

Analysis of Impact of Covid19 Pandemic on Stock Prices of Large Tech Video Communication Companies



A REPORT PRESENTED AS A PART OF THE COURSE
FINANCIAL ECONOMETRICS, SPRING 2021

Rounak Ghosh, MB1904
Nadim Akhtar, MB1906
Subhrajyoty Roy, MB1911
Monitirtha Dey, MB1921

Programme : Master of Statistics (M.Stat. Year II)
Institution: Indian Statistical Institute, Kolkata
Date: July 2, 2021

1 Introduction

Forecasts are an important tool for investors to close successful deals. This gives traders the confidence to know in a timely manner when to buy and sell stocks. Forecasting stock prices is important for market participants because sufficiently accurate forecasts can also provide huge financial benefits for policymakers who are analyzing the market’s impact on national or international events. Despite the pressuring need for stock price forecasting and numerous statistical, machine learning, deep learning and computation techniques available [15], various economic and social factors create uncertainty and volatility in those predictions.

Infectious diseases have always been a threat to mankind, especially those that are little or not known about. The World Health Organization (WHO) describes a pandemic as “the spread of a new disease worldwide” [12] and while at such times the greatest concern is saving lives, the next first goal is to save the economy and maintain prosperity. During recent times of the ongoing COVID-19 pandemic, it has become apparent that the study of the pandemic is an influential component of the behaviour of the stock market. From 24 to 28 February 2020, stock markets around the world experienced their biggest one-week decline since the 2008 financial crisis [13], with traders starting to sell stocks out of fear of imminent recession and market liquidity crisis. As a result, the four-way switch was triggered four times in March. The effect of COVID-19 was not only direct, but indirect effects associated with healthcare policies and lockdown strategies [6], miscommunication and spread of fake news [10] were also some of the social factors that cause the stock market crash in 2020.

When the Coronavirus pandemic began to hit the world, both the academic and corporate industries adopted a study-from-home and work-from-home culture. Consequently, the flexible e-learning tools, video conferencing applications relying on the Internet have proven to be the key to teacher-student interactions [2] in academics, employer-employee relationships in the corporate industry. It has become a very effective solution for everyone involved, keeping the public safe while promoting the useful use of technology, as well as allowing various companies to save budgets of office maintenance costs. In academics, despite various social issues pertaining to internet access inequalities, many agree that continuing online classes proved beneficial by allowing the students to avoid knowledge gaps that could have arisen if students did not have access to education for a long time due to lockdown. Many researchers even suspect that this temporary change to counter the threat of disease spread could become a permanent culture [1], as most of the world has already adapted to the emerging system.

This radical change in social interactions has significantly boosted the usage of online video conferencing applications over the past year. Among different video conferencing tools, Google Meet, Microsoft Teams, Zoom and Skype are more popular than the rest. Clearly, the associated companies, namely Alphabet Inc (Google) and Microsoft Corporation both have been able to generate enormous revenue by selling their products, which should be reflected in the prices of these particular stocks. Hence, we expect that the intensity of worldwide confirmed covid cases should well be correlated with their stock prices. Therefore, in this project, we aim to examine the effects of the COVID-19 pandemic on the stock prices of Alphabet Inc (Google) and Microsoft Corporation.

2 Description of the Data

The World Health Organization (WHO) declared the outbreak of COVID-19 a Public Health Emergency of International Concern on 30 January 2020 [14]. So, we restricted our dataset starting from 30th January, 2020 to 30th April, 2021. Among stock prices, our dataset includes daily closing prices of the stocks Alphabet Inc (symbol GOOGL) and Microsoft Corporation (symbol MSFT) collected from Yahoo Finance [4]. Both of these stocks are registered in the American stock exchange market NASDAQ [9]. Figure 1 shows the line diagram of these two

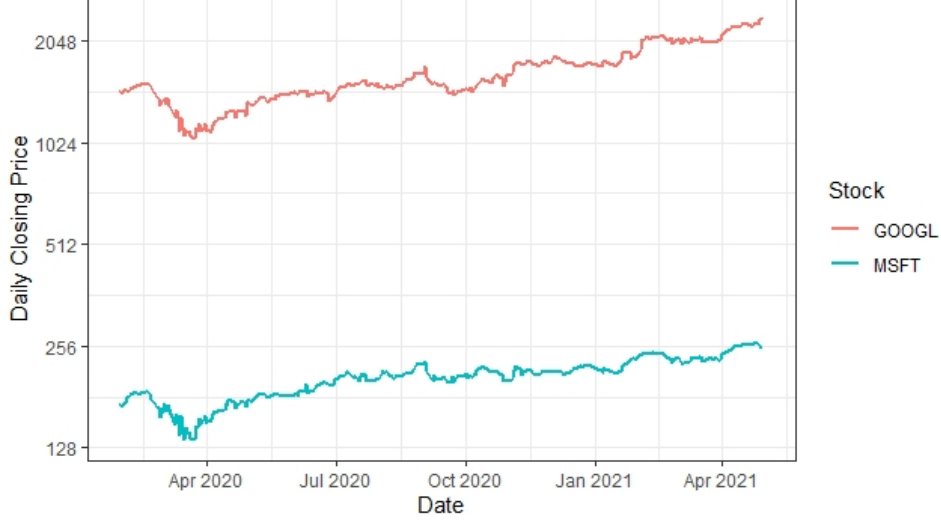


Figure 1: Daily closing stock prices for the stocks of Alphabet Inc (GOOGL) and Microsoft Corporation (MSFT)

