

Eichenauer, Vera; Indergand, Ronald; Martínez, Isabel; Sax, Christoph

Working Paper

Constructing daily economic sentiment indices based on Google Trends

KOF Working Papers, No. 484

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Eichenauer, Vera; Indergand, Ronald; Martínez, Isabel; Sax, Christoph (2020) : Constructing daily economic sentiment indices based on Google Trends, KOF Working Papers, No. 484, ETH Zurich, KOF Swiss Economic Institute, Zurich, <https://doi.org/10.3929/ethz-b-000423817>

This Version is available at:

<https://hdl.handle.net/10419/235098>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Paper

Constructing Daily Economic Sentiment Indices Based on Google Trends

Author(s):

Eichenauer, Vera; Indergand, Ronald; Martínez, Isabel Z.; Sax, Christoph

Publication Date:

2020-06

Permanent Link:

<https://doi.org/10.3929/ethz-b-000423817> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

KOF Swiss Economic Institute

Constructing Daily Economic Sentiment Indices Based on Google Trends

Vera Z. Eichenauer, Ronald Indergand, Isabel Z. Martínez and Christoph Sax

KOF Working Papers, No. 484, June 2020

KOF

ETH Zurich
KOF Swiss Economic Institute
LEE G 116
Leonhardstrasse 21
8092 Zurich, Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Constructing Daily Economic Sentiment Indices Based on Google Trends

Vera Z. Eichenauer^{a,*}, Ronald Indergand^{b,**}, Isabel Z. Martínez^{a,**}, Christoph Sax^{c,**}

^a*ETH Zürich, KOF Swiss Economic Institute, Leonhardstrasse 21, 8092 Zürich, Switzerland*

^b*State Secretariat for Economic Affairs SECO, Holzikofenweg 36, 3003 Bern, Switzerland*

^c*University of Basel; cynkra LLC, Badenerstrasse 21, 8004 Zürich, Switzerland*

Abstract

Google Trends have become a popular data source for social science research. We show that for small countries or sub-national regions like U.S. states, underlying sampling noise in Google Trends can be substantial. The data may therefore be unreliable for time series analysis and is furthermore frequency-inconsistent: daily data differs from weekly or monthly data. We provide a novel sampling technique along with the R-package *trendecon* in order to generate stable daily Google search results that are consistent with weekly and monthly queries of Google Trends. We use this new approach to construct long and consistent daily economic indices for the (mainly) German-speaking countries Germany, Austria, and Switzerland. The resulting indices are significantly correlated with traditional leading indicators, with the advantage that they are available much earlier. *JEL classification: E01, E32, E37*

Keywords: Google Trends, measurement, high frequency, forecasting, Covid-19

1. Introduction

Google search volumes (GSV), commonly known as Google Trends, provide an interesting source of real-time data.¹ In economics, many studies have shown that GSV can be useful to forecast private consumption (e.g., Vosen and Schmidt, 2011, Woo and Owen, 2019) or unemployment (e.g., Smith, 2016, for the U.K.; González-Fernández and González-Velasco, 2018, for Spain; and Maas, 2019, for

*Corresponding author

**Equal contribution

Email addresses: eichenauer@kof.ethz.ch (Vera Z. Eichenauer), *ronald.indergand@gmx.ch* (Ronald Indergand), *martinez@kof.ethz.ch* (Isabel Z. Martínez), *christoph@cynkra.com* (Christoph Sax)

¹Researchers across disciplines, including political science (Stephens-Davidowitz, 2014) sociology (Swearingen and Ripberger, 2014, Gross and Mann, 2017), or health sciences (Tkachenko et al., 2017) have used this data to answer a series of questions. See Jun et al. (2018) for a recent overview.

the U.S.). Compared to data traditionally used in business cycle analysis, GSV data has the great advantage that it is available in real-time and on a daily basis. Furthermore, it is available for many countries, regions and even some large cities. These features are especially important in times of crisis, which are usually characterized by rapid changes in the real economy. A recent example is the outbreak of the COVID-19 pandemic, during which macroeconomic conditions deteriorated from day to day.

