Using Consumer Survey Data to Track Households Consumption in Belgium

Keywords: *lasso, glmnet, consumer surveys, media indices*

1 Introduction

The use of qualitative surveys to track macroeconomic quantities is common practice due to their timeliness. For instance, the Economic Sentiment Indicator in Belgium has been reported to be useful in nowcasting models that track GDP growth. In this paper, we focus on the subset of survey questions that are addressed to the consumers and discuss whether they have potential to be used in models that track consumption. Our analysis is based on the lasso method [1], which makes it feasible to estimate large linear regressions[1].

It is not straightforward that qualitative consumer surveys have predictive power for households consumption, since the time series based on each question are the sum of balances of positive/negative/neutral responses. The Consumer Confidence Indicator, which is a composite indicator based on a simple average of four questions, does not seem to be strongly correlated with households consumption in Belgium after the Great Recession period (see for example [3]). Nevertheless, we show that consumption is to some extent predictable when the complete set of questions is used. The forecasting simulation results based on the consumer survey are compared with alternatives, such as the 20 indices underlying the so-called Economic Policy Uncertainty (EPU) index [4] extracted from an automatic read of Belgian newspapers.

Figure 1: Overview of the Survey Questions [5]

			MICRO	MACRO	EXPECT	Old CCI	New CCI
Economic situation in Belgium	Assessment of the last twelve months	Q3		■			
	Outlook for the next twelve months	Q4		■	■	■	■
Development of prices in Belgium	Evolution during the last twelve months	Q5		■	■		
	Outlook for the next twelve months	Q6		■			
Unemployment	Unemployment Belgium for the next twelve months	Q7		■	■	■	
Important purchases	Is current time good for important purchases?	Q8	■	■			
	Important purchases next twelve months	Q9	■	■	■		■
Financial situation of the households	Evolution during the last twelve months	Q1	■				■
	Current financial situation of households	Q12	■				
	Outlook for the next twelve months	Q2	■		■	■	■
Save by the families	Outlook for savings in the next twelve months	Q11	■		■	■	
	Assessment of whether or not current time is favorable for saving	Q10	■				

Beyond its properties as a shrinkage estimator, lasso can also be used as a variable selection device and inform practitioners regarding subset of questions that could be used to determine composite indicators. However, there is no guarantee that the number of indicators selected by lasso will be small.

[1] In this paper we use the R implementation glmnet [2]

2 Methods

Our regression model can be written as follows:

$$y_t = \beta_0 + \beta_1 x_t^{q1} + \beta_2 x_t^{q2} + \ldots + \beta_n x_t^{qn} + e_t \tag{1}$$

where y_t is the year-on-year growth rate of consumption at a quarterly frequency and x_t^{qi} refers to the ith survey question averaged over a quarter. The regression parameters in lasso are estimated by minimizing the following expression over a grid of values of λ

$$\min_{\beta_0,\beta} \frac{1}{N} \sum_{i=1}^{N} w_i l(y_i, \beta_0 + \beta^T x_i) + \lambda ||\beta||_1 \tag{2}$$

From a Bayesian point of view, the minimization problem stated above would be equivalent to the estimation of the parameters $\beta_1, \ldots, \beta_n$ under a double exponential prior (see for example the discussion in Hastie et al. [6]). This prior provides us with a strong shrinkage towards zero, so non-informative or weakly informative explanatory variables will be forced towards zero.

