we aim at designing a survey-based unemployment sentiment indicator (USI). While macroeconomists have long ago used survey data to build confidence indicators, these are mostly designed to track households' financial decisions (Białowolski, 2019) or macroeconomic aggregates such as inflation or economic growth (Batchelor, 1982; Hansson et al., 2005). The construction of unemployment expectations indexes seems to be quite neglected in the literature. Hutter and Weber (2015) and Claveria (2019a,b) are among the very few existing studies addressing this issue. We use qualitative survey data from the Consumer survey conducted by the European Commission (EC) to compute households' expectations and design a consumer USI.

Consumer surveys ask respondents about a very wide set of economic issues ranging from their spending intentions to the general economic situation influencing those decisions. The EC builds upon this abundance of survey data and constructs a set of widely utilized survey-based indicators by choosing particular survey questions and weighting their respective response balances (Sorić et al., 2016). We use an assumption-free approach based on empirical modelling to inspect whether the information content of the survey mostly reflects consumers' perceptions about the past or their future expectations. We follow Claveria et al.'s (2017) approach to quantify qualitative survey data on the direction of change. This data-driven procedure based on evolutionary computation allows us to find the combination of survey questions that best tracks unemployment rates. Evolutionary computation can be regarded as a subfield of artificial intelligence, and is being increasingly applied to automated problem-solving in economics (Álvarez-Díaz, 2019; Claveria et al. 2019a,b).

Some of the features of empirical modelling are particularly indicated to deal with the problem at hand. On the one hand, this approach is especially suitable for finding patterns in large data sets with little or no prior information about the system. On the other hand, empirical modelling allows to simultaneously evolve both the structure and the parameters of the models without imposing any assumptions regarding agents' expectations. In this study we apply a metaheuristic based on symbolic regression (SR) via genetic programming (GP). This modelling approach is characterised by the implementation of evolutionary algorithms for the search of the space of mathematical expressions that best fit a given dataset. See Dabhi and Chaudhary (2015), O'Neil et al. (2010), Poli et al. (2010), and White et al. (2013) for a detailed review of the main issues of GP.

Empirical studies of this sort mostly resort to quantifying population-wide consumer sentiment or expectations (Manski, 2018). However, even the early empirical studies (Zagórski and McDonnell, 1995), as well as the more recent ones (Sorić and Čižmešija, 2013; Easaw et al., 2013; Toussaint-Comeau and McGranahan, 2006) have shown that economic sentiment is remarkably heterogeneous across socio-economic or socio-demographic groups. We consider the observed heterogeneity by segregating the data provided by the EC consumer survey into strata according to age, income, education, gender, and occupation. For each consumer category within individual stratification variables, we apply GP and calculate a group-specific USI and assess its out-of-sample forecasting performance. The purpose of this empirical strategy is to scrutinize whether socio-demographics indeed do play a role in governing consumers' unemployment sentiment.

Finally, our last contribution stems from a more recent line of research, investigating how agents respond to economic news (Nguyen and Claus, 2013). We measure the reactions of consumer unemployment sentiment to various economic news at a disaggregate level, taking into account each individual socio-demographic group. This approach allows us to test for possible asymmetries in consumers' reaction to positive and negative news, building upon the renowned prospect theory (Kahneman and Tversky, 1979).

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

2. Literature review

Economic agents' capability to accurately assess and predict micro and macroeconomic variables has long been one of the focal points of applied economic research (Abberger, 2007; Billstam et al., 2017; Bruestle and Crain, 2015; Claveria et al., 2007; Clements, 2019; Graff, 2010; Hanson et al., 2005; Lehmann and Wohlrabe, 2017; Martinsen et al., 2014; Mitchell et al., 2005a,b; Vermeulen, 2014; Wilms et al., 2016). When focused on consumer sentiment, this strand of research has provided an abundance of empirical evidence in favour of the Katonian hypothesis, which refers to the leading properties of consumer confidence, or economic sentiment in general, with regards to various forms of economic activity. See e.g. Blood and Phillips (1995) for an overview of the concept. Consumer confidence is found to feed into aggregate household consumption (Nguyen and Claus, 2013), stock market returns (Chen, 2011), and the overall economic activity (Van Aarle and Kappler, 2012). However, several voices have been raised that different socio-economic or demographic groups generate immensely diverse perceptions and expectations of main macroeconomic indicators. The purpose of this section is to pinpoint socio-demographic characteristics that appear to be the driving force behind the heterogeneity of consumer confidence between various population sub-groups, and to briefly revise some of the main applications of GP in economic research.

2.1. Heterogeneity in consumers' assessments of the economy

The heterogeneity between the consumer confidence levels of different socio-demographic cohorts has been recognized even by the early contributions in this field of research. Coleman (1983) lays out the theoretical foundations of the interrelationships between socio-economic classes, their set of beliefs and consumption patterns. However, the obtained empirical evidence has remained mostly descriptive or based on unpretentious methodological frameworks such as simple t-tests. For example, Zagórski and McDonnell (1995) revealed that the composite US Consumer Confidence Index (CCI) significantly diminished with respondents' age. They also reported that CCI heavily depended on gender (men being more optimistic), occupational characteristics (white collar vs. blue collar) and work status (working vs. not working). Zagórski and McDonnell (1995) found that socio-economically deprived consumers tended to be less optimistic about the overall economic climate in the country and their own financial situation. It should be noted that these findings were not inherent solely to the aggregate concept of consumer confidence. Similar findings were also found for the notion of inflation sentiment. For example, Sorić and Čižmešija (2013) found that economically disadvantaged consumers tended to be more upward biased both in terms of inflation perceptions and expectations than the well-off consumers.

Toussaint-Comeau and McGranahan (2006) obtained very similar results, proving that the obtained inference has stood the test of time. They provide additional evidence by finding that race and schooling also played an important role in the process of generating economic perceptions and expectations. Toussaint-Comeau and McGranahan (2006) explained socio-economically driven differences by attaching the role of a moderation variable to unemployment. They reverted to the arguments of Ludvigson (2004), who asserted that the broad concept of consumer confidence firmly co-moves with unemployment. The latter variable thus affecting the strength of the relationship between consumers' socio-demographic characteristics and their judgements of the economy. Toussaint-Comeau and McGranahan (2006) corroborated such a moderation effect by finding that consumers' survey response patterns were highly correlated to their group-based unemployment experience rather than to the unemployment figures for the entire population. Lahiri and Zhao (2016) empirically proved the same rationale by finding that unemployment expectations explained a considerable part of the consumer sentiment variation.

It is also worthwhile mentioning a study by Tahal et al. (2016). The authors performed a meticulous stratification of Czech consumers into socioeconomic classes. When inspecting the levels of economic optimism within individual classes, the most pronounced differences aroused in terms of fear of poverty or finding suitable jobs.

To summarize the referenced literature, it is obvious that the information set contained in CS is largely related to the construct of unemployment sentiment. At the same time, similar empirical studies are mostly focused on the utilization of survey data in tracking business cycles (Cristiansen et al., 2014), economic activity (Eickmeier and Ng, 2011) or inflation (Sorić and Čižmešija, 2013). In this study, we aim to fill this gap by quantifying unemployment sentiment.

2.2. The news – Economy nexus

Whether or not news have an influence on economic decision-making depends on agents' cognitive capacity of processing information (Reis, 2006), their level of economic literacy (Burke and Manz, 2014) and their knowledge of the underlying models needed to properly assess economic phenomena, as well as on the saturation levels of news messages (de Vreese and Boomgaarden, 2006).

Overall, it has been widely documented that news have a certain effect on the behaviour of economic agents, their economic perceptions and expectations, as well as on the functioning mechanisms of the national economic system. Macroeconomic news announcements were reported to increase the volatility of government bond prices (Green, 2004). It was also found that they add to the conditional variance of both the exchange rate and the long-term interest rates (Fornari et al., 2002). Soroka (2006) found that media reports about the economy significantly fed into citizens' prevailing economic sentiment. In a similar vein, Ju (2008) even documented that news media had a larger role in shaping consumers' economic sentiment than do hard economic data. To summarise, the literature is unequivocal in the finding that news play a certain role in governing both economic sentiment and actual economic activity.

Many voices have also been raised that the nexus between news and the economy is not so straightforward and linear. On the contrary, there seem to be asymmetric effects, i.e. economic agents seem to react more intensively to bad news than to good news. The sole idea of possible asymmetries dates back to the well-known *prospect theory*. Kahneman and Tversky (1979), based on a series of experiments, found that economic agents are loss averse. They seem to react more strongly to a loss in their utility function than to a gain of an exact same magnitude.

Over the years, this theoretical notion has been translated into a large number of studies corroborating the stated asymmetry. The largest body of evidence refers to the asymmetric reaction of financial markets to bad vs. good news (Čižmešija et al., 2017; De Bondt and Thaler, 1985; Koutmos and Booth, 1995).