Although GSV data is in principle available on a daily basis for any country or region of the world, much of the prior research focused on large countries and used monthly rather than weekly or daily GSV data. This is due to the combination of the following limitations of available GSV data: i) For privacy reasons, Google only provides an index of search volumes rather than the actual number of searches, and scaling of the index varies with the chosen time window. ii) For time spans longer than 90 days (5.25 years), GSV default to weekly (monthly) data. Business cycle analysis and forecasting models, however, typically require data spanning more than a decade. iii) For a chosen time frame, GSV data are based on a random sample drawn by Google. We show that in small countries or sub-national regions, or for keywords with relatively little search volume, the sampling variation of returned GSV series turns out to be substantial. For example, in an area with 1 M inhabitants, the standard deviation of the results for ‘recession’ can be as large as its mean value.

To overcome these issues, we present a novel method to construct consistent daily, weekly, and monthly series of GSV over a long period.² Our method allows us to build daily economic sentiment indices (DESI) for the (mainly) German-speaking countries Germany, Austria, and Switzerland over the period 2007–2020, including the beginning of the COVID-19 pandemic. These indices are based on GSV for several keywords, including ‘economic crisis’ (‘Wirtschaftskrise’) and ‘unemployed’ (‘arbeitslos’). To aggregate the series into a DESI, we extract the common signal in the keyword series.³ We show that our indices are significantly correlated with other leading indices in all three countries. The correlations for the much smaller economies Austria and Switzerland are of similar magnitude as for Germany, where GSV data suffer significantly less from sampling variation. We conclude that GSV data can be useful to monitor economic activity at a high frequency, also in smaller countries or regions. Moreover, as our approach to construct consistent, long-run, daily GSV series allows for improved use of GSV data, our contribution expands well beyond economic forecasting and can be applied more generally in a variety of settings and across disciplines.

²We provide the accompanying R-package *trendecon* on <https://trendecon.github.io/trendecon/>.

³We continue updating our indices daily and make them available on www.trendecon.org.

2. Related Research

Our paper makes an important methodological contribution to a growing body of research that uses GSV data. Given that most internet searches are performed by private households, the focus of economic studies usually lies on target variables related to the labor market or consumption. Vosen and Schmidt (2011), for example, show that Google search data outperforms survey-based indicators in forecasting U.S. private consumption. Building on this, Woo and Owen (2019) further document the usefulness of incorporating online search data in forecasting models for U.S. private consumption. Several studies have explored using Google Trends for forecasting unemployment, including Smith (2016) for the U.K., González-Fernández and González-Velasco (2018) for Spain, and Maas (2019) for the U.S. Götz and Knetsch (2019) and Ferrara and Simoni (2019) show that Google search data can also be useful for nowcasting quarterly GDP of Germany and the Euro Area, respectively. Narita and Yin (2018) show that GSV data significantly correlate with macroeconomic variables such as real GDP, inflation, and capital flows in low-income developing countries. A comprehensive overview of the literature using GSV can be found in Götz and Knetsch (2019) or Maas (2019).

To our knowledge, the previous literature has largely neglected the sampling issue we address with our method. A notable exception are Vosen and Schmidt (2012), who note that in the case of Germany “there is some non-negligible variation between samples drawn on different days” They address this problem by using averages of 50 samples of monthly data drawn on different days. That procedure, however, does not allow for the use of daily or weekly GSV data in real-time and hence one of the great advantages Google Trends offers is lost. Furthermore, we have not found any reference to the inconsistency between monthly, weekly, and daily time series obtained from Google Trends.

Since we use GSV data to construct economic sentiment indicators for a set of German-speaking countries, our paper further relates to other alternative indices to monitor the state of the economy. The most prominent example is Baker et al. (2016), who have developed a new index of economic policy uncertainty (EPU) based on newspaper coverage frequency for the U.S. and a number of other countries. Castelnuovo and Tran (2017) use monthly GSV data for the U.S. and Australia to construct an uncertainty index based on uncertainty-related keywords such as ‘bankruptcy’, ‘stock market’, ‘economic reforms’, and ‘debt stabilization’.

