Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Google search volumes and the financial markets during the COVID-19 outbreak

Michele Costola\textsuperscript{a}, Matteo Iacopini\textsuperscript{b}, Carlo R.M.A. Santagiustina \textsuperscript{*,a}

\textsuperscript{a} Ca' Foscari University of Venice, Cannaregio 873, 30123 Venice Italy
\textsuperscript{b} Vrije Universiteit Amsterdam, De Boelelaan 1169, 1081 HV Amsterdam The Netherlands

\textbf{ARTICLE INFO}

\textbf{JEL classification:}
G41
G15
C32

\textbf{Keywords:}
COVID-19
Financial Markets
Google Trends

\textbf{ABSTRACT}

During the outbreak of the COVID-19, concerns related to the severity of the pandemic have played a prominent role in investment decisions. In this paper, we analyze the relationship between public attention and the financial markets using search engine data from Google Trends. Our findings show that search query volumes in Italy, Germany, France, Great Britain, Spain, and the United States are connected with stock markets. The Italian Google Trends index is found to be the main driver of all the considered markets. Furthermore, the country-specific market impacts of COVID-19-related concerns closely follow the Italian lockdown process.

\section{1. Introduction}

Search engine data has been successfully used for tracking collective attention and public concerns, which are often correlated with social, environmental and economic events. Google Trends (GT) provides indexes based on the relative web-search volumes of a specific topic over time. These indexes can be retrieved for selected geographic areas or on a worldwide scale. The interpretation of GT indexes is straightforward: the higher the value of a given GT index, the more public attention there is on that topic. In recent years, it has been shown that the informational content of GT data has explanatory and forecasting power in several fields of economics and finance.

GT data has been successfully used for disease surveillance purposes for MERS (Shin et al., 2016), chickenpox (Bakker et al., 2016) and flu (Yang et al., 2015). In relation to economics and finance, there are two main strands of research concerning Google search data. A first strand of literature studies how web-search volumes are related to financial markets in terms of returns, trading volume, and liquidity. Those papers make use of keywords related to the listed companies and to the stock market indexes. For instance, Joseph et al. (2011), Ding and Hou (2015), and Bijl et al. (2016) search for tickers and names of companies included in the S&P 500. Da et al. (2011) consider the stock names listed in the Russell 3000 index while Takeda and Wakao (2014) study those ones included in the Nikkei 225. Dimpfl and Jank (2016) and Tantaopas et al. (2016) make use of keywords related to the stock market indexes. Overall, these papers are all concerned with a specific research question, that is, whether web-searches volumes on company names or stock market indexes can predict returns for a given individual firm. The second part of literature focuses on the use of GT indexes for the construction of economic uncertainty indicators that can explain several macroeconomic variables (Donadelli, 2015; Castelnuovo and Tran, 2017; Donadelli and Gerotto, 2019).

This paper takes a different perspective and relates to the literature that studies the impact of exogenous shocks on the financial
market, such as natural disasters, terrorist events, and infection diseases (Worthington and Valadkhani, 2004; Chesney et al., 2011; Bourdeau-Brien and Kryzanowski, 2017). As noted in Gong et al. (2020), the latter has received limited attention so far. We consider the relationship between stock markets and web-searches volumes related to the COVID-19 pandemic, a topic that, unlike the previously mentioned studies on Google search volumes, is not a priori related to any financial and economic activity. To our knowledge, this paper is the first to analyze this type of relationship.

We believe that the outbreak of the pandemic on the financial markets represents an interesting and unprecedented case to investigate. No previous disease outbreak has affected financial markets as the COVID-19 pandemic (Baker et al., 2020). For instance, the previous H1N1 pandemic that occurred in the middle of the Global financial crisis in 2009 has left almost no traces in the stock market (see, for a comparative study, Schell et al., 2020). We also note that the massive amount of aggregate volumes of web data was different at that time, partially due to the facts that the number of internet users was significantly lower and the mobile internet was in its infancy.

In this respect, we retrieve GT country indexes for the topic of coronavirus (GT-COVID-19) from January 2020 to April 2020. We use them as proxies for country-level public attention and investigate their impact on the financial markets during the outbreak of the coronavirus disease. The COVID-19 pandemic is an interesting case to investigate since, due to its virulence and infectiousness, it represents a major exogenous shock to the economic and financial systems. Even if a priori a new pandemic is not an unlikely event (Goodell, 2020), its origin, severity and impact cannot be reasonably foreseen and hence priced by markets.