Some authors have found similar asymmetric effects in shifts of consumers' confidence. Nguyen and Claus (2013) obtained a significant effect of bad news on consumer sentiment for different socio-demographic groups of consumers, while the impact of good news was inconsequential. The authors also found that consumer sentiment reductions decreased the volume of private consumption, while sentiment increases did not have a significant effect.

Ju (2008) found that consumers adapted their sentiment much more intensively to an intensification of negative than of positive news reports. Moreover, Soroka (2006) showed that the prevailing economic sentiment did not react significantly at all to positive news reports on the economy, while negative stories did reflect in the growth of economic pessimism.

It seems, therefore, that there is plentiful evidence of asymmetries in the way consumers' expectations adapt to positive and negative news. In this research we perform such an analysis for unemployment expectations of different socio-demographic groups.

2.3. Applications of GP in economic research

Data-driven methods are being increasingly applied in economic research. One of the most versatile data-driven procedures is GP, which is a soft computing technique based on genetic algorithms (GAs). This approach for model approximation applies Darwinian principles during an evolution process in which an initial population of computer programs are bred through generations in order to find a set of analytical functions that best fit the data. One of the main advantages of this approach is that it does not assume an a priori model, allowing the model structure to vary during the evolution, and finally providing a mathematical expression that approximates the dynamics of the time series. This feature provides a quick overview of the most relevant interactions between variables and can help to identify new unknown links. As a result, due to its suitability to find patterns in large datasets and to handle complex modelling tasks, this empirical modelling approach is attracting researchers from different areas (Alexandridis et al., 2017; Álvarez-Díaz et al., 2009; Pan et al., 2019).

Koza (1992) used GP to estimate SR. One of the first applications of GP in economic research is that of Koza (1995), who used this procedure to assess the non-linear interactions

between the price level, gross national product, money supply, and the velocity of money. Other seminal works are those of Szpiro (1997a,b), Allen and Karjalainen (1999) and Beenstock and Szpiro (2002).

In recent years there has been a growing number of applications of evolutionary computing, especially in finance (Acosta-González and Fernández, 2014; Chen and Kuo, 2002; Fogel, 2006; Lawrenz and Westerhoff, 2003; Vasilakis et al., 2013; Wei, 2013; Yu et al., 2004), and for stock price forecasting in particular (Chen et al., 2008; Kaboudan, 2000; Larkin and Ryan, 2008; Sheta et al., 2015; Wilson and Banzhaf, 2009). See Drake and Marks (2002) for a review of the applications of GAs in financial forecasting.

By contrast, applications in economics have been scarcer, and mostly oriented to forecasting. See Chen and Kuo (2002) for a classification of the literature on the application of evolutionary computation to economics and finance. Regarding trade, Chen et al. (2010) applied GP in a vector error correction model to forecast US imports from China, and Klůčik (2012) estimated total exports and imports to Slovakia by means of SR. Regarding prices, Kronberger et al. (2011) identified interactions between economic indicators via SR in order to estimate the evolution of prices in the US. Álvarez-Díaz and Álvarez (2003, 2005) used GP to forecast exchange rates.

With regards to economic activity, Kotanchek et al. (2010) combined SR and GP to model GDP per capita using a wide range of countries and attributes. Duda and Szydło (2011) applied a version of GP to develop a set of empirical models to forecast GDP in Poland. Yang et al. (2015) used a data-driven approach based on SR to predict oil production in the US, using data from the 48 lower states since 1859. Marković et al. (2017) made use of 10 science and technology factors as inputs for GDP growth prediction in 28 EU countries. More recently, Claveria et al. (2018) applied a GP procedure to derive a set of building blocks used to estimate economic activity in the main European countries, finding that the proposed economic indicator outperformed a random walk model used as a benchmark in most cases.

In this study we implement SR via GP to find the relationship between a wide range of expectational variables and unemployment rates. By combining a SR approach with GP, we derive mathematical functional forms that can be regarded as the optimal combinations of survey variables that best fit the actual evolution of unemployment.

3. Data and methodology

This section briefly introduces the dataset and describes the employed methodology. On the one hand, we describe the experimental setup designed to extract consumers' unemployment expectations. On the other hand, we present the model used to construct economic news and assess their influence on the quantified unemployment expectations.

3.1. Data

As we aim to formalise the optimal interactions between consumers' expectations that best estimate the evolution of unemployment in the EA, we use two types of information: qualitative survey data and quantitative official statistics from 1998:04 to 2018:08. Regarding the former, we make use of seasonally adjusted balances from the joint harmonised EU consumer survey conducted by the European Commission (https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys_en). We employ the information from all available monthly questions (see Table 1). With regards to the quantitative information used as the target variable, we employ the standardised unemployment rates provided by the European Central Bank (ECB), which are also freely available (http://sdw.ecb.europa.eu/quickview.do?SERIES_KEY=132.STS.M.I8.S.UNEH.RTT000.4.000).

Table 1
Consumer Survey – Monthly variables

Financial situation over last 12 months	$X1_{it}$
Financial situation over next 12 months	$X2_{it}$
General economic situation over last 12 months	$X3_{it}$
General economic situation over next 12 months	$X4_{it}$
Price trends over last 12 months	$X5_{it}$
Price trends over next 12 months	$X6_{it}$
Unemployment expectations over next 12 months	$X7_{it}$
Major purchases at present	$X8_{it}$
Major purchases over next 12 months	$X9_{it}$
Savings at present	$X10_{it}$
Savings over next 12 months	$X11_{it}$
Statement on financial situation of household	$X12_{it}$

Table 1 contains the twelve survey variables used in the study, denoted as X_{it} , where i refers to each group of consumers and t to the time period. See Batchelor (1982, 1986) for an appraisal of the data from the consumer survey. Survey variables can be divided in judgements about the present ($X_{8_{it}}$, $X_{10_{it}}$ and $X_{12_{it}}$), perceptions about the past ($X_{1_{it}}$, $X_{3_{it}}$ and $X_{5_{it}}$), and expectations about the future ($X_{2_{it}}$, $X_{4_{it}}$, $X_{6_{it}}$, $X_{7_{it}}$, $X_{9_{it}}$ and $X_{11_{it}}$). We group judgements and perceptions, on the one hand, and expectations on the other in order to generate present and future expectations, which we average to construct the USI.

Table 2
Socio-demographic groups – Survey sub-categories

Income of the household	
1 st quartile included	RE1
2 nd quartile included	RE2
3 rd quartile included	RE3
4 th quartile included	RE4
Education of respondent	
Primary	ED1
Secondary	ED2
Further	ED3
Age of the respondent	
16-29	AG1
30-49	AG2
50-64	AG3
65+	AG4
Sex of the respondent	
Male	MAL
Female	FEM
Occupation of respondent	
Work full-time	PR8
Work part-time	PR9
Unemployed	PR0

The analysis is carried out for 16 sub-categories of consumers regarding the income of the household (four quartiles; RE4 being the lowest income quartile), the level of education (primary, secondary and further), the age and gender of the respondents (16-29, 30-49, 50-64, and 65 or more) and the occupation (full-time, part-time, unemployed). See Table 2.

3.2. Genetic programming as a tool for extracting expectations from survey data

GP can be regarded as a soft computing technique, particularly suited to automated problem solving. In this study, we use GP to estimate symbolic regression (SR). We design two independent SR experiments for each socio-demographic group. In the first, we link unemployment rates to judgements and perceptions so as to obtain a coincident USI. In the second, we link unemployment rates in $t+12$ to expectations in order to obtain a prospective USI. In both cases, genetic algorithms search the space of mathematical expressions that best fit our dataset, finding patterns across survey variables. This approach is particularly suitable, as there is no a priori information regarding the interactions between the different survey variables. GP simultaneously evolves the structure and the parameters of the expressions that formalize the interactions between survey variables that best fit the evolution of unemployment rates. As there is an arbitrary functional relationship between the set of survey variables, we link them to the unemployment rates in the EA by means of two SR models:

$$y_t = f(x1_{it}, x3_{it}, x5_{it}, x8_{it}, x10_{it}, x12_{it}) \quad (1)$$

$$y_{t+12} = f(x2_{it}, x4_{it}, x6_{it}, x7_{it}, x9_{it}, x11_{it}) \quad (2)$$

Where x_{it} are the different survey variables, and y_t is a scalar referring to the monthly rate of unemployment in the EA at time t . Expression (1) is used for judgements about the present and perceptions about the past, and expression (2) for expectations for the next 12 months.

