WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Vienna University of Economics

Master's Thesis

The Predictive Ability of Google Search Volumes for Macroeconomic Variables

Author
Felix Zangerl, BSc

Supervisor
Univ.Prof.Dr. Jesus Crespo Cuaresma

November 23, 2021

WIRTSCHAFTSUNIVERSITÄT WIEN
Vienna University of Economics and Business

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

EQUIS ACCREDITED AACSB ACCREDITED AMBA ACCREDITED

Master's Thesis

Title of Master's Thesis:	The Predictive Ability of Google Search Volumes for Macroeconomic Variables
Author (last name, first name):	Zangerl, Felix
Student ID number:	01154814
Degree program:	Master in Economics
Examiner (degree, first name, last name):	Univ.Prof.Dr. Jesus Crespo Cuaresma

I hereby declare that:

1. I have written this master's thesis myself, independently and without the aid of unfair or unauthorized resources. Whenever content has been taken directly or indirectly from other sources, this has been indicated and the source referenced. I am familiar with the regulations specified in the Directive on Plagiarism and Other Types of Academic Fraud in Academic Theses.

2. This master's thesis has not been previously presented as an examination paper in this or any other form in Austria or abroad*.

3. This master's thesis is identical with the thesis assessed by the examiner.

4. (only applicable if the thesis was written by more than one author): this master's thesis was written together with

   The individual contributions of each writer as well as the co-written passages have been indicated.

23/11/2021 F. Zangerl

Date Signature

*This does not apply to master's theses written as part of WU cooperation programs (joint or double degrees).

I dedicate this work to my family and my beloved partner Sandra, who have supported me throughout my studies to becoming a first generation academic. Without you this thesis paper would not have been possible.
Thank you for everything.

Abstract

This paper demonstrates the use of Google search data and the construction of an index of the perceived economic situation (PES) in Austria which is used to make forecasts based on a VAR model. The VAR model includes real data on inflation and unemployment and the subsequent forecasts predict the development of those variables. The PES indicator serves as a good proxy for GDP growth and therefore its forecasts may give information about future GDP growth. The results from the in-sample forecasts show that the PES indicator outperforms the consumer confidence index during crises periods in terms of forecasting accuracy. During the current crisis the VAR model is not able to outperform simpler univariate processes (e.g. AR(1) or Random Walks).

Keywords: Google Trends, forecasting, vector autoregression, COVID-19, econometrics, macroeconomics

Contents

Introduction

In this paper we examine the forecasting accuracy of an indicator of the perceived economic situation (PES) constructed with Google Search Volumes. Based on a substantial methodological contribution from the trendEcon project in this field of research, we use the PES indicator in a VAR model together with real economic variables. The results from the model are used to produce in sample forecasts and the accuracy measures are compared to benchmark models in order to assess the quality of the forecasts. The results of our analysis indicate that forecasts using the PES indicator together with real economic data are as accurate as the benchmark model where PES is replaced by the consumer confidence index.

According to Choi and Varian who wrote two seminal papers concerning the use of Google data, Choi and Varian (2009) and Choi and Varian (2012), the first paper that applied web search data was only written in 2005 by Ettredge, Gerdes, and Karuga (2005). Since then a large body of research has been conducted using search data for economic research, particularly Google Search Volumes (GSV). Vosen and Schmidt (2011) as well as Woo and Owen (2019) forecast private consumption with GSV and Castelnuovo and Tran (2017) use the data to proxy uncertainty.

Two main reasons for the use of search data are that GSV are free of cost and easily accessible, so that researchers can easily reproduce results and contribute to the work of others as well as the fact that GSV are available in real time. The latter contributed to a surge of research papers in the time of the present SARS-CoV-2 pandemic which aimed to fill the gap when no hard data (e.g. quarterly GDP) was available to make credible statements about the current state of the economy. Ferrara and Simoni (2020) showed that GSV deliver useful information when there is no macroeconomic data available. Quite intuitively though, the relative nowcasting power vanishes as soon as official data is released. Eichenauer et al. (2020) introduced a novel sampling technique that overcomes the substantial sampling noise in GSV when applied to small countries. They made their work freely and openly available in form of an R package called `trendecon`, which is available on Github, as well as an online Dashboard where they interactively display their work. Our results show that the PES indicator disposes a high relative nowcasting power.

In this section we gave an overview of the literature, in section Data we summarize the data used in the study and describe the construction of the main indicator in detail. In section Narrative timeline we explain the influence of major events, namely the World Financial Crisis and the COVID-19 pandemic, within the timeline. Section Forecasting Framework outlines the model and results of the forecasts, compared against benchmark models and concludes with limitations. The conclusion of this paper is drawn in section Conclusion.