3. Querying Consistent Long-Run Daily GSV Data

Queries of GSV are based on a random subsample of all searches.⁴ This introduces variation between the samples drawn for different time periods. This

⁴Google does not share further details about the sampling method. All official information on Google Trends can be found in their FAQ: <https://support.google.com/trends/answer/4365533?hl=en>.

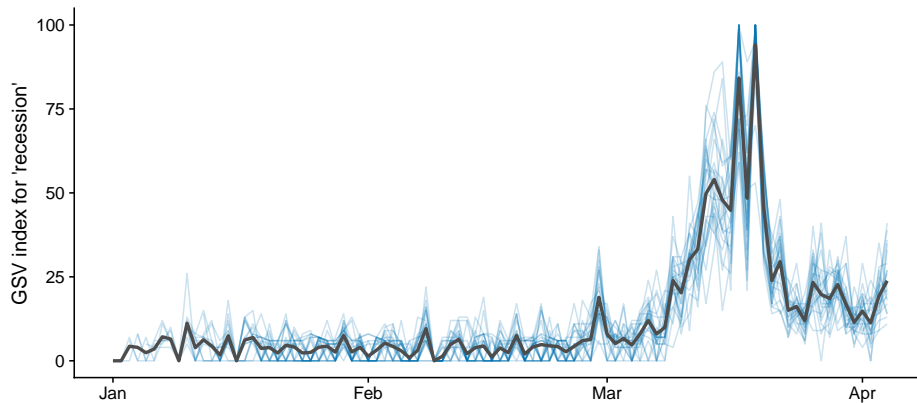


Figure 1: GSV query results for the term ‘recession’ during overlapping periods between January 1 and April 24, 2020, for the state of Massachusetts. The bold black line is the average of all 24 draws of daily data, pale lines show the individual draws.

problem is even present for larger countries (e.g. Germany, see Vosen and Schmidt, 2012). Although the use of Google as a search engine has increased dramatically over the past decade, sampling still is a cause for concern. The problem worsens for less popular keywords, for smaller geographic regions, and for higher data frequency. Indeed, for most U.S. states or smaller countries like Switzerland, results from search queries can differ heavily between draws.

Figure 1 illustrates the problem. The pale lines show 24 draws of daily query results for the term ‘recession’ in the state of Massachusetts. The first draw runs from January 1 to April 4, 2020, the second from January 2 to April 4, 2020, and so forth. The last draw runs from January 24 to April 4, 2020. The resulting series show strikingly different results, the average of all 24 draws is summarized in the bold line. This illustrates how the usage of the raw Google Trends data, for example in a forecast exercise, may lead to misleading results.

Figure 2 provides further systematic evidence of the sampling problem. For all 50 U.S. states and the (mainly) German-speaking countries Austria, Germany, and Switzerland we query monthly, weekly, and daily GSV time series for ‘recession’ (‘Rezession’ in German, respectively) during overlapping periods between January 1, 2007, until April 4, 2020. Panel A of Figure 2 shows how the mean of the standard deviation of the draws decreases as the population of a U.S. state or a country increases. For this monthly data, visual inspection suggests that individual monthly queries are particularly noisy if the population is smaller than 10 million inhabitants. Panel B of Figure 2 shows the the corresponding mean of the standard deviation for queries at daily frequency. Again the series show more randomness for smaller regions. Furthermore, the comparison between Panels A and B shows that for any given region the instability problem is more severe in daily than in monthly data. For weekly data, the standard deviations are almost identical in magnitude as for daily data.