For each socio-demographic group we implement GP to evolve a symbolic expressions (1) and (2) until a stopping criterion is reached. Given the trade-off between accuracy and simplicity, we have chosen a maximum number of 100 generations as the stopping criterion. For each simulation, a first random population of 75000 functions is generated. Afterwards, the best 7500 elements are selected for the evolutionary phase, where genetic operators (crossover and mutation) are applied to raise the fitness of the population generation after generation. We use the mean squared error (MSE) as the fitness function. See Claveria et al. (2017) for a detailed description on the implementation of GP.

3.3. A methodological framework for assessing asymmetries in news effects

There are valid reasons for the hypothesis that consumers respond to various inputs from their surroundings by revising their expectations in an asymmetrical manner (Kahneman and Tversky, 1979). Numerous signals come to consumers, and the sources vary from their family and friends, to media and changes in the real economy (store purchases, job situation, industry changes, etc.). Given the wide range of inputs, consumers do not necessarily react in the same manner to all of them. To fully understand and quantify how news about the real economy influence consumers' unemployment expectations, we employ a framework developed by Nguyen and Claus (2013).

First, we define news in our analysis, and then explain how news are linked to expectations of various heterogenous consumer groups. Secondly, we inspect the mapping between news and unemployment expectations.

As we are interested in observing how consumers' expectations about unemployment react to new economic information, we focus on information about several relevant variables depicting the current state of the EA with regards to stock markets, production, and inflation. Specifically, we extract news about (i) the Eurostoxx50 index (source: Thomson Reuters), (ii) industrial production in the manufacturing sector (source: Eurostat), (iii) the Harmonized Index of Consumer Prices (source: Eurostat), (iv) the index of retail turnover (source: Eurostat), and (v) unemployment rates in the EA (source: ECB).

In this study, we define news in a particular variable as the first difference of the variable:

$$\varepsilon_t = \Delta s_t \quad (3)$$

where s_t refers to the natural logarithm of one of the following four monthly variables: *retail* (retail turnover), *ind* (industrial production), *unem* (unemployment rate), and *hicp* (inflation). Since stock market data is available on a daily basis, we use its daily figures to match the CS conduction time precisely. Stock market news are then defined as the average of differences in the first 21 days as only working days are included:

$$\varepsilon_t = \sum_{k=2}^{21} \Delta s_k \quad (4)$$

where s_k is the logarithm of the stock market index.

In order to analyse the effect of news on expectations, we decompose ε_t into positive and negative parts in the following way:

$$\varepsilon_t^+ = \max(\varepsilon_t, 0) \text{ and } \varepsilon_t^- = \min(\varepsilon_t, 0) \quad (5)$$

where ε_t^+ denotes *good* news, linked to a positive change in the economy, and ε_t^- denotes *bad* news. We want to note that news about unemployment and inflation have an inverse interpretation.

The second part is built upon mapping lagged news with unemployment expectations in the following manner:

$$unem_exp_t = c + \sum_{i=0}^k (\beta_i^+ \varepsilon_{t-i}^+ + \beta_i^- \varepsilon_{t-i}^-) + e_t \quad (6)$$

where $unem_exp_t$ is the logarithm of unemployment expectations (USI) as computed in the previous section, c is a constant, β_i^+ and β_i^- are the corresponding good/bad news coefficients, and e_t is the error term. The number of lags $k = 0, \dots, 6$ is determined by minimizing the Akaike criterion. To inspect for possible asymmetries in the model, we perform two Wald tests ($Wald^+$ and $Wald^-$) with the corresponding hypothesis respectively $H_0: \beta_0^+ = \beta_1^+ = \dots = \beta_k^+ = 0$ and $H_0: \beta_0^- = \beta_1^- = \dots = \beta_k^- = 0$. If the null is rejected for the $Wald^+$ test, then good news in the corresponding variable s_t has a significant effect. However, if the null is not rejected, then there is not enough evidence to support the influence of good news. The same reasoning is valid for the $Wald^-$ test.

Special care is taken of the lag choice in equation (6). Unemployment expectations are calculated using survey data, which are published by the end of each month, so we cannot match them with the corresponding monthly figures of *ind*, *retail*, *hicp*, and *unem*, as consumers do not yet have the opportunity to react to them. The detailed schedule is available at ECB's website (<https://ec.europa.eu/eurostat/news/release-calendar>). So, instead of using lags $i = 0, 1, \dots, k$ (maximum k is six) in the equation (6) we use $i = 1, 2, \dots, k$ (maximum k is seven). The rationale behind this mapping is to assess only the data that consumers could have heard about and accordingly could have used to adjust their expectations. Stock market data is available daily, so we use lags as in equation (6).

We apply this framework to investigate if there are differences between the reactions of the various socio-demographic sub-groups: four income cohorts, four age groups, three different employment status groups, three education levels, and male vs. female comparison.

4. Empirical results

In this section we first assess the forecasting performance of the symbolic expressions obtained through the evolution process for each group of consumers. These expressions can be regarded as USI for the present and for the future respectively.

Tables 3 and 4 contain the evolved expressions that link survey information to unemployment rates for each socio-demographic group, both for perceptions and judgements about the present (Table 3) and for expectations about the future (Table 4). The symbolic expressions obtained for each category in the first round of experiments can be regarded as coincident sentiment indicators of unemployment, while the expressions evolved in the second round are leading sentiment indicators of unemployment.

Next, we reserve the last 32 periods of the dataset ranging from 2016:01 to 2018:08 to assess the out-of-sample forecasting performance of the evaluation of the USI obtained by aggregating present and future expectations according to the period they are referred to. The splitting strategy of the sample period followed in this study is the one recommended in the literature (Bishop, 1995; Hastie et al., 2009). In Table 5 we present the root mean squared forecasting error (RMSFE) for each socio-demographic group.

In Table 3 we can see that variables *X1* (financial situation over the last 12 months) and *X5* (price trends over the last 12 months) are the most relevant survey variables, followed by *X3* (general economic situation over the last 12 months) and *X10* (savings at present). Regarding future economic expectations (Table 4), variables *X9* (major purchases over the next 12 months) and *X11* (savings over the next 12 months) are the ones that appear more frequently in most consumer groups.

Table 3

Evolved expressions for each socio-demographic group – Coincident unemployment indicators

RE1	$-0.03 * X5 - 0.01 * X8 - 0.27 * X12 + 9.73$
RE2	$-\frac{X3 + X5}{0.1 * X10 - (X3 + 2) * \frac{X10 - 2.88}{X5}} + 10$
RE3	$10.5 + \frac{8.11}{X12} + \frac{X5 - 0.55}{X1 - 10}$
RE4	$\frac{X5 + \frac{10}{0.1 * \frac{X10}{X11}}}{-X1 + X12} + 10$
ED1	$\frac{-X3 + 10 * (X12) * (-X1 + X12) - (X12) * (X1 + X5 + X12 - 9.5)}{(X12) * (-X1 + X12)}$
ED2	$-\frac{X1 + X10}{-0.24 * X1 - 0.24 * X5 + 0.76 * X10 + \frac{1}{X10 - 1.03}} + 10$
ED3	$\frac{-0.07 * X1 * X1 + 0.32 * X5 - 96.63}{-10}$
AG1	$\frac{-3.83 * X1 - 1.1 * X5 + 10 * X10 + 1}{X10}$
AG2	$\frac{12 * X1 + X5 - 10 * X12 + 15.35}{X1 - X12 + 1.53}$
AG3	$10.7 - \frac{0.1}{\left(0.5 * \frac{X1}{X5 - 13.12}\right) * \left(0.5 * \frac{X1}{X5}\right)}$
AG4	$\frac{10 * (X3) * (-X3 + X12 + 3.63) - (X3) * (X3 + 2 * X5 - 20) + 2}{(X3) * (-X3 + X12 + 3.63)}$
MAL	$-2 * \frac{X1}{X10} + 3.8 * \frac{X8}{4.12 * X10} + 9$
FEM	$10 + \frac{X5}{X3} \left(\frac{X3 + X5 - 1.84}{X10}\right)$
PR8	$\frac{-5.05 * X1 - 1.1 * X5 + 8.91 * X10 - 0.11}{X10}$
PR9	$-\frac{X5}{-X3 + \frac{X12}{X10}} - 0.07 * X10 + 12$
PR0	$\frac{11 * X1 + 10 * X3 + 11 * X5 + X10 - 205}{X1 + X3 + X5 - 20.5}$

Table 4

Evolved expressions for each socio-demographic group – Leading unemployment indicators

