Proof Central

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. First, return to the online proofing interface by clicking "Edit" at the top page, then insert a Comment in the relevant location. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof.

Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

AUTHOR QUERIES

Note: *It is crucial that you NOT make direct edits to the PDF using the editing tools as doing so could lead us to overlook your desired changes. Edits should be made via the 'Comments' feature.*

AUTHOR PLEASE ANSWER ALL QUERIES

AQ1—Please confirm that given name(s) [yellow] and surname(s) [green] have been identified correctly and are presented in the desired order, and please carefully verify the spelling of all authors' names, including all special characters and accents. Please note that no changes can be made to authorship at this point in the publication process.

AQ2—Please ensure the accuracy of the affiliation(s) of each author. Affiliations cannot be changed once the article has been published online.

AQ3—If you wish to edit the affiliation details, please consider the following house style requirements to avoid any further queries and to speed up publication: (a) city and country names, in English, are mandatory; (b) state name is required only for the USA and not to be included for any other countries; (c) address lines, such as Avenue, Boulevard, Street, should not be included; (d) postal codes are not allowed; (e) degrees, roles, and titles (e.g. Dr., Professor, student, scholar, etc.) are not allowed. Please note that any details that do not conform to the house style requirements will not be retained in the modified/newly included affiliations.

AQ4—Please be advised that affiliations are fixed at submission. If you wish to note a changed affiliation since submission but prior to acceptance, please advise of this affiliation in full (including department, city, state if based in the US, and country) **as well as the date your association with this new institution began. Failure to provide this date may cause significant delay to the publication of your article**. Once this change has been confirmed, this affiliation will either be added as a dual affiliation or as part of the author biography information, using the wording "The author is also affiliated with [AFFILIATION DETAILS]".

AQ5—Please note the addition of "Department of" to the affiliation "Department of Economic and Business Administration, Facultad de Comercio y Gestión, Universidad de Málaga, Málaga, Spain". We request you to verify the accuracy and appropriateness of this.

AQ6—Please confirm that the provided email "acasado@uma.es" is the correct address for official communication.

AQ7—References "Ratanen *et al*., 2020; Ratanan *et al*., 2020" are cited in the text but not provided in the reference list. Please provide them in the reference list or delete these citations from the text.

AQ8—The citation “de Casado-Molina *et al.*, 2020” has been changed to match the author name in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

AQ9—Reference(s) that are not cited in the text are moved to the “Further reading” section. Please check.

International Journal of Bank Marketing

AQ: 1
Banking reputation and its impact on stock markets: a big data analysis through online comments

AQ: 2 Ana-María Casado-Molina
Department of Economic and Business Administration, Facultad de Comercio y Gestión, Universidad de Málaga, Málaga, Spain
AQ: 3 María-Mercedes Rojas-de-Gracia and Pilar Alarcón-Urbistondo
AQ: 4 *Universidad de Málaga, Málaga, Spain, and*
AQ: 5 Célia M.Q. Ramos
ESGHT, University of the Algarve, Faro, Portugal

Received 16 July 2024
Revised 1 December 2024
12 December 2024
Accepted 12 December 2024

Abstract

Purpose – Given the substantial number of social conversations on the Internet, companies must remain vigilant about protecting their reputations and businesses. The purpose of this research is to measure the impact on share prices of reputational variables, measured through online comments, at the banking sector level as well as patterns of behavior between these variables and fluctuations in share prices.

Design/methodology/approach – Using big data and business intelligence techniques, bank reputation was analyzed through online social comments. The sample includes seven Spanish banks. To measure the impact of reputational variables on share prices, an Online Reputation Index was created. These variables were then correlated and filtered with the share price variation rate of each bank on a daily basis under two scenarios: using all share price data and also focusing on times of sharp fluctuations. Finally, multiple linear regression analysis was used to identify patterns in these relationships.

Findings – The findings reveal that negative comments focusing on attitude, emotions and governance experiences – particularly regarding ethical performance – are the only reputational variables with a clear impact on share price fluctuations in the banking sector. Strong correlations between reputation variables and share prices were observed only when online comments were extremely negative. These behaviors were more frequent during periods of significant price fluctuations.

Originality/value – This is the first study to analyze the impact of the reputation of the banking sector measured in the online environment on share price, taking into account its multidimensional construct.

Keywords Online perceptions, Reputational intelligence, Big data, Banking reputation, Stock market fluctuation, Share price

Paper type Research paper

1. Introduction

In today's highly competitive and changing world, opinions and social judgments about firms exert a growing influence on their business (Etter *et al.*, 2018). One measure of a firm's success in the social context is its corporate reputation (CR). Consequently, firms need to monitor their reputation since it largely determines their sustainability in the market in which they operate (Casado-Molina *et al.*, 2020; Etter *et al.*, 2018). One of the sectors in which reputation analyses are particularly important is banking (Carè *et al.*, 2024). This sector is more sensitive for reasons such as: its services are only offered by a small number of large brands; and widespread consumption of these services by most members of society means that the banking industry generates a higher volume of public perceptions than other sectors (Zaby and Pohl, 2019). Furthermore, banks need to increase the trust they generate among customers and society given the negligence associated with their corporate governance models during the last financial crisis (Bravo-Urquiza and Moreno-Ureba, AQ: 7 2021; Carè *et al.*, 2024; Ratanen *et al.*, 2020).

This work is funded for open access charge by the Universidad de Málaga/CBUA.

International Journal of Bank Marketing
© Emerald Publishing Limited
e-ISSN: 1758-5937
p-ISSN: 0265-2323
DOI 10.1108/IJBM-07-2024-0419

However, although the literature describes the advantages to firms of managing their reputation, it also warns of the difficulty involved in measuring this (De Castro *et al.*, 2006). This difficulty comes from the following four aspects. Firstly, the current confusion about the definition of reputation has resulted in there being very diverse models for measuring it, relating to the different attributes that comprise the construct. Therefore, before we decide what to measure, we first need to define what is meant by CR. Much of the literature defines reputation as a multidimensional construct that encompasses emotional and rational dimensions (Dollinger *et al.*, 1997). Thus, Fombrun *et al.* (2015) defined it as different public perceptions generated by a firm's past actions and future prospects, which can be compared at different points in time or with the reputation of its sector. Secondly, it should be noted that its intangible nature makes it difficult to measure because what can in reality be measured are its effects, such as its impact on financial performance (Dowling, 2001). In relation to the high volume and diversity of sources (third aspect) that must be examined to extract social opinions, academics assume that the media, the views of accredited bodies and rankings all influence collective judgments. However, all these sources relate to institutional bodies with specific interests (Shoemaker and Reese, 2013) and only provide a limited indication of society's opinions. That is why we need to broaden the range of sources to include the social media so that we can have a more diverse and democratic public debate (Etter *et al.*, 2018; Etter *et al.*, 2019; Rojas-de-Gracia *et al.*, 2021).

Finally, the new world of online social interaction (fourth aspect) has empowered public opinion, forcing companies to monitor comments in this environment. However, in this scenario, organizations experience even more difficulty in measuring their reputation due to three aspects (Ravasi *et al.*, 2019): sources mutually influencing each other in the debate and in opinions appearing on them; the range of data, which hampers aggregation and processing; and the need to develop new methodologies to process online data. Thus, the very nature of these quantitative methods involving offline analysis (structural equations, PLS, and regression and correlation approaches) renders them impractical for analyzing and measuring online perceptions. That is why the literature has highlighted the need to apply big data and business intelligence (BI) methodologies that facilitate the creation of metrics with robust data that can then be analyzed (Casado-Molina *et al.*, 2020; Ratanan *et al.*, 2020). The absence of such studies creates a gap in the tools that connect online opinions with managing an organization's tangible economic resources effectively (Dayan *et al.*, 2017). In fact, few reputation scales and indices have been proposed for banks, and these always refer to the offline environment and, therefore, do not consider the added difficulty of the online environment (Zaby and Pohl, 2019).

But why do banks need to measure their reputation? A good reputation creates a sustainable competitive advantage and increases profits, as it helps firms achieve a higher return on their assets and even boost daily returns (Dowling, 2001). In this sense, the aim of this study is to understand how, in the online context and for the banking sector, the reputation construct affects share price at the sector level, and whether there are behavioral patterns involved. To achieve this, the behavior of the seven largest Spanish banks has been analyzed in terms of online comments from their stakeholders and fluctuations in their share prices during a pivotal year for the Spanish banking sector and its reputation.