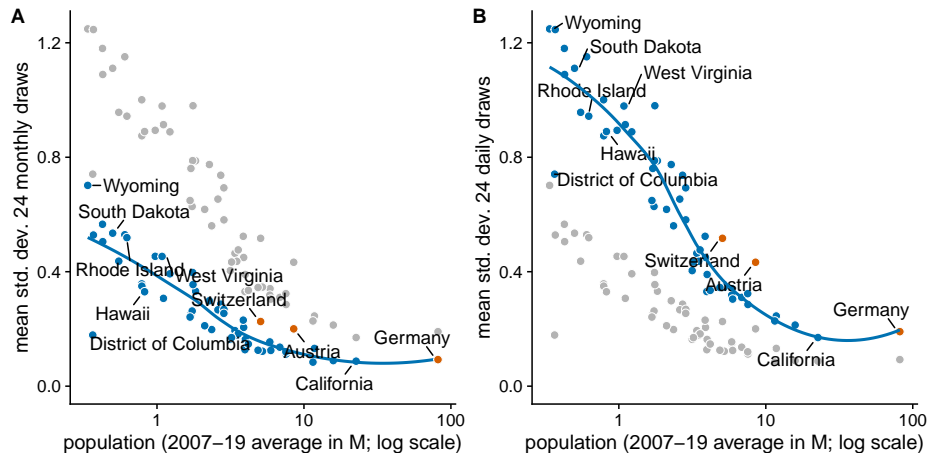


Figure 2: Average of the within-month and within-day sampling standard deviations, respectively, of the GSV query results for ‘recession’ by population size across U.S. states, Austria, Germany, and Switzerland. For German-speaking countries we use the translation ‘Rezession’. Panel A highlights monthly results in blue, results for daily data are shown in gray. B highlights daily results in blue and monthly data are shown in gray. For weekly data, the standard deviations are almost identical in magnitude as for daily data. We omit them from the figure for readability.

3.1. Sampling

To obtain consistent, long-run daily series, we draw at least 12 samples for each keyword and each frequency. Monthly series are drawn exactly 12 times; daily (weekly) series are drawn by applying a rolling window that shifts by 15 days (11 weeks). We choose a minimum of 12 draws as this reduces the variance of the mean values of the series at a given point in time by approximately 90 percent. It is a compromise that balances the need not to overstress the Google API, as Google blocks the IP after too many queries within a short period. As we add 2 updated draws per day to our economic indicators, the number of draws on the margin keeps increasing over time.

3.2. Harmonization of daily, weekly and monthly series

The obtained averaged daily, weekly, and monthly series still lead to different results across frequencies for any given keyword. For example, the weekly average of the daily series obtained after resampling and averaging as described in Section 3.1. does not correspond to the resampled and averaged weekly series. To combine the information from the three series into a single daily series that is consistent with the weekly and monthly series, we use a two-step procedure based on Chow and Lin’s (1971) disaggregation routine. The method can be used to disaggregate low frequency to high frequency series, while ensuring that the aggregation of the resulting high frequency series is consistent with the low frequency. Our procedure is based on the following assumptions:

1. Monthly data captures the long term trend in search activity in the most accurate way.
2. Weekly data is best to analyze the searches over a few weeks.
3. Daily data is best to analyze short term behavior over several days.

In a first step, we temporally disaggregate the weekly data by using the daily series as an indicator in the Chow and Lin (1971) framework implemented in Sax and Steiner (2013). This preserves the movement of the daily series and ensures that weekly averages are identical to the original (i.e., resampled and averaged) weekly series.

In a second step, we apply the same procedure using these adjusted weekly series from the first step as an indicator to temporally disaggregate the monthly values. This produces a series that maintains most of the movement of the daily and the weekly series, but has the same monthly averages as the original monthly series.

4. Robust Daily Economic Sentiment Indices

We use the method described above to obtain consistent GSV data for several keywords, and construct daily economic sentiment indices (DESI) for the (mainly) German-speaking countries Germany, Austria, and Switzerland over the period 2007–2020. Below we describe how we choose the keywords and time frame, how we seasonally adjust the robust GSV series, how we aggregate them into a daily index, and how the resulting indices compare to existing economic indicators.

4.1. Time frame and keyword choice

GSV data is in principle available since 2004, yet internet usage was not as widespread in early years and Google made improvements to their database over time, too. Thus, data quality may be worse in earlier years. Our index starts in January 2007 for two reasons. First, this is the year the first iPhone was introduced. This innovation triggered a dramatic increase in mobile internet use. Second, it covers the 20 months before the collapse of Lehman Brothers which marks the onset of the subsequent recession in 2008/2009. This will allow us to evaluate our DESI over a long time span which contains two major economic downturns.