Data

For our analysis, we built on top of the work of trendEcon and replicated their main indicator for Austria, which we will refer to as Perceived Economic Situation (PES).

The data that we use can be divided into three categories: GSV - PES (main indicator), soft data (consumer confidence indicator) and hard data (gdp, unemployment, inflation). Loosely speaking one can say, if we are interested in consumer spending behaviour, the first and second category captures consumers' desire and willingness to pay, while the hard indicators reflect their ability to pay.

We use our PES indicator together with unemployment and inflation in a VAR model. We expect the unemployment rate to have a negative impact on GDP growth and the perception of the job market drives our main indicator. Inflation and its relationship to the unemployment rate is an important factor that should be included in the VAR model. Our forecasts are computed from the results of the VAR model. For benchmarking and comparison we compare the main model with one where PES is replaced by the consumer confidence index and in order to inspect structural changes caused by the current COVID-19 crisis, we estimate the model for current and past timeframes.

Perceived Economic Situation

The indicator is available on my Github[1] as well as on my online Dashboard[2] and will be updated regularly. trendEcon refers to the indicator both as daily economic sentiment indices (DESI) in the paper and perceived economic situation (PES) on their website. We only use the term PES because we work with both the daily and the monthly aggregated series, but in our model framework we only use the monthly frequency in order to have consistent frequencies across all variables.

We use four keywords which trendEcon found work best in capturing the perceived economic situation. These are the German words for economic crisis
("Wirtschaftskrise"), insolvency ("Insolvenz"), and unemployed ("arbeitslos"). The selected keywords must be general enough to consistently deliver non-zero results. The literature suggests that adding more keywords does not necessarily improve the results. An advantage of using fewer keywords is that it speeds up the scraping process, because Google frequently blocks the IP after too many queries in a short period of time. Another insight into search data in general is that people look up negatively connoted words during periods of downturns while positively connoted words do not coincide adequately with upswing periods. Because of the use of negative connoted keywords, the main indicator is inverted in a last step.

[1] https://github.com/FelixZangerl/gsv_data
[2] https://felixzangerl.github.io/gsv_dashboard/

Sampling process

Google does not share all the search data generated, instead they draw random samples[3] . This is due to privacy concerns and to speed up the process of data retrieval. The data is available on a daily basis for every region in the world and goes back as far as 2004. The trendEcon project uses the data from 2007 onwards, because the first iPhone was introduced in this year and because there is still enough time left to cover the prelude of the World Financial Crisis in 2008/09. The drawback when selecting a sample is, that as you go farther back the data default to weekly and then to monthly time series, which are not consistent with one another. Another problem is the sampling variation being substantially high for small regions, with a standard deviation that can be as high as the mean value in an area with 1 million inhabitants (Eichenauer et al., 2020). The authors of the trendEcon project overcome the sampling variation issue by forcing Google to draw multiple samples for every frequency through the use of a rolling time window. They report that by computing the average of all windows they reduce the variance of the mean values by approximately 90 percent. The method for overcoming the problem of inconsistent series of varying frequencies is explained in subsection Disaggregation.

Index scale

GSV data is normalized with respect to time and location of a query. The values for the indicators are scaled on a range of 0 to 100 based on the query's proportion to all searches on all topics. Reporting proportions instead of absolute search numbers is done in order to protect the privacy of individuals in smaller regions who could otherwise potentially be identified.

[3] Google does not share all the information about the sampling process. The official information can be found in their FAQ: https://support.google.com/trends/answer/4365533?hl=en.

Disaggregation

To overcome the problem of combining varying frequencies of time series, temporal disaggregation is needed. This is done by applying the disaggregation routine by Chow and Lin (1971). The underlying assumption is that the monthly frequency is best suited to reflect the long-term trend while the higher frequencies capture their respective trends. This process is applied stepwise, from the highest to the lowest frequency, and ensures that the series are consistent.

Seasonal adjustment

For removing the seasonal component from a daily time series a method is needed that can account for holidays and other irregular events. The Prophet procedure by Taylor and Letham (2018) is best suited in this case and fortunately the authors supply an R package called `prophet`. A central assumption of the Prophet procedure is that weekly and yearly seasonality is constant over time and it disregards monthly seasonality which is not captured by weekly seasonal effects.

Principal component

In order to reduce the time series of multiple keywords into one common signal, principal component analysis (PCA) is applied. The first principal component is extracted and normalized so that the mean of the time series equals zero and the standard deviation equals one. As mentioned above in Index scale due to Google merely reporting scores instead of absolute search numbers, this solution is adequate in order to form the index.

