LUISS Università Guido Carli

Department of Business and Management
Bachelor degree in Management and Computer Science

Using Brand Reputation as predictor for brand value and stock price

Relator: Francisco Villarroel Ordenes

Student: Emanuele Carosi

Table of Contents

Acknowledgments

In the first place, I would like to thank my relator, Francisco Villaroel Ordenes, for the opportunity to develop my thesis in the field of Marketing and Data Analytics, and for all the support showed during the its realisation.

I would also like to express my gratitude to my first professor of Python, Luca Alulli, for his incredible kindness during these years.

Finally, I would like to thank my friends, Matteo Antonazzo for constantly believing and challenging me during the whole three years of bachelor's degree, and Lorenzo Giarè for his immense moral support during the development of this thesis.

Abstract

Through the creation of the Brand Reputation Framework, Rust et al. (2021) highlighted the existence of a relation between brand reputation and brand value, and managed to develop a tracker for the brand reputation based on the stakeholders' sentiment on social media. Using a sample of comments addressed to Facebook and Google on Twitter, I replicated the tracker and developed a regression model able to highlight a relation between the brand reputation and the stock price. Although the range of scope of my study is narrower than the original, all findings are consistent with the ones of Rust et al. and corroborate therefore the reliability of their framework.

Introduction

The fact that nowadays social media are unlimited sources of valuable information for companies is not a mystery. An entire new business model rose with the birth of social media and now more than ever the ownership of information can be so profitable. The phenomenon that I will study in this paper deals with the exploitation of social media in order to collect valuable information about specific brands. Specifically, I will use the stakeholders' underlying sentiment present on social media to create a tracker for the value of the brand; furthermore, I will use the values obtained by the tracker to try to predict corresponding stock price.

I will base the assumptions of my work on the article “Real-Time Brand Reputation Tracking Using Social Media” published in 2021 by Rust et al., trying to replicate its findings. In the article the authors created a Brand Value tracker from the stakeholders’ comments towards the brand on social media, able to monitor the brand performances through the evaluation of the “Brand Reputation”, a construct made of several components which in the article is the unit of measure for brand value.

The managerial implications of the original article are multiple, but the most important derive from the characteristics of the aforementioned tracker. Its division in actionable, individual factors, in fact, is the key for managers to evaluate and intervene directly on the sensitive areas of the company. Moreover, the tracker can also be used to monitor the competition, allowing the company to have control over the dynamics of the market.

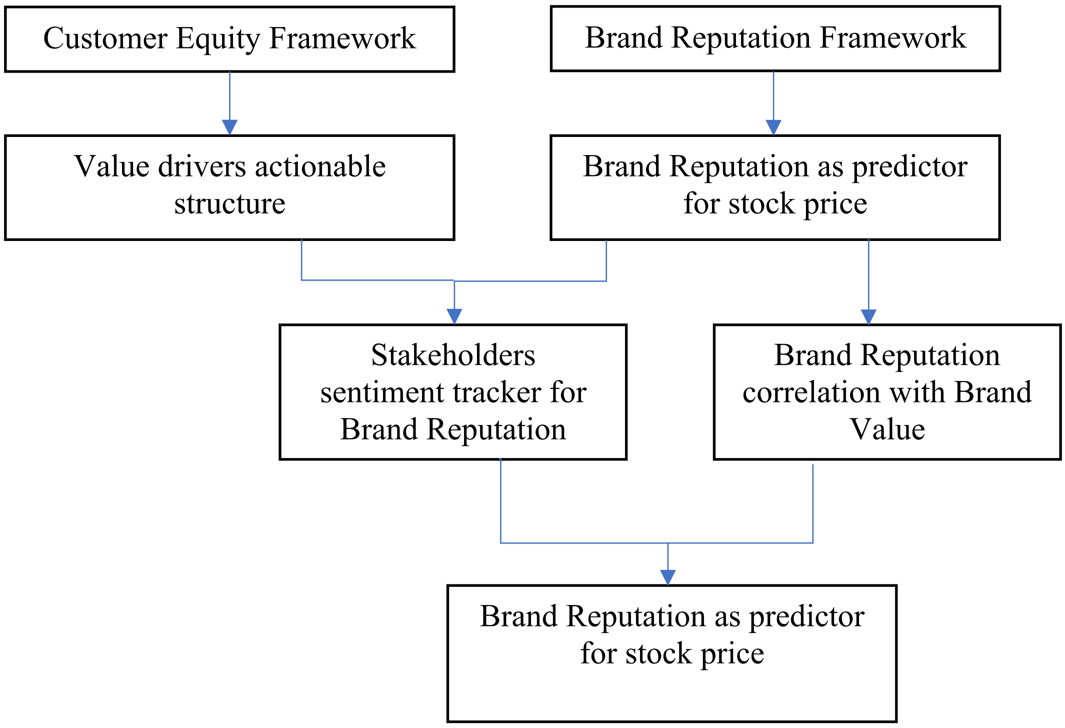
Another key characteristic of the tracker that further increases its managerial relevance is to research in the absence of similar metrics for the evaluation of brand value. So far, the only measure to estimate the value of a brand has been through the evaluation of survey-based techniques, which are usually available only once a year and that cannot describe efficiently the underlying reasons of such results, lacking in seizing the same granularity of the model of Rust et al. (2021).