We start with a set of German keywords related to the state of the economy. These keywords are commonly used in the economic policy debate, as well as in daily life. As GSV data default to zero for small search volumes (due to low popularity of a search term, the small size of the country or region, or a combination of both) to protect user’s privacy, we discard keywords that frequently return zeros. The keywords must therefore be general enough to be used frequently over the whole time period. We only consider keywords that show significant changes over time which coincide with changes in economic activity, such as the financial crisis.

As it turns out, search terms with a positive connotation, such as ‘economic recovery’ (‘Wirtschaftsaufschwung’), ‘economic growth’ (‘Wirtschaftswachstum’)

or ‘invest’ (‘investieren’) do either not coincide with changes in economic activity or have very low frequencies, such that they either mostly default to zero or return no results at all. We conclude that people’s interest in the state of the economy reflected in internet searches is asymmetric: it goes up during busts but not during booms. Google search activity related to the state of the economy therefore carries especially valuable, timely information during downturns, which in turn is reflected by the rise in search terms with a negative connotation.

Finally, we also consider the factor loadings in a principal component analysis when aggregating the keywords into an index (see Section 4.2.). We end up with four keywords: economic crisis (‘Wirtschaftskrise’), short-time work⁵ (‘Kurzarbeit’), unemployed (‘arbeitslos’), and bankruptcy (‘Insolvenz’). GSV for four keywords may not seem like a lot to build an economic index. However, as Baker et al. (2016) show in their seminal contribution of economic uncertainty indices, adding more words does not always significantly increase the content of information.⁶

4.2. Seasonal adjustment and principal component

For each keyword, we construct daily indices using the sampling method described in Section 3. These daily series feature weekly, monthly and yearly seasonalities. Moreover, there are irregular holidays, such as Easter, that occur at different dates each year. We use the Prophet procedure (Taylor and Letham, 2018) to seasonally adjust each time series on a daily basis. Framing the forecasting problem as a curve-fitting exercise, the fully automated procedure uses an additive model with non-linear trends, periodic changes (e.g., yearly and weekly seasonality) and irregular events such as holidays. However, the Prophet procedure assumes that yearly and weekly seasonal effects are constant over time and the procedure disregards any possible monthly seasonality patterns not captured by the weekly seasonal effects.

To aggregate the seasonally adjusted series into a DESI, we perform a principal component analysis to extract the first principal component as common signal in the four keyword series. Each DESI shows the relative change in search volumes over time. Because Google only shares relative search term frequency due to privacy concerns, we normalize all series such that the long-term average equals zero and the standard deviation is one. An index value of 2 therefore means that the search volume for the index is two standard deviations above its long-run average.

4.3. DESI for German-speaking countries

Using the procedure described above, we construct daily economic sentiment indices for Germany, Austria, and Switzerland. We use the same four German

⁵This is a well-established public program which allows firms to temporarily reduce working hours of their staff. The state covers 60-80% of foregone wages. The scheme exists in all three countries under the name ‘Kurzarbeit’ and has been in place for many decades.

⁶Their main index is based on three out of eight pre-defined terms. Similar to our index, their terms also tend to bear a negative connotation.

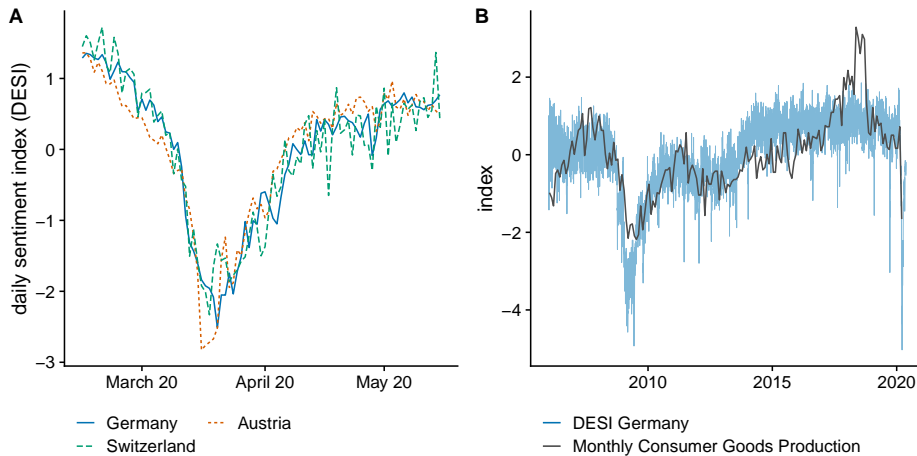


Figure 3: Daily Economic Sentiment Indices (DESI)