Real economic data

Table 1 gives an overview about the variables that are being used in the model. All of them are, together with the PES index, on a monthly frequency.

Variables used in VAR model		
Data	Unit	Source
Unemployment	international Definition in % (YoY)	Statistics Austria
Inflation	national CPI in % (YoY)	Statistics Austria
Consumer Confidence	Business and consumer survey in %	European Commission

Table 1: Description and sources of the real economic data used in the VAR forecasting framework (monthly frequencies)

Consumer confidence indicator

The consumer confidence indicator serves as a benchmark indicator. Analogous to the trendEcon project we show that also for Austria the main indicator (PES) shows a strong comovement with the consumer confidence index and with GDP growth and therefore serves as an adequate proxy in the VAR forecasting framework. See Figure A1 in Section Appendix

Narrative timeline

World Financial Crisis

The main indicator (PES) unfortunately shows a less significant drop in 2008-09 for Austria during the World Financial Crisis (WFC) than for Switzerland and Germany. It is still significant on its own but the WFC did not impact our neighbouring countries significantly harder and therefore the explanation for the less pronounced drop must lie in the data. After inspection of the components of PES, namely the four keywords, it can be seen that the keyword 'arbeitslos' was not being searched more frequently during the WFC than at other times but overall contributes strongly to the PES indicator.

Such idiosyncrasies are of course a drawback when using search data and have to be kept in mind as one would expect that searches for unemployment would have gone up during the WFC in Austria, when in fact they did not. There is no obvious explanation for this. On the other hand the keywords 'Wirtschaftskrise' (economic crisis) and 'Insolvenz' (insolvency) did show significant spikes and therefore maybe should receive more weight in the composition of the indicator.

COVID-19 crisis

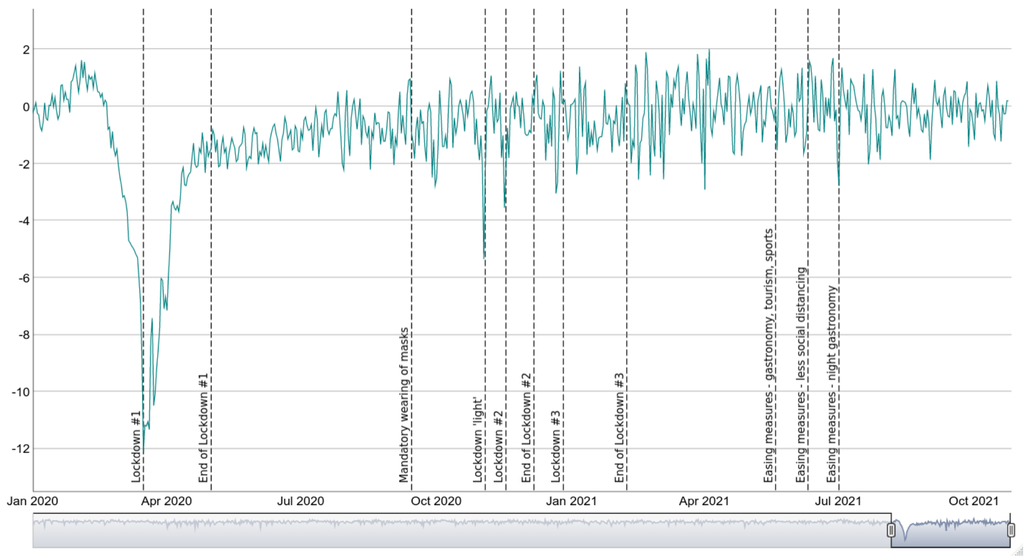

Figure 1: Narrative timeline PES indicator Austria, Jan. 2020 until Nov. 2021

The recent and still ongoing COVID-19 pandemic gave rise to various Nowcasting projects like the trendEcon project and consequently this Master's Thesis. Usually real economic data moves sluggishly and therefore suffices to be available on a quarterly basis but during times of crisis the speed in which the economic activity fluctuates increases. The PES indicator was one of many indicators that could deliver just-in-time information for policy makers and researchers, because it coincided well with actual economic development. The PES indicator in Figure 1 tracks the events during the pandemic on a real time daily basis. For the whole period it has a mean value of 0 and standard deviation 0.62. For the period starting in 2020 the mean value falls to -0.87 and the standard deviation drastically increases to 1.54. The biggest drop occurred exactly on the day of the announcement of the first lockdown on March 13th and then recovered slightly in the following days until it reverted back to a stable yet slightly sub par value on the day the first lockdown ended. The lockdown "light" shows a similar effect and could possibly be amplified by the events of the Vienna terrorist attack on November 2. Two days after the terrorist attack the indicator was back to the stable but sub par level that we still see at the time of writing. The day the second lockdown became effective was the next significant downturn and the prelude of the third lockdown also showed a significant decrease in the PES indicator. During the on-going pandemic other search words concerning sectors like Travel Abroad, Food Delivery and other fields of consumption also drew an interesting picture which is not included in this work but can be seen on my Dashboard.