In this paper, the optimal λ is determined via cross-validation. Figure 2 illustrates how we select a slightly more restrictive λ than the one that minimizes the Mean Squared Error (MSE)

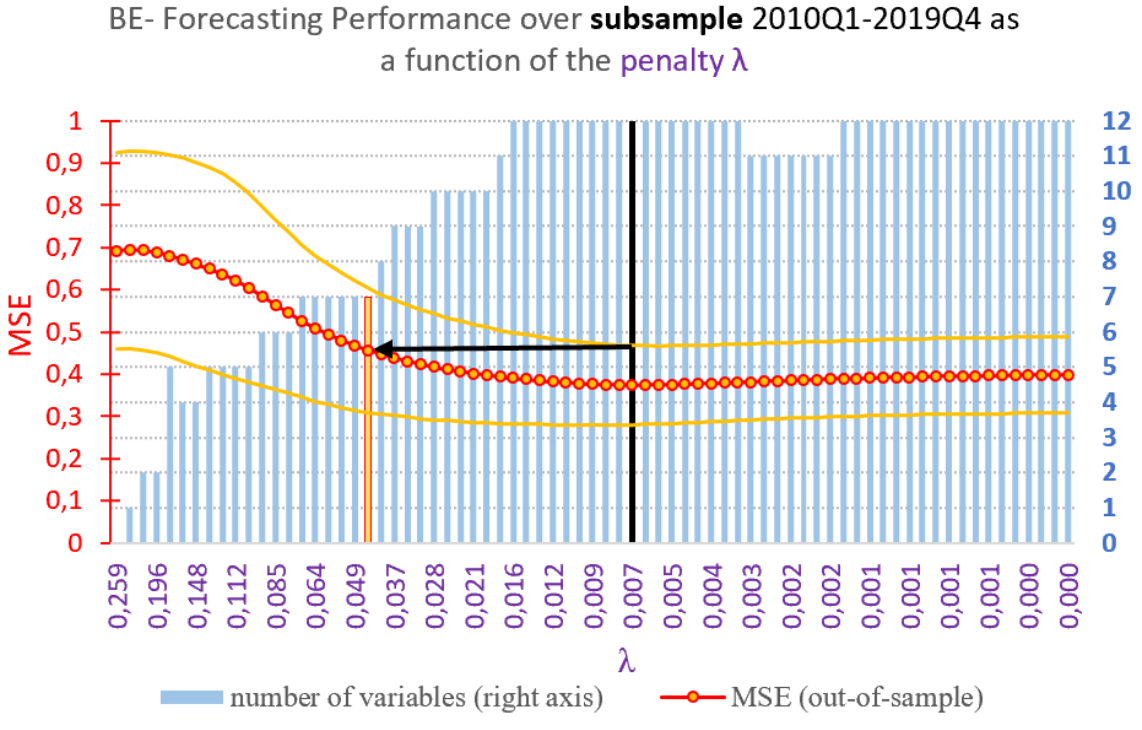


Figure 2: Selecting λ

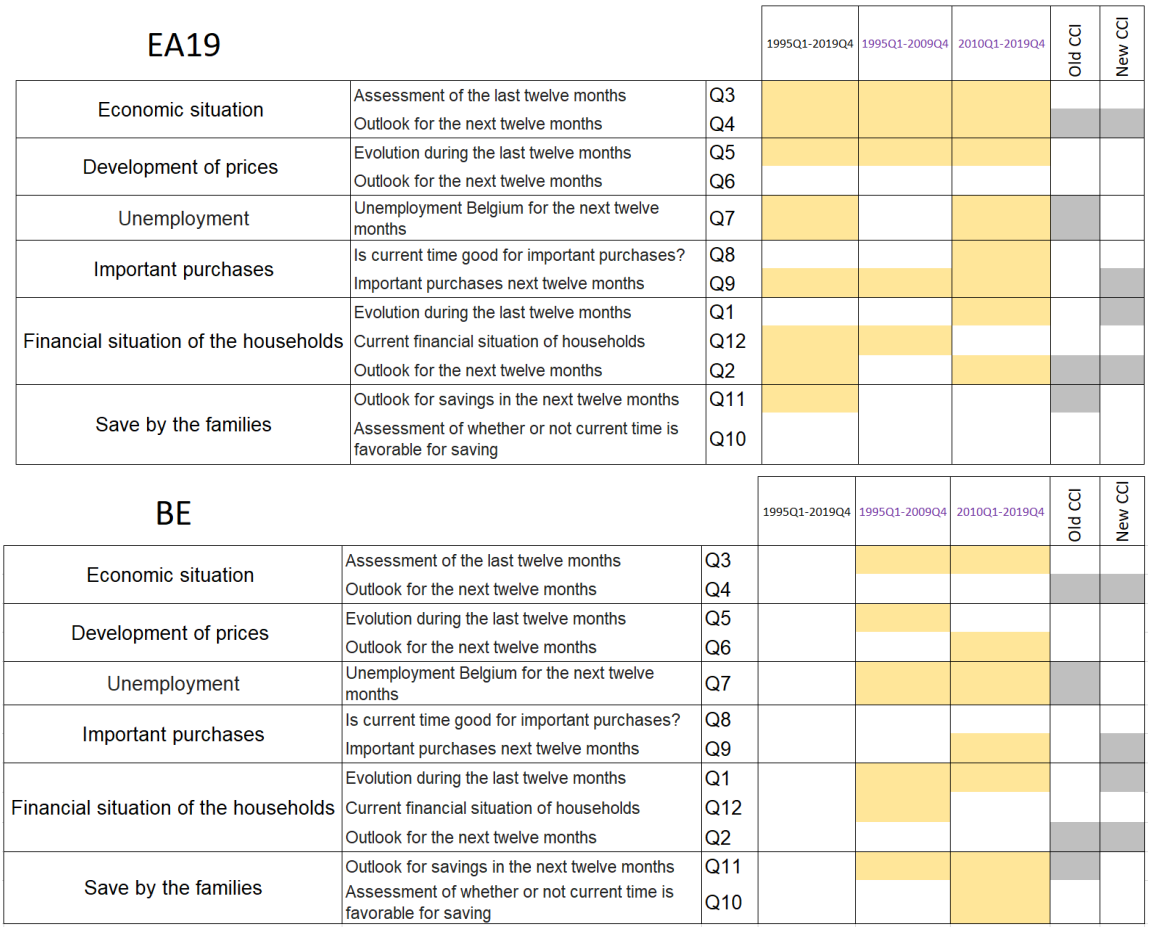
Figure 2 shows the number of variables picked by lasso as a function of the penalty λ and at the same time, the corresponding MSE together with confidence bounds given by its standard deviation. The chosen λ is such that the corresponding MSE is equal to the optimal MSE plus one standard deviation.

3 Results

- Figure 3 shows with a yellow shade the variables selected by lasso when the regression model 1 is applied to the euro area (19) and to Belgium. The analysis

is done for the full sample and for the subsamples. Note that for Belgium, using the full sample implies that the resulting lasso does not select any variable, i.e. the penalty λ resulting from the cross-validation exercise is so large that the resulting model has all parameters equal to zero, except for the constant. This result suggests that the link between consumption and the surveys has changed after the Great Recession.

Figure 3: Lasso selection for the EA19 and Belgium



- For the second subsample, Figure 4 provides a more detailed perspective, where it is also shown how the MSE changes when we expand the regression model 1 with the 20 indices underlying the so-called Economic Policy Uncertainty (EPU) index extracted from an automatic read of Belgian newspapers. Thus, for each one of the three information sets considered, Figure 4 displays the MSE as a function of the penalty λ. As discussed above, the selected λ is not the one that minimizes the MSE, but we take into account the uncertainty around it, which leads us to select a higher value of λ. The resulting parameter estimates that are strictly different from zero define the subset of variables selected by lasso, which is also displayed in the graph.

