Using Consumer Survey Data to Track Households Consumption in Belgium

Machine learning techniques for variable selection

David De Antonio Liedo

Data Science Center
National Bank of Belgium

January 10, 2021

Overview

To David

David De Antonio Liedo (1980-2021)

- Working at the National Bank of Belgium since 2011 (Data Science Center)
- Member of the JDemetra+ core team
- Specialist in nowcasting
- Last work: NTTS 2021

- And much more...

Consumer survey

			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 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					

New Consumer Confidence Indicator (CCI)

- New CCI of the European Commission in January 2019
  - Tracking private consumption growth
  - Solid theoretical background
- Lack of evidence of the CCI for Belgium
  - Poor predictive power
  - Large differences before and after the Great Recession period (2008-2009)
- Reasons for differences between BE/EU and before/after 2009 not investigated $\Longrightarrow$ focus on statistical treatment

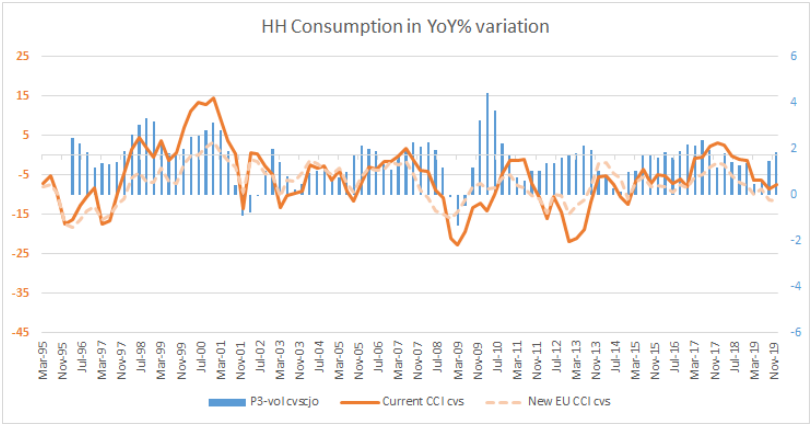
Consumption and consumer survey (BE): naive regressions

Link between consumer survey and (partial) private consumption (YoY growth, quarterly figures)

	1996-2019	1996-2009	2010-2019
Old CCI	0.10	0.26	0.00
New CCI	0.01	0.02	0.00
All questions	0.26	0.69	0.47
All questions and "cleaned" C	0.46	0.78	0.52

Table: Naive regression: adjusted R^2

Consumer survey



Overview of the methods

We consider models with all questions + extensions like:

- Leads, lags
- Additional information: Economic Policy Uncertainty (EPU) index (www.policyuncertainty.com)

Need for variable selection/shrinkage $\Longrightarrow$ Focus on specific statistical methods

- Lasso, with cross-validation (glmnet R-package)
- Bayesian regression with sparse priors (bayesreg R-package)

The Lasso method (glmnet)

Regression model:

$$y_t = \beta_0 + \beta_1 x_t^{q_1} + \cdots + \beta_1 x_t^{q_1} + \epsilon_t$$

For fixed λ, the estimated coefficients correspond to:

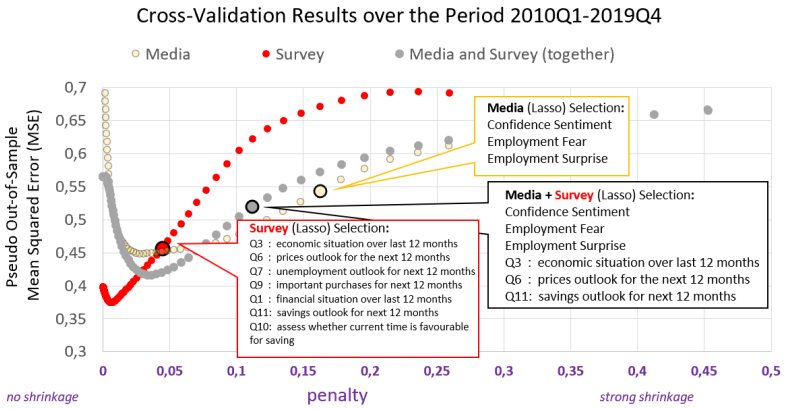
$$\min_{\beta} \frac{1}{N} \sum_{i=1}^{N} w_i l(y_i, \beta_0 + \beta' x_i) + \lambda \|\beta\|_1$$

- Estimation of the optimal regularization parameter λ through cross-validation.
- Choice of the smallest model (Diebold-Mariano tests on equal accuracy).

Selection of the consumer survey questions using Lasso

BE			1995Q1-2019Q4	1995Q1-2009Q4	2010Q1-2019Q4	Old CCI	New CCI
Economic situation	Assessment of the last twelve months	Q3					
	Outlook for the next twelve months	Q4					
Development of prices	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					

Integration of (media) EPU using Lasso



Bayesian regularized regression (bayesreg)

- Bayesian hierarchical models with sparse priors
  - Laplace (or double exponential) prior
  - Horseshoe prior
- No separate estimation of the regularization parameter (full bayesian approach)
- See bayesreg for further details

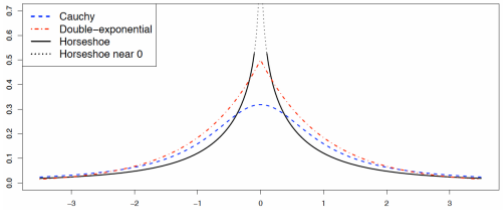
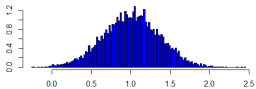
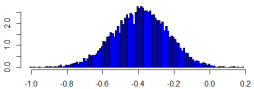
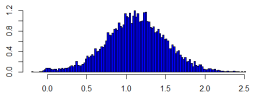
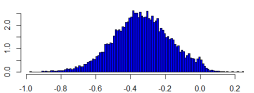


Figure: from H.F.Lopes

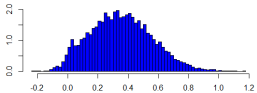
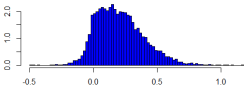
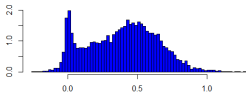
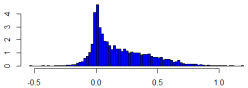
Selection of questions using Bayesian regression

BE (Bayesian approach)			1995Q1-2019Q4	1995Q1-2009Q4	2010Q1-2019Q4	Old CCI	New CCI
Economic situation	Assessment of the last twelve months	Q3	+	+			
	Outlook for the next twelve months	Q4					
Development of prices	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					

Bayesian regression. Posterior samples (I)

1995-2009	Economic situation	Unemployment
Laplace prior		
Horseshoe prior		

Bayesian regression. Posterior samples (II)

2010-2019	Savings outlook	Important purchases
Laplace prior		
Horseshoe prior		

Conclusions

- Obvious break in 2009 (BE). To be investigated
- Consider all questions and not only the CCI for prediction
- Lasso = useful shrinkage / variable selection in large models but...
- Limits of "naive" cross-validation (more information $\not\Rightarrow$ better forecasts)
- Bayesian regularized regression is an appealing alternative

- More new questions than answers

Thank you for your attention

Thank you David