RE1	$\frac{-0.1 * X7 + 0.1 * X9 + (X7 - 0.57) * (-0.2 * X9 + 4.97)}{X7 - 0.57}$
RE2	$\frac{(-0.26 * X9 + 3.08) * (X9 + X11) - 34.14}{X9 + X11}$
RE3	$\frac{-0.22 * X9 * X11 + 10 * X11 + (-0.25 * X9 + 4.52) * (X9 * X9)}{X9 * X9}$
RE4	$\frac{64.28 * X2 - (X2 - 5.93) * (X2 + 1) - 352.68}{6.43 * X2 - 35.27}$
ED1	$\frac{X2 + (X7) * (-0.24 * X9 + 4.69)}{X7}$
ED2	$-0.38 * X9 + 0.09 * X11 + 2.89$
ED3	$\frac{-0.18 * X2 * (X2 - 5.99) + (-0.25 * X9 + 5.49) * (X9)}{X9}$
AG1	$0.32 * X2 * \frac{X2}{X9 * X9} - 0.29 * X9 + 4.79$
AG2	$-0.41 * X9 - \frac{10}{X9 + 10} - \frac{X9 - X11}{(0.41 * X9 + 0.35) * (X9)}$
AG3	$\frac{-0.78 * X9 * X9 - 10.34 * X9}{2 * X9 + 27.09}$
AG4	$\frac{(-0.22 * X9 + 4.37) * (0.88 * X9 + 10.1) - 1.86}{0.88 * X9 + 10.1}$
MAL	$\frac{-X11 + (-0.23 * X9 + 4.38) * (X9 + 3.4)}{X9 + 3.4}$
FEM	$\frac{-X11 + (-0.28 * X9 + 4) * (X9) - 9.31}{X9}$
PR8	$0.04 * X7 + 8.53$
PR9	$-0.07 * X2 - 0.1 * X9 + 7.34$
PR0	$4.66 - \frac{(X4 - 1.40) * \frac{X6 - 4.72}{X9} - 0.5}{X11} + \frac{0.49}{-\frac{2.92}{X9}}$

Table 5

Out-of-sample forecast accuracy – RMSFE (2016:M1–2018:M8)

RMSFE	RE1	RE2	RE3	RE4	ED1	ED2	ED3	M	F
Present	1.52	1.79	4.23	0.28	0.72	0.87	0.29	10.20	23.62
Future	0.60	0.58	0.58	0.41	0.34	0.35	0.53	0.30	0.35
Aggregate expectations	0.88 (8.40)	0.59 (2.59)	1.12 (5.71)	0.28 (2.94)	0.45 (5.38)	0.27 (3.27)	0.22 (3.23)	2.26 (4.96)	4.94 (2.66)

RMSFE	AG1	AG2	AG3	AG4	PR0	PR8	PR9
Present	0.25	0.43	1.59	0.66	-	-	-
Future	1.06	0.49	0.29	0.48	0.16	0.78	0.44
Aggregate expectations	0.33 (2.63)	0.09 -	0.36 (2.39)	0.37 (5.17)	-	-	-

Note: Due to data availability for PRO, PR8 and PR9 the RMSFE can only be computed for the future up until 2017:04. Diebold-Mariano test statistic with Newey-West estimator between brackets. Null hypothesis: the difference between the two competing series is non-significant. A negative sign of the statistic implies that the second series has bigger forecasting errors. Critical value at the 5% level: 2.028.

Regarding the out-of-sample performance of USI in the *future* setting, Table 5 shows that the results are similar across the different socio-demographic groups. The group where we observe the bigger differences among categories is the occupation, where the lowest forecast errors are obtained for the unemployed respondents. This reveals a very interesting, but quite unexpected result. Economic and psychological studies of consumer behaviour suggest that higher socio-economic classes tend to be future oriented, while the less fortunate are mainly occupied with building short-term prosperity of their families and their inner social circle (Henry, 2005; Piff et al., 2010). However, Table 5 strongly contradicts that for the case of unemployment sentiment.

Regarding the *present* setting, the results seem much more heterogeneous. Quite curiously, when it comes to income, the disadvantaged consumers (RE4 category) anticipate future unemployment rates with the highest accuracy. This can be linked to the fact that the unemployed consumers also exhibit multiple times lower RMSFEs than the employed ones (PR8 and PR9). Both of these findings are explicable through the availability heuristic (Kahneman and Tversky, 1974). Consumers obviously generate more accurate predictions of actual unemployment rates if they themselves have the experience of being unemployed, or if they are currently facing job insecurity.

The present setting also reveals that education is a significant predictor of forecasting accuracy. As expected, consumers with the highest education level (ED3) provide remarkably lower RMSFEs than the less-educated respondents. This result is deeply rooted in the concept

of economic literacy (Burke and Manz, 2014). It is obvious that formal economic knowledge, willingness to gather information, as well as the cognitive abilities for organizing and processing them, are highly related to the overall precision of unemployment expectations.

Males also produce significantly lower RMSFEs than females. This again comes as no surprise since women are in general found to be more pessimistic (Jacobsen et al., 2014) and more biased when it comes to assessing hard macroeconomic data (Zagórski and McDonnell, 1995).

When expectations are aggregated, the lowest RMSFE is obtained for expectations of respondents aged between 30 and 49 years (AG2). In order to evaluate whether the reduction in forecast accuracy between each group of consumers and the group with the best forecast results (AG2) is statistically significant, we 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 series has bigger forecasting errors. Results presented in Table 5 show that the reduction in forecast accuracy obtained by respondents aged between 30 and 49 years when compared with the rest of consumer groups is always significant.

Next, we explore the reaction of the different groups of consumers to a range of economic news and we evaluate whether there is asymmetry in the response to news where consumers react to bad news but not to good news. Equations (5) and (6) are estimated using the Newey-West Heteroskedasticity-and autocorrelation-consistent (HAC) standard errors. In doing so, the coincident (Table 3) and leading indicators of unemployment (Table 4) are averaged for each socio-demographic group to obtain the dependent variable in equation (6). Table 6 summarises the impact of stock market news on the expectations of the assessed socio-demographic subgroups of consumers. Estimations for the three unemployment statuses (PR0, PR8 and PR) categories are done for the period 1999:04 – 2016:04. The time span for all other subgroups is 1999:04 - 2018:08.

In Table 6, we can observe that consumers' unemployment expectations seem to respond to stock market news in the case of low-income consumers, the youngest and the oldest consumer groups, as well as women and the unemployed respondents. This is in line with the findings of Zagórski and McDonnell (1995).

Table 6

Impact of news on the expectations of heterogeneous consumer groups

News	p-value	RE1	RE2	RE3	RE4	ED1	ED2	ED3	M	F
Stock market	<i>Wald</i> ⁺	0.64	0.02	0.00	0.24	0.52	0.10	0.37	0.57	0.07
	<i>Wald</i> ⁻	0.35	0.71	0.04	0.07	0.23	0.16	0.27	0.12	0.74
Industrial production	<i>Wald</i> ⁺	0.19	0.17	0.35	0.00	0.48	0.21	0.69	0.09	0.81
	<i>Wald</i> ⁻	0.09	0.27	0.02	0.00	0.14	0.69	0.97	0.06	0.68
Unemployment	<i>Wald</i> ⁺	0.50	0.00	0.25	0.00	0.02	0.89	0.02	0.58	0.40
	<i>Wald</i> ⁻	0.92	0.01	0.00	0.02	0.05	0.06	0.11	0.89	0.07
Inflation	<i>Wald</i> ⁺	0.00	0.02	0.07	0.01	0.00	0.01	0.06	0.06	0.00
	<i>Wald</i> ⁻	0.14	0.86	0.64	0.63	0.43	0.58	0.42	0.37	0.83
Retail turnover	<i>Wald</i> ⁺	0.74	0.71	0.00	0.23	0.91	0.57	0.91	0.92	0.47
	<i>Wald</i> ⁻	0.68	0.98	0.00	0.02	0.88	0.77	0.57	0.81	0.60
	p-value	AG1	AG2	AG3	AG4	PR0	PR8	PR9		
Stock market	<i>Wald</i> ⁺	0.00	0.26	0.42	0.19	0.07	0.10	0.24		
	<i>Wald</i> ⁻	0.03	0.38	0.07	0.05	0.15	0.72	0.59		
Industrial production	<i>Wald</i> ⁺	0.75	0.56	0.83	0.63	0.55	0.59	0.43		
	<i>Wald</i> ⁻	0.93	0.52	0.67	0.68	0.67	0.31	0.77		
Unemployment	<i>Wald</i> ⁺	0.00	0.07	0.15	0.05	0.28	0.00	0.02		
	<i>Wald</i> ⁻	0.02	0.04	0.04	0.01	0.06	0.00	0.00		
Inflation	<i>Wald</i> ⁺	0.07	0.04	0.01	0.00	0.00	0.01	0.10		
	<i>Wald</i> ⁻	0.05	0.01	0.03	0.00	0.05	0.02	0.24		
Retail turnover	<i>Wald</i> ⁺	0.36	0.93	0.93	0.86	1.00	0.84	0.45		
	<i>Wald</i> ⁻	0.81	0.62	0.80	0.89	0.49	0.93	0.98		

Notes: *Wald*⁺ is the p-value for Wald test with $H_0: \beta_0^+ = \beta_1^+ = \dots = \beta_k^+ = 0$. *Wald*⁻ is the p-value for Wald test with $H_0: \beta_0^- = \beta_1^- = \dots = \beta_k^- = 0$. ‘Bold’ numbers denote significant estimates at the 10% significance level. Estimated parameters are available in the Appendix Table A1 – A5.