2. Conceptual background

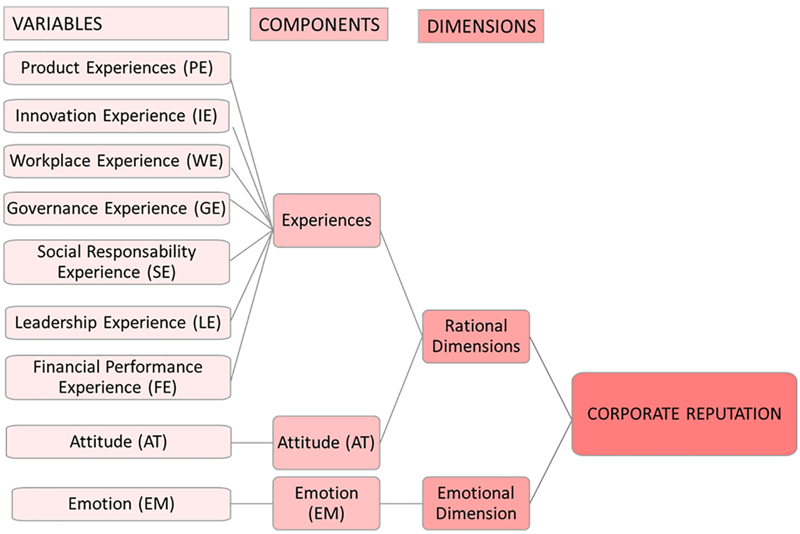
2.1 Key characteristics of reputation and the online environment

To measure reputation, we need to consider the four characteristics of the construct: its slow accumulation, perspectives of the different stakeholders, the complex social process, and its multidimensionality. In relation to the first, it should be noted that building a reputation takes time, because it is not based solely on the perceptions on a single day or in the present, but also on perceptions formed as a result of the firm's history and past performance (De Castro *et al.*, 2006). Therefore, the concept of reputation is dynamic (Etter *et al.*, 2018). It requires methodologies that consider both public information flows on the social media over a period, especially relevant today, and the cumulative stock of information resulting from previous publications and exchanges (Etter *et al.*, 2019). One of the techniques that best analyzes these online reputational dynamics is automated sentiment analysis (Etter *et al.*, 2018).

As regards the second characteristic, different stakeholders may perceive the dimensions of a firm's reputation differently. This is since they all have different needs, expectations and knowledge about the firm because of their own specific physical or legal nature (Fombrun *et al.*, 2015; Dowling, 2001). As a result, reputation involves a complex social measurement process (third characteristic). It is extracted from the subjective perceptions of the very diverse stakeholders that are hard to explain (De Castro *et al.*, 2006). In the online environment, these perceptions add to the complexity as they are linked to emotional responses (Toubiana and Zietsma, 2017), resulting in debates ranging from the creation of public arenas to so-called echo chambers (Crane *et al.*, 2008). In addition, collective perceptions are the result of reciprocal flows between the messages issued by powerful institutions (e.g. the media), those of social media influencers and the conversations involving the public on social media (Ravasi *et al.*, 2019). This social complexity creates the need to consider more than one source of opinion (Etter *et al.*, 2019).

In relation to the fourth characteristic, nowadays there is general consensus on the multidimensional nature of reputation and the variables that comprise it (Dollinger *et al.*, 1997; De Castro *et al.*, 2006; Ponzi *et al.*, 2011). The Reptrak model (Fombrun *et al.*, 2015) is an example of that consensus. In fact, it is the most commonly used model in the business sector. This model includes the emotional aspect of the reputation construct and incorporates rational dimensions that help create a more complete definition of the construct and how these dimensions correlate with a series of favorable attitudes toward the firm. These rational dimensions are: products or services, innovation, work environment, social responsibility, leadership of the managers, economic-financial performance, and company governance.

Considering the two dimensions proposed by the Reptrak model, we consider that in order to measure the reputation construct, it is necessary to expressly and independently incorporate the variables for both dimensions. The emotional dimension is explained through a component called emotion. Meanwhile, the rational dimension consists of two components: experience
F1 and attitude. In other words, CR is made up of three components (see Figure 1): experiences, emotions, and the public's attitudes toward the brand.



Source(s): Authors' own work

Figure 1. Variables, components and dimensions of corporate reputation

Experiences have been studied from all the different stakeholder perspectives. Fombrun *et al.* (2015) stated that the public mainly expresses an opinion on their experience with companies in relation to seven variables: products (PE); innovative nature of the company (IE); workplace experience (WE); governance experience (GE); social responsibility (SE); leadership (LE); and financial performance (FE). Attitude (AT), the second component of the rational dimension of reputation, is defined as a predisposition toward a brand, which can be acquired through one's own personal experiences and/or emotions or those of another person. This generates behavior that affects the firm's business (Ajzen and Fishbein, 2005). Attitude is largely determined by recommendations made by influential people, such as financial analysts, and can generate favorable or unfavorable behavior toward a brand (Fombrun *et al.*, 2015). Finally, emotions consist of feelings felt by different stakeholders as a subjective response to experiences in their environment. Schwaiger (2004) has termed these emotions "affective reputation". Emotions are measured by their valence and intensity (Scherer, 2005). These contain a large amount of information on the relationship between a firm and an individual, as well as on inner thoughts about this relationship which in turn are highly influenced by experiences with the firm (Scherer, 2005).

In the online environment, the literature on the measurement of reputation highlights the analysis of online conversations and debates, categorizing them by topic and sentiment (Etter *et al.*, 2019). However, in their analyses, these studies do not categorize the topics according to the dimensions, components and variables that comprise the CR construct. Only Rantanen *et al.* (2020) create a multidimensional classification framework for measuring online CR, but without analyzing the interrelationship between these variables and their impact on organizations.

2.2 Relation between reputation dimensions and tangible assets in online contexts

In the banking sector, only Gómez-Carrasco and Michelon (2017) have analyzed the impact of reputation on share price in the online context. These authors found that tweets from two institutions (consumer associations and unions) had a negative impact on the stock market value of the shares. In the case of consumer associations, this impact was determined by their ability to disseminate protests. The authors concluded that when the protest disseminated by these stakeholders reached a wide Twitter audience, there was a significant negative impact on share prices in bearish and bullish markets alike. However, this study was limited to analyzing online protests, meaning that it only examined negative emotions. Furthermore, the classification of the content of online comments did not reflect the reputation dimensions identified in the literature.

Focusing on the literature relating to the online environment and dealing with reputation in general, and outside the banking sector, there are authors who associate the publication and dissemination of online comments with an impact on share price. Cakra and Trisedya (2015) studied the rational and emotional aspects of reputation jointly, relating them to decisions made about shares in the stock market. Other authors have conducted more specific studies, relating the emotional aspect to share price. In this respect, there are two opposing schools of thought: those that find a relationship between emotion and share price (Barakat *et al.*, 2019; Bollen *et al.*, 2011; Duan *et al.*, 2021; Hu and Tripathi, 2016) and those that do not (Kim and Kim, 2014). To a lesser extent, attempts have been made to relate rational aspects to share price. Such studies have reported a relationship between EE and share price (van Den Broek *et al.*, 2017) and between PE and firm value (Luo and Zhang, 2013).

In short, in an online context, the studies identified have separately measured the impact of some elements of these reputation dimensions on some tangible assets. The main studies on reputation addressing the online environment have primarily explored the relationship between emotion and share price in various sectors. However, these studies do not select reputation variables based on the full set of dimensions and variables considered in the literature for this construct. The emotional dimension is the only one where they have come close to doing this.

Furthermore, most of the generic studies identified which associate reputation with an impact on share price have focused on a specific group of stakeholders, namely investors, financial experts or the community, and their data is extracted from the specialized information sources and platforms that these stakeholders use and consult or the social media profiles in which this community specifically interacts (Yang *et al.*, 2015). Only the aforementioned studies by Bollen *et al.* (2011) and Gómez-Carrasco and Michelon (2017) do so for non-financial stakeholders (unions and consumer associations) and the population in general. Therefore, the literature has mainly focused on a single type of stakeholder or, when considering several of them, has treated them as closed groups, meaning that the opinions of one group cannot affect or modify the opinions of the other groups.

In addition, the literature on non-financial stakeholders or the population in general has focused on the relationship between negative comments and share price, and rarely on the impact of positive ones (Gómez-Carrasco and Michelon, 2017). There is no literature that considers the experiences of the public with the analyzed brands and how these influence share price.

To fill these gaps and gain better insights, the aim is to understand how, in the online context and for the banking sector, the reputation construct affects share price and whether there are behavioral patterns between all the reputation variables and share price. Specifically, the research questions (RQ) for this study are:

RQ1. Using the data extracted from the online environment, what is the relationship between reputation variables and fluctuations in share prices in the banking sector?

RQ2. Are there patterns of behavior between the reputational variables for the banking sector, measured in the online environment, and share price fluctuations?