What I plan to do in my paper is to investigate whether it is possible or not to replicate the findings of the original article, disclosing in addition if even a smaller dataset with respect to the extensive one used by Rust et al., can achieve the same level of granularity in the results.

The main purposes of this thesis, therefore, will be the following: first, I will verify if it is possible to track the brand value and highlight its main events during the year, using only the social media stakeholders' sentiment towards the brand; secondly, I will investigate the possibility of using the Brand Reputation calculated in the former analysis for the development of a regression model able to predict the stock price.

The contributions that I expect to make are to validate the use of the tracker, corroborating its managerial relevance, and assessing the existence of a correlation between the Brand Reputation and the stock price through a specific back testing process based on observed data.

Below it is possible to visualise the conceptual framework of my paper (Lindgreen et al. 2020).



Literature Review

Brand Reputation Framework

Starting from the foundation that the reputation of a brand is a valuable asset influencing the financial performances of the company, the Brand Reputation Framework presented by the original authors aims to grasp all the aspects that may influence the reputation of a brand, and therefore of its value. Therefore, the attention will revolve on the whole stakeholders rather than a simple customer centric approach. Furthermore, the article focuses specifically on the possibility for managers to actively influence the brand reputation through marketing events, corroborating the brand reputation as an actionable lever for the financial success of the firm (Sabate and Puente 2003).

Let us now analyse the components of the Brand Reputation Framework according to the original authors, which would be: Corporate Reputation, Brand Equity, and Customer Equity. Corporate reputation represents how the brand is perceived by stakeholders, accounting for the brand's history and its foreseeable future (Lange, Lee, and Dai 2011). Brand reputation is the same concept applied when more firms of the same brand adopt a branded-house strategy (e.g., Sony). As for Brand Equity, it ensembles the overall impression of stakeholders towards the brand and takes into consideration aspects such as knowledge (Keller 1993), and emotions, which may also alter the objective perceived value of the service/good in the eyes of the customer (Rust et al. 2004). Among the components of the Brand Reputation, Brand Equity is the component that relates the most with the stakeholders' sentiment. For the Customer Equity concept the authors relied on the interpretation present in "What drives Customer Equity" (Lemon et al. 2001), which defined Customer Equity as "the discounted lifetime values summed over all of the firm's current and potential customers". Customer Equity is particularly important for the Brand Reputation Framework, because both the brand reputation and the customer equity are influenced by the same actionable drivers, which directly affect the financial performances of the company.

Customer Equity Framework

As mentioned in the Introduction, the work of Rust et al. relies on the Customer Equity framework to justify the actionable drivers for the brand reputation, and for extension for the brand value as well. More in detail, the Customer Equity Framework is composed of three components that drive the financial performance of the company (Lemon, Rust, and Zeithaml 2001; Rust, Lemon, and Zeithaml 2004), which are: Value Equity, Brand Equity, and Relationship Equity. Each driver seizes a specific characteristic of the brand, which in turn, can be influenced by different levers. Value Equity is defined as the customer's objective assessment of the utility of a brand, based on the trade-off between what is given up and what is received. It can be influenced by three key levers: quality, price and convenience. Brand Equity is interpreted likewise in both this and Brand Reputation framework, representing therefore the customer's subjective sentiment towards the brand. Its key actionable levers are brand awareness, attitude toward the brand, and corporate ethics. Finally, Relationship Equity is the driver related to the development of the customers relationship as a means to keep choosing their brand over the others. Levers that the firm can develop to exploit this driver are loyalty programs, special recognition and treatment, affinity programs, community-building programs, and knowledge-building programs (Lemon et al. 2001).

Brand Reputation Tracker

Starting from the levers of each driver of the Customer Equity Framework, Rust et al. individuated eleven individual and actionable subcomponents, called subdrivers, each of which seizes a different aspect of the driver of belonging. Regarding Value Equity, the authors individuated three subdrivers: Price, Service Quality and Goods Quality. For Brand Equity and Relationship equity it was possible to define four sub drivers each: Cool, Exciting, Innovative, Social responsibility for the former; Community, Friendly, Personal relationships and Trustworthy for the latter. The structure just described will be crucial for the tracker, which will make the eleven subdrivers the predictors for the Brand Reputation.

Research questions

In this section I will highlight two important aspects of this paper explaining the research questions to which this study wants to answer, and the rationale behind their development. The first research question, and probably the most subtle, aims to demonstrate, through a rational empirical analysis, such an intuitive correlation as it is that between a brand's reputation and its financial performances. The second research question wants to clarify if the tracker is reliable enough to provide meaningful results using a smaller dataset with the same integrity as the original article. The rationale behind this research question is that if the Framework ideated by the authors is solid, it should be able to give consistent findings also on a smaller scale. Finally, under the assumption that the relationship between brand reputation and its financial performance should be able to be reflected, the third research question investigates the possibility of showing a relationship between brand reputation and the company's stock price through the framework devised by the authors.