stock prices for the studied timespan. Our dataset also includes the number of daily new confirmed cases of covid on a worldwide basis which are collected from JHU CSSE Github repository [11]. In addition to these, we believe that the number of google searches about certain keyword terms such as “google meet”, “microsoft teams”, “skype”, “zoom” and “covid test”, may be useful to quantify the current trend of popularity of video conferencing applications among the masses. Google publishes these aggregated searching data through the Google Trends platform [8].

Google Trends [8] provides access to a largely unfiltered sample of actual search requests made to Google. It’s anonymized (no one is personally identified), categorized (determining the topic for a search query), and aggregated (grouped together). This allows us to display interest in a particular topic from around the globe or down to city-level geography. Instead of providing the actual number of searches about a particular topic, it normalizes the search data to make comparisons between terms easier. Search results are normalized to the time and location of a query by the following process:

- Each data point is divided by the total number of searches in the geography and time range to represent relative popularity. Otherwise, places with the most search volume would always be ranked highest.
- The resulting numbers are then scaled on a range of 0 to 100 based on a topic’s proportion to all searches on all topics.
- Different regions that show the same search interest for a term do not always have the same total search volumes.
- Finally, the normalized scores of popularity are rounded to the nearest integer between 0 to 100.

We found that Google Trends data containing the whole time range and all the five search terms together have little use since most of the terms have relatively low popularity compared to that of covid related searches, resulting in a normalized score of 0 for most of the time. Hence, we segmented the whole timespan into 4 overlapping segments of 5-6 months, with one month of data common between them. That allowed us to have more descriptive popularity scores, but with different normalization factor for different timespan. For instance, let x_1, x_2, \dots, x_n be

the original time series. However, for the time span $\{1, 2, \dots, k, k+1, \dots, r\}$, let the observed scores are y_1, y_2, \dots, y_r and for the time span $\{k, k+1, \dots, r, r+1, \dots, n\}$, let the observed scores are z_k, z_{k+1}, \dots, z_n . In general, $y_i = x_i/c_1$ and $z_i = x_i/c_2$, where c_1 and c_2 are possibly different normalization factors to scale the corresponding series in the range 0-100. Since $y_i c_1 = x_i = z_i c_2$ for $i = k, k+1, \dots, r$, the ratio c_1/c_2 can be approximated by a simple average of the values $z_k/y_k, z_{k+1}/y_{k+1}, \dots, z_r/y_r$. This ratio can then be multiplied with the complete z -series to obtain a properly scaled version of it which can be easily merged with the y -series (Note that after the multiplication, the normalization constant becomes the same as c_1). This procedure allowed us to combine the segmented time series of search trends into combined search trends for the whole study period.

3 Methodologies

As mentioned earlier, our study focuses on quantifying the effect of worldwide confirmed covid cases on the stock prices of Alphabet Inc (Google) and Microsoft Corporation. It is standard practice in stock price analysis and economics to perform a logarithmic transformation of the prices, which is also done in this case. Figure 2 represents the relationship between the two log-transformed stock prices and daily new confirmed covid cases on a worldwide basis. We observe non-linear relationships between the log-transformed stock prices and daily new confirmed covid cases. To mitigate this, square root transformation and log transformation are two popular options on the covid cases variable. We found that square root transformation performs better for our dataset. This revised variable is then found to have a better linear relationship with the log-transformed stock prices; see Figure 3 for illustration. Therefore, instead of using the number of new confirmed covid cases in our analysis, we use the square root of this variable as an explanatory variable. Also, both the stock prices of Alphabet Inc and Microsoft Corporation should be dependent on the popularity ratings (which are measured by the search indices provided by Google Trends) of their competitors, i.e Skype and Zoom. Naturally, these ratings are also included as explanatory variables.

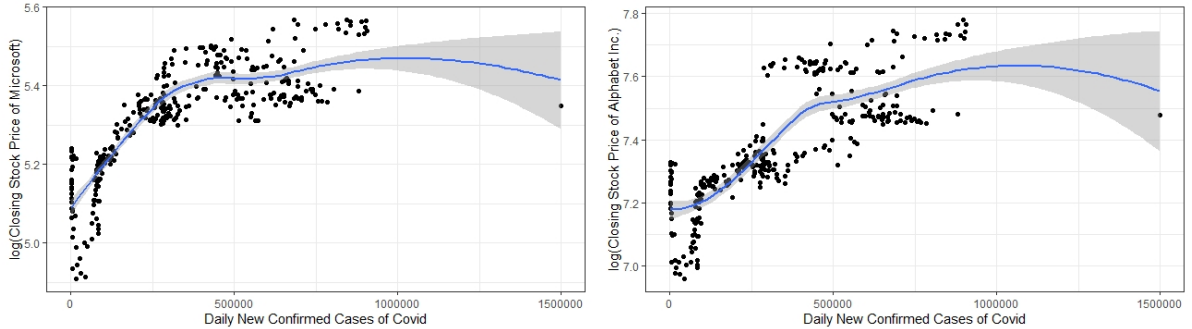


Figure 2: Plots of Daily closing stock prices of Alphabet Inc (GOOGL) and Microsoft Corporation (MSFT) vs daily new confirmed covid cases