3. Methodology

3.1 Data sample

To achieve the study objectives, we work with the reputation variables indicated in the literature (see Figure 1). These variables are extracted from online comments, online news, and opinions among the general population and from the financial analysts in digital ecosystems (hereinafter, inputs). These inputs relate to seven banks in the Spanish market (BBVA, Sabadell, Popular, Santander, Bankia, Bankinter and Caixabank) that were in operation in 2015. Due to the limited number of firms in this sector, these seven banks represented more than 90% of total industry assets in Spain (Sánchez, 2016) and were operating around the world. We consider 2015 to be an ideal timeframe for our study for two reasons: firstly, the sector was feeling the effects of the reputational crises caused by banking malpractice, which allowed us to identify different reputational issues; secondly, in relation to purely financial behavior in the sector, it reflected a more stable scenario than in previous years, when the mergers prompted by the financial crisis of 2008 came to a standstill. In fact, the Spanish banking sector contracted from 62 banks and savings banks in 2008 to only 15 at the beginning of 2015 (Escudero, 2017).

The inputs were extracted in text form and in Spanish. The comments on the banks being analyzed were collected for two types of Spanish-speaking online users: the general online population and financial analysts. These two groups were chosen in order to have an overview of how reputation dimensions relate to share price. The inputs extracted for financial analysts were used to work on the "attitude" component, because these analysts are key opinion leaders in terms of encouraging people to buy or sell shares, and attitude is considered to be a stock market recommendation (van Riel and Fombrun, 2007). The inputs relating to general public opinion were used to analyze the other two components of reputation (experiences and emotions), since the intention is to identify the perceptions of those individuals who have had experience with a bank.

There were 268,174 inputs relating to the analyzed banks, giving a total of 198,425 net inputs after filtering to ensure that they correspond exactly to the banks analyzed. For all of them, there was an analysis of all inputs over the 250 working days in which the Madrid Stock Exchange was active in 2015. In addition to these inputs, daily share prices were gathered for the seven banks analyzed on the closure date during that same period.

3.2 Data collection, transformation and storage

Data collection and cleaning. To collect the inputs related to reputation, it was necessary to use big data and BI techniques. With these procedures, a large volume of qualitative data can be extracted from the online environment and transformed into quantitative data for subsequent processing and analysis.

As is usual in this type of study, prior to the collection of reputational inputs, the sources used were selected based on tracking using web crawlers (Chau and Chen, 2003). The sources were carefully selected based on the recommendations of relevant institutions, such as the study by the AIMC (association for media research) and finance experts (university professors). In summary, the online sources used to extract inputs comprised forums, chats, Twitter, blogs, online mass media publications, and stock exchange platforms. The source for extracting share price data was the Spanish Securities Market Commission website (www.cnmv.es).

Subsequently, the information gathering process began and this varied depending on the type of information in question, the share price and the reputation inputs. The extraction of data on share prices at the close of the day did not result in any significant issues, as this is public data. The extraction of reputational inputs was performed using public and private APIs. For tweets, we used Twitter's API; for online news, we used news service APIs; and for the remaining sources, we used web scraping processes.

Because not all inputs were valid for analysis the next phase was to filter them, following the methodological recommendations of Batrinca and Treleaven (2015) and Casado-Molina *et al.* (2020). To do this, these inputs were subjected to an analysis using the NLTK python library and logic rules (Tobergte *et al.*, 2013), eliminating any elements that do not refer to the bank and/or reputation input being analyzed. For example, if we wanted to analyze communications relating to PE with Bank A, such as an online commentary like "Bank A loans are the best", we would first define the rules that allowed us to extract it, through the appearance of keywords and their connectors: "Bank A", "Bank A loans", "Bank A" and "loans", etc.

Next, again following the recommendations of Batrinca and Treleaven (2015) and Casado-Molina *et al.* (2020), we conducted a semantic analysis in order to obtain the first metadata and accelerate subsequent analyses. To do this, we used a text corpus endorsed by the Spanish Society for Natural Language Processing (Spanish initials: SEPLN), which contains 70,000 phrases and uses the TensorFlow open−source library. As a result of this first semantic analysis, two issues were observed. Firstly, public comments on innovation refer in the vast majority of cases to banking products. In other words, there are hardly any comments on the innovative nature of banks in general, since they refer to new types of loans or new services, for example. Secondly, when referring to the governance of banks, users focus on the ethics behind the practices of these firms. For this reason, the decision was made to, on the one hand, merge the variables IE and PE into a single variable, which we have simply called "PE"; and on the other hand, change the name of GE and call it "ethic experience (EE)".

Data evaluation. Once the inputs had been collected, filtered, and subjected to this initial semantic analysis, they were given a numeric value to enable the processing of the information (transformation phase). This allowed us to associate a given valence to each input, using a classification algorithm based on an artificial neural network (ANN) with peer-supervised training.

On the basis of this value, labels were created for each reputation entry (see Table 1). In the T1
case of experience and emotion, the labels used were: admiration, acceptance, indifference,

Table 1. Labels of reputation variables

Labels	Values
Experiences and emotions	
Admiration	X ≥ 8
Acceptance	6 ≤ X < 8
Indifference	4 ≤ X < 6
Rejection	2 ≤ X < 4
Hate	X < 2
Attitudes	
Buy	10
Buy/Hold	7.5
Hold	5
Hold/Sell	2.5
Sell	0

Note(s): X means the value assigned to the item valued
Source(s): Authors' own work

rejection, and hate (Herrera and Herrera–Viedma, 2000). The intensity with which each label appears in an input is measured using a discreet scale. For example, the input "Those dishonest folks from Bank X charge €2 for taking money out from their ATMs! Boycott!" (online comment extracted from Twitter, 03/May/2015), corresponding to the product variable, was assigned the hate label with a value of 1. In the case of attitudes (see Table 1), we also used five labels (buy, buy/hold, hold, sell/hold, and sell) that correspond to the way in which financial experts talk about their stock market recommendations. This time, the values were of a discreet nature on a scale of 0–10 (with 10 corresponding to the label buy). In order to generate the final database (which includes each input with its label and value) training of the corpus data had first to be performed, using a Bayes classifier (Batrinca and Treleaven, 2015).

Datawarehouse. The data warehouse was constructed using MongoDB, a NoSQL database system that saves data in tables, as with relational databases. This database features a dynamic schema, rendering data integration easier, simpler, and more dynamic, since each document in a collection can have different fields.

Obtaining the daily value of the variables. The next step was to reduce the large number of entries for each reputation variable to a single opinion value per bank per day. Each of the days analyzed was assigned a unique numerical value that was then categorized by label according to its value. Therefore, first, we calculated the average daily value of all comments relating to each reputational variable. The database that includes the average values for each reputational variable by bank and day comprises 1,750 records (multiplying the 250 days analyzed by the seven banks). This database was supplemented with the corresponding share price at the close of each day.

3.3 Data analysis

3.3.1 Preliminary considerations when preparing the data for analysis. The research questions for this paper involve framing the study around how share price is affected by comments posted by different stakeholders on social media. To do this, it is necessary to take into account two prior considerations: (1) Given that we are examining the online world, should we talk about a cumulative effect of comments on social media?; (2) How can we reflect their impact on price?

When answering question one, we should not forget that brand reputation is a construct measured from the accumulated opinions of each user over time (Fombrun *et al.*, 2015) and that in the online world these opinions behave like quicksand where the snowball effect means that we have to consider them not as specific opinions but as being built up over time. (Casado *et al.*, 2020; Etter *et al.*, 2019). Therefore, in a paper like this one, the cumulative effect of these comments must be considered in relation to their influence on share price. In this sense, our study considers the general and continued perception of the public and not the impact of isolated events on banking reputation. Therefore, for the variables that comprise experiences and emotions, a second cumulative value was calculated, considering the recommendations from Etter *et al.* (2019). Inherent to this question is the issue of how long the influence period is for comments. This is because what was said in the recent past (a few days ago) influences today's price. In line with the above, we calculated the average over a time horizon of 15 days, following the recommendation Casado-Molina *et al.* (2020) and Rojas-de-Gracia *et al.* (2021). To give more weight to the most recent comments, this average was linearly weighted. As the study is conducted over an entire year, it was necessary to start collecting inputs on reputational variables from the last 14 working days of the previous year. If there were no comments for a variable on a particular day, the value for the previous day was used, based on the assumption that in this case perceptions had not changed. The cumulative value was not considered when calculating the attitude variable, since financial analysts issue online recommendations at specific moments. Instead, only the daily ratio was used. Thus, on the days when financial analysts did not issue opinions, the value for attitude was assigned the label "uninformed." Finally, for both online comments and the attitude variable, the values of the previous day (N_{-1}) were taken as the influential value for the price on day N. This allowed us to gather all the comments produced during the 24 h of a day for the impact on share price. AQ: 8

This cumulative effect of comments on the price discussed in the previous paragraphs also reflects the fact that the influence of comments on price is not immediate but rather gradual. If all investors were traders, the price change would probably be (almost) immediate, since they are usually up to date with rumors that affect share prices. However, there are distinct types of investors participating in the stock market. In this case, the impact of comments about price is not necessarily immediate, but instead may be felt with a lag.