Method

Data Description

As with the original article, I analysed the time frame covering the 52 weeks of 2018. For the selection of the brands on the other hand, it would have not been feasible taking into account all the 100 brands analysed by Rust et al., so I decided to focus only on two sample brands: Facebook and Google. For their selection I took inspiration from the authors' paper, in which they used these two brands as examples to highlight the performances of the tracker. The reasons to highlight specifically Facebook and Google are several in fact: in the first place both brands are among the most successful in the world, implying an important engagement with their stakeholders, characteristic that is crucial to retrieve a proper dataset for the analyses of my study; secondly, they are both 'service' brands operating in the same sectors, allowing for a fair comparison of their performances; lastly, in 2018 they both had events with an important social impact on their stakeholders, which very likely altered the brands reputation. The presence of these main events is very important, because it will permit to evaluate if the tracker is able to grasp them. More specifically, Facebook was affected by two major events, the first took place in March 2018, when it was disclosed to the public the Cambridge Analytica scandal, and the second in September 2018, when the accounts of more than 50 million users were hijacked. For Google, on the other hand, only one major event occurred, namely the update of its major algorithms on its 20th anniversary in September.

As already mentioned, the analysis focuses on two separate Brands for two different datasets, composed by approximately 90,000 comments each. All comments are entirely collected on Twitter through the 'Advanced Research' feature that the social media provides. Their selection has been tailored to keep the integrity of the model through a specific framework of data collection, based on: 'English' as language, 2018 as time frame, only comments addressed to the specific brand (@Facebook and @Google), and only tweets without nested comments or links. It is important to highlight that, differently from the original article, I take into consideration only Twitter as a source of data. This choice has been mostly forced, since Facebook and Instagram (the other two data sources of the article of Rust et al.) do not allow users to narrow down the research of contents, denying any possibility to download consistent and specific datasets.

Measurement Development

Through the implementation of the tracker, it has been possible to calculate the polarities towards the 11 subdrivers of each comment in the datasets. Specifically, it measured the performances of each brand gathering the findings in brand weekly observations, where for observation is the number of comments with positive or negative polarity, towards a specific subdriver. For both Facebook and Google, in addition to the absolute number of positive and negative observations, the scores calculated for each subdriver are the delta and the ratio between negative and positive observations. It is possible to aggregate these scores, in turn, to calculate the three drivers (Value Driver, Brand driver and Relationship Driver) and finally the Brand Reputation. The value to which we commonly refer to as Brand Reputation is the sum of drivers' deltas. However, it is important to underline that also the Brand Reputation ratio is a crucial score, since it represents the sentiment towards the brand of a specific week.

To verify the capability of the tracker to seize the important events and to properly visualise the weekly brand performances, are present eight graphs showing three main drivers and the Brand Reputation. The first four graphs show the weekly deltas for both drivers and brand reputation. The other four, instead, show how the ratios change during time, highlighting the variation of sentiment towards the brands.

Finally, to summarise the overall annual performances of each brand, the respective averages and standard deviations of the 11 subdrivers, the 3 drivers and the Brand Reputation are calculated.

Modelling

This section will describe the collection of methods used in my analysis so as to predict the stock price from the Brand Reputation estimated by the tracker.

Before proceeding, it is important to give a definition of stock price in order to understand why it is so valuable and difficult to predict. The stock price of a company, also called share price when referring to one stock unit, is defined as a security representing the right of ownership over a fraction of a corporation, entitling the owner of the stock to a part of the same assets and profit directly proportional to the fraction of stock owned. Stock price is therefore a relative and proportional value for the company's worth, implying that a variation in the performances of the firm would result in a variation in the stock price as well, and vice versa (Hayes 2022).

Introduction to linear regression models

In general, a regression model describes a cause-effect relationship where the independent variables, or predictors, explain a dependent variable, also called response variable. It is rather important to underline that the regression alone cannot evaluate the relationship of the causation between the variables; in order to evaluate it and create a well-designed regression model it is necessary to investigate the theory and the relations between the variables of the specific case.

Graphically, what a linear regression model does is evaluate the straight line that best represents the observed data. To do that, it is chosen the regression line which minimises the sum of the squares of the differences between the observed dependent variable (values of the variable being observed) in the given dataset and those predicted by the linear function of the independent variable, also called Ordinary Least Squares method. The function of a simple linear regression would be given by:

$$y_t = \beta_0 + \beta_1 x_t$$

where y_t is the response variable at time t, t is the week, β_0 is the intercept with the y-axis, β_1 is the slope of the regression line calculate trough the Ordinary Least Squares (OLS) method and x_t is the independent variable at time t. Respect to the two models implemented for Facebook and Google y_t would be the stock price and x_t the Brand Reputation.

OLS assumptions and Panel Data Analysis introduction

However, reducing the prediction of stock price to such a basic model would be too simplistic. In order to design a proper regression model, in fact, it is necessary to investigate two important aspects: the type of correlation between the dependent and independent variable(s), and whether or not the data meet the assumptions for an OLS regression to work properly.

As for the first aspect, the models that would seize the characteristics of the variables of this study, are the ones relying on Panel Data Analysis. In fact, the data frame of the observed values can be described as a combination of cross-sectional (i= 1, 2) and time-series data(t = 1, ..., 52), or in other terms, a Panel Data framework.