In our analysis, we consider the six most impacted countries worldwide in terms of confirmed cases, as of May 1 2020: the United States, Spain, Italy, Great Britain, Germany, and France. Our findings are three-fold. First, the GT-COVID-19 index for Italy is found to anticipate the GT indexes for all the other countries considered. This lead-lag relationship is primarily due to the fact that Italy was the first European country to experience an outbreak of COVID-19, as well as being the first European country to implement lockdown measures since World War II. The ensuing spread of the coronavirus disease in other countries provided a similar dynamic on the new cases and for the implemented lockdown measures. Therefore, the delayed reaction of those indexes is likely due to the shock of the aforementioned events.

Second, we analyze whether GT-COVID-19 indexes explain the stock market returns. Given that an epidemic disease is, by definition, an adverse event, GT indexes can be interpreted as a measure of coronavirus-related uncertainty and perceived risk. Our findings show that GT-COVID-19 indexes contribute to explaining the dynamic of stock market returns for Italy, Spain, and Germany. Interestingly, substituting country-specific GT indexes with the Italian index magnifies the relationship with all considered markets, bursting the explained variance of stock index returns. This highlights that the outbreak in Italy may have exerted a role in the general perception of the pandemic’s severity.

Third and finally, we perform time-varying analysis to investigate the impact of the GT-COVID-19 indexes on the financial markets.

---

**Fig. 1. Google Trends data extraction process.** Flow chart of Google Trends data extraction process. Balls represent queries, black color represents queries matching the topic parameter q.
over time. Interestingly, we identify the most severe impacts at the beginning of each phase of the Italian lockdown process and find almost no impact prior to the beginning of the pandemic in Europe, even though the presence of COVID-19 in China was already known by the end of December 2019.

2. GT-COVID-19 indexes

The construction of the GT-COVID-19 indexes

The construction of indexes of collective attention based on Google Trends data follows the assumption that economic agents, as well as market participants, search for online information as part of their decision process. This implies that the higher is the web-search volume related to COVID-19, the higher is the level of public concern on this topic. Google Trends offers the possibility to generate indexes either by term(s) or by topic, which are groups of terms that have equivalent meanings in different languages. Even though most of the literature is concerned with searches by term(s), searching by topic has several advantages. For example, it makes it possible to take into account searches done in the selected countries in languages other than the official one(s). As a result, the GT indexes used in this study are also representative of searches relating to COVID-19 made by linguistic minorities/foreigners (e.g., searches in Chinese, made from Italy).

We have collected Google Trends for Germany (DE), France (FR), Great Britain (GB), United States (US), Italy (IT), and Spain (ES) from January 1 to April 14 2020. Those countries were the most impacted worldwide in terms of confirmed cases, as of May 1 2020.

For each country, we have retrieved GT daily indexes for Youtube, Google News, and Google Search, matching the “coronavirus” topic (q=m/01cpyy). Google Trends (GT) provides rescaled values of relative search volumes per unit-of-time. The data-generation procedure of GT can be summarized as follows:

- Extract a random sample of queries corresponding to the searched geographical area (geo), source (gprop) and time-span (date);
- For every unit-of-time (whose length/duration depends on the time-span) divide the count of the number of searches matching the query term(s)/topic parameter (q), by the total number of searches in the same unit-of-time;
- Rescale the obtained time series in the [0,100] interval by removing the minimum value of the series to all values, dividing all values of the series by the maximum value of the series, multiplying all values by 100, and rounding to the nearest integer.

The data extraction process for GT is represented in detail in Figure 1 and the parameters values are listed in Table 1. Once downloaded, the GT indexes have been rescaled in the [0,1] interval.

Financial market data

For the same period (75 market days), we consider the closing price of the stock market indexes to obtain the corresponding log-returns. Using Bloomberg as data provider, we have downloaded the S&P500 (US), the FTSE MIB 30 (IT), the DAX (DE), the CAC 40 (FR), the IBEX (ES) and FTSE 100 (GB).

Figure 2 shows the time series of log-returns and the three groups of GT indexes, subdivided by source. For each country $i \in \{\text{DE, FR, GB, US, IT, ES}\}$, we observe the log-return series, $y_{i,t}$, and three GT indexes $GT_{i,t}$, where $j \in \{Y, N, S\}$, corresponding to YouTube, Google News and Google Search, respectively. The country-specific GT series show a common dynamic, while for each source there is cross-country heterogeneity. It is worth noting that high values of either GT index correspond to intervals that exhibit turbulent dynamics for stock market returns. This is particularly evident throughout March. By looking at the beginning of the period, we observe that both the GT indexes and the financial series do not appear to react to epidemics-related facts that originated in China.