keywords in all three countries.⁷ Panel A in Figure 3 shows the resulting DESI during the period of the COVID-19 outbreak in 2020. Broadly in line with the spread of the virus and lock-down policy measures, the DESI for Austria declined and bottomed out slightly earlier than the indices for the other two countries. Subsequently, it took somewhat longer for the German and Swiss DESI to recover, as Austria was earlier in easing the lock-down measures. Switzerland—the smallest of the three countries—evolves almost identically to the German index but still reveals somewhat more volatility on a daily basis.

Panel B in Figure 3 depicts the German DESI in comparison with the monthly consumer goods production index during the full time horizon. This first visual inspection of our German DESI shows that it correlates well with economic fluctuations over the past 15 years for which reliable GSV data are available. Albeit being much more volatile due to its daily frequency, the DESI clearly mirrors changes in the medium term such as the Great Recession in 2008/2009.

4.4. Correlations with other main economic indicators

Finally, we assess the goodness of fit of our indices compared with other, established economic indicators. Following Neusser (2016), we fit an $AR(p)$ process to each series in order to extract autocorrelation and use the residuals to test for significant cross-correlations with existing economic indicators. To determine the lag length p , we use the AIC criterion.

Cross-correlation tests of these prewhitened series show that for all three countries the respective DESI are indeed significantly correlated with other leading economic indicators at a monthly frequency. Both, the German and the

⁷Strictly speaking, the index for Switzerland is therefore only representative of the German-speaking part of the country. However, German is the main language of 63% of the total population in Switzerland and of 73% of Swiss citizens.

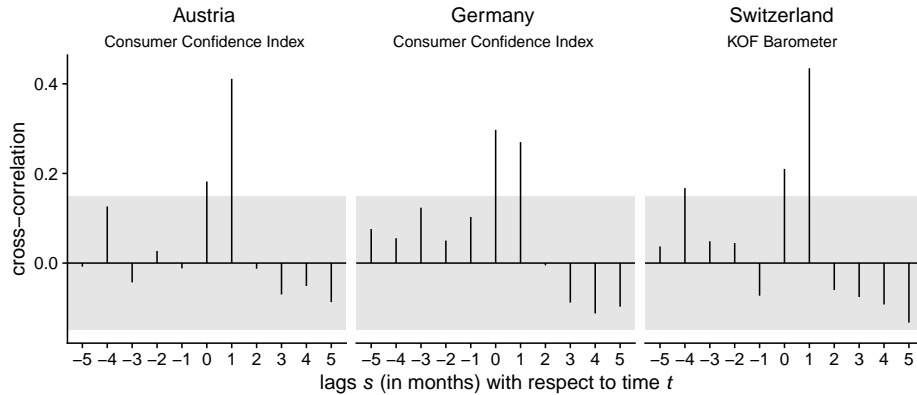


Figure 4: Cross-correlation of DESI (t) with other leading indicators ($t+s$). Gray areas mark 95%-confidence intervals. In all three countries, correlations between our DESI based on GSV and other leading monthly indicators are statistically significant in t_0 and t_1 .

Austrian DESI, feature a significant lead as well as concurrent correlation with the consumer confidence index of the respective country (Panels A and B in Figure 4). For Switzerland, where no monthly data of consumer confidence exist, the DESI also shows a significant concurrent and leading correlation with the most prominent leading composite indicator for the Swiss economy, the KOF Barometer, as Panel C in Figure 4 shows. While the correlations are lower during the much shorter time window 2011–2019 which excludes substantial economic downturns, they become very high during crisis episodes. In Germany and Austria, the DESI also shows significant concurrent correlations with real variables, such as industrial production.⁸ This is a valuable feature for real-time business cycle analysis, since our index can be calculated at the end of the month, while German industrial production, for example, is published with a lag of around 40 days.