Let us denote the day-level variables as

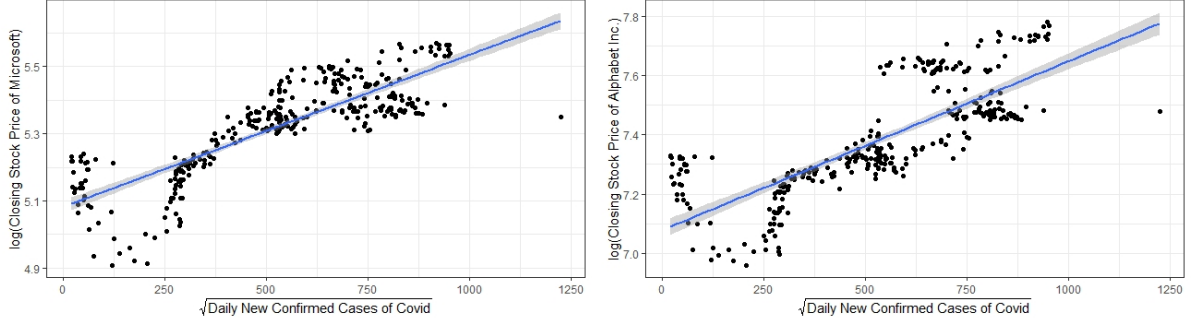


Figure 3: Plots of Daily closing stock prices of Alphabet Inc (GOOGL) and Microsoft Corporation (MSFT) vs square root of daily new confirmed covid cases

$$\begin{aligned}
SP_G &= \text{Log-transformed Stock Price of Alphabet Inc (Google)} \\
SP_M &= \text{Log-transformed Stock Price of Microsoft Corporation} \\
C_{new} &= \sqrt{\text{New confirmed covid cases}} \\
PR_G &= \text{Popularity Rating of Google} \\
PR_M &= \text{Popularity Rating of Microsoft} \\
PR_S &= \text{Popularity Rating of Skype} \\
PR_Z &= \text{Popularity rating of Zoom}
\end{aligned} \tag{1}$$

The popularity ratings are essentially the normalized scores of Google search of a search term as provided by Google Trends [8]. As discussed above, working with the segmented timespans and merging them allowed the rating to scale proportionately outside the usual range of 0-100, allowing much clear variations in the ratings. Figure 4 presents a heat-plot of the Pearson's product moment correlation coefficient between the above variables. Most of the variables turned out to be very much linearly correlated, hence linear regression is a possible choice to model the effect of the variables on the stock prices. The two stock prices are also strongly dependent on each other. This is supported by the scatterplot depicted in Figure 5.

Therefore we can see that there is a circular relationship between SP_G and SP_M . From this and earlier considerations, we start with the following two structural equations,

$$SP_G = \alpha_0 + \alpha_1 SP_M + \alpha_2 C_{new} + \alpha_3 PR_Z + \alpha_4 PR_S + \alpha_5 PR_G + \epsilon_1 \tag{2}$$

$$SP_M = \beta_0 + \beta_1 SP_G + \beta_2 C_{new} + \beta_3 PR_Z + \beta_4 PR_S + \beta_5 PR_M + \epsilon_2 \tag{3}$$

where ϵ_1 and ϵ_2 are zero-mean errors. These two equations represent a system of simultaneous equation, having 12 structural parameters, namely $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ and β_5 . We shall obtain reduced form equations from here. Substituting the value of SP_M from Eq. (3) in Eq. (2), we obtain

$$\begin{aligned}
SP_G &= [\alpha_0 + \alpha_1 \beta_0] + [\alpha_1 \beta_1] SP_G + [\alpha_2 + \alpha_1 \beta_2] C_{new} + [\alpha_3 + \alpha_1 \beta_3] PR_Z + \\
&\quad [\alpha_4 + \alpha_1 \beta_4] PR_S + \alpha_5 PR_G + [\alpha_1 \beta_5] PR_M + [\epsilon_1 + \alpha_1 \epsilon_2]
\end{aligned}$$

which yields,

$$\begin{aligned}
SP_G &= \left[\frac{\alpha_0 + \alpha_1 \beta_0}{1 - \alpha_1 \beta_1} \right] + \left[\frac{\alpha_1 \beta_5}{1 - \alpha_1 \beta_1} \right] PR_M + \left[\frac{\alpha_2 + \alpha_1 \beta_2}{1 - \alpha_1 \beta_1} \right] C_{new} + \\
&\quad \left[\frac{\alpha_3 + \alpha_1 \beta_3}{1 - \alpha_1 \beta_1} \right] PR_Z + \left[\frac{\alpha_4 + \alpha_1 \beta_4}{1 - \alpha_1 \beta_1} \right] PR_S + \left[\frac{\alpha_5}{1 - \alpha_1 \beta_1} \right] PR_G + \left[\frac{\epsilon_1 + \alpha_1 \epsilon_2}{1 - \alpha_1 \beta_1} \right] \\
&= G_0 + G_1 PR_M + G_2 C_{new} + G_3 PR_Z + G_4 PR_S + G_5 PR_G + \eta_1
\end{aligned}$$