The second question raised is how to measure the impact of comments on share price. This means not talking about share prices *per se*, but about changes to them. To answer this question, our work focuses on how comments on social media cause changes in share prices. To do this, we work with the concept of the "share price variation rate" which is calculated for each day and bank. The aim of this rate is to highlight whether the price variation for a given day is relevant or not when compared to the share price variation on the following days. To calculate this, the first step is to identify the difference between the closing share prices on subsequent days (Closing price on day N – Closing price on day N_{-1}). To identify how important the variation is on that day, it needs to be put into the context of a nearby time interval. Therefore, the daily variation in the closing price was compared with the average share price variation over the previous 15 days.

In addition, since working with all the daily rates could create noise by incorporating periods during which there had been a lot of price stability but significant variation rates, the decision was made to work on a second scenario which included only those days on which the price variation was significant when compared to the price variations on the previous days. In this more restrictive scenario, those days in which the "share price variation rate" is above 1 were considered. That is, it refers to those days on which the share price experienced a variation (increase or decrease) greater than the average rate of price variations within the considered interval. In the study, these days are referred to as T moments.

In short, to answer the research questions, two scenarios are used: a general one that includes price variations and cumulative values of the reputation variables for all days of the year and another more restrictive one that includes this information only for T moments, which are those subperiods in which significant prices changes were recorded. In both, the data is used to examine the relationship with the reputational variables in the previous period. In our study, following the literature, this nearby period is specified as the preceding 15 days.

These considerations led to the need for prior preparation of the data set to be analyzed. First, each reputational variable was associated with its respective label and with the share

price and share price variation rate for each of the banks and days, as shown in the example in
F2 Figure 2.

Second, since our study considers that comments are made throughout the 24 h of the day, the following scenario was developed to capture their full impact on share prices, also based on Figure 2, which already reflects the cumulative effect of the reputation variables: "what was said yesterday influences today's share price."

F3 Third, as shown in Figure 3, the T moments are selected, representing those days on which the share price variation exceeded the average share price variation over the preceding 15 days (share price variation rate greater than 1).

3.3.2 Data analysis techniques. Once the database had been constructed, a descriptive analysis of the online reputation of each bank and of the sector was produced as a starting point. For this, an evaluation of the reputation performance of each bank by variable, as well as the average for the banking sector, was extracted from the large number of online inputs. To do this, what we have called the Online Reputation Index (ORI) was calculated (1):

$$ORI = \frac{\sum_{i=1}^{n}(\alpha_i * x_i)}{\sum_{i=1}^{n}\alpha_i} \qquad (1)$$

Where is the scale level; is the weighted reputation scale level measured $i = 1, \ldots, n$ αi as the number of days on which the average of all one day's comments was in scale level i; and x is the scale value for the i level. In our study, we have decided that the value difference between each consecutive scale value is one point, with the value scale going from 1 to 5, as shown in
T2 Table 2. In short, the ORI assesses a company's reputation based on the proportion of days of the year on which each reputational variable is assigned a certain label (in the case of

Date	Bank	Product Experience	Work Experience	Ethical Experience	Social Responsability Experience	Leadership Experience	Finacial Perfomance Experience	Emotion	Attitude	Share Price	Share price variation rate	Share price variation rate previous day
01/16/2015	Bank #2	EP Indifference	EL Acceptance	EE Indifference	ES Indifference	ED Indifference	EF Indifference	EM Indifference	AT to Hold/Bu	2,28	0.081	0
01/17/2015	Bank #2	EP Indifference	EL Acceptance	EE Indifference	ES Indifference	ED Rejection	EF Rejection	EM Indifference	AT to Hold/Bu	2,28	0	0
01/18/2015	Bank #2	EP Indifference	EL Acceptance	EE Acceptance	ES Indifference	ED Indifference	EF Acceptance	EM Acceptance	AT to Hold/Bu	2,28	0	0.377
01/19/2015	Bank #2	EP Indifference	EL Acceptance	EE Indifference	ES Indifference	ED Indifference	EF Indifference	EM Indifference	AT to Hold/Bu	2,308	0.377	0.473
01/20/2015	Bank #2	EP Indifference	EL Acceptance	EE Rejection	ES Acceptance	ED Indifference	EF Acceptance	EM Indifference	AT to Hold/Bu	2,349	0.473	0.416
01/21/2015	Bank #2	EP Indifference	EL Acceptance	EE Acceptance	ES Acceptance	ED Indifference	EF Acceptance	EM Indifference	AT to Hold/Bu	2,39	0.416	0.544
01/22/2015	Bank #2	EP Acceptance	EL Acceptance	EE Rejection	ES Indifference	ED Acceptance	EF Acceptance	EM Indifference	AT to Hold/Bu	2,449	0.544	–0.837
01/23/2015	Bank #2	EP Indifference	EL Acceptance	EE Hate	ES Indifference	ED Acceptance	EF Acceptance	EM Indifference	AT to Hold/Bu	2,361	–0.837	0
01/24/2015	Bank #2	EP Indifference	EL Acceptance	EE Hate	ES Acceptance	ED Acceptance	EF Acceptance	EM Indifference	AT to Hold/Bu	2,361	0.000	0
01/25/2015	Bank #2	EP Indifference	EL Indifference	EE Hate	ES Indifference	ED Indifference	EF Indifference	EM Indifference	AT to Hold/Bu	2,361	0.000	0.134
01/26/2015	Bank #2	EP Indifference	EL Acceptance	EE Hate	ES Acceptance	ED Acceptance	EF Acceptance	EM Acceptance	AT to Hold/Bu	2,373	0.134	–0.663
01/27/2015	Bank #2	EP Acceptance	EL Indifference	EE Hate	ES Indifference	ED Indifference	EF Indifference	EM Indifference	AT to Hold/Bu	2,331	–0.663	–1.625

Source(s): Authors' own work

Figure 2. Example of research data organization (general scenario)

Date	Bank	Product Experience	Work Experience	Ethical Experience	Social Responsability Experience	Leadership Experience	Finacial Perfomance Experience	Emotion	Attitude	Share Price	Share price variation rate	Share price variation rate previous day
01/28/2015	Bank #2	EP Acceptance	EL Acceptance	EE Rejection	ES Acceptance	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,225	–1,625	1,788
02/12/2015	Bank #2	EP Acceptance	EL Indifference	EE Acceptance	ES Indifference	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,231	0.735	1,107
02/17/2015	Bank #2	EP Acceptance	EL Acceptance	EE Acceptance	ES Acceptance	ED Acceptance	ER Indifference	EM Indifference	AT Hate	2,283	–0.288	1,142
03/05/2015	Bank #2	EP Acceptance	EL Indifference	EE Acceptance	ES Acceptance	ED Indifference	ER Acceptance	EM Indifference	AT Hate	2,410	0.280	1,038
03/09/2015	Bank #2	EP Rejection	EL Indifference	EE Acceptance	ES Rejection	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,443	0.333	1,022
03/27/2015	Bank #2	EP Indifference	EL Indifference	EE Indifference	ES Acceptance	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,132	–0.118	2,753
04/23/2015	Bank #2	EP Acceptance	EL Acceptance	EE Hate	ES Acceptance	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,257	0.141	2,120
04/29/2015	Bank #2	EP Acceptance	EL Indifference	EE Indifference	ES Acceptance	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,325	–0.827	1,042
05/08/2015	Bank #2	EP Acceptance	EL Acceptance	EE Rejection	ES Indifference	ED Acceptance	ER Acceptance	EM Indifference	AT Hate	2,305	0.400	1,142
05/16/2015	Bank #2	EP Indifference	EL Indifference	EE Rejection	ES Rejection	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,342	0	1,438
05/29/2015	Bank #2	EP Acceptance	EL Acceptance	EE Acceptance	ES Acceptance	ED Acceptance	ER Indifference	EM Indifference	AT Hate	2,258	–0.113	1,105
06/09/2015	Bank #2	EP Indifference	EL Indifference	EE Indifference	ES Acceptance	ED Acceptance	ER Indifference	EM Indifference	AT Hate	2,280	0.332	2,127
06/21/2015	Bank #2	EP Acceptance	EL Indifference	EE Indifference	ES Indifference	ED Indifference	ER Indifference	EM Acceptance	AT Hate	2,250	0.000	2,098
07/09/2015	Bank #2	EP Acceptance	EL Indifference	EE Rejection	ES Indifference	ED Acceptance	ER Indifference	EM Indifference	AT Hate	2,157	0.837	1,012
07/22/2015	Bank #2	EP Indifference	EL Acceptance	EE Acceptance	ES Indifference	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,205	–0.377	1,154
08/12/2015	Bank #2	EP Indifference	EL Acceptance	EE Indifference	ES Indifference	ED Indifference	ER Indifference	EM Indifference	AT Hate	2,026	–0.769	1,034
09/16/2015	Bank #2	EP Indifference	EL Rejection	EE Rejection	ES Indifference	ED Indifference	ER Acceptance	EM Indifference	AT Hate	1,714	0.433	2,588

Source(s): Authors' own work

Figure 3. Example of research data organization (T moments)

Table 2. Scale values of reputation variables

Scale values	Experiences	Emotions	Attitudes
1	Hate	Hate	Sell
2	Rejection	Rejection	Hold/Sell
3	Indifference	Indifference	Hold
4	Acceptance	Acceptance	Buy/Hold
5	Admiration	Admiration	Buy

Source(s): Authors' own work

experiences and emotions: hate, rejection, indifference, acceptance or admiration; and in the case of attitudes: sell, hold/sell, hold, buy/hold or buy).