Cross-Correlation among the GT-COVID-19 indexes

Looking at the GT indexes included in Figure 2, we notice that the Italian ones (green line) seem to anticipate those of the other countries. In fact, the GT-COVID-19 indexes for Italy reach the last peak around March 11 and then start to decrease prior to all the others. To test for the presence of lead-lag relations, we perform a cross-correlation analysis of each country against Italy for Google Search, YouTube and Google News. Table 2 reports the peak in the cross-correlation, with negative values implying that the series for Italy is leading. We find that Italy leads the other countries, with a lead ranging from 3 days (Spain and Germany) to 8 days (Great Britain).

3. Impact of GT-COVID-19 indexes on stock markets

We investigate the relationship between stock market returns and public concern proxied by the GT-COVID-19 indexes. The aim is to detect whether those indexes contribute in explaining the financial market stress during the COVID-19 outbreak. In this respect, we perform a constant parameter and a time varying parameter analyses.
Constant parameters analysis

We estimate for each country $i$ the following model:

$$y_{i,t} = \beta_0 + \beta_1 GT_{i,t} + \beta_2 y_{i,t-1} + \epsilon_{i,t}$$

**Table 1**

| Param. | Description | Value(s) |
|--------|-------------|----------|
| geo    | filters by geographic area | IT, FR, DE, US, GB, ES |
| gprop  | filters by source, defaults to (all) google searches | (all), youtube, news |
| date   | filters data through the time dimension | 01-01-2020 to 14-04-2020 |
| q      | filters data by term(s) OR topic | /m/01cpyy ('coronavirus' topic) |

**Table 2**

Lead-lag relationships for GT indexes versus Italy. Peak of the daily cross correlation of the GT-COMOD-19 indexes (in levels) versus Italy. Negative values mean that the corresponding series for Italy are leading.

|        | DE  | FR  | GB  | US  | ES  |
|--------|-----|-----|-----|-----|-----|
| YouTube| -3  | -4  | -6  | -4  | -3  |
| Google News| -3  | -5  | -7  | -5  | -3  |
| Google Search| -4  | -6  | -8  | -6  | -3  |

**Table 3**

Results of the constant parameter regressions, using the country-specific GT-COMOD-19 index for YouTube.

|        | DE  | FR  | GB  | US  | IT  | ES  |
|--------|-----|-----|-----|-----|-----|-----|
| $\beta_0$ | 0.006** | 0.002 | -0.000 | 0.005 | 0.011*** | 0.004 |
| (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) | |
| $\beta_1$ | -0.032** | -0.024 | -0.014 | -0.030* | -0.058*** | -0.050* |
| (0.017) | (0.021) | (0.017) | (0.021) | (0.021) | (0.033) | |
| $\beta_2$ | -0.078 | -0.059 | -0.070 | -0.469*** | -0.309* | -0.225 |
| (0.139) | (0.166) | (0.159) | (0.171) | (0.189) | (0.224) | |
| $R^2$ | 0.047 | 0.012 | -0.007 | 0.200 | 0.189 | 0.107 |

Notes: HAC standard errors are in parentheses; *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively.

**Constant parameters analysis**

We estimate for each country $i$ the following model:
Here, we present the results on YouTube indexes, i.e. $GT_{i,t} = GT_{Y,t}$. Table 3 highlights that the impact of the country-specific GT-COVID-19 index is significant and negative for Germany ($-0.032$), United States ($-0.03$), Italy ($-0.058$) and Spain ($-0.05$). The GT-COVID-19 index behaves similarly to sentiment indicators, e.g., the Financial and Economic Attitudes Revealed by Search (FEARS) based on GT data and proposed by Da et al. (2015). For instance, an increase of one unit for Italy (here 0.01 with the rescaled index) corresponds to a loss of 0.058% in the related Italian stock market. Conversely, the relation is not significant for France and Great Britain suggesting that the domestic public concern and the market in those countries have reacted differently.

Following the insights from the lead-lag relations, we check the exposition of stock markets to public concern in Italy, estimating the model using $GT_{i,t} = GT_{IT,t}$ for all $i$. Interestingly, as shown in Table 4, the Italian GT-COVID-19 index results a key explanatory variable for all country-level stock index returns, and its use remarkably increases the adjusted $R^2$ and the statistic significance (less than 1%) compared to using country-specific GT indexes. For instance, the adjusted $R^2$ moves from 0.047 (0.107) to 0.122 (0.186) for Germany (Spain), as shown in the first (last) column of the two tables. Also, the magnitude of the coefficients considerably increases (e.g., Germany, US). Finally, the French and British markets are insensitive to the domestic public concern, but are significantly related to the Italian GT-COVID-19. This highlights that the severity of the outbreak perceived from Italy represents a timely indicator of the destabilizing effect of the pandemic on financial markets.