Forecasting Framework

In order to test the forecasting accuracy of the main indicator (PES), we set up a VAR model and report the model diagnostics. With the results from the VAR in-sample forecasts are generated. The results of the in-sample forecasts are presented in a table and a figure in order to compare different time-series accuracy measures against benchmark models.

VAR

The first step in the setup of the VAR is the investigation of the cross correlation of the variables used in the model. They can be seen in Figure A2 in Section Appendix. The cross correlations show that, as expected, inflation and unemployment are significantly negatively correlated and that the main indicator (PES) is significantly positively correlated with the consumer confidence index and significantly negatively correlated with unemployment.

The VAR model is represented by the following formula:[4]

$$\mathbf{y}_t = A_1\mathbf{y}_{t-1} + A_2\mathbf{y}_{t-2} + \ldots + A_{12}\mathbf{y}_{t-12} + \mathbf{u}_t$$

where

$$\mathbf{y_t} = [PES_t, u_t, \Delta cpi_t]$$

where PES_t is the main indicator, u_t is the unemployment rate and cpi_t is the price index (CPI/VPI). Inflation is constructed by taking the year-on-year percent changes of the consumer price index.

The lags are determined using the Hannan–Quinn information criterion which suggests twelve lags. Twelve lags are sufficient in order to account for serial correlation in the model. Table A3 in Section Appendix shows the results of the VAR.

Model diagnostics

The conclusion from the Portmanteau- and Breusch-Godfrey test are ambigious. While the Protmanteau test concludes that the residuals from the VAR are not serially correlated, i.e. that the autocorrelation and correlation between the variables and their lags is captured by the inclusion of a sufficient number of lags (p-value = 0.47), the Breusch-Godfrey test rejects the null of no serial correlation (p-value = 0). The Jarque-Bera test for normality tests if the residuals are coming from a normal distribution which

[4]The estimation of the VAR model, forecasting and accuracy measures are calculated in R through the 'vars' package.

is rejected (p-value = 0). This is apparently due to the spike in the main indicator (PES) during the COVID-19 crisis. While interpreting the results from impulse response functions and forecast error decompositions would be biased, the ability to make forecasts is less distorted by the violation of the normality assumption. The ARCH-LM test fails to reject the null hypothesis of homoskedasticity (p-value = 0.76). Therefore no correction of the standard errors is needed.

Forecasting

In sample Forecast

The forecasting horizon of the VAR is calculated recursively. In Figure 2 the in sample forecast is shown for three periods ahead. In the validation period, that is August until October 2021, inflation slightly goes up in the first month and stabilizes in the subsequent month, unemployment goes up and then slightly down in the last month and the PES indicator rises sharply.

Accuracy

The in-sample forecast is calculated in order to compare the forecasts to the actual data. Table 2 reports the accuracy measures for the three months ahead rolling in-sample forecast.

var	model	ME	RMSE	MAE	MPE	MAPE
pes	PES	-1.66	1.67	1.66	176.02	2854.59
une	PES	4.76	4.77	4.76	74.71	74.71
cpi	PES	0.55	0.82	0.60	20.74	23.86

Table 2: 3-month ahead rolling in-sample forecast. Displayed here are several accuracy measures for the PES forecasting model.