5. Conclusion

Google Trends queries suffer from a small sample problem on daily, weekly and monthly frequencies. We show how this problem can be solved and provide an open source code which allows to query sample-robust GSV data even for small countries. We apply this method to construct DESI for German-speaking countries, including small countries such as Switzerland or Austria. A comparison to other, well-established leading economic indicators shows that our DESI are significantly correlated with these leading indicators. Our indices seem to be particularly useful during crisis episodes such as the recent COVID-19. This makes them especially useful for policy analysis: they capture large downturns

⁸Results available upon request. For Switzerland, monthly industrial production data does not exist prior to October 2010.

in real-time, and crisis episodes are typically marked by an increased need of timely data.

Acknowledgments

We thank Angelica Becerra, Daniel Kaufmann, Stefan Legge, Nina Mühlebach, Klaus Neusser, Uwe Thümmel as well as participants at the VersusVirus hackathon for valuable input.

Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636, 2016. doi: 10.1093/qje/qjw024.

Efrem Castelnuovo and Trung Duc Tran. Google it up! A Google Trends-based Uncertainty index for the United States and Australia. *Economics Letters*, 161:149–153, 2017.

Gregory C. Chow and A. Lin. Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series. *The Review of Economics and Statistics*, 53(4):372–375, 1971.

Laurent Ferrara and Anna Simoni. When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage, 2019.

Marcos González-Fernández and Carmen González-Velasco. Can Google econometrics predict unemployment? Evidence from Spain. *Economics Letters*, 170: 42–45, 2018.

Thomas B Götz and Thomas A Knetsch. Google data in bridge equation models for German GDP. *International Journal of Forecasting*, 35(1):45–66, 2019.

Neil Gross and Marcus Mann. Is There a “Ferguson Effect?” Google Searches, Concern about Police Violence, and Crime in U.S. Cities, 2014–2016. *Socius*, 3, 2017. doi: 10.1177/2378023117703122.

Seung-Pyo Jun, Hyoung Sun Yoo, and San Choi. Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change*, 130:69–87, 2018.

Benedikt Maas. Short-term forecasting of the US unemployment rate. *Journal of Forecasting*, 39(3):394–411, 2019.

Futoshi Narita and Rujun Yin. In Search of Information: Use of Google Trends’ Data to Narrow Information Gaps for Low-income Developing Countries. IMF Working Paper No. 18/286, 2018.

Klaus Neusser. *Time Series Econometrics*. Springer, Cham, Switzerland, 2016.

- Christoph Sax and Peter Steiner. Temporal Disaggregation of Time Series. *The R Journal*, 5(2):80–87, 2013. doi: 10.32614/RJ-2013-028. URL <https://doi.org/10.32614/RJ-2013-028>.
- Paul Smith. Google’s MIDAS Touch: Predicting UK Unemployment with Internet Search Data. *Journal of Forecasting*, 35(3):263–284, 2016.
- Seth Stephens-Davidowitz. The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, 118:26–40, 2014.
- C Douglas Swearingen and Joseph T Ripberger. Google Insights and U.S. Senate Elections: Does Search Traffic Provide a Valid Measure of Public Attention to Political Candidates? *Social Science Quarterly*, 95(3):882–893, 2014.
- Sean J Taylor and Benjamin Letham. Forecasting at Scale. *The American Statistician*, 72(1):37–45, 2018.
- Nataliya Tkachenko, Sarunkorn Chotvijit, Neha Gupta, Emma Bradley, Charlotte Gilks, Weisi Guo, Henry Crosby, Eliot Shore, Malkiat Thiarai, Rob Procter, et al. Google Trends can improve surveillance of Type 2 diabetes. *Scientific Reports*, 7(1):1–10, 2017.
- Simeon Vosen and Torsten Schmidt. Forecasting private consumption: Survey-based indicators vs. Google trends. *Journal of Forecasting*, 30(6):565–578, 2011.
- Simeon Vosen and Torsten Schmidt. A monthly consumption indicator for Germany based on internet search query data. *Applied Economics Letters*, 19(7):683–687, 2012.
- Jaemin Woo and Ann L Owen. Forecasting private consumption with Google Trends data. *Journal of Forecasting*, 38(2):81–91, 2019.