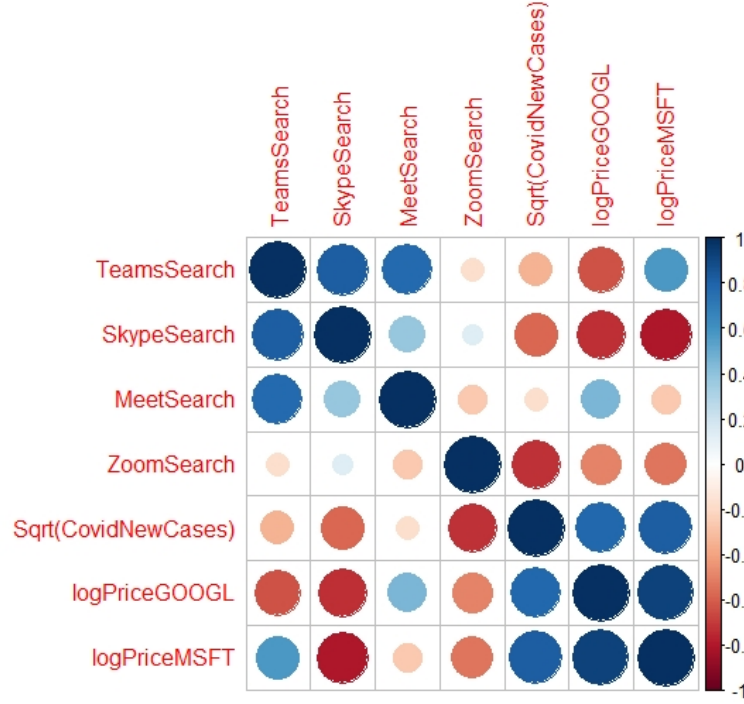


Figure 4: Correlations between the seven variables described in Eq. (1)

where G_i 's are the expressions that it is replacing (quantity inside the $(i + 1)$ -th bracket) for $0 \leq i \leq 5$ and η_1 is the error term inside the last parenthesis. Similarly, putting value of SP_G from Eq. (2) into Eq. (3), we obtain

$$\begin{aligned}
 SP_M &= \left[\frac{\beta_0 + \beta_1 \alpha_0}{1 - \alpha_1 \beta_1} \right] + \left[\frac{\beta_1 \alpha_5}{1 - \alpha_1 \beta_1} \right] PR_G + \left[\frac{\beta_2 + \beta_1 \alpha_2}{1 - \alpha_1 \beta_1} \right] C_{new} + \\
 &\quad \left[\frac{\beta_3 + \beta_1 \alpha_3}{1 - \alpha_1 \beta_1} \right] PR_Z + \left[\frac{\beta_4 + \beta_1 \alpha_4}{1 - \alpha_1 \beta_1} \right] PR_S + \left[\frac{\beta_5}{1 - \alpha_1 \beta_1} \right] PR_M + \left[\frac{\epsilon_2 + \beta_1 \epsilon_1}{1 - \alpha_1 \beta_1} \right] \\
 &= M_0 + M_1 PR_G + M_2 C_{new} + M_3 PR_Z + M_4 PR_S + M_5 PR_M + \eta_2
 \end{aligned}$$

Combining these two derivations together, we obtain **two reduced form equations** as

$$SP_G = G_0 + G_1 PR_M + G_2 C_{new} + G_3 PR_Z + G_4 PR_S + G_5 PR_G + \eta_1 \quad (4)$$

$$SP_M = M_0 + M_1 PR_G + M_2 C_{new} + M_3 PR_Z + M_4 PR_S + M_5 PR_M + \eta_2 \quad (5)$$

which now incorporates 12 reduced form parameters. Since there are 12 structural parameters and 12 reduced form parameters, the simultaneous equation system is *exactly identified* and we can get back the estimates of structural parameters from the estimates of reduced form parameters.