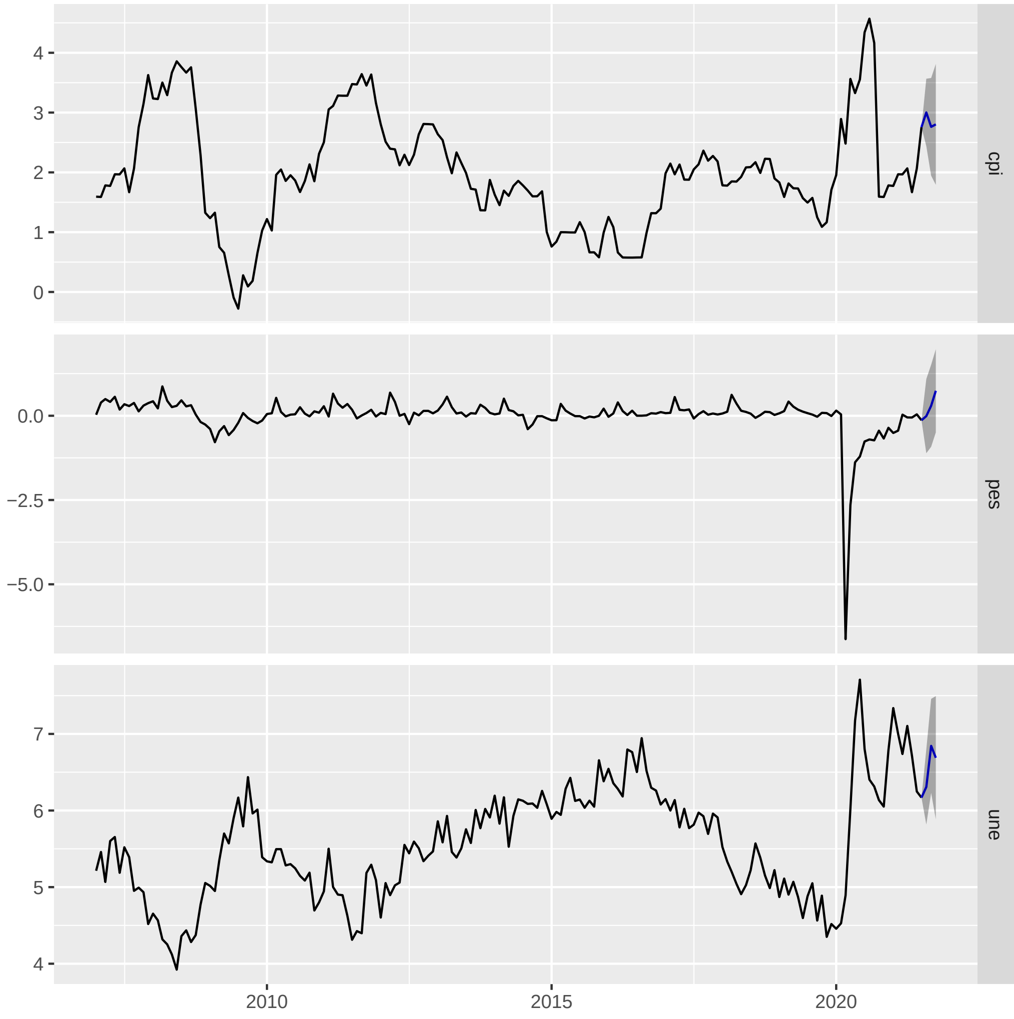

Figure 2: 3-month ahead in sample forecasts. The blue line depicts the forecasting horizon and the gray areas mark the 95%-confidence intervals.

Benchmarks

For comparison we have provided five benchmark models depicted in Figure 3. These models are: "PES" - the main model, "CC" where we replace the main indicator by the consumer confidence index, "COV19 Crisis" and "Before COV19", where the main model forecasts the before- and respectively the during-crisis timeframe[5]. Two univariate models "AR(1)" and "Random Walk" are used to forecast the model variables during (Panel A) and before the actual COVID-19 crisis. Figure 3 plots the root mean squared errors (RMSE) of the benchmark models. In Panel A of Figure 3 it can be seen that the PES model produces slightly better forecasts than the model with the consumer confidence index. In Panel B, the crisis model poses as a lower bound benchmark as we would not have expected the model to accurately capture the volatility of the COVID-19 crisis timeframe. However, coincidentally, unemployment rate forecasts are more accurate during this time period. The before Covid period produces slightly better forecasts than the main model, while the current (PES) forecast can't hold up to the univariate benchmark specifications. It can be seen that both univariate models outperform the VAR models in the current period, whereas only the "AR(1)" outperforms the "Before COV19" VAR model.

Conclusion

Google Search Volumes have become a widely used data source and even more so during the COVID-19 pandemic. The trendEcon project has constructed an indicator that can capture the perceived economic situation via internet searches and in the process they built a technique to produce these search results consistently, so that the results are prone to sampling issues for smaller countries and subregions. On that basis we were able to reproduce an economic indicator for Austria which shows strong comovement with real GDP growth and the consumer confidence index. While the down-

[5] The before crisis forecasting horizon starts in January 2019 and the during crisis period in March 2020.