As for the assumptions an OLS regression model should meet in order to work properly, they are the following: Linearity, Exogeneity, Homoscedasticity, Non-autocorrelation, No Multicollinearity. Linearity means that the response variable is a linear combination of the parameters (regression coefficients) and the predictor variables. Exogeneity instead, is verified when the predictors are uncorrelated with the error. More in detail, this assumption is rebutted whenever the model presents endogeneity, which can be verified for three different reasons. The first cause is the omission of one or more independent variables in the regression model. Let us call 'z' an independent variable omitted in a fictitious model, the expectation of the model is based on the response of both x and z, but since z is not present it is not observable either, implying that the regression's error will not be given only by itself, but also by z. If z is correlated with x the error will therefore be correlated with x leading to an endogenous model. The second case in which the model can manifest endogeneity is due to simultaneous relation between the dependent and independent variables. If x is related to y and y is related to x and it is not possible to determine which implies the other it follows that the errors will be correlated with the predictors, leading again to an endogeneity. Lastly, the third case verifies when it is not possible to discern the real observation by the error in measuring it. Since the error is correlated with the predictor, once again the model is endogenous. Let us now continue to explain the other conditions that an OLS regression model should meet: Homoscedasticity means that the error term has uniform variance; for what regards non autocorrelation instead, this assumption is met when the variable at time t have no correlation with the observations at time t-1. Lastly, the model should not present multicollinearity, which means that none of the predictors might be expressed as a linear combination of the others (Brugger 2021; Landstrom 2020).

Panel Data regression models

Let us now present the characteristics of the Panel Data regression models. As already anticipated, one of the main risks for a regression model is to be endogenous. Since the models I will implement foresee only the Brand Reputation as independent variable, the hazard caused by the omission of independent variables could really be a crucial problem. However, if the effect caused by the omitted variable(s) is either time constant or cross-sectionally constant, the use of panel data models can cancel the effect(s) and evaluate a consistent regression estimate. If either one between time-invariant and cross-sectional unobserved effect exists, the panel regression model is called a one-way model; on the contrary if both effects are jointly present, it is referred to as a two-way model. Depending on the nature of the unobserved effect and its relationship with the other independent variables it is possible to implement three different types of Panel Data Regression models: Pooled Ordinary Least Squares regression is the most indicated in case of unobserved effect(s); Fixed effect (FE) is used when the unobserved effect(s) is(are) correlated with the independent variables; and Random effect (RE) for the models in which the unobserved effect is uncorrelated with the independent variables (Landstrom 2020).

The pooled OLS regression assumes that neither of the two aforementioned unobserved effects exist and it performs a standard OLS regression to all data across all cross sections and all periods. It can be described by the following function:

$$y_{it} = x_{it}\beta + \epsilon_{it}$$

where i is the index of the entity, which in our case can be Facebook or Google; t is the index for time, or weeks in our model; x_{it} is the row of the design matrix describing the observation of cross section i and time t; β is the estimated vector of the coefficients; and ϵ_{it} is the idiosyncratic error (Sheppard 2022).

The Fixed Effect (FE) model determines the individual effects of unobserved, independent variables as fixed over time, with the possibility to account both for time-invariant unobserved effect t and the cross-sectional invariant unobserved effect i. To understand how this model works, it basically evaluates the fixed effect for the unobserved variable, and it applies it to the prediction of every observation. The function of a Fixed Effect regression model accounting for time invariant effect and cross-sectional invariant unobserved effects is the result of the prediction of an observation minus the fixed effects affecting the model. The formula is given by:

$$y_{it} - \bar{y}_i = (x_{it} - x_i)\beta + (\epsilon_{it} - \bar{\epsilon}_i)$$

where the final prediction, $y_{it} - \bar{y}_i$, is given by the estimation of the value minus the fixed effect; $(x_{it} - x_i)$ would be the predictor adjusted for the fixed effect; β is the estimated vector of the coefficients; and $(\epsilon_{it} - \bar{\epsilon}_i)$ would be the idiosyncratic error adjusted for the fixed effect (Sheppard 2022).

The FE model is also possible do be expressed as a Least Squares Dummy Variable (LSDV) estimator with the following function:

$$y_{it} = x_{it}\beta + d_{1,it}\gamma_1 + d_{2,it}\gamma_2$$

where $d_{1,it}\gamma_1$ and $d_{2,it}\gamma_2$ are the dummy variables for the first and second effect.

As opposed to the Fixed Effect, the Random Effect model (RE) assumes that the unobserved effect is unrelated to the independent variables within a cross section, resulting in a model described by the following function:

$$y_{it} - \hat{\theta}_i\bar{y}_i = (1 - \hat{\theta}_i)\alpha_i + (x_{it} - \hat{\theta}_i\bar{x}_i)\beta + (\epsilon_{it} - \hat{\theta}_i\bar{\epsilon}_i)$$

where θ_i is a function of the variance of $\in_{it}$, the variance of α_i and the number of observations for entity i (Sheppard 2022).

How to choose the proper Panel Data regression model

In order to evaluate which of the three Panel Data Regression models would be the best to predict the stock price I had to further research the characteristics of the variables involved in the model. Among the aforementioned assumptions we will focus specifically on Homoscedasticity, Non autocorrelation and Exogeneity.

To verify homoscedasticity I provided three different methods: the first was through the graph tracking the predicted values (x-axis) vs. residuals (y-axis), in which a high dispersion of the observations would have rebutted the homoscedasticity assumption; the other two methods were respectively, the White-test and the Breusch-Pagan-test, which in case of an outcome of the p values smaller than .05 reject homoscedasticity assumption.