Towards this, we first express the reduced form parameters in terms of structural parameters

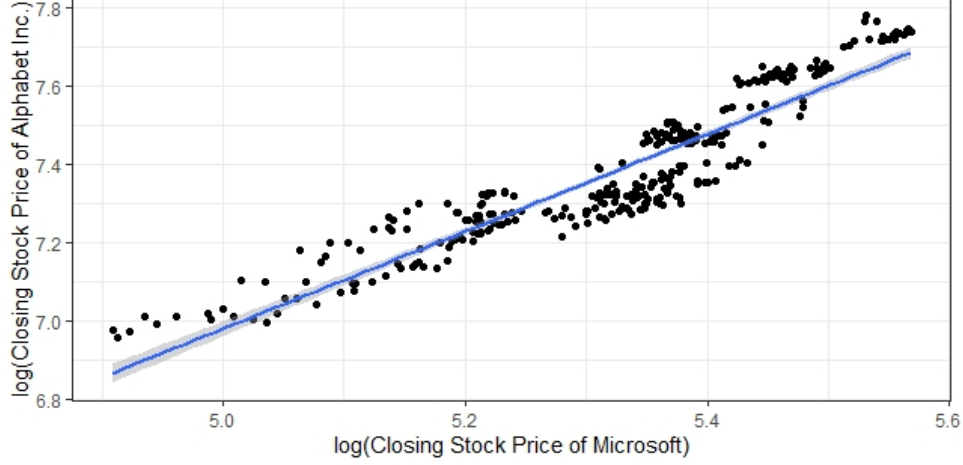


Figure 5: Scatterplot of Microsoft (MSFT) stock price vs Alphabet Inc (GOOGL) stock price

$$G_0 = \frac{\alpha_0 + \alpha_1 \beta_0}{1 - \alpha_1 \beta_1} \quad (\text{A1}) \quad M_0 = \frac{\beta_0 + \beta_1 \alpha_0}{1 - \alpha_1 \beta_1} \quad (\text{B1})$$

$$G_1 = \frac{\alpha_1 \beta_5}{1 - \alpha_1 \beta_1} \quad (\text{A2}) \quad M_1 = \frac{\beta_1 \alpha_5}{1 - \alpha_1 \beta_1} \quad (\text{B2})$$

$$G_2 = \frac{\alpha_2 + \alpha_1 \beta_2}{1 - \alpha_1 \beta_1} \quad (\text{A3}) \quad M_2 = \frac{\beta_2 + \beta_1 \alpha_2}{1 - \alpha_1 \beta_1} \quad (\text{B3})$$

$$G_3 = \frac{\alpha_3 + \alpha_1 \beta_3}{1 - \alpha_1 \beta_1} \quad (\text{A4}) \quad M_3 = \frac{\beta_3 + \beta_1 \alpha_3}{1 - \alpha_1 \beta_1} \quad (\text{B4})$$

$$G_4 = \frac{\alpha_4 + \alpha_1 \beta_4}{1 - \alpha_1 \beta_1} \quad (\text{A5}) \quad M_4 = \frac{\beta_4 + \beta_1 \alpha_4}{1 - \alpha_1 \beta_1} \quad (\text{B5})$$

$$G_5 = \frac{\alpha_5}{1 - \alpha_1 \beta_1} \quad (\text{A6}) \quad M_5 = \frac{\beta_5}{1 - \alpha_1 \beta_1} \quad (\text{B6})$$

Now, from Eq. (A2) and Eq. (B6), we obtain $\alpha_1 = \frac{G_1}{M_5}$. Similarly, from Eq. (B2) and Eq. (A6), we obtain $\beta_1 = \frac{M_1}{G_5}$. Putting the expression of α_1 in Eq. (A2), we get $\beta_5 = (1 - \alpha_1 \beta_1) M_5 = M_5 - \beta_1 G_1$. In the same way, we obtain $\alpha_5 = (1 - \alpha_1 \beta_1) G_5 = G_5 - \alpha_1 M_1$ putting the expression of β_1 in Eq. (B2). Also, we note that Eq. (A1) and Eq. (B1) can be expressed as

$$\alpha_0 + \alpha_1 \beta_0 = c_1 \quad \text{and} \quad \beta_0 + \beta_1 \alpha_0 = d_1$$

where $c_1 = (1 - \alpha_1 \beta_1) G_0$ and $d_1 = (1 - \alpha_1 \beta_1) M_0$. Solving these, we get

$$\alpha_0 = \frac{c_1 - \alpha_1 d_1}{1 - \alpha_1 \beta_1} \quad \text{and} \quad \beta_0 = \frac{d_1 - \beta_1 c_1}{1 - \alpha_1 \beta_1}$$

After some algebraic manipulations, these lead to

$$\alpha_0 = G_0 - \alpha_1 M_0 \quad \text{and} \quad \beta_0 = M_0 - \beta_1 G_0$$

Similarly, we can obtain expressions of (α_2, β_2) using Eq. (A3) and (B3); expressions of (α_3, β_3) using Eq. (A4) and (B4); expressions of (α_4, β_4) using Eq. (A5) and (B5).