We checked the robustness of our findings by controlling for (i) the Italian weekend effects (We$_{IT,t}$), (ii) the variation of the Oxford stringency index ($\Delta OX_{it}$), (iii) the country-specific implied volatility ($IV_{i,t}$), and (iv) the country-specific growth rate of new COVID-19 cases ($\Delta^{\text{new}}Cc_{i,t}$).

In the latter case, we estimated the model

\[ y_{it} = \alpha_i + \beta_i GT_{i,t} + \delta_i y_{i,t-1} + \epsilon_i \sim \mathcal{N}(0, \sigma_i). \]

Results are included in Table 5 and show that the magnitude of the coefficients for the Italian GT-COVID-19 increases even further after controlling for those variables. The growth rate of new COVID-19 cases is statistically significant for Italy, Spain and Germany. As expected, we find no statistically significant impact of the weekend effect ($We_{IT,t}$). This is likely due to the fact that the GT-COVID-19 index time series are quite persistent in the period under analysis, hence the weekend variable does not provide any additional information as compared to the contemporaneous GT-COVID-19 index. For example, on Monday the value of weekend effect contains the information about the GT index on Saturday and Sunday, while the contemporaneous GT-COVID-19 index carries also the "new" information of Monday itself. This is clearly evident when the GT index is increasing (decreasing) path, since in this situation the very last observation has a higher value than the before last ones.

Additionally, we have extended our analysis to the BRICS countries and controlled for Google searches that occurred before the closure of the markets using the lagged GT-COVID-19 indexes. The analyses confirm our results and are included in the Supplementary material.

**Time-varying parameters analysis**

To further analyze the relationship between the GT-COVID-19 indexes and market returns, we estimate a time-varying parameters model where the impact of the GT-COVID-19 indexes is allowed to evolve according to:

\[ y_{i,t} = \alpha_i + \beta_{i,t} GT_{i,t} + \delta_i y_{i,t-1} + \epsilon_{i,t} \sim \mathcal{N}(0, \sigma_i), \]

\[ \beta_{i,t} = \alpha_{\beta_{i,t}} + \eta_{\beta_{i,t}} n_{i,t} \sim \mathcal{N}(0, \delta), \]

where $\epsilon_{i,t}$ and $\eta_{i,t}$ are mutually independent. Estimation of $\beta_{i,t}$ is performed using the Kalman smoother.

Figure 3 plots the country-specific time-varying parameters of the GT indexes on the corresponding stock index returns. The estimated coefficients share a common dynamic and are very close to zero before the first confirmed case in Germany (January 23). Interestingly, the lowest peaks occur in correspondence of the enactment of the Italian lockdowns: (i) the lockdown of northern Italian provinces (February 24); (ii) the lockdown for all Italian citizens (March 9); and (iii) the lockdown of most Italian economic activities (March 12). This highlights that the public attention has no impact on financial markets before the spreading of the pandemic to Europe, even though the presence of COVID-19 in China has been known since the end of December 2019. Therefore, the perception of...
the economic consequences due to the severity of the COVID-19 pandemic almost certainly has changed after the European outbreak.

4. Conclusions

Recent literature has shown that Google Trends data can successfully explain current and future patterns of the state of the
economy, especially during unfavorable events (Heiberger, 2015; Yu et al., 2019; Zhong and Raghib, 2019). Unlike recent financial crises, COVID-19 has been an exogenous shock to the system that has affected several countries with different timings that relate to the spread of the disease. We have investigated the exposure of the stock index returns of Italy, France, Germany, Great Britain, the United States, and Spain to GT-COVID-19 indexes based on search-engine query volumes.

Our findings show that most of these indexes have significant explanatory power on stock market returns. Interestingly, the Italian GT-COVID-19 index acts as forerunner and better explains other countries’ market returns. Moreover, the greatest impact of GT indexes occur in correspondence of the different phases of the lockdown in Italy, despite public awareness of the contagion in China existing since January.

The disruptive effect of COVID-19 on financial markets is well described by the public concern perceived in Italy, which has been the first Western country to experience a virulent outbreak and to adopt drastic measures.