To assess if the model showed or not autocorrelation I also implemented a visual proof and a statistical one, showing the autocorrelation coefficients of the Brand Reputation, and implementing the Durbin-Watson-Test. This test returned an output between 0 and 4, where 2 means no autocorrelation, 0 extremely positive autocorrelation, and 4 extremely negative autocorrelation.

Exogeneity will not be proven statistically because all the assumptions made so far rebut its presence in the model. The main reason is that since the model foresees only Brand Reputation as predictor for the Stock price, the endogeneity caused by omitted variables seems to be integral part of the model without further investigation.

If all these assumptions are violated, FE and RE may be preferable to pooled OLS. More specifically, to decide whether to implement FE or RE, if the assumption of exogeneity is confirmed, RE should be preferred; on the other hand, if exogeneity is rebutted in favour of endogeneity, then FE would be more indicated.

Finally, to evaluate the performance of each model will be calculated the coefficient of determination (R^2), which is the universal recognised statistical measure for accuracy, describing how close the real values are with respect to the fitted regression line; it can range from 0 to 1, where one indicates a perfect prediction of the data.

Results

The structure of this section will follow the one of the Method, so I will dedicate the first part to talk about the tracker and its findings, while in the second one I will go through the results of the regression models predicting the stock price from the Brand Reputation of the two firms.

Tracker Findings

As one may expect, since the brands analysed are among the most successful in the entire world, the ratios of the subdrivers, drivers and Brand Reputation are all strongly positive. With an average of 5.36 for the Value Driver, 12.37 for the Brand Driver and 18.01 for the Relationship Driver, Google Reputation totalised 7.77 as final score. The results highlight that, although the Value Driver is very positive, the real strength of Google seems to be enclosed in the ability of the other two drivers to develop exceptional customer relationships and a very appreciated environment for its customers. The success of these two drivers is testified by the peaks in three out of the four subdrivers of the Relationship Driver, and for the Brand Driver there is the peak of the 'Innovation' subdriver, which outperformed all the others with an outstanding score of 46.92 (**Figure 1**).

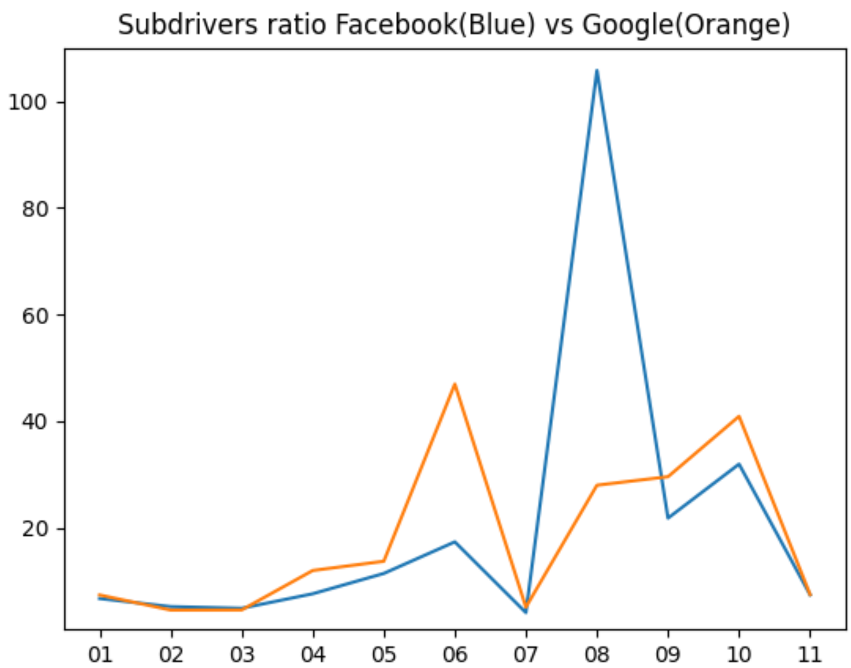


Figure 1

In the graph above it is possible to visualise the ratio of each subdriver for Facebook and Google; on the x-axis there are the subdrivers codified with numbers from 01 to 11, while on the y-axis is represented the ratio between positive and negative observations. For sake of clarity they are codified according to the following order: Price(01), Service

Quality(02), Goods Quality(03), Cool(04), Exciting(05), Innovative(06), Social Responsibility(07), Community(08), Friendly(09), Personal Relationship(10), Trustworthy(11).

Regarding Facebook instead, the tracker evaluated the Brand Reputation to be 8.68, slightly higher than Google. Its drivers instead totalised a score of 5.66 for the Value Driver, 8.96 for the Brand Driver and an outstanding 24.4 for the Relationship Driver. As for Google, the Value Driver is quite good, but it is still distant from the Brand Driver and almost four times smaller than the Relationship Driver. Although the first two drivers are very positive, as for Google the real dragger seems to be the Relationship Driver. Regarding the subdrivers of the Relationship Driver it is rather interesting to highlight the score of the 'Community Subdriver': 105.73 (**Figure 1**). This is probably an example of how a scarcity in the number of observations for this driver led to a clear error in its evaluation.

Let us compare their time series now. The first four graphs that we are going to analyse investigate the ratios of the Drivers and the Brand Reputation (**Figure 2**).

Figure 2

The four graphs above show the timeseries of the ratios of Brand Reputation (top left), Value Driver(top right), Brand Driver (down left) and Relationship Driver (down right), for Facebook (Blue) and Google (Orange). It is important to highlight that the ratios represent the stakeholders' sentiment towards the brand of a specific week.