  The results of Figure 4 suggest that the NBB Consumer survey helps to achieve a smaller MSE than the Media indicators that automatically read Belgian newspapers. Still, it is worth to mention that our exercise is based on quarterly aggregates while those Media indicators are available at a daily frequency, so their potential to predict consumption could be improve in real-time exercises where timeliness is crucial. When both the NBB Survey and the Media indices are combined, lasso is not able to deliver a lower MSE than the one obtained

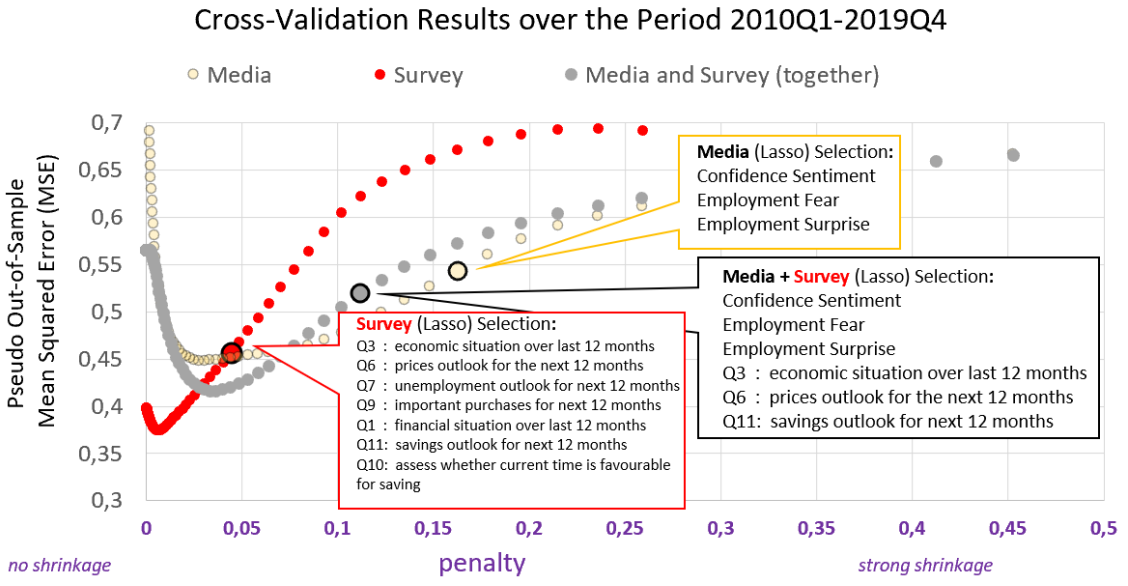


Figure 4: Overview of the Survey Questions

with the NBB surveys alone. This result suggests that a pre-selection of the indicators may have potential to further improve forecasting accuracy.

4 Conclusions

This paper applies lasso as a shrinkage device that makes it feasible to estimate large regressions intended to predict households consumption in Belgium on the basis of multiple qualitative survey questions and alternative indicators. The results suggest that using the whole set of questions is more useful than exploiting the consumer confidence composite indicator.

Although we focus on Belgium, we compare the results with those resulting for the euro area (19), where it is well known that the surveys are very useful to track consumption. However, the selection of questions for Belgium is very different, and it crucially depends on the sample considered.

Finally, this exercise also shows an important limitation of the lasso approach. When both the NBB Survey and the Media indices based on Belgian newspapers are combined, lasso is not able to deliver a lower MSE than the one obtained with the NBB surveys alone. Thus, the larger information set does not guarantee a better forecasting accuracy.

References

[1] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996. ISSN 00359246. URL http://www.jstor.org/stable/2346178.

[2] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, Articles*, 33(1):1–22, 2010. ISSN 1548-7660. DOI: 10.18637/jss.v033.i01. URL https://www.jstatsoft.org/v033/i01.

[3] KBC. Consumer confidence not always a reliable predictor of consumer spending. *Economic Opinions*, 2017. URL https://multimediafiles.kbcgroup.eu/uploadpdf/EO20170920E.pdf.

[4] Andres Algaba; Samuel Borms; Kris Boudt and Jeroen Van Pelt. The economic policy uncertainty index for flanders, wallonia and belgium. 2020. DOI: https://ssrn.com/abstract=3580000. URL https://www.jstatsoft.org/v033/i01.

[5] European Commission. A revised consumer confidence indicator. 2018. URL https://ec.europa.eu/info/sites/info/files/new_cci.pdf.

[6] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2017.