The assessed unemployment expectations seem to be quite non-responsive to news about industrial production. We find a significant effect, again, for the two lowest income categories. A noteworthy relationship is found for men, both for positive and negative news. Although the latter finding contradicts the hypothesis that the expectations of economically deprived categories should be more responsive to economic news, this comes as no surprise. Namely, men historically generate a lion’s share of total employment in the industrial sector (Aragón et al., 2018; United Nations, 2015), so it is no wonder that their overall unemployment expectations react to shifts in that sector. Female expectations, on the other hand, do not.

As it might be expected, Table 6 elucidates that consumers’ unemployment sentiment reacts the strongest to news in unemployment itself. The stated link is significant for the majority of assessed socio-economic categories. There does not seem to be any specific pattern when it comes to socio-demographic characteristics as potential moderators of the observed relationships.

A very similar effect is also observed for the inflation news. The trade-off between unemployment and inflation is widely recognized in theoretical literature. A classical Keynesian response to an increase of cyclical unemployment would be stimulating aggregate demand, assuming that the labour supply curve is infinitely elastic. However, Phillips (1958) postulates that raising aggregate demand comes at a cost. It puts upward pressure on the aggregate price level since employers have to counteract by paying higher wages to their workers. The specified trade-off between unemployment and inflation is also widely empirically confirmed. See Bhattarai (2016) for a recent literature review and empirical verification. Having in mind the empirical results presented in Table 6, it is evident that consumers easily recognise these economic concepts and translate them into their unemployment expectations. Finally, retail news seem to be significant only for the expectations of the two low-income categories, again corroborating the role of socio-economic deprivation in governing expectations.

A glance at Table 6 reveals that there are indeed some asymmetries in consumers' responses to economic news. In general, consumers tend to react more strongly to negative news (falling stock prices/industrial production/retail trade volume and increasing prices) than to positive ones. The asymmetry is most pronounced for inflation news. Out of the 12 found significant relationships (at the 5% significance level), 9 of them refer to negative news (about growing prices).

The stated asymmetries are treated in various ways by different authors. The literature usually recognises that output contractions and expansions necessitate monetary policy responses of different intensity (Clark et al. 1996; Dolado et al., 2005; Schaling 2004). However, the findings from Table 6 echo another line of research: the one dealing with asymmetric preferences of economic agents (Santoro et al. 2014; Surico, 2007). This strand of literature postulates that the monetary authorities might be more concerned about overshooting the targeted inflation rate than about undershooting it (Surico, 2007). It should be noted that the empirical verification of the asymmetric preferences hypothesis heavily depends on the geographical scope of the studies, as well as on their methodological frameworks. Surico (2007) found no evidence of asymmetry for the United States, while the results of Dolado et al. (2005) spoke in favour of it for Germany, France, Spain and the EA. Moreover, Sznajderska (2014) also documented asymmetric reactions in case of the National Bank of Poland. The results presented in Table 6 are more inclined to the asymmetric preferences literature, only focused on consumers instead of central bankers. News on rising inflation are documented to resonate considerably more intensively when it comes to consumers' unemployment expectations.

5. Conclusion

In this study we have evaluated the effect of economic news on consumer unemployment expectations for sixteen socio-demographic groups. To this end, we have designed an unemployment sentiment indicator. By means of genetic programming we have evolved the expressions that optimally link unemployment rates in the Euro Area to qualitative expectations about a wide range of economic variables. Then we have assessed the out-of-sample forecast accuracy of the evolved indicators. We have obtained better forecasts for the leading unemployment sentiment indicator than for the coincident one.

Regarding the different survey variables, we have found that the financial situation and the price trends over the last 12 months are the survey variables with the highest predictive power for the present, while major purchases and savings for the next 12 months are the survey variables with the highest predictive power for the future.

Finally, we have extracted economic news about inflation, industrial production, and stock markets and have linked them to unemployment expectations in order to analyse the reaction of the different groups of consumers. To test whether there is asymmetry in the response to news, where consumers react to bad news but not to good news, we have matched positive and negative news with unemployment sentiment using a distributed lag regression model for each news item. We have found that the responses of consumers' unemployment expectations to economic news tend to be stronger in the case of negative news, especially for inflation.

There are still several limitations to be addressed. Given that we used a data-driven method to estimate consumers' unemployment expectations, our quantitative estimates lack any theoretical background. Extending the analysis to different European countries would allow us to examine the similarities between the obtained functional forms across countries. Another issue left for further research is the implementation of alternative evolutionary algorithms to analyse if they could improve the forecast accuracy of the empirically-generated estimates of expectations.

References

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

- Acosta-González, E. and Fernández, F. (2014). Forecasting financial failure of firms via genetic algorithms. *Computational Economics*, 43(2): 133–157.
- Alexandridis, A. K., Kampouridis, M. and Cramer, S. (2017). A comparison of wavelet networks and genetic programming in the context of temperature derivatives. *International Journal of Forecasting*, 33(1): 21–47.
- Allen F. and Karjalainen, R. (1999). Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51(2): 245–271.
- Álvarez-Díaz, M. (2019). Is it possible to accurately forecast the evolution of Brent crude oil prices? An answer based on parametric and nonparametric forecasting methods. *Empirical Economics*. Forthcoming.
- Álvarez-Díaz, M. and Álvarez, A. (2003). Forecasting exchange rates using genetic algorithms. *Applied Economics Letters*, 10(6): 319–322.
- Álvarez-Díaz, M. and Álvarez, A. (2005). Genetic multi-model composite forecast for non-linear prediction of exchange rates. *Empirical Economics*, 30(3): 643–663.
- Álvarez-Díaz, M., Mateu-Sbert, J. and Rosselló-Nadal, J. (2009). Forecasting tourist arrivals to Balearic Islands using genetic programming. *International Journal of Computational Economics and Econometrics*, 1(1): 64–75.
- Aragón, F., Rud, J. P. and Toews, G. (2018). Resource shocks, employment, and gender: Evidence from the collapse of the UK coal industry. *Labour Economics*, 52: 54–67.
- Batchelor, R. A. (1982). Expectations, output and inflation. *European Economic Review*, 17(1): 1–25.
- Batchelor, R. A. (1986). The psychophysics of inflation. *Journal of Economic Psychology*, 7(3): 99–120.
- Beenstock M. and Szpiro, G. (2002). Specification search in nonlinear time-series models using genetic algorithms. *Journal of Economic Dynamic and Control*, 26(5): 811–835.
- Bhattacharai, K. (2016). Unemployment-inflation trade-offs in OECD countries. *Economic Modelling*, 58: 93–103.
- Białowolski, P. (2019). Economic sentiment as a driver for household financial behavior. *Journal of Behavioral and Experimental Economics*, 80: 59–66.
- Billstam, M., Frändén, K., Samuelsson, J. and Österholm, P. (2017). Quasi-real-time data of the Economic Tendency Survey. *Journal of Business Cycle Research*, 13(1): 105–138.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.
- Blood, D. J. and Phillips, C. B. (1995). Recession headline news, consumer sentiment, the state of the economy, and presidential popularity: A time series analysis 1989-1993. *International Journal of Public Opinion Research*, 7(1): 2–22.
- Bruestle, S. and Crain, W. M. (2015). A mean-variance approach to forecasting with the consumer confidence index. *Applied Economics*, 47(23): 2430–2444.
- Burke, M.A. and Manz, M. (2014). Economic literacy and inflation experiments: evidence from a laboratory experiment. *Journal of Money, Credit and Banking*, 46(7): 1421–1456.
- Chen, S. (2011). Lack of consumer confidence and stock returns. *Journal of Empirical Finance*, 18(2): 225–236.