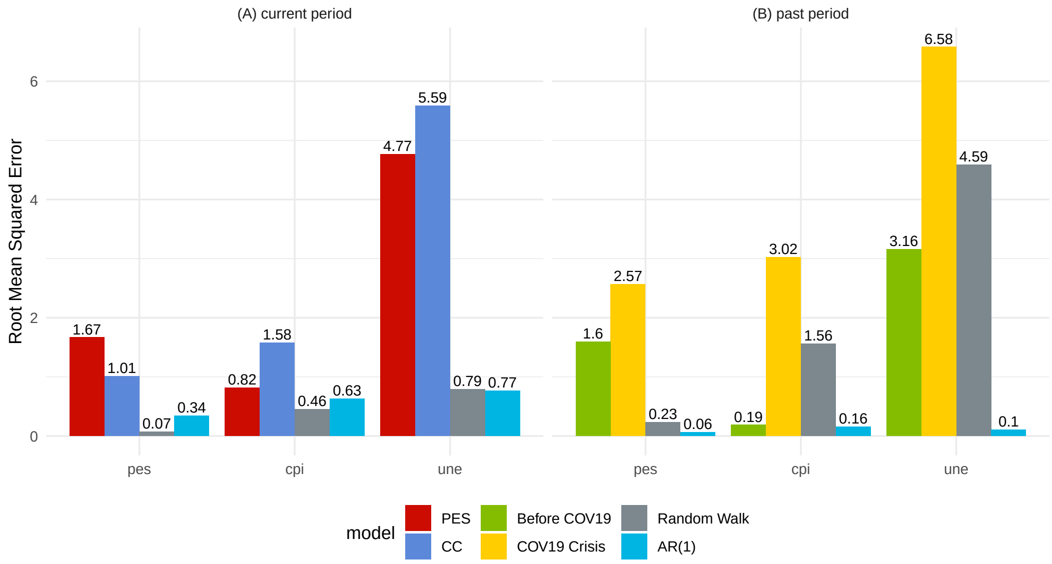

Figure 3: Accuracy of Benchmark forecasting models (varying models and timeframes)

turn during the World Financial Crisis is not captured well enough, we are positive that this could be dealt with by replacing or applying (country) specific weights for individual search terms, under the assumption that certain keywords may reflect (regional) idiosyncrasies.

In this study we have shown that the perceived economic situation indicator, scraped from Google Trends, can be used for forecasting with an accuracy as good as the consumer confidence index. The model is not able to outperform a simpler univariate model in neither the current nor the before crisis period. Periods of crisis are characteristic for higher variances in the parameters and unobserved measurement errors, especially in regards to inflation post March 2020[6], can also play an important role.

The VAR forecasting framework is one of the many possible models that can be used in order to forecast the future economic development. The inclusion of real economic variables helps to investigate how well the perceived situation relates to actual economic activity. In the forecasting literature models are constantly improved upon with a focus on holding the balance

[6]During the lockdowns the survey of consumer prices was only possible to a limited extent (Beckmann and Rumler, 2020).

between specificity and generality. While the PES indicator itself delivers useful insights as a Nowcasting indicator, especially when quarterly GDP data is not available, its combination with econometric models could also lead to more accurate early warning indicators.

Forecasting in a multivariate model is only possible if enough explanatory variables are available just-in-time to build a sound model. It is also a challenge to retrieve important variables in a high frequency like monthly or even daily data. There are however methods to compensate for this problem like mixed-data sampling (MIDAS) or applications of the Kalman filter. We are looking forward to continue working with such models and reading such papers alike.

References

Beckmann, Elisabeth and Fabio Rumler (2020). *Schwerpunktthema – Inflation aktuell Q4/20: Inflationserwartungen in Österreich während der COVID-19-Pandemie*. URL: https://www.oenb.at/dam/jcr:6df69915-8c2e-4fb0-915e-b20688ec6a9d/schwerpunktthema_Q4-20.pdf (visited on 10/27/2021).

Castelnuovo, Efrem and Trung Duc Tran (2017). "Google it up! A google trends-based uncertainty index for the United States and Australia". In: *Economics Letters* 161, pp. 149–153.

Choi, Hyunyoung and Hal Varian (2009). "Predicting initial claims for unemployment benefits". In: *Google Inc* 1, pp. 1–5.

— (2012). "Predicting the present with Google Trends". In: *Economic record* 88, pp. 2–9.

Chow, Gregory C and An-loh Lin (1971). "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series". In: *The review of Economics and Statistics*, pp. 372–375.

Eichenauer, Vera et al. (2020). *Constructing daily economic sentiment indices based on Google trends*. Tech. rep. KOF Working Papers.

Ettredge, Michael, John Gerdes, and Gilbert Karuga (2005). "Using web-based search data to predict macroeconomic statistics". In: *Communications of the ACM* 48.11, pp. 87–92.

Ferrara, Laurent and Anna Simoni (2020). "When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage". In: *arXiv preprint arXiv:2007.00273*.