Combining everything, the structural parameters can be written in terms of the reduced form parameters using a set of 12 equations,

$$\begin{aligned}
\alpha_1 &= G_1/M_5, \quad \beta_1 = M_1/G_5, \\
\alpha_0 &= G_0 - \alpha_1 M_0, \quad \beta_0 = M_0 - \beta_1 G_0, \\
\alpha_2 &= G_2 - \alpha_1 M_2, \quad \beta_2 = M_2 - \beta_1 G_2, \\
\alpha_3 &= G_3 - \alpha_1 M_3, \quad \beta_3 = M_3 - \beta_1 G_3, \\
\alpha_4 &= G_4 - \alpha_1 M_4, \quad \beta_4 = M_4 - \beta_1 G_4, \\
\alpha_5 &= G_5 - \alpha_1 M_1, \quad \beta_5 = M_5 - \beta_1 G_1.
\end{aligned} \tag{6}$$

This whole exercise suggests that we can solve the simultaneous equation system in Eq. (2) and (3) by performing the solving the linear regression problem in Eq. (4) and (5) and estimating the reduced form parameters. Then, an application of the formulas shown in Eq. (6) can be used to recovered the structural parameters. However, before proceeding with the simultaneous linear regression, we notice that the variables in concern are time series variables. In other words, the observed samples are $SP_{G,t}, SP_{M,t}$ where t is the time index, hence the errors are actually η_{1t}, η_{2t} (The time index has been dropped in Eq. (4) and Eq. (5)). Therefore, a temporal dependence is expected to be present within each variable and the samples may be correlated, i.e. η_{1t} and $\eta_{1(t-1)}$ may be correlated. Assuming that the errors are stationary, we consider an autoregressive model of order 1 with $\eta_{it} = \phi_i \eta_{i(t-1)} + w_{it}$ for $i = 1, 2$, where w_{it} is a white noise process with mean 0 and variance $\sigma_{w,i}^2$. In this case,

$$\text{Cov}(\eta_{it}, \eta_{i(t-s)}) = \phi_i^s \sigma_{w,i}^2 / (1 - \phi_i^2), \quad \text{for } i = 1, 2$$

where $\phi_i, \sigma_{w,i}^2$ are estimable parameters. Clearly, an ordinary least squares procedure would be inadequate for this, and hence a generalized least squares procedure can be used to estimate these parameters [5]. In particular, if we consider Σ_i as the variance covariance matrix of $(\eta_{i1}, \eta_{i2}, \dots)$, then pre-multiplying each variable in Eq. (4) and (5) by $\Sigma_i^{-1/2}$ turns the system into usual linear regression model with independent samples.

However, inclusion of number of covid cases as an explanatory variable also creates another problem. Many scientists have warned time and again that covid cases and deaths are being undercounted throughout the world [3]. Countries like France, Italy, the United States of America, Iran and Spain have extremely high numbers of undetected and underreported cases [7]. Thus, the variable C_{new} involves measurement error and hence is correlated with the error terms η_{1t}, η_{2t} . This induces an endogeneity problem. Consequently, the parameters present in the two reduced form equations (4) and (5) can not be estimated unbiasedly and consistently. To overcome this problem, we introduce an instrumental variable namely covid search, denoted by C_s . This is a similar normalized popularity rating of the search term ‘‘covid’’ or related topics which is collected from Google trends dataset along with other search terms. Since square root transformation was applied to form C_{new} , C_s also denotes the square root transformation of the normalized popularity rating. We find that this has a correlation 0.9605317 with C_{new} . On the other hand, intuitively C_s should not be directly related with the two stock prices. Therefore, it can be used as an instrumental variable to estimate the parameters in Eq. (4) and (5). Using C_s as the instrumental variable, we may use IV regression method to estimate these reduced form parameters.

Although the problem of endogeneity through measurement error in number of covid cases worldwide is suspected, one may use Hausman test of endogeneity to confirm the suspicion. Let \hat{G}_{IV} and \hat{G}_{OLS} be the IV and the OLS estimates obtained. Note that after transforming the data into independent samples through pre-multiplying estimated $\hat{\Sigma}_i$, the obtained OLS estimate \hat{G}_{OLS} is actually a GLS estimate. To test the endogeneity hypothesis

$$H_{0G} : C_{new} \text{ is not endogeneous in Eq. (4)}$$

$$H_{1G} : C_{new} \text{ is endogeneous in Eq. (5)}$$

we can use the Hausman test statistic

$$n \left(\hat{G}_{IV} - \hat{G}_{OLS} \right)^\top \left(\text{Var}(\hat{G}_{IV}) - \text{Var}(\hat{G}_{OLS}) \right)^{-1} \left(\hat{G}_{IV} - \hat{G}_{OLS} \right)$$

which under the null hypothesis H_{0G} , follows a chi-squared distribution χ_p^2 with degrees of freedom p equal to the number of parameters. In our case, $p = 6$. We shall reject the null hypothesis and conclude that there is an endogeneity problem if the value of the observed statistic exceeds the critical value obtained from the $100(1 - \alpha)$ -th quantile of χ_6^2 distribution at α level of significance. In a very similar manner, test for the presence of endogeneity in Eq. (5) can also be performed.