Observing the graph of the Brand Reputation it is possible to see that even if Facebook is averagely overperforming Google, the gap between the two brands is very close. Focusing on the graphs of the Drivers, in fact, it is possible to observe that Facebook dominates Google only regarding the Relationship Driver. For the Value Driver it would be almost impossible to declare a winner, while for the Brand Driver graph Google absolutely wins over Facebook.

In each graph of Google is present a major peak in correspondence of the 38th week, which would be the week from September 17th to September 24th. This temporary and incredible increase demonstrates the ability of the tracker to detect an important event for the brand. In that specific week, in fact, Google celebrated its 20th anniversary, announcing the updating of its main algorithms. Evidently, the news was very welcome by Google users, who must have reacted with a lot of positive tweets addressing the company.

Figure 3

The four graphs above show the timeseries of the deltas of Brand Reputation (top left), Value Driver (top right), Brand Driver (down left) and Relationship Driver (down right), for Facebook (Blue) and Google (Orange). It is important to clarify that "delta" and "net" are terms indicating the same scores.

Also analysing the other four graphs(**Figure 3**), which show the Net Brand Reputation of both companies, it is possible to highlight, specifically in the Brand Driver graph, the same huge peak in correspondence of the 38th week. Moreover, the fact that among the three Drivers, the Brand Driver was the one highlighting it the most is very important, because it is the Driver that is supposed to grasp the brand's ability to innovate and excite its stakeholders.

Referring to Facebook, although the graphs showing the ratios (**Figure 2**) do not highlight particular variations in correspondence of the 12th week, which would be when the Cambridge Analytica scandal was diffused, the graphs showing the Net brand Reputation (**Figure 3**) highlights heavily this event. There is a big discrepancy, though. Since the scandal had an important negative effect on the public, the peak that should have been highlighted by the tracker was supposed to be downward, highlighting a decrease of the brand reputation and not an increase. At the same time, even if the ratios of those two weeks partially indicate that the sentiment towards the brand was not the best, those oscillations are not enough to justify a news of that magnitude.

What the tracker totally missed instead, was any kind of reaction towards the 50 million profiles hijacked from Facebook in September. Despite the reasons behind this missed detection may be several, as the comments regarding that event being later deleted, or the dataset lacking in accurately describing the importance of what happened in that specific time frame, the error made by the tracker is relevant and deserved to be highlighted.

Before proceeding with the final conclusions regarding the performances of the tracker, I would like to compare my graph showing the results of the Brand Reputation timeseries for Facebook and Google with the one of the original article.

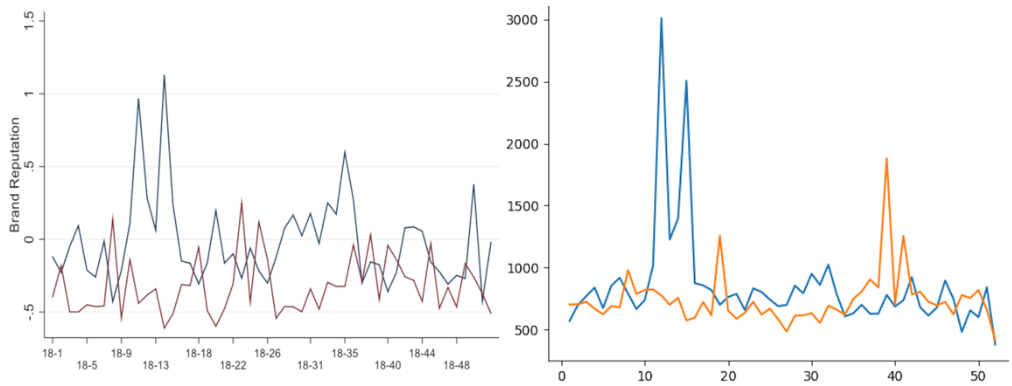


Figure 4

Both the graphs above show the Brand Reputation timeseries for Facebook(in Blue) and Google(in Orange). On the right it is shown my graph, while on the left is reported the one of the original article (Rust et al. 2021, 29).

As it is possible to see in **Figure 4** the graphs seem very similar, confirming therefore the correct implementation of the tracker. Comparing the two graphs in fact, although the intensity of the oscillations does not always coincide, the direction is hardly ever different.

Also for what regards the detection of the main events both graphs achieve very similar results. Analysing the scores in correspondence of the weeks affected by the aforementioned main events, it is possible to observe the following outcome: both present important peaks for the weeks in correspondence of the Cambridge Analytica Scandal that highlight the social impact of this event; for the 20^{th} anniversary of Google, in correspondence of the 38^{th} week (September 2018), is very interesting to notice that the event is clearly detected by my tracker, while it is very emphasized by the graph of the authors; finally, is also very interesting to notice that in neither of the two are highlighted any significative variations in correspondence of the 50 million Facebook accounts hijacked on September, implying that probably this event did not have a real impact on the overall stakeholders' sentiment.

I would like to remark that all inferences regarding the comparison formerly done, are limited to the graphs representing the timeseries of the overall score of the Brand Reputation. Since all the article is based on that specific score as representative of the brand value I believe that it was the most significative comparison as well.