To establish the relationship between reputation variables and fluctuations in share prices at the banking sector level (first research question), a descriptive analysis was performed for both scenarios. In addition, to answer the second research question, which involves identifying repeated patterns in these relationships, multiple linear regression was used, also applied to both scenarios. Multiple linear regression is a statistical technique designed to analyze why things happen and identify the main explanations for a phenomenon. In our study, multiple linear regression allows us to understand how reputation variables are related and their influence on share prices.

In summary, we have developed the methodology phases shown in Figure 4. F4

4. Results

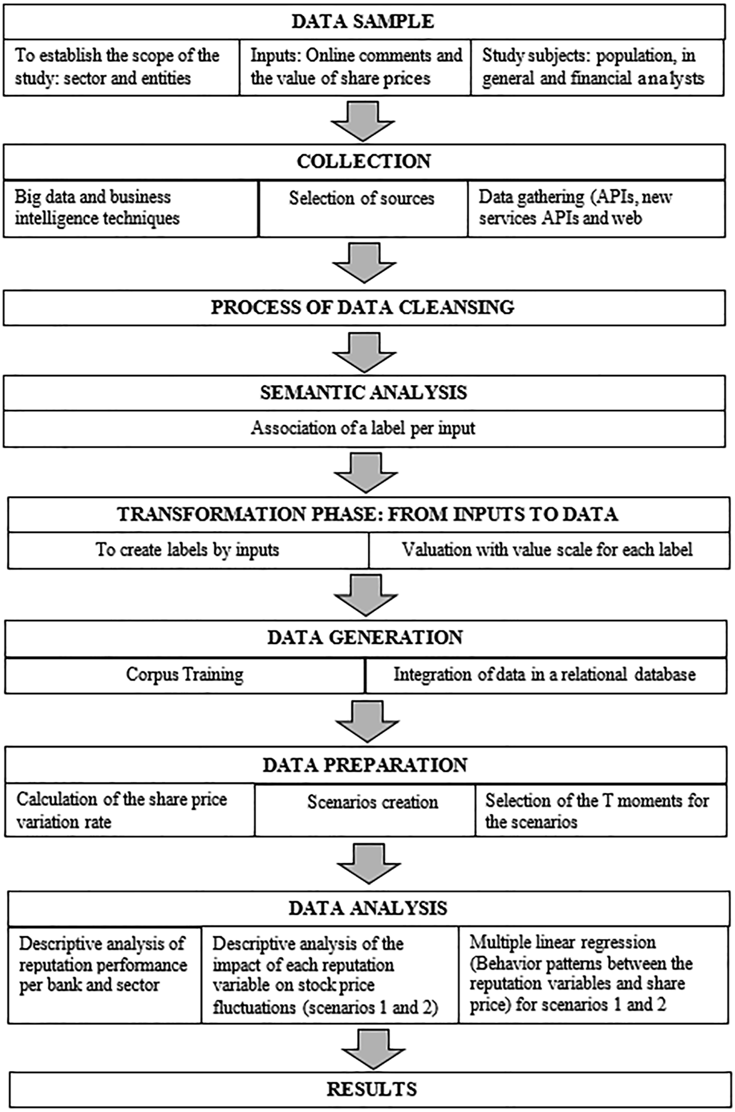
4.1 Description of the reputation performance of the banking sector in the online environment

In Figure 5, the result of performing the calculation to extract reputation (ORI) from online sources, we can see the values corresponding to each of the reputation variables for each bank and for the sector, this being the average for the seven banks analyzed. As regards the labels (admiration, acceptance, indifference, reject, hate) for the different experiences with the banks, the most positive opinions were related to EE and LE, which have been labeled as acceptance and admiration. In contrast, WE, PE, and FE were more negative for the banks. An intermediate value was obtained for the general emotion inspired by the banking sector, albeit with notable differences between banks. Particularly striking was the case of bank 5, which had by far the worst performance for experiences and emotions alike. However, for attitudes, measured as analysts' recommendations on the advisability of buying its shares, the same bank exceeded the scores obtained by banks 2, 3, and 6. F5

4.2 Impact of each reputation variable on fluctuations in share prices in the banking sector

To understand the impact of each of the reputation variables on share prices at the sector level (RQ1), this is analyzed considering the two scenarios discussed: (1) general, considering all the days of the year in which there was a quote; and (2) the days for which the share price variations were more notable and above average (what we have called T moments).

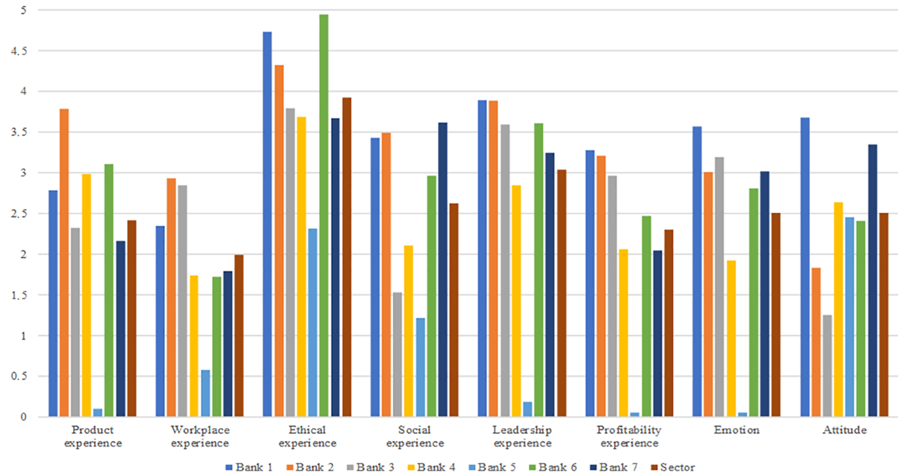
For both scenarios, the situations that can be considered logical given the descriptive analysis are shown in italic (see Tables 3 and 4). A situation is considered logical when on the days when the information observed on the Internet about a reputational variable was positive (or negative), the share price variation ended up being positive (or negative). In fact, in the second scenario, since only the T moments are considered, this requires those days to have shown even greater increases (or decreases) than the preceding days. Obviously, this concept of logic cannot be applied to situations where the comments were indifferent, or the attitude was simply one of maintenance or misinformed. T3, 4



Source(s): Authors' own work

Figure 4. Phases of methodology

Considering the general scenario, it can be observed that for all variables, the intermediate evaluation levels (acceptance and rejection) are identified with both increases and decreases, which would indicate that these levels of the reputational variables are not clearly related to price variations. Only four variables (PE, WE, EE, SE) behave logically, although there are no significant differences in the number of days with comments leading to increases versus decreases in the share price. In the case of Attitude, the days that appear at "Hold/Sell" and "Sell" recommendation, produce, albeit not clearly, logical relationships with the positive behavior, respectively, of the shares (i.e. increases in stock prices).



Source(s): Authors' own work

Figure 5. Performance of the reputation variables of the banking sector

Interestingly, in the scenario of T moments, all positive "acceptance" experiences exhibited an "inverse rule" of behavior. That is, they were associated with a downward trend in share prices. In contrast, the strongest logical relationships were associated with negative sentiments. Specifically, for all the reputational variables, days when the sentiment was "rejection" corresponded with a fall in share prices and, for four of the seven variables (PE, WE, EE, LE), this negative sentiment was associated with falls in share prices when it related to "hate" sentiments. Finally, it can be observed that for the variables (PE, WE, LE) there is a logical relationship, both for increases and decreases in share price.

Lastly, in the case of the attitude variable (Table 4), the recommendations associated with hold ("Buy/Hold", "Hold" and "Hold/Sell") lead to significant declines in prices.