4 Results and Discussions

Applying GLS along with IV estimation technique for estimate the parameters, we obtain

$$\begin{aligned} \hat{G}_{OLS} &= (7.30212, -3.67621 \times 10^{-4}, 1.22134 \times 10^{-4}, 8.74257 \times 10^{-4}, -8.40212 \times 10^{-4}, \\ &\quad -11.65843 \times 10^{-4}) \text{ and;} \\ \hat{G}_{IV} &= (7.29918, -3.82965 \times 10^{-4}, 1.34618 \times 10^{-4}, -8.49866 \times 10^{-4}, \\ &\quad -8.19565 \times 10^{-4}, 11.45836 \times 10^{-4}) \end{aligned}$$

We observe that in \hat{G}_{IV} coefficient of google meet search is positive. This tells that the stock price of Alphabet Inc increases as google meet search increases. It is also noted that the estimated coefficient of zoom search is negative in \hat{G}_{IV} . This can be interpreted as evidence in favour of the fact that increased popularity of rival video conferencing service results into decrease in stock price. The signs of coefficients estimated using OLS are difficult to interpret. The value of the Hausman test statistic is 40.60228 with a p-value = 3.467842×10^{-7} . Hence, we reject the null hypothesis that C_{new} is not endogenous at 0.01 level of significance. This confirms our earlier assertion that C_{new} is endogeneous in the regression equation (4). Finally, it yields a multiple R-squared value equal to 0.869, which suggests a very good fit of the linear regression.

Similarly, for the linear regression given in Eq. (5), the estimates turned out to be

$$\begin{aligned} \hat{M}_{OLS} &= (5.22482, -8.58101 \times 10^{-5}, 1.68968 \times 10^{-4}, 4.05142 \times 10^{-4}, -8.77877 \times 10^{-4}, \\ &\quad -1.91413 \times 10^{-4}) \text{ and;} \\ \hat{M}_{IV} &= (5.2204, -5.56586 \times 10^{-5}, 1.87783 \times 10^{-4}, -3.68382 \times 10^{-4}, -8.46759 \times 10^{-4}, \\ &\quad 2.14539 \times 10^{-4}) \end{aligned}$$

The signs of coefficients in \hat{M}_{IV} are similar to those of \hat{G}_{IV} . For example, the estimated coefficient of zoom search is negative in \hat{M}_{IV} too. We also observe that in \hat{M}_{IV} , coefficient of microsoft teams search is positive. This can be interpreted as evidence in favour of the fact that increased popularity of microsoft teams generally results into increased stock price of Microsoft corporation. Again, the signs of components of \hat{M}_{OLS} are difficult to interpret in this manner. The observed value of the Hausman test statistic turns out to be 99.17337 with a p-value 3.732403×10^{-19} . Therefore, we reject the null hypothesis that C_{new} is not endogeneous again at 0.01 level of significance and confirm that the data shows sufficient evidence to suspect that C_{new} is endogeneous in the linear regression equation (5) as well. Finally, it yields a multiple R-squared value equal to 0.742, which suggests a very good fit of the linear regression considered here.

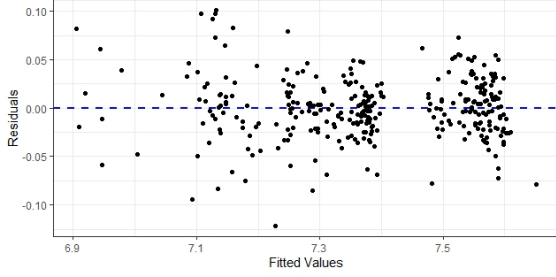
We have seen in Section 3 how we can get back the estimates of structural parameters from the estimates of the reduced form parameters. Using the relationships derived there, we obtain

the estimates of $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ from \hat{G}_{IV} and \hat{M}_{IV} . The resulting estimates, if put back into Eq. (2) and (3) yields the simultaneous regression equations as

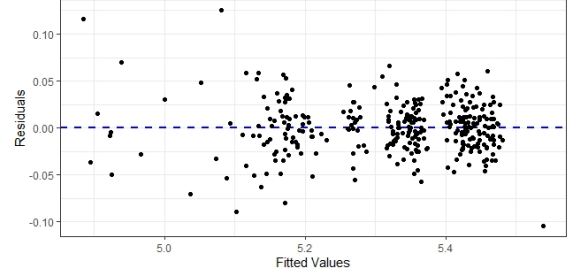
$$SP_G = 16.6178 - 1.785SP_M + 0.00047C_{new} - 0.0015PR_Z - 0.0023PR_S + 0.00105PR_G \quad (7)$$

and

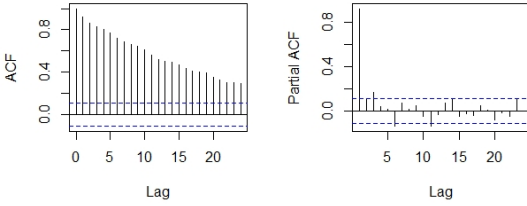
$$SP_M = 5.57495 - 0.0485SP_G + 0.000194C_{new} - 0.00041PR_Z - 0.00089PR_S + 0.000195PR_M \quad (8)$$



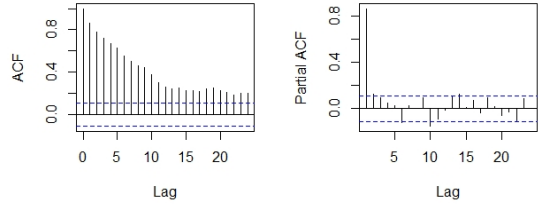
(a) Fitted values vs Corrected Residuals for Regression Eq. (4)