Taylor, Sean J and Benjamin Letham (2018). "Forecasting at scale". In: *The American Statistician* 72.1, pp. 37–45.

Vosen, Simeon and Torsten Schmidt (2011). "Forecasting private consumption: survey-based indicators vs. Google trends". In: *Journal of forecasting* 30.6, pp. 565–578.

Woo, Jaemin and Ann L Owen (2019). "Forecasting private consumption with Google Trends data". In: *Journal of Forecasting* 38.2, pp. 81–91.

List of Figures

List of Tables

A Appendix

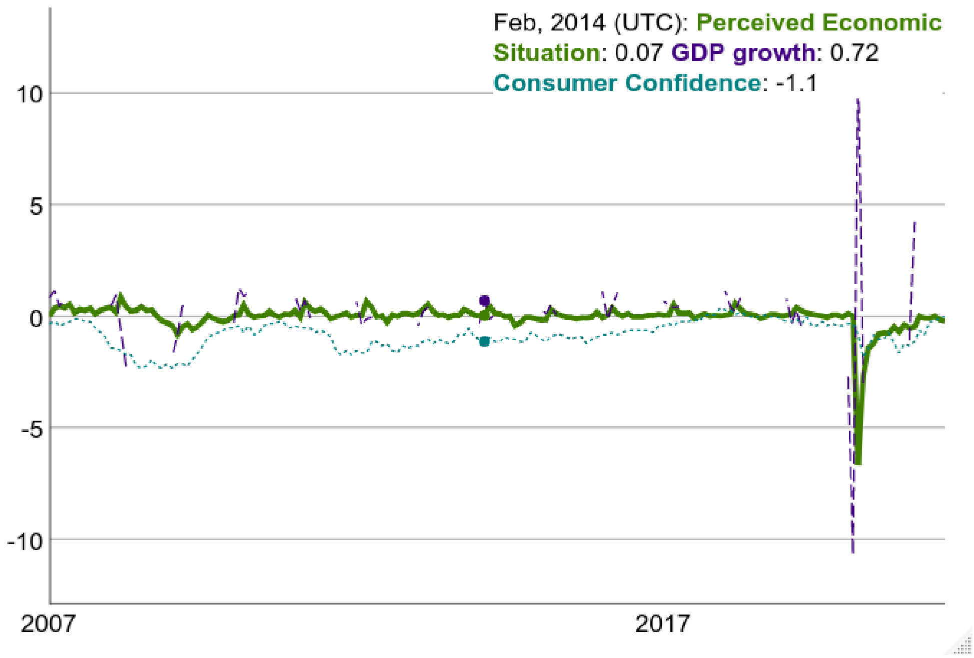

Figure A1: Comovement of economic sentiment variables. The solid green line depicts the PES index, the blue dotted line the consumer confidence index and the purple dashed line displays quarterly GDP growth.

Table A3: Vector Autoregression(12). The results from the estimation are displayed in the table. While the signifcance of the coefficients can be checked, the values are not trivial to interpret due to the nature of a VAR model.