4.3 Behavior patterns between the reputation variables and share price

To answer the second research question, in order to find out how the reputation variables explain the change in share price, a multiple linear regression analysis study was carried out for the two scenarios (considering, on the one hand, all the days on which there was a stock market price and, on the other, only the days on which the shares showed a significant rise or fall, the so-called T moments). As shown in Table 5, in scenario 1, the reputation variables that best explain price variability are EE, EM, SE, LE, EM and AT, with special emphasis on AT, EM and EE due to their high coefficients and significance. Opinions on SE and, to a lesser extent, on Leadership Experience also predict variations in share price. T5

To be more precise about the patterns of behavior between the reputational variables and share price, we carried out a second analysis to identify those that best explain the change in price during T moments. That is, the price on the next day was taken as the dependent variable in the subperiods for which significant changes in prices were recorded (rises and falls above the average). As shown in Table 6, the reputation variables that most influence the prices that fluctuate the most during T moments are Attitude and Emotion. Opinions on Ethics and, to a lesser extent, on Product also predict significant variations in share price. T6

However, in both scenarios, even though the model as a whole is significant, Table 5 [F-Statistic = 17.42; $p < 2.2e-16$] and Table 6 [F-Statistic = 17.66; $p < 2.2e-16$] present a moderate fit, with the reputation variables explaining 27.9% (Table 5) and 28.13% (Table 6) of the variation in prices, respectively.

Table 3. Impact of reputation variables on the share price variation rate of the banking sector (year 2015)

Variable	Category	Days with the share price variation rate Positive	Negative
PE	*Admiration*	*4*	2
	Acceptance	387	388
	Indifference	964	961
	Rejection	61	*64*
	Hate	*2*	2
WE	Admiration	*7*	10
	Acceptance	*575*	541
	Indifference	733	753
	Rejection	97	*104*
	Hate	6	*9*
EE	Admiration	15	15
	Acceptance	*243*	221
	Indifference	581	622
	Rejection	514	497
	Hate	8	*62*
SE	Admiration	25	27
	Acceptance	*639*	616
	Indifference	667	682
	Rejection	85	*88*
	Hate	2	*4*
LE	*Admiration*	*10*	4
	Acceptance	423	438
	Indifference	914	910
	Rejection	69	63
	Hate	2	2
FE	Admiration	12	14
	Acceptance	*362*	355
	Indifference	1,009	1,007
	Rejection	35	*41*
	Hate	0	0
EM	Admiration	0	0
	Acceptance	125	135
	Indifference	1,289	1,273
	Rejection	4	*9*
	Hate	0	0
AT	Buy	37	37
	Buy/Hold	16	17
	Hold	36	38
	Hold/Sell	*14*	11
	Sell	*32*	27
	Uninformed	1,586	1,597

Note(s): The values are expressed for the number of days on which each situation occurred. The situations considered logical appear in italic
Source(s): Authors' own work

5. Discussion

This study shows that EE and LE are those that generate the most favorable comments, which in turn improve the reputation of the banking sector measured in online contexts. In contrast, WE and PE are perceived most negatively by the general population on the Internet. This result shows that studies that focus on analyzing banking reputation through online protests do not offer a true image of the reputation in these contexts, as not all the variables that explain reputation or its values are treated with the same importance. Furthermore, the fact that the variables with the largest number of negative comments are WE and those that refer to PE

Table 4. Impact of reputation variables on the share price variation rate of the banking sector (T moments)

Variable	Category	Days with the share price variation rate Positive	Negative
PE	*Admiration*	*1*	0
	Acceptance	39	43
	Indifference	115	152
	Rejection	6	*12*
	Hate	0	*1*
WE	*Admiration*	*1*	0
	Acceptance	65	70
	Indifference	87	119
	Rejection	8	*18*
	Hate	0	*1*
EE	Admiration	1	1
	Acceptance	21	34
	Indifference	67	94
	Rejection	64	*70*
	Hate	8	*9*
SE	Admiration	0	5
	Acceptance	70	95
	Indifference	80	94
	Rejection	10	*14*
	Hate	1	0
LE	*Admiration*	*2*	0
	Acceptance	45	56
	Indifference	108	141
	Rejection	6	*10*
	Hate	0	*1*
FE	Admiration	0	3
	Acceptance	35	51
	Indifference	125	145
	Rejection	1	*9*
	Hate	0	0
EM	Admiration	0	0
	Acceptance	14	20
	Indifference	147	185
	Rejection	0	*3*
	Hate	0	0
AT	Buy	4	4
	Buy/Hold	1	*2*
	Hold	4	*10*
	Hold/Sell	*2*	1
	Sell	2	2
	Uninformed	171	219

Note(s): The values are expressed in the number of days on which each situation occurred. The situations that are considered logical appear in italic
Source(s): Authors' own work

confirms the trend shown by the general population to display their discontent on social media and write negative comments on aspects that affect them in a personal and direct way (Hu and Tripathi, 2016).

We can now examine the impact that each reputational variable has on fluctuations in share prices at the banking sector level. In the case of the general scenario, while price increases or decreases are not caused by reputational variables with an intermediate sentiment (acceptance and rejection), in the scenario of T moments we do observe interesting relationships. In addition, although positive experiences of "acceptance" were found with an "inverse rule"

Table 5. Multiple linear regression analysis results related to share prices (all quotes for the year 2015)

Variables	Estimate	Std. error	t-value	Pr(>|t|)
(Intercept)	−1.70143	0.65926	−2.581	0.00992**
PE	0.05820	0.06697	0.869	0.38492
WE	−0.08295	0.05004	−1.658	0.09751
EE	0.14265	0.03559	4.008	6.32e−05***
SE	0.10254	0.04859	2.110	0.03494*
LE	0.13984	0.06163	2.269	0.02336*
FE	−0.14032	0.07429	−1.889	0.05906
EM	0.86528	0.14033	6.166	8.32e−10***
AT	0.56908	0.02689	21.160	<2e−16***

Residual Standard error = 2.102	Multiple R-squared = 0.279	Adjusted R-squared = 0.263	F-statistic = 17.42	Df 360	p-value = <2.2e−16

Note(s): Signif. codes: 0 “***” 0.001 “**” 0.01 “*” 0.05 “.” 0.1 “ ” 1
Source(s): Authors’ own work

Table 6. Multiple linear regression analysis results related to share prices (T moments)

Variables	Estimate	Std. error	t-value	Pr(>|t|)
(Intercept)	−2.9321387	1.3872076	−2.114	0.035227*
PE	0.3331093	0.1383371	2.408	0.016543*
WE	−0.1496660	0.1131502	−1.323	0.186766
EE	0.2351934	0.0801961	2.933	0.003574
SE	−0.0004085	0.1041925	−0.004	0.996874
LE	0.0914726	0.1352898	0.676	0.499395
FE	−0.1582806	0.1616603	−0.979	0.328189
EM	1.0471371	0.2983235	3.510	0.000505
AT	0.5708933	0.0629169	9.074	<2e−16***

Residual standard error = 2.095	Multiple R-Squared = 0.2813	Adjusted R-squared = 0.2654	F-statistic = 17.66	Df 361	p-value = <2.2e−16

Note(s): Signif. codes: 0 “***” 0.001 “**” 0.01 “*” 0.05 “.” 0.1 “ ” 1
Source(s): Authors’ own work

of behavior, the best logical relationships were with negative feelings. Along these same lines, the work of Gómez-Carrasco and Michelon (2017) reaches the same conclusion that negative comments are related to a fall in share price, although that study is more restrictive because they only focus on online protests. In our case, on days when PE, WE, EE and LE were “hate,” there were falls in the share prices, i.e. the rate of change in the price was negative.

In relation to the behavioral patterns between all the reputational variables and share price, the variables AT, EM and EE are those that best explain the variability of today’s prices. Similarly, when we look at the T moments in which large fluctuations in share prices occur, the reputational variables that best explain these are again the same.

As far as attitude is concerned, our results seem logical since analysts, as experts, are usually listened to by investors. Given that in our study, attitudes are expressed by financial analysts and experiences and emotions by the general public and customers, this result is in line with those of the studies, in both offline and online environments, which suggest that certain analysts do not necessarily make their recommendations based on what users say about their experiences and emotions (Bollen *et al.*, 2011). One possible explanation for this is that although analysts use online publications and comments to take the pulse of public opinion about the company's activity, they also consider other, more specialized sources of information when making their recommendations (Klimczak and Dynel, 2018). In other words, they interact with and are influenced by other public sources (Crane *et al.*, 2008) such as the messages coming from powerful institutions (Toubiana and Zietsma, 2017).

For the relationship between the emotional aspect and share price, our results lie halfway between the two opposing schools of thought. On the one hand, positive comments are often not associated with high share prices, in keeping with some authors (Barakat *et al.*, 2019; Bollen *et al.*, 2011; Duan *et al.*, 2021; Hu and Tripathi, 2016),. However, on the other hand, very negative comments reflecting a feeling of hate were often associated with a fall in share price, in keeping with other studies (Barakat *et al.*, 2019; Duan *et al.*, 2021). One explanation for the absence of clearer results that were generalizable to positive comments and all the reputation dimensions may be that, as Yang *et al.* (2015) have shown, the emotion identified from posts aimed at the financial community is more closely associated with share performance than that extracted from posts aimed at the general public, as was the case in our study.