(b) Fitted values vs Corrected Residuals for Regression Eq. (5)



(c) ACF and PACF plots of the difference between true stock price of Google and Prediction from linear regression model



(d) ACF and PACF plots of the difference between true stock price of Microsoft and Prediction from linear regression model

Figure 6: Residual Diagnostics

Finally, to test whether the assumptions of linearity and autoregressive residuals were correct, we can simply plot the corrected residuals and the fitted values from the transformed regression models with independent samples (The one obtained by pre-multiplying with $\Sigma_i^{-1/2}$). As shown in Figures 6a and 6b, both of the residual plots show a random distribution of residuals to both sides of the x-axis. Since no visible pattern is found in the residuals, it seems the linear regression model was an adequate choice. Also, Figures 6c and 6d plots the ACF and PACF plots of the errors in prediction of the final stock prices, i.e. the difference between the true stock prices and the prediction from the two linear regression models (exponential of the fitted values from Eq. (4) and (5)). These plots support our assumption that η_{it} 's have an autoregressive model of order 1, as indicated in Section 3.

5 Conclusion

From the simultaneous equations obtained in Eq. (7) and Eq. (8), we observe the following;

1. Google and Microsoft stock price are contrasting in nature. Each one has a decreasing effect on the other companies stock prices.

2. In both equations, the number of new covid cases has positive coefficients. In other words, both companies' stock prices boomed during the covid19 outbreak.
3. Zoom and Skype search have a negative impact. Skype although owned by Microsoft is not boosted for sale during the covid period.
4. Google meet search has positive feedback for Alphabet Inc. stock price, and similar for Microsoft teams search and Microsoft Corporation.

The results of our study also substantiate the claim that countries have grossly understated covid caseloads. Underreporting the caseload has far-reaching consequences for how a country responds. Governments could fail to direct funding and other resources to an area in need if the virus seems to have a relatively low threat there. Moreover, accurate counts provide scientists with crucial information about the virus. On the contrary, erroneously low tallies—due to a lack of testing or, in some cases, pressure or policies by government officials—can affect individual behavior [3]. If people feel there is a low risk of contracting COVID-19 in their community, they could choose to interact more with others, resulting in an increased caseload.

Finally, in answer to the objective of this study, we find that there is a significant effect of covid on these two stock prices. Although most of the economy suffered heavily during the start of the pandemic, these two companies could produce their respective online learning and video conferencing tools meeting their demands, and in turn, earning a significant boost in their stock prices.

References

- [1] Alexander Bick, Adam Blandin, and Karel Mertens. Work from home after the covid-19 outbreak. 2020.
- [2] Tinggui Chen, Lijuan Peng, Bailu Jing, Chenyue Wu, Jianjun Yang, and Guodong Cong. The impact of the covid-19 pandemic on user experience with online education platforms in china. *Sustainability*, 12(18):7329, 2020.
- [3] Claire Felter. By how much are countries underreporting covid-19 cases and deaths? Accessed from <https://www.cfr.org/in-brief/how-much-are-countries-underreporting-covid-19-cases-and-deaths>.
- [4] Yahoo Finance. Daily historical stock prices. Accessed from <https://in.finance.yahoo.com/> on 25th May, 2021.
- [5] John Fox. Time-series regression and generalized least squares. *An R S-PLUS Companion to Appl. Regres. Thousand Oaks, CA*, pages 1–8, 2002.
- [6] Biplab Kumar Guru and Amarendra Das. Covid-19 and uncertainty spillovers in indian stock market. *MethodsX*, 8:101199, 2021.
- [7] H Lau, T Khosrawipour, P Kocbach, J Ichii, H Bania, and V Khosrawipour. Evaluating the massive underreporting and undertesting of covid-19 cases in multiple global epicenters. *Pulmonology*, 6 2020.
- [8] Google LLC. Google trends. Accessed from <http://trends.google.com/trends> on 25th May, 2021, 2021.
- [9] NASDAQ. Daily stock market overview, data updates, reports & news. <https://www.nasdaq.com/>.

- [10] Afees A Salisu and Xuan Vinh Vo. Predicting stock returns in the presence of covid-19 pandemic: The role of health news. *International Review of Financial Analysis*, 71:101546, 2020.
- [11] Johns Hopkins University. Covid-19 data repository by the center for systems science and engineering (csse).
- [12] World Health Organization (WHO). What is a pandemic? Published on https://www.who.int/csr/disease/swineflu/frequently_asked_questions/pandemic/en/ on 24 February, 2010.
- [13] Wikipedia. 2020 stock market crash. Accessed from https://en.wikipedia.org/wiki/2020_stock_market_crash on 1st June, 2021.
- [14] Wikipedia. Severe acute respiratory syndrome coronavirus 2. Accessed from https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome_coronavirus_2 on 1st June, 2021.
- [15] Wikipedia. Stock market prediction. Accessed from https://en.wikipedia.org/wiki/Stock_market_prediction on 1st June, 2021.