CRediT authorship contribution statement

Michele Costola: Conceptualization, Methodology, Data curation, Writing - review & editing. Matteo Iacopini: Conceptualization, Methodology, Data curation, Writing - review & editing. Carlo R.M.A. Santagiustina: Conceptualization, Methodology, Data curation, Writing - review & editing.

Acknowledgments

Matteo Iacopini acknowledges financial support from the EU Horizon 2020 programme under the Marie Skłodowska-Curie scheme (grant agreement no. 887220).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.frl.2020.101884.

References

Baker, S.R., Bloom, N., Davis, S.J., Kost, K., Sammon, M., Viratyosin, T., 2020. The unprecedented stock market reaction to covid-19. The Review of Asset Pricing Studies.
Bakker, K.M., Martinez-Bakker, M.E., Helm, B., Stevenson, T.J., 2016. Digital epidemiology reveals global childhood disease seasonality and the effects of immunization. Proceedings of the National Academy of Sciences 113 (24), 6689–6694.
Bijl, L., Krinngaauw, G., Molnar, P., Sandvik, E., 2016. Google searches and stock returns. International Review of Financial Analysis 45, 150–156.
Bourdeau-Brien, M., Kryzanowski, L., 2017. The impact of natural disasters on the stock returns and volatilities of local firms. The Quarterly Review of Economics and Finance 63, 259–270.
Castelnuovo, E., Tran, T.D., 2017. Google it up! a Google Trends-based uncertainty index for the United States and Australia. Economics Letters 161, 149–153.
Chesney, M., Reshetar, G., Karaman, M., 2011. The impact of terrorism on financial markets: An empirical study. Journal of Banking & Finance 35 (2), 253–267.
Da, Z., Engelberg, J., Gao, P., 2015. The sum of all fears investor sentiment and asset prices. The Review of Financial Studies 28 (1), 1–32.
Dimpfl, T., Jank, S., 2016. Can internet search queries help to predict stock market volatility? European Financial Management 22 (2), 171–192.
Ding, R., Hou, W., 2015. Retail investor attention and stock liquidity. Journal of International Financial Markets, Institutions and Money 37, 12–26.
Donadelli, M., 2015. Google search-based metrics, policy-related uncertainty and macroeconomic conditions. Applied Economics Letters 22 (10), 801–807.
Donadelli, M., Gerotto, L., 2019. Non-macro-based google searches, uncertainty, and real economic activity. Research in International Business and Finance 48, 111–142.
Gong, D., Jiang, T., Lu, L., 2020. Pandemic and bank lending: Evidence from the 2009 h1n1 pandemic. Finance Research Letters 101627.
Goodell, J.W., 2020. Covid-19 and finance: Agendas for future research. Finance Research Letters 101512.
Heiberger, R.H., 2015. Collective attention and stock prices: Evidence from Google Trends data on Standard and Poor’s 100. PloS one 10 (8), e0135311.
Joseph, K., Wintoki, M.B., Zhang, Z., 2011. Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. International Journal of Forecasting 27 (4), 1116–1127.
Scheh, D., Wang, M., Huyhn, T.L.D., 2020. This time is indeed different: A study on global market reactions to public health crisis. Journal of Behavioral and Experimental Finance 100349.
Shin, S.-Y., Seo, D.-W., An, J., Kwak, H., Kim, S.-H., Gwack, J., Jo, M.-W., 2016. High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. Scientific reports 6, 32920.
Takeda, F., Wakao, T., 2014. Google search intensity and its relationship with returns and trading volume of Japanese stocks. Pacific-Basin Finance Journal 27, 1–18.
Tantaqas, P., Padungsaksawadi, C., Treepongkaruna, S., 2016. Attention effect via internet search intensity in asia-pacific stock markets. Pacific-Basin Finance Journal 38, 107–124.
Wooding, A., Tahdikhani, A., 2004. Measuring the impact of natural disasters on capital markets: an empirical application using intervention analysis. Applied Economics 36 (19), 2177–2186.
Yang, S., Santilanna, M., Kou, S.C., 2015. Accurate estimation of influenza epidemics using Google search data via ARGO. Proceedings of the National Academy of Sciences 112 (47), 14473–14478.
Yu, L., Zhao, Y., Tang, L., Yang, Z., 2019. Online big data-driven oil consumption forecasting with Google Trends. International Journal of Forecasting 35 (1), 213–223.
Zhong, X., Raghib, M., 2019. Revisiting the use of Web search data for stock market movements. Scientific reports 9 (1), 1–8.