The results found in relation to the significance of EE are in keeping with those found in other studies based on offline and online data (Rojas-de-Gracia *et al.*, 2021; van Den Broek *et al.*, 2017). Similarly, other significant variables explain the behavior of share prices, albeit to a lesser extent. SE and LE influence the change in the price (general scenario). This can be considered an indication of the importance of these reputational variables, as other authors have already found in the offline context (Dollinger *et al.*, 1997; van Riel and Fombrun, 2007). In addition, PE influences significant variations in prices (at T moments), confirming the importance of this variable in the value of an organization, in keeping with the work of Luo and Zhang (2013).

As shown in the results, these variables vary with the scenario analyzed. Thus, while all reputational variables except PE influence the change in price, to a greater or lesser extent, at T moments, when large price increases or decreases occur, only four variables exert an influence, with WE, LE, SE and FE not showing such an effect. One of these variables, PE, was not significant in the general scenario. However, despite finding highly significant common variables in both scenarios (AT, EM and EE), the percentage fit of the model (28.13%) suggests that price variations are influenced by other variables in addition to reputational ones, which is a logical result.

6. Conclusions

Our study goes one step further than the existing literature on the impact of reputation on share price, in terms of both its results and methodology. It is the first to analyze the impact of the banking sector's reputation, measured in the online environment, on share price fluctuations, considering its multidimensional construct. Compared to other studies using online data that simply conduct an analysis of emotions following a unidimensional approach, our study offers a more exhaustive measurement of reputation. This is partly because it considers all of the construct's variables mentioned in the literature as being necessary for its measurement, and partly because of its multistakeholder approach.

As a result, our methodological design considers three key aspects. First, the classification of reputational data is not arbitrary; the semantic analysis of online comments is in keeping with the variables and dimensions of the theoretical construct of reputation. Second, the intrinsic nature of reputation is reflected, namely its long-term cumulative characteristic. For this, the study includes comments over an entire year and in an aggregated way. Furthermore, following the recommendation of Etter *et al.* (2019) regarding the need to examine more plural and democratic debates, our study includes a broader range of sources for the main public opinions that could have an influence on variations in share prices in the banking sector: the general population and financial analysts.

Our methodological contribution is enhanced by establishing a robust process based on the use of big data and BI techniques. These techniques have proven to be ideal for extracting a large amount of qualitative data on all the reputational variables from the online environment and transforming this into quantitative data. This allows us to work with large data sets and conduct a rigorous analysis of the effects and relationships between each reputational variable and share price, resolving the measurement difficulties suggested by Ravasi *et al*. (2019).

This study has allowed us to understand how each of the reputational variables affects fluctuations in share prices, in the online environment, in the two scenarios: for all days of the year and at T moments, that is, when large price variations occur. We mainly find interrelationships in an extremely negative online context. In other words, in both scenarios, when the general population's experiences and emotions are very negative, and the attitude of analysts is to hold or sell shares, the share price variation rate for the banking sector falls.

When we look at the behavioral patterns between the reputation construct and all its variables and share price, AT, EM and EE emerge as the only variables with a clear impact on the variation in share price at a sector level. These variables have more of an influence in negative cases than in positive ones. In the same way, albeit to a lesser extent, the leadership and social variables also affect share price. In addition, for large fluctuations in the price at T moments, the product experience will also have an impact.

Lastly, our results suggest that when we talk about the reputation of the banking sector in these contexts with different audience types, in our case the general public and financial analysts, there is no influence of one type of stakeholder on the perceptions of the other. The only correlation observed is when experiences and emotions are very negative.

The main practical implication of our findings is that given the correlations between posts on different aspects of bank performance, banks should bear in mind that negative posts on one aspect of a bank will probably prompt subsequent negative posts on other aspects. This reaction is further compounded by the snowball effect found in the online environment (Bekkers *et al*., 2011), indicating that managers should be on the alert in order to take rapid action to manage their CR. Monitoring EE, AT and EM related comments on the banks, and to a lesser extent LE and PE related ones, is therefore of great importance due to their impact on the fluctuation of the share price. Regardless of the topic the public is expressing a view about online, those in charge can find extremely negative opinions among public opinion, which is a warning indicator of behavior related to a variation in share price.

This study has a limitation common to much of the research using internet publications: although automated semantic analysis is effective and widely accepted (Liu and Li, 2019), there is still a minimum classification error. In the same way that studies using offline information miss online data, our study faced the inverse limitation. To overcome this, future research should combine both offline and online data. Additionally, replicating this study in countries with different banking cultures would enhance our understanding of banking reputation and its impact on global markets. Finally, to improve the fit of the model, future research could enrich the results by adding new variables on relevant topics extracted from online debates involving public opinion. Furthermore, based on our results, it would be interesting to conduct a more detailed study of specific scenarios, comparing the behavior of reputational variables with respect to small and large price variations. Another element that could be analyzed is whether there is an inverse influence, i.e. whether price variations in turn influence online comments over subsequent days. Finally, new hypotheses could be proposed to study whether the effects of reputational variables on price could be subject to a lag, meaning that their influence on price may not be seen on the next day but rather on subsequent days. For all these suggestions, it would also be worth considering distinct behavior depending on the reputational variable we are talking about or depending on the valence labels involved.

References

Ajzen, I. and Fishbein, M. (2005), “The influence of attitudes on behavior”, in Albarracín, D., Blair, T.J. and Mark, P.Z. (Eds), *The Handbook of Attitudes*, Lawrence Erlbaum, Mahwah, pp. 173-222.

Barakat, A., Ashby, S., Fenn, P. and Brice, C. (2019), “Operational risk and reputation in financial institutions: does media tone make a difference?”, *Journal of Banking and Finance*, Vol. 98, pp. 1-24, doi: 10.1016/j.jbankfin.2018.10.007.

Batrinca, B. and Treleaven, P.C. (2015), “Social media analytics: a survey of techniques, tools and platforms”, *AI and Society*, Vol. 30 No. 1, pp. 89-116, doi: 10.1007/s00146-014-0549-4.

Bekkers, V., Beunders, H., Edwards, A. and Moody, R. (2011), “New media, micromobilization, and political agenda setting: crossover effects in political mobilization and media usage”, *The Information Society*, Vol. 27 No. 4, pp. 209-219, doi: 10.1080/01972243.2011.583812.

Bollen, J., Mao, H. and Zeng, X. (2011), “Twitter mood predicts the stock market”, *Journal of Computational Science*, Vol. 2 No. 1, pp. 1-8, doi: 10.1016/j.jocs.2010.12.007.

Bravo-Urquiza, F. and Moreno-Ureba, E. (2021), “Does compliance with corporate governance codes help to mitigate financial distress?”, *Research in International Business and Finance*, Vol. 55, 101344, doi: 10.1016/j.ribaf.2020.101344.

Cakra, Y.E. and Trisedya, B.D. (2015), “Stock price prediction using linear regression based on sentiment analysis”, in *ICACSIS 2015*, IEEE, USA, pp. 147-154.

Carè, R., Fatima, R. and Lèvy, N. (2024), “Assessing the evolution of banking reputation literature: a bibliometric analysis”, *International Journal of Bank Marketing*, Vol. 42 No. 5, pp. 1059-1091, doi: 10.1108/ijbm-07-2023-0417.

Casado-Molina, A.M., Ramos, C.M.Q., Rojas-de-Gracia, M.M. and Peláez, J.I. (2020), “Reputational Intelligence: innovating brand management through social media data”, *Industrial Management Systems*, Vol. 120 No. 1, pp. 40-56, doi: 10.1108/imds-03-2019-0145.

Chau, M. and Chen, H. (2003), “Comparison of three vertical search spiders”, *IEEE Computer*, Vol. 36 No. 5, pp. 56-62.

Crane, A., McWilliams, A., Matten, D., Moon, J. and Siegel, D.S. (2008), *The Oxford Handbook of Corporate Social Responsibility*, Oxford University Press, New York, NY.

Dayan, R., Heisig, P. and Matos, F. (2017), “Knowledge management as a factor for the formulation and implementation of organization strategy”, *Journal of Knowledge Management*, Vol. 21 No. 2, pp. 308-329, doi: 10.1108/jkm-02-2016-0068.

De Castro, G.M., Navas, J.E. and López, P. (2006), “Business and social reputation: exploring the concept and main dimensions of corporate reputation”, *Journal of Business Ethics*, Vol. 63 No. 4, pp. 361-370, doi: 10.1007/s10551-005-3244-z.

Dollinger, M.J., Golden, P.A. and Saxton, T. (1997), “The effect of reputation on the decision to joint venture”, *Strategic Management Journal*, Vol. 18 No. 2, pp. 127-140, doi: 10.1002/(sici)1097-0266(199702)18:2<127::aid-smj859>3.0.co;2-h.