- Čižmešija, M., Sorić, P. and Matošec, M. (2017). Zagreb stock exchange and the (a)symmetric effects of news. Proceedings of the 14th International Symposium on Operational Research, Bled, Slovenia, Zadnik-Stirn, L. et al. (Eds.), pp. 403–408.
- Chen, X., Pang, Y. and Zheng, G. (2010). Macroeconomic forecasting using GP based vector error correction model. In J. Wang (Ed.), *Business Intelligence in Economic Forecasting: Technologies and Techniques* (pp. 1–15). Hershey, PA: IGI Global.
- Chen, S. H. and Kuo, T. W. (2002). Evolutionary computation in economics and finance: A bibliography. In S. H. Chen (Ed.), *Evolutionary Computation in Economics and Finance* (pp. 419–455). Heidelberg: Physica-Verlag.
- Chen, S. H., Kuo, T. W. and Hoi, K. M. (2008). Genetic programming and financial trading: How much about “what we know”. In C. Zopounidis et al. (Eds.), *Handbook of financial engineering* (pp. 99–154). New York: Springer.
- Clark, P., Laxton, D. and Rose, D. (1996). Asymmetry in the U.S. output-inflation nexus. *IMF Staff Papers*, 43(1): 216–251.
- Claveria, O. (2019a). A new consensus-based unemployment indicator. *Applied Economics Letters*, 26(10): 812–817.
- Claveria, O. (2019b). Forecasting the unemployment rate using the degree of agreement in consumer unemployment expectations. *Journal for Labour Market Research*, 53(3): 1–10.
- Claveria, O., Monte, E. and Torra, S. (2017). Using survey data to forecast real activity with evolutionary algorithms. A cross-country analysis. *Journal of Applied Economics*, 20(2): 329–349.
- Claveria, O., Monte, E. and Torra, S. (2018). A data-driven approach to construct survey-based indicators by means of evolutionary algorithms. *Social Indicators Research*, 135(1): 1–14.
- Claveria, O., Monte, E. and Torra, S. (2019a). Evolutionary computation for macroeconomic forecasting. *Computational Economics*, 53(2): 833–849.
- Claveria, O., Monte, E. and Torra, S. (2019b). Empirical modelling of survey-based expectations for the design of economic indicators in five European regions. *Empirica*, 46(2): 205–227.
- Claveria, O., Pons, E. and Ramos, R. (2007). Business and consumer expectations and macroeconomic forecasts. *International Journal of Forecasting*, 23(1): 47–69.
- Clements, M. P. (2019). *Macroeconomic survey expectations*. Palgrave Texts in Econometrics. Palgrave Macmillan.
- Coleman, R. P. (1983). The continuing significance of social class to marketing. *Journal of Consumer Research*, 10(3): 265–280.
- Cristiansen, C., Eriksen, J. N. and Moller, S. V. (2014). Forecasting US Recessions: The Role of Sentiments. *Journal of Banking & Finance*, 49: 459–468.
- Dabhi, V. K. and Chaudhary, S. (2015). Empirical modeling using genetic programming: a survey of issues and approaches. *Natural Computing*, 14(2): 303–330.
- De Bondt, W. and Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, XI(3): 793–805.
- de Vreese, C. and Boomgaarden, H. (2006). Media effects on public opinion about the enlargement of the European Union. *Journal of Common Market Studies*, 44(2): 419–436.

- Dias, J. (2017). Unemployment and sovereign debt crisis in the Eurozone: A k-means-r analysis, *Physica A*, 482: 108–117.
- Diebold, F. X. and Mariano R. S. (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13(3): 253–263.
- Dolado, J. J., María-Dolores, R. and Naveira, M. (2005). Are monetary policy reaction functions asymmetric? The role of nonlinearity in the Phillips curve. *European Economic Review*, 49: 485–503.
- Duda, J. and Szydło, S. (2011). Collective intelligence of genetic programming for macroeconomic forecasting. In P. Jędrzejowicz et al. (Eds.), *Computational Collective Intelligence. Technologies and Applications* (pp. 445–454). Berlin: Springer.
- Easaw, J., Golinelli, R. and Malgarini, M. (2013). What determines households inflation expectations? Theory and evidence from a household survey. *European Economic Review*, 61: 1–13.
- Eickmeier, S. and Ng, T. (2011). Forecasting national activity using lots of international predictors: An application to New Zealand. *International Journal of Forecasting*, 27(2), 496-511.
- European Central Bank (2012). Euro Area Labour Markets and the Crisis. Occasional Paper Series No. 138.
- European Commission (2016). The Joint Harmonized EU Programme of Business and Consumer Surveys, User Guide.
- European Commission (2018). Standard Eurobarometer 89. Spring 2018, Report: Public Opinion in the European Union.
- Fogel, D. B. (2006). *Evolutionary computation. Toward a new philosophy of machine intelligence* (Third Edition). Hoboken, NJ: John Wiley & Sons.
- Fornari, F., Monticelli, C., Pericolli, M. and Tivegna, M. (2002). The impact of news on the exchange rate of the lira and long-term interest rates. *Economic Modelling* 19(4): 611–639.
- Graff, M. (2010). Does a multi-sectoral design improve indicator-based forecasts of the GDP growth rate? Evidence from Switzerland. *Applied Economics*, 42(21): 2759–2781.
- Green, T. C. (2004). Economic news and the impact of trading on bond prices. *The Journal of Finance*, 59(3): 1201–1233
- Hansson, J., Jansson, P. and Löf, M. (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting*, 30(1): 65–77.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). The elements of statistical learning. New York: Springer-Verlag.
- Henry, P.C. (2005). Social class, market situation, and consumers' metaphors of (dis)empowerment, *Journal of Consumer Research*, 31(4): 766–778.
- Hutter, C. and Weber, E. (2015). Constructing a new leading indicator for unemployment from a survey among German employment agencies. *Applied Economics*, 47(33): 3540–3558.
- International Labour Office (2010). *Global Employment Trends*. January 2010 (Geneva: ILO).
- Jacobsen, B., Lee, J. B., Marquering, W. and Zhang, C. Y. (2014). Gender differences in optimism and asset allocation. *Journal of Economic Behavior & Organization*, 107B: 630–651.

- Ju, Y. (2008). The asymmetry in economic news coverage and its impact on public perception in South Korea. *International Journal of Public Opinion Research*, 20(2): 237–249.
- Kaboudan, M. A. (2000). Genetic programming prediction of stock prices. *Computational Economics*, 16(3): 207–236.
- Kahneman, D. and Tversky, A. (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, New Series, 185(4157): 1124–1131.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2): 263–292.
- Klůčik, M. (2012). Estimates of foreign trade using genetic programming. *Proceedings of the 46 the scientific meeting of the Italian Statistical Society*.
- Kotanchek, M. E, Vladislavleva, E. Y. and Smits, G. F. (2010). Symbolic regression via genetic programming as a discovery engine: Insights on outliers and prototypes. In R. Riolo et al., (Eds.), *Genetic Programming Theory and Practice VII, Genetic and Evolutionary Computation* Vol. 8 (pp. 55–72). Springer Science+Business Media, LLC.
- Koutmos, G. and Booth, G. G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance*, 14(6): 747–762
- Koza, J. R. (1992). Genetic programming: On the programming of computers by means of natural selection. Cambridge, MA: MIT Press.
- Koza, J. R. (1995). Genetic Programming for econometric modeling. In S. Goonatilaje and P. Treleaven (Eds.), *Intelligent Systems for Finance and Business* (pp. 251–269). London: John Wiley and Sons.
- Krištić, I. R., Dumančić, L. R. and Arčabić, V. (2018). Persistence and stochastic convergence of euro area unemployment rates. *Economic Modelling*, 76: 192–198.
- Kronberger, G., Fink, S., Kommenda, M. and Affenzeller, M. (2011). Macro-economic time series modeling and interaction networks. In C. Di Chio et al. (Eds.), *EvoApplications*, Part II (pp. 101–110). LNCS 6625.
- Lahiri, K. and Zhao, Y. (2016). Determinants of Consumer Sentiment over business cycles: Evidence from the US surveys of consumers. *Journal of Business Cycle Research*, 12(2): 187–215.
- Lane, P. R. (2012). The European Sovereign Debt Crisis. *Journal of Economic Perspectives*, 26(3): 49–68.
- Larkin, F. and Ryan, C. (2008). Good news: Using news feeds with genetic programming to predict stock prices. In M. O’Neil et al. (Eds.), *Genetic Programming* (pp. 49–60). Berlin: Springer-Verlag.
- Lawrenz, C. and Westerhoff, F. (2003). Modeling exchange rate behaviour with a genetic algorithm. *Computational Economics*, 21(3): 209–229.
- Lehmann, R., and Wohlrabe, K. (2017). Experts, firms, consumers or even hard data? Forecasting employment in Germany. *Applied Economics Letters*, 24(4): 279–283.
- Lucarelli, B. (2011). German neomercantilism and the European sovereign debt crisis. *Journal of Post Keynesian Economics*, 34(2): 205–224.
- Ludvigson, S.C. (2004). Consumer Confidence and Consumer Spending. *Journal of Economic Perspectives*, 18(2): 29–50.