	dependent variables: $pes, \Delta une, cpi$		
	pes	Δune	cpi
	(1)	(2)	(3)
pes_{t-1}	0.457*** (0.091)	−0.187*** (0.041)	−0.126*** (0.047)
Δune_{t-1}	−0.062 (0.200)	0.560*** (0.089)	−0.066 (0.102)
int_{t-1}	−0.462** (0.180)	0.031 (0.080)	1.067*** (0.092)
pes_{t-2}	−0.113 (0.110)	−0.183*** (0.049)	0.098* (0.056)
Δune_{t-2}	−0.108 (0.227)	0.208** (0.101)	−0.070 (0.116)
int_{t-2}	0.518* (0.269)	−0.069 (0.120)	−0.144 (0.138)
pes_{t-3}	0.145 (0.116)	−0.108** (0.052)	−0.095 (0.059)
Δune_{t-3}	−0.028 (0.229)	−0.067 (0.102)	−0.083 (0.117)
int_{t-3}	−0.278 (0.274)	−0.001 (0.122)	0.131 (0.140)
pes_{t-4}	−0.104 (0.119)	0.075 (0.053)	−0.077 (0.061)
Δune_{t-4}	0.497** (0.229)	0.112 (0.102)	0.105 (0.117)
int_{t-4}	0.421 (0.274)	0.001 (0.122)	−0.154 (0.140)
pes_{t-5}	0.039 (0.119)	0.050 (0.053)	−0.087 (0.061)
Δune_{t-5}	−0.291 (0.236)	−0.092 (0.105)	−0.086 (0.121)
int_{t-5}	−0.056 (0.266)	−0.002 (0.119)	0.109 (0.136)
pes_{t-6}	0.086 (0.118)	−0.016 (0.053)	0.099 (0.061)
Δune_{t-6}	0.294 (0.236)	0.182* (0.105)	0.144 (0.121)
int_{t-6}	0.147 (0.228)	−0.051 (0.102)	−0.189 (0.117)
pes_{t-7}	0.134 (0.116)	0.026 (0.052)	0.267*** (0.059)
Δune_{t-7}	−0.256 (0.236)	0.027 (0.105)	−0.096 (0.121)
int_{t-7}	−0.363* (0.215)	0.121 (0.096)	0.219** (0.110)
pes_{t-8}	0.150 (0.126)	0.028 (0.056)	−0.129** (0.064)
Δune_{t-8}	−0.070 (0.234)	0.133 (0.104)	0.098 (0.120)
int_{t-8}	0.176 (0.221)	0.014 (0.099)	−0.276** (0.113)
pes_{t-9}	−0.056 (0.125)	−0.001 (0.056)	−0.066 (0.064)
Δune_{t-9}	0.592** (0.233)	0.018 (0.104)	0.014 (0.119)
int_{t-9}	−0.062 (0.227)	0.012 (0.101)	0.095 (0.116)
pes_{t-10}	0.124 (0.127)	−0.053 (0.057)	−0.023 (0.065)
Δune_{t-10}	−0.161 (0.232)	−0.008 (0.103)	−0.035 (0.119)
int_{t-10}	0.021 (0.259)	0.067 (0.115)	−0.102 (0.132)
pes_{t-11}	−0.050 (0.129)	0.139** (0.057)	−0.039 (0.066)
Δune_{t-11}	−0.098 (0.225)	−0.042 (0.100)	0.010 (0.115)
int_{t-11}	−0.105 (0.268)	−0.204* (0.119)	0.272** (0.137)
pes_{t-12}	−0.052 (0.124)	0.024 (0.055)	−0.027 (0.064)
Δune_{t-12}	−0.090 (0.188)	−0.027 (0.084)	0.012 (0.096)
int_{t-12}	0.123 (0.186)	0.159* (0.083)	−0.146 (0.095)
constant	−1.377* (0.805)	−0.153 (0.359)	0.525 (0.412)
Observations	163	163	163
R^2	0.396	0.910	0.928
Adjusted R^2	0.223	0.884	0.907
Residual Std. Error (df = 126)	0.562	0.251	0.288
F Statistic (df = 36; 126)	2.295***	35.207***	44.805***

Note: *p<0.1; **p<0.05; ***p<0.01

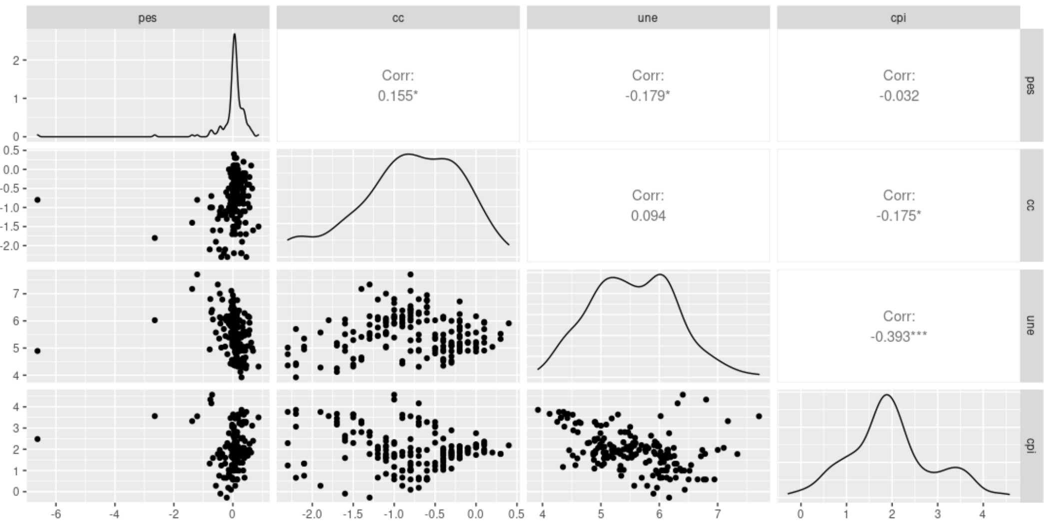

Figure A2: Cross correlations between model variables. This figure depicts the Pearson correlation coefficients as well as the scatterpoints on the opposite diagonal.