In conclusion, it is possible to say that the tracker successfully captures the main events of a brand, but in doing so it is not always accurate, especially in representing the correct sentiment of stakeholders toward the brand.

Regression models findings

Regression Model	**Coefficient of determination (R^2)**
Simple Linear Regression (Facebook)	.0131
Simple Linear Regression (Google)	.0147
Fixed Effect (FE)	.1029
Random Effect (RE)	.0242

Table 1

In the table above it is possible to visualise the models implemented in my study and the correspondent coefficient of determination.

As anticipated, the two simple linear regression models using the Brand Reputation as explanatory variable for the stock price performed quite poorly for both Facebook and Google, scoring coefficients of determination respectively of .013 for the first and .015 for the second.

Let us now focus on the Panel Data regression models and on the results of the several tests to evaluate the best one for our purpose. According to the White-Test and the Breusch-Pagan-Test, which gave as outcomes respectively .0005 and .07, the assumption of homoscedasticity has been rejected in favour of heteroscedasticity. It is also possible to visually confirm this conclusion through the two graphs showed below (**Figure 5**).

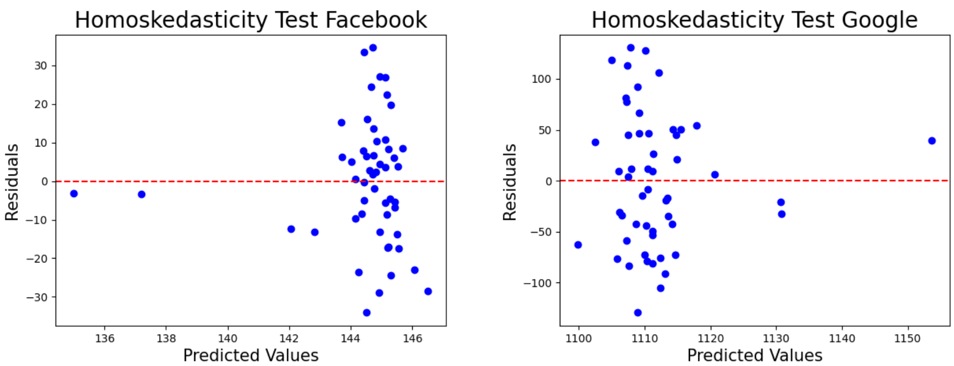


Figure 5

The two graphs above show on the x-axis the predicted values for the stock price and the y-axis the correspondent residuals; from these graphs it is possible to visually ascertain if the data of both Facebook and Google present homoskedasticity.

In order to verify it we must see if the data points plotted tend to be aligned in correspondence of the dotted line. If this is the case, it means that the variance tends to zero and the assumption of homoscedasticity is verified.

Since the data points spread a lot along the y-axis we can confirm, as also revealed by the other tests, that the homoskedasticity assumption is rebutted.

For what regards the assumption of non-autocorrelation, the Durbin-Watson-Test outcome was .086, rebutting it and highlighting instead a strong autocorrelation of the stock price. Also this characteristic can be visually confirmed through graphs. Below are reported the so called "Correlograms", which represent the autocorrelation coefficients of each of the 52 weekly stock prices.

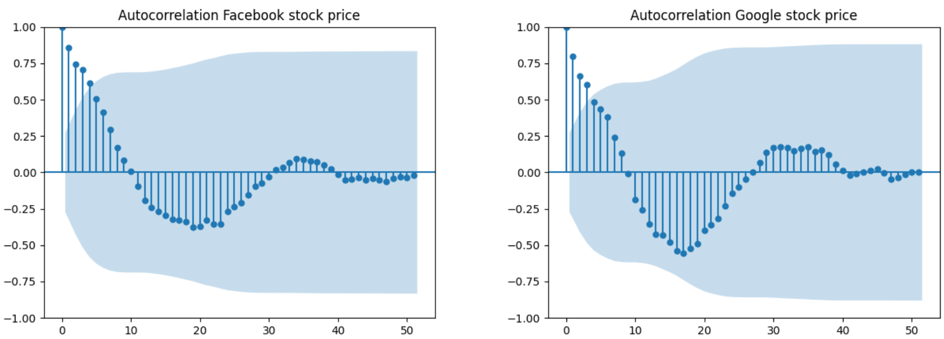


Figure 6

The two "Correlograms" above show the separate lags of the stock price for Facebook and Google. Lag would be the correlation between values that are one time period apart. On the x-axis are represented the 52 weeks of 2018, on the y-axis is shown the correspondent coefficient of autocorrelation for the specific weekly stock price.

In order to visually assess if the models present autocorrelation we have to check if the points align in order to show a pattern. Since both graphs emphasizes the existence of a pattern it is possible to rebut the assumption of non-autocorrelation.

Considering all the results of the test we can agree that, at least theoretically, the Fixed Effect model would be the most appropriate. However, in order to confirm that the model is at least partially endogenous I developed both the FE and the RE to check empirically that FE was the most appropriate. The coefficient of determination of RE gave a very poor result, .0242, only slightly better than the simple linear regression models analysed formerly. The FE model scored a coefficient of determination of .1029, confirming to be the best one.

Before drawing the conclusions, it is important to remark that the stock price is subject to a countless number of variables that can influence its outcome, so I never expected to be able to design a regression model able to accurately predict the stock price. What I wanted to reveal was if it is possible to empirically demonstrate the existence of a relation between the Brand Reputation and the stock price.