- Manski, C. F. (2018). Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise, In: *NBER Macroeconomics Annual 2017*, Vol. 32, Eichenbaum, M. and Parker, J. A. (Eds.), pp. 411–471.
- Marković, D., Petković, D., Nikolić, V., Milovančević, M. and Petković, B. (2017). Soft computing prediction of economic growth based in science and technology factors. *Physica A*, 465: 217–220.
- Mitchell, J., Smith, R. and Weale, M. (2005a). Forecasting manufacturing output growth using firm-level survey data. *The Manchester School*, 73(4): 479–499.
- Mitchell, J., Smith, R. and Weale, M. (2005b). An indicator of monthly GDP and an early estimate of quarterly GDP growth. *The Economic Journal*, 115(501): F108–F129.
- Martinsen, K., Ravazzolo, F. and Wulfsberg, F. (2014). Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting*, 30(1): 65–77.
- Nguyen, V. H. and Claus, E. (2013). Good news, bad news, consumer sentiment and consumption behavior. *Journal of Economic Psychology*, 39: 426–438.
- O’Neill, M., Vanneschi, L., Gustafson, S. and Banzhaf, W. (2010). Open issues in genetic programming. *Genetic Programming and Evolvable Machines*, 11(3-4): 339–363.
- Pan, X., Uddin, M. K., Ai, B., Pan, X. and Saima, U. (2019). Influential factors of carbon emissions intensity in OECD countries: Evidence from symbolic regression. *Journal of Cleaner Production*, 220: 1194–1201.
- Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom. *Economica* 1861-1957: 283–299.
- Piff, P. K., Kraus, M. W., Cote, S., Cheng, B. H. and Keltner, D. (2010). Having less, giving more: the influence of social class on prosocial behavior. *Journal of Personality and Social Psychology*, 99(5): 771–784.
- Pissarides, C. A. (2013). Unemployment in the Great Recession. *Economica*, 80(319): 385–403.
- Poli, R., Vanneschi, L., Langdon, W. B. and Mcphee, N. F. (2010). Theoretical results in genetic programming: The next ten years?. *Genetic Programming and Evolvable Machines*, 11(3): 285–320.
- Reis, R. (2006). Inattentive consumer. *Journal of Monetary Economics*, 53(8): 1761–1800.
- Santoro, E., Petrella, I., Pfajfar, D. and Gaffeo, E. (2014). Loss aversion and the asymmetric transmission of monetary policy. *Journal of Monetary Economics*, 68: 19–36.
- Schaling, E. (2004). The nonlinear Phillips curve and inflation forecasting: symmetric versus asymmetric monetary policy rules. *Journal of Money, Credit and Banking*, 36(3): 361–386.
- Sheta, A. F., Ahmed, S. E. M. and Faris, H. (2015). Evolving stock market prediction models using multi-gene symbolic regression genetic programming. *Artificial Intelligence and Machine Learning*, 15(1): 11–20.
- Sorić, P. and Čižmešija, M. (2013). Price sentiment of Croatian consumers: The upward bias of collective memories. *Društvena Istraživanja: Časopis Za Opća Društvena Pitanja*, 22(1): 1–21.
- Sorić, P., Lolić, I. and Čižmešija, M. (2016). European economic sentiment indicator: an empirical reappraisal. *Quality & Quantity*, 50(5): 2025–2054.

- Soroka, S. N. (2006). Good news and bad news: asymmetric responses to economic information. *The Journal of Politics*, 68(2): 372–385.
- Surico, P. (2007). The FED's monetary policy rule and U.S. inflation: The case of asymmetric preferences. *Journal of Economic Dynamics & Control*, 31: 305–324.
- Sznajderska, A. (2014). Asymmetric effects in the Polish monetary policy rule. *Economic Modelling*, 36: 547–556.
- Szpiro, G. G. (1997a). Forecasting chaotic time series with genetic algorithm. *Physical Review E*, 55(3): 2557–2568.
- Szpiro G. G. (1997b). A search for hidden relationships: Data mining with genetic algorithms. *Computational Economics*, 10(3): 267–277.
- Tahal, R., Chytkova, Z. and Novinsky, M. (2016). The effect of socioeconomic classes on the subjective perception of economic situation. *Studia Commercialia Bratislavensia*, 9(33): 102–112.
- Toussaint-Comeau, M. and L. McGranahan (2006). Variations in consumer sentiment across demographic groups. *Economic Perspectives*, 30(1), 19–37.
- United Nations (2015). *The World's Women. Trends and Statistics*. Department of Economic and Social Affairs. United Nations, New York.
- Van Aarle, B. and Kappler, M. (2012). Economic sentiment shocks and fluctuations in economic activity in the euro area and the USA. *Intereconomics*, 47(1):44–51.
- Van Aarle, B. and Moons, C. (2017). Sentiment and uncertainty fluctuations and their effects on the Euro Area business cycle. *Journal of Business Cycle Research*, 13(2): 225–251.
- Vasilakis, G. A., Theofilatos, K. A., Georgopoulos, E. F., Karathanasopoulos, A. and Likothanassis, S. D. (2013). A genetic programming approach for EUR/USD exchange rate forecasting and trading. *Computational Economics*, 42(4): 415–431.
- Vermeulen, P. (2014). An evaluation of business survey indices for short-term forecasting: Balance method versus Carlson-Parkin method. *International Journal of Forecasting*, 30(4): 882–897.
- Wei, L. Y. (2013). A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. *Economic Modelling*, 33: 893–899.
- White, D. R., McDermott, J., Castelli, M., Manzonei, L., Goldman, B. W., Kronberger, G., Jaśkowski, W., O'Reilly, U. and Luke, S. (2013). Better GP benchmarks: community survey results and proposals. *Genetic Programming and Evolvable Machines*, 14(1): 3–29.
- Wilms, I., Gelper, S. and Croux, C. (2016). The predictive power of the business and bank sentiment of firms: A high-dimensional Granger Causality approach. *European Journal of Operational Research*, 254(1): 138–147.
- Wilson, G. and Banzhaf, W. (2009). Prediction of interday stock prices using developmental and linear genetic programming. In M. Giacobini et al. (Eds.), *Applications of Evolutionary Computing* (pp. 172–181). Berlin: Springer-Verlag.
- Yang, G., Li, X., Wang, J., Lian, L. and Ma, T. (2015). Modeling oil production based on symbolic regression. *Energy Policy*, 82(1): 48–61.
- Yu, T., Chen, S. and Kuo, T. W. (2004). A genetic programming approach to model international short-term capital flow. *Applications of Artificial Intelligence in Finance and Economics*: 19, 45–70.
- Zagórski, K. and McDonnell (1995). "Consumer confidence" indexes as social indicators. *Social Indicators Research*, 36(3): 227–246.

Appendix

Table A1. Effects of stock market news on the expectations of heterogeneous consumer groups

Coeff.	RE1	RE2	RE3	RE4	ED1	ED2	ED3	AG1	AG2	AG3	AG4	M	F	PR0	PR8	PR9
<i>const</i>	2.26	2.25	2.33	2.23	2.26	2.24	2.24	2.19	2.23	2.24	2.24	2.27	2.08	2.24	2.21	2.25
β_0^+	5.09	6.15	4.58	6.89	5.55	5.48	4.27	6.71	6.97	6.62	4.1	6.01	9.22	5.32	9.04	5.06
β_0^-	1.74	0.89	-1.24	0.6	1.75	1.01	2.41	-0.38	1.42	4.39	2.09	1.78	-3.6	1.86	0.99	3.34
β_1^+	5.67	6.85	4.6	5.84	5.13	6.43		7.08	6.07	4.86	5.31	4.48	9.26	5.34	7.44	3.61
β_1^-	1.61	0.2	-0.99	1.14	2.61	0.62		-0.53	1.52	2.74	0.71	2.72.	-3.64	0.5	1.55	2.55
β_2^+	4.17		4.99	4.61	3.79	3.47		6.78	4.65	2.96	5.3.	2.3	8.94	2.18	5.45	3.57
β_2^-	3.49		-0.78	2.01	3.24	3.64		1.01	2.46	3.81	1.6	4.33	-2.24	2.18	2.09	2.95
β_3^+	3.22		3.25	5.02	3.32			5.27	5.00				10.12	4.56	6.5	
β_3^-	3.14		0.02	0.67	3.09			0.16	1.65				-1.44	2.07	0.99	
β_4^+	2.37		3.27					3.02					10.7		5.12	
β_4^-	3.05		-1.15					0.67					-1.66		1.52	
β_5^+			2.93					5.11					14.32		4.99	
β_5^-			-1.24					-1.1					-2.89		1.59	
β_6^+			5.87					8.6					17.34		4.82	
β_6^-			-2.75					-2.14					-3.75		1.38	

Note: Estimated parameters are obtained using Newey-West HAC standard errors and covariance. 'Bold' numbers denote significant estimates at the 10% significance level.