Dowling, G. (2001), *Creating Corporate Reputations*, Oxford University Press, Oxford.

Duan, Y., Liu, L. and Wang, Z. (2021), “COVID-19 sentiment and the Chinese stock market: evidence from the official news media and sina weibo”, *Research in International Business and Finance*, Vol. 58, 101432, doi: 10.1016/j.ribaf.2021.101432.

Escudero, J. (2017), “El gráfico que resume 10 años de fusiones de bancos y cajas en el sector financiero español”, available at: https://www.elconfidencial.com/empresas/2017-06-07/grafico-fusiones-sector-bancario-espana-banco-popular-santander_1395273/ (accessed 20 June 2022).

Etter, M., Colleoni, E., Illia, L., Meggiorin, K. and D’Eugenio, A. (2018), “Measuring organizational legitimacy in social media: assessing citizens’ judgments with sentiment analysis”, *Business and Society*, Vol. 57 No. 1, pp. 60-97, doi: 10.1177/0007650316683926.

Etter, M., Ravasi, D. and Colleoni, E. (2019), “Social media and the formation of organizational reputation”, *Academy of Management Review*, Vol. 44 No. 1, pp. 28-52, doi: 10.5465/amr.2014.0280.

Fombrun, C.J., Ponzi, L.J. and Newburry, W. (2015), “Stakeholder tracking and analysis: the RepTrak® system for measuring corporate reputation”, *Corporate Reputation Review*, Vol. 18 No. 1, pp. 3-24, doi: 10.1057/crr.2014.21.

Gomez-Carrasco, P. and Michelon, G. (2017), “The power of stakeholders’ voice: the effects of social media activism on stock markets”, *Business Strategy and the Environment*, Vol. 26 No. 6, pp. 855-872, doi: 10.1002/bse.1973.

Herrera, F. and Herrera-Viedma, E. (2000), “Linguistic decision analysis: steps for solving decision problems under linguistic information”, *Fuzzy Sets and Systems*, Vol. 115 No. 1, pp. 67-82, doi: 10.1016/s0165-0114(99)00024-x.

Hu, T. and Tripathi, A. (2016), “Impact of social media and news media on financial markets”, *Paper Presented at the Thirty Seventh ICIS*, Dublin, 11-14 December.

Kim, S.H. and Kim, D. (2014), “Investor sentiment from internet message postings and the predictability of stock returns”, *Journal of Economic Behavior and Organization*, Vol. 107, pp. 708-729, doi: 10.1016/j.jebo.2014.04.015.

Klimczak, K.M. and Dynel, M. (2018), “Evaluation markers and mitigators in analyst reports in light of market response to stock recommendations”, *International Journal of Business Communication*, Vol. 55 No. 3, pp. 310-337, doi: 10.1177/2329488417738082.

Liu, X. and Li, Z. (2019), “Grouping tourist complaints: what are inbound visitors’ problems with Chinese destinations?”, *Asia Pacific Journal of Tourism Research*, Vol. 24 No. 4, pp. 348-364, doi: 10.1080/10941665.2018.1564682.

Luo, X. and Zhang, J. (2013), “How do consumer buzz and traffic in social media marketing predict the value of the firm?”, *Journal of Management Information Systems*, Vol. 30 No. 2, pp. 213-238, doi: 10.2753/mis0742-1222300208.

Ponzi, L.J., Fombrun, C.J. and Gardberg, N.A. (2011), “ReptrakTM pulse: conceptualizing and validating a short-form measure of corporate reputation”, *Corporate Reputation Review*, Vol. 14 No. 1, pp. 15-35, doi: 10.1057/crr.2011.5.

Rantanen, A., Salminen, J., Ginter, F. and Jansen, B.J. (2020), “Classifying online corporate reputation with machine learning: a study in the banking domain”, *Internet Research*, Vol. 30 No. 1, pp. 45-66, doi: 10.1108/intr-07-2018-0318.

Ravasi, D., Etter, M. and Colleoni, E. (2019), “Why would the rise of social media increase the influence of traditional media on collective judgments?”, *Academy of Management Review*, Vol. 44 No. 1, pp. 222-226, doi: 10.5465/amr.2018.0065.

Rojas-de-Gracia, M.M., Casado-Molina, A.M. and Alarcón-Urbistondo, P. (2021), “Relationship between reputational aspects of companies and their share price in the online environment”, *Technology in Society*, Vol. 64, 101500, doi: 10.1016/j.techsoc.2020.101500.

Sánchez, E. (2016), “Evolución del sector de la banca”, available at: https://www.elmundo.es/grafico/economia/2016/06/11/575c582d46163ffc138b4571.html (accessed 19 June 2022).

Scherer, K.R. (2005), “What are emotions? And how can they be measured?”, *Social Science Information*, Vol. 44 No. 4, pp. 695-729, doi: 10.1177/0539018405058216.

Schwaiger, M. (2004), “Components and parameters of corporate reputation. An empirical study”, *Schmalenbach Business Review*, Vol. 56 No. 1, pp. 46-71, doi: 10.1007/bf03396685.

Shoemaker, P.J. and Reese, S.D. (2013), *Mediating the Message in the 21st Century: A Media Sociology Perspective*, Routledge, New York, NY.

Tobergte, D.R., Curtis, S., Repasky, M. and Beuming, T. (2013), “Improving neural networks with Dropout”, *Journal of Chemical Information and Modeling*, Vol. 53 No. 9, pp. 1689-1699, doi: 10.1021/ci400128m.

Toubiana, M. and Zietsma, C. (2017), “The message is on the wall? Emotions, social media and the dynamics of institutional complexity”, *Academy of Management Journal*, Vol. 60 No. 3, pp. 922-953, doi: 10.5465/amj.2014.0208.

van Den Broek, T., Langley, D. and Hornig, T. (2017), "The effect of online protests and firm responses on shareholder and consumer evaluation", *Journal of Business Ethics*, Vol. 146 No. 2, pp. 279-294, doi: 10.1007/s10551-017-3653-9.

van Riel, C.B. and Fombrun, C.J. (2007), *Essentials of Corporate Communication: Implementing Practices for Effective Reputation Management*, Routledge, New York, NY.

Yang, S.Y., Mo, S.Y. and Liu, A. (2015), "Twitter financial community sentiment and its predictive relationship to stock market movement", *Quantitative Finance*, Vol. 15 No. 10, pp. 1637-1656, doi: 10.1080/14697688.2015.1071078.

Zaby, S. and Pohl, M. (2019), "The management of reputational risks in banks: findings from Germany and Switzerland", *Sage Open*, Vol. 9 No. 3, 2158244019861479, doi: 10.1177/2158244019861479.

Further reading

AQ: 9

Baselga-Pascual, L., Trujillo-Ponce, A., Vähämaa, E. and Vähämaa, S. (2018), "Ethical reputation of financial institutions: do board characteristics matter?", *Journal of Business Ethics*, Vol. 148 No. 3, pp. 489-510, doi: 10.1007/s10551-015-2949-x.

Forcadell, F.J. and Aracil, E. (2017), "European banks' reputation for corporate social responsibility", *Corporate Social Responsibility and Environmental Management*, Vol. 24 No. 1, pp. 1-14, doi: 10.1002/csr.1402.

Karjaluoto, H., Mäkinen, H. and Järvinen, J. (2016), "A firm's activity in Social Media and its relationship with corporate reputation and firm performance", in D'Ascenzo, F., Magni, M., Lazazzara, A. and Za, S. (Eds), *Blurring the Boundaries through Digital Innovation*, Springer, Cham, pp. 161-172.

Ray, S. and Turi, R.H. (1999), "Determination of number of clusters in k-means clustering and application in colour image segmentation", *Proceedings of the 4th IAPR Conference*, New Delhi, Narosa Publishing House, pp. 137-143.

About the authors

Ana-María Casado-Molina earned her PhD degree from University of Málaga. She is a Professor of the Department of Economics and Business Administration, University of Málaga (Spain). Her research focuses on the study of intangible assets, mainly corporate reputation and brand on offline and online environments. Ana-María Casado-Molina is the corresponding author and can be contacted at: acasado@uma.es AQ: 6

María-Mercedes Rojas-de-Gracia earned her PhD degree from the University of Málaga. She is a professor of the Department of Economics and Business Administration, University of Málaga (Spain). Her research focuses on the study of consumer decision making and stakeholder behavior on social media.

Pilar Alarcón-Urbistondo earned her PhD degree from the University of Málaga. She is a professor of the Department of Economics and Business Administration, University of Málaga (Spain). She is working on the consumer and stakeholder behavior, the social image of the tourism destinations, and brand image.

Célia M.Q. Ramos earned her PhD degree from University of Algarve. She is a Professor of the Department of Technology and Information Systems of the School of Management, Hospitality and Tourism, University of Algarve (Portugal). Her research focuses on the study of business intelligence techniques applied to analyze data extracted from social media.