Having made the due considerations, a coefficient of determination of .1029 can be considered as a confirmation for the existence of the aforementioned dependency.

Discussion

Contributions

In the first place, with this paper I provided evidence of the intuitive dependency between a brand's reputation and its financial performances through rational empirical analysis. Secondly, I demonstrated the reliability of the Brand Reputation Framework devised by Rust et al.(2021) by replicating the functions of its Brand Reputation tracker and verifying its ability to obtain consistent results also on a small scale.

Finally, through the development of the various statistical models, I contributed to highlight the existence of the relation between the Brand Reputation, estimated by the aforementioned tracker, and the stock price.

Managerial Relevance

As I made it clear several times during the development of my paper, the results of the original article, and by extension of this one, are very important for managerial implications. Through its actionable drivers and subdrivers, the framework developed by Rust et al. is built specifically for the use of companies and their managers. Specifically, the tracker for the Brand Reputation is a real tool, not only it can be used to manage company reputation and monitor competitive dynamics, but it allows managers to evaluate the critical aspects of the firm and specifically intervene on them through the actionable drivers, influencing the financial performances.

Limitations and Further Research

Concerning the limitations of this article, as explained in the Data Description section, the range of scope of the paper is limited only to two brands, Facebook and Google, with an average of 90,000 observations for each dataset. The fact that the study investigate only two brands can be considered as an important limitation of the generalisability of the results of this work, both with regard to the ability of the tracker to seize the important events of the brand, and with respect to the consistency of the predictions of the regression models. The number of observations can be considered as a smaller limitation as well; although it was a discrete number of observations, it is undeniable that a bigger dataset with a consistent distribution of the observations among the weeks would have conferred a higher level of granularity and reliability of the results.

With regard to further research, I would like to encourage the development of one or more paper focusing specifically on the prediction of the stock price from the Brand Reputation. In the first place, it would be rather interesting the development of a Fixed Effect regression model using a collection of datasets of multiple brands having the characteristics formerly mentioned. The results of that study would be particularly relevant because they would allow to generalise the results of my work. On the other hand, another important aspect that could be developed, would be the investigation of a regression model using Brand Reputation as one among multiple predictors for the Brand Reputation, instead of the only one.

References

Brugger, Bernhard (2021), “A Guide to Panel Data Regression: Theoretics and Implementation with Python,” Towards Data Science. Last accessed June 6, 2022. https://towardsdatascience.com/a-guide-to-panel-data-regression-theoretics-and-implementation-with-python-4c84c5055cf8

Hayes, Adam (2022), ”Stock Definition”. Last accessed June 6, 2022. https://www.investopedia.com/terms/s/stock.asp

Keller, Kevin L. (1993), “Conceptualizing, Measuring and Managing Customer-Based Brand Equity,” Journal of Marketing, 57 (1), 1–22. Last accessed June 6, 2022. https://journals.sagepub.com/doi/10.1177/002224299305700101

Landstrom, Joachim (2020), “Regression Analysis and Panel Data – A Survival Guide to Using R Applied to Cross Sections, Time Series, and Panels”. Last accessed June 6, 2022. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3487658

Lange, Donald, Peggy M. Lee, and Ye Dai (2011), “Organizational Reputation: A Review,” Journal of Management, 37 (1), 153–84. Last accessed June 6, 2022. https://journals.sagepub.com/doi/10.1177/0149206310390963

Lemon, Katherine N., Roland T. Rust, and Valarie A. Zeithaml (2001), “What drives Customer Equity?,” Marketing Management, 10(1), 20-25. Last accessed June 6, 2022. https://www.researchgate.net/publication/284404248_What_drives_customer_equity

Lindgreen, Adam, C. Anthony Di Benedetto, Roderick J. Brodie, and Elina Jaakkola (2020), “How to develop great conceptual frameworks for business-to-business marketing,” Industrial Marketing Management, 94(1), 2-10. Last accessed June 6, 2022. https://www.researchgate.net/publication/340650202_How_to_develop_great_conceptual_frameworks_for_business-to-business_marketing

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), “Return on Marketing: Using Customer Equity to Focus Marketing Strategy,” Journal of Marketing, 68 (1), 109–27. Last accessed June 6, 2022.

https://www.researchgate.net/publication/248764523_Return_on_Marketing_Using_Customer_Equity_To_Focus_Marketing_Strategy

Rust, Roland T., William Rand, Ming-Hui Huang, Andrew T. Stephen, Gillian Brooks, and Timur Chabuk (2021), "Real-Time Brand Reputation Tracking Using Social Media," Journal of Marketing, 85(4), 21-43. Last accessed June 6, 2022. https://journals.sagepub.com/doi/abs/10.1177/0022242921995173

Sabate, Juan, and Esther Q. Puente (2003), "Empirical Analysis of the Relationship Between Corporate Reputation and Financial Performance: A Survey of the Literature," Corporate Reputation Review, 6(2), 161-177. Last accessed June 6, 2022. https://www.researchgate.net/publication/31956895_Empirical_Analysis_of_the_Relationship_Between_Corporate_Reputation_and_Financial_Performance_A_Survey_of_the_Literature

Sheppard, Kevin (2022), "Linear Models," Github repository. Last accessed June 7, 2022. https://bashtage.github.io/linearmodels/index.html