Table A2. Effects of industrial production news on the expectations of heterogeneous consumer groups

Coeff.	RE1	RE2	RE3	RE4	ED1	ED2	ED3	AG1	AG2	AG3	AG4	M	F	PR0	PR8	PR9
<i>const</i>	2.28	2.3	2.41	2.29	2.28	2.26	2.25	2.25	2.26	2.25	2.26	2.29	2.21	2.25	2.26	2.25
β_0^+	-1.43	-1.1	-0.52	-0.48	-0.71	-1.48	-0.45	0.31	-0.66	-0.29	-0.48	-1.51	0.85	-0.62	-0.71	-0.95
β_0^-	1.83	0.89	0.06	0.79	1.45	0.43	0.04	-0.08	0.66	0.54	0.39	1.65	-1.39	0.4	1.07	0.32
β_1^+		-0.78	-0.67	-0.54												
β_1^-		0.43	-0.22	0.17												
β_2^+		-1.25	-1.22	-0.92												
β_2^-		-0.28	-0.07	-0.02												
β_3^+		-1.97	-1.22	-1.55												
β_3^-		-0.31	-0.37	0.25												
β_4^+		-3.07	-1.63	-1.93												
β_4^-		-0.5	-0.7	-0.94												
β_5^+		-1.5	-1.62	-2.35												
β_5^-		-0.88	-1.25	-1.35												
β_6^+			-1.22	-1.77												
β_6^-			-0.62	-1.26												

Note: Estimated parameters are obtained using Newey-West HAC standard errors and covariance. ‘Bold’ numbers denote significant estimates at the 10% significance level.

Table A3. Effects of unemployment news on the expectations of heterogeneous consumer groups

Coeff.	RE1	RE2	RE3	RE4	ED1	ED2	ED3	AG1	AG2	AG3	AG4	M	F	PR0	PR8	PR9
<i>const</i>	2.26	2.3	2.42	2.27	2.29	2.28	2.26	2.28	2.27	2.25	2.26	2.28	2.32	2.25	2.27	2.26
β_0^+	-0.97	-1.53	-0.3	-1.18	-2.45	-0.77	-1.82	-2.2.	-2.01.	-2.08.	-1.53	-0.64	-2.65	-0.52	-2.07	-0.78
β_0^-	-0.22	-1.78	0.36	-2.29	-2.6	-2.21	-1.41	0.47	-2.38	-4.85	-2.86	0.3	-0.76	-2.83	-3.64	-4.12
β_1^+		-0.64	-0.03	-0.26	-0.44	-0.19	-0.59	0.01	-0.54	-0.08	0.09		-0.33	-0.11	-0.74	-0.52
β_1^-		0.39	2.16	-0.17	-0.58	-0.16	-0.52	-0.26	-0.61	-1.74	-0.86		0.70	-1.32	-0.57	-0.4
β_2^+		2.26	0.39	1.48	1.04	0.74	1.67	1.45	1.82	2.05	1.24		2.50	1.04	1.61	2.11
β_2^-		2.1	1.74	1.11	1.17	1.9.	1.02	1.63	0.91	0.91	0.57		0.53	-0.41	1.22	0.19
β_3^+		-0.2	-0.01	0.14	0.04	-0.35	0.79	-0.42	0.13	0.6	-0.08		-0.38	-0.5	0.3	-0.64
β_3^-		-0.11	1.61	1.41	0.34	1.05	1.41.	1.77	0.98	-0.28	0.12		4.05	0.62	1.33	1.1
β_4^+		-0.58	0.15	-1.20	-1.1	-0.55	-1.44	-0.97	-1.33	-1.32	-0.95		-1.26	-0.62	-1.86	-1.58
β_4^-		0.24	1.52	0.73	0.37	0.77	0.79	0.78	0.74	-0.18	0.97		2.89	1.25	1.45	1.87
β_5^+		0.35	0.11	1.42	0.75	0.65	1.34	1.55	1.35	1.35	1.16		0.00	1.32	1.43	1.71
β_5^-		3.93	2.25	2.83	2.26	3.15	2.18	2.98	3.18	3.53	2.05		6.94	2.07	3.4	2.77
β_6^+		0.720	1.2	3.56	2.73	1.06	4.18	3.38	3.37	3.12	2.31		0.62	1.84	3.48	2.4
β_6^-		3.95	2.2	4.18	3.79	4.34	3.19	2.52	4.47	6.20	3.64		10.33	4.99	5.74	3.93

Note: Estimated parameters are obtained using Newey-West HAC standard errors and covariance. ‘Bold’ numbers denote significant estimates at the 10% significance level.

Table A4. Effects of HICP news on the expectations of heterogeneous consumer groups

Coeff.	RE1	RE2	RE3	RE4	ED1	ED2	ED3	AG1	AG2	AG3	AG4	M	F	PR0	PR8	PR9
<i>const</i>	2.39	2.36	2.4	2.33	2.39	2.37	2.33	2.31	2.36	2.39	2.36	2.37	2.27	2.35	2.36	2.34
β_0^+	-12.53	-15.66	-7.34	-12.59	-11.47	-14.00	-11.36	-6.18	-11.60	-17.46	-10.66	-8.69	-14.39	-11.23	-9.48	-9.96
β_0^-	3.33	2.51	-3.12	-0.29	7.04	3.17	-4.05	-4.19	2.3	2.86	0.16	8.12	-22.35	1.14	4.48	-4.74
β_1^+	-12.87	-6.4		-8.78	-9.95	-11.92	-9.16	-6.1	-10.44	-14.86	-10.62	-9.5	-20.34	-8.47	-10.11	-10.74
β_1^-	4.75	5.84		1	2.84	3.68	3.33	-3.93	0.83	8.49	-0.18	1.62	-15.07	6.37	6.45	-1.88
β_2^+	-7.56	-10.02		-6.8	-9.14	-9.79	-8.64	-7.65	-9.00	-12.66	-9.25	-7.95		-10.08	-9.57	-10.27
β_2^-	1.5	-2.47		-2.02	5.67	-2.61	-2	-0.14	-0.11	-1.07	-1.46	-6.59		3.88	5.03	-3.61
β_3^+	-11.63	-8.52		-7.19	-11.41	-8.58	-9.08	-6.79	-10.03	-13.47	-10.0	-8.13		-8.72	-11.24	-7.87
β_3^-	3.78	-0.8		-7.99	3.96	-2.18	-1.76	-2.97	-1.99	-1.78	0.08	-8.76		-3.1	-2.5	-3.49
β_4^+	-9.29	-5.32		-4.68	-10.02	-7.94	-5.86	-4.55	-9.37	-10.89	-7.9	-10.19		-8.95	-9.59	-7.76
β_4^-	-2.28	4.48		-6.58	-1.14	-0.58	-7.05	0.05	-4.79	-3.93	-9.1	-4.55		-1.21	-4.09	-8.69
β_5^+	-10.77	-9.15		-8.27	-10.72	-8.22	-6.72	-5.93	-7.83	-11.25	-8.67	-5.41		-8.03	-9.52	-7.31
β_5^-	-7.82	-3.17		-9.35	-3.61	-4.06	-9.24	-10.78	-8.43	-9.14	-8.39	-1.49		-5.59	-7.17	-11.07
β_6^+	-14.85				-11.85	-9.84	-6.8	-6.24	-8.59	-9.38	-7.64	-11.46		-9.18	-7.77	-7.12
β_6^-	-4.080				-4.52	-2.71	-12.93	-10.62	-12.49	-7.66	-11.31	-10.13		-6.78	-15.01	-11.84

Note: Estimated parameters are obtained using Newey-West HAC standard errors and covariance. 'Bold' numbers denote significant estimates at the 10% significance level.

Table A5. Effects of retail turnover news on the expectations of heterogeneous consumer groups

Coeff.	RE1	RE2	RE3	RE4	ED1	ED2	ED3	AG1	AG2	AG3	AG4	M	F	PR0	PR8	PR9
<i>const</i>	2.27	2.27	2.42	2.28	2.27	2.26	2.24	2.25	2.25	2.24	2.26	2.27	2.26	2.24	2.25	2.25
β_0^+	-0.56	-0.64	-1.56	-1.19	-0.16	-1.14	-0.2	-1.55	-0.15	0.19	-0.27	0.18	-6.86	0.01	-0.51	-1.5
β_0^-	0.65	0.04	-0.91	-0.11	0.2	-0.54	-1.04	-0.43	-0.92	-0.54	0.21	-0.35	5.49	-1.01	-0.17	0.03
β_1^+			-3.39	-2.08												
β_1^-			-2.29	-1.24												
β_2^+			-4.39	-3.3												
β_2^-			-2.94	-2.06												
β_3^+			-4.09	-3.08												
β_3^-			-4.12	-3.9												
β_4^+			-5.23	-4.07												
β_4^-			-3.71	-4.52												
β_5^+			-4.05	-3.93												
β_5^-			-3.46	-3.9												
β_6^+			-2.54	-2.8												
β_6^-			-1.46	-2.73												

Note: Estimated parameters are obtained using Newey-West HAC standard errors and covariance. 'Bold' numbers denote significant estimates at the 10% significance level.