

Article

The Impact of YouTube on Present and Future Firm Value: Using Unstructured Text Analysis

Hyung Jong Na ¹, Yong Ha Kim ² and Hyun Jin Jo ^{3,*}¹ Department of Accounting and Taxation, Semyung University, Jecheon 27136, Republic of Korea² Department of Financial Economics, Soongsil University, Sangdo-ro 369, Dongjak-gu, Seoul, 06978, Republic of Korea³ Department of Big Data Analytics, KyungHee University, 26, Kyungheedaero, Dongdaemun-gu, Seoul 02447, Republic of Korea

* Correspondence: rkwhrdb95@khu.ac.kr; Tel.: +82-10-9040-8146

Abstract: This study aims to provide research results through empirical analysis on how customers' reactions on social media affect the present and future value of a company. This research selected Korean KOSPI-listed companies that actually own and operate YouTube channels, and collected data through text mining the comments on YouTube videos with high views. In addition, the TF-IDF value was calculated, keywords were extracted, and keywords were classified into three groups through topic modeling. The characteristics of the three groups could be transformed into a "current-oriented topic" as advertising promotion content focused on fun or interest; a "future-oriented topic" as critical content pointing out problems, and a "neutral topic" as content of a neutral attitude toward companies. This study uses a regression analysis model to perform an empirical analysis by setting a company's YouTube-related variable as an independent variable and setting a company's current value and future value-related variable as a dependent variable. The results of this research are as follows. First, this paper found that companies that directly operate and manage YouTube accounts currently have lower corporate value than those that do not. Second, this study also found that companies which directly operate and manage YouTube accounts have higher future corporate value than those that do not. Third, the results showed that if a customer simply mentions interesting content or advertising/promotion-related content through corporate YouTube comments, the current corporate value may be improved in the short term, but in the long term, it has a negative effect on future corporate value. Fourth, the results of this research also presented that if a customer criticizes a company or points out a company's problems through YouTube comments, the current corporate value decreases due to damage to the company's image, but it was found that the future corporate value increases. Fifth, this paper found that neutral content, not just for fun and interest, nor for constructive criticism or dissatisfaction with the company, was not related to the company's current and future corporate value. The contributions and expected effects of this paper are as follows. First of all, this paper provides useful information through research results which shows that companies are more advantageous in improving future corporate value from a long-term perspective by strategically operating social media directly. In addition, the research results of this study objectively demonstrated through YouTube channels that it is more helpful for companies in the long run to respond well to customer complaints and negative opinions, and to implement policies that continuously manage customer opinions. Finally, the research method used in this paper, that is, the research methodology that conducted empirical analysis through quantification of unstructured tax data, is expected to provide guidelines for many scholars to expand the scope of data available for empirical research in the future.

Citation: Na, H.J.; Kim, Y.H.; Jo, H.J. The Impact of YouTube on Present and Future Firm Value: Using Unstructured Text Analysis. *Sustainability* **2023**, *15*, 4346. <https://doi.org/10.3390/su15054346>

Academic Editor: Gaetano della Corte

Received: 23 January 2023

Revised: 24 February 2023

Accepted: 24 February 2023

Published: 28 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: YouTube; YouTube comments; text mining; current firm value; future firm value

1. Introduction

The development of social media has changed the business environment of companies [1]. In order for a company to continue to grow, it is necessary to quickly grasp the changing business environment and improve its ability to adapt to it. This study will provide important information on how companies should cope with the spread of social media.

In the recent business environment, consumers' opinions are immediately delivered to companies, and public opinion about companies is formed in real time. Accordingly, companies must respond to consumer reactions and manage consumer public opinion [2].

Global social media such as Facebook, Twitter, Instagram, and YouTube have accounts not only for individuals but also companies who interact with consumers [3]. Recently, people tend to use and prefer YouTube, which posts videos, rather than Facebook, Twitter, and Instagram, which simply post photos and writings [4].

Therefore, this study focuses on YouTube accounts among various social media. It expects that there will be a difference between the present value and the future value between a company that directly operates a YouTube account and a company that does not operate a YouTube account, and an empirical analysis will be conducted on this. Consumers say positive things about companies through social media and negative things about companies [5]. In other words, they expect that just because a company operates a YouTube account directly and communicates with consumers will not necessarily have a positive impact on current and future corporate values.

Park et al. (2016) reported empirical analysis results that show companies operating Facebook accounts have higher future corporate value than companies that do not [6]. In the short term, this does not immediately affect corporate value because it is a mixture of positive and negative reactions to companies through Facebook, but in the long run, it can be interpreted that the future firm value can increase by reflecting consumer criticism and improving the problems of the company. In addition, the results of Park et al. (2016) show that there is no significant relationship between the 'likes numbers' and 'shared numbers' on Facebook and current and future corporate values [6]. These results may have been derived because even if consumers show a positive response to the company through Facebook, it is actually an action from an advertising company to increase the number of likes or shares. In other words, it is interpreted that consumers' positive responses on social media may lack authenticity and may be simply focused on interest or fun, so there may be no significant relationship with current and future corporate value.

Therefore, this paper aims to provide research results through empirical analysis on how customers' reactions on social media affect the present and future value of a company. Comments on YouTube channels operated by a company are composed of unstructured text data, and the company's financial data are quantitative data. In other words, in order to analyze the subject of this study, it is impossible with the existing research methodology. After quantifying the comments on the YouTube channel operated by the company, empirical analysis should be performed along with the company's financial data. In other words, an innovative interdisciplinary convergence research methodology should be introduced [7,8].

Information on companies does not only contain quantitative data, but also qualitative data. However, due to the limitations of the research methodology, in the accounting and financial fields, which mainly study companies, they have mainly studied using only quantitative data and corporate financial data.

The research model used in this study differs from existing research methodologies in that it quantified unstructured text data called corporate YouTube comments and conducted empirical analysis through a regression analysis model along with corporate financial data. Namely, this paper is a convergence study that combines text mining techniques and regression analysis and performs empirical analysis by quantifying qualitative text data. The results of this study are expected to contribute to further expanding the scope of the study by presenting a methodology that quantifies text data.

This study selected Korean KOSPI-listed companies that actually own and operate YouTube channels, and collected data through text mining on comments in YouTube videos with high views. In addition, the TF-IDF value was calculated, keywords were extracted, and keywords were classified into three groups through topic modeling. The characteristics of the three groups could be transformed into a “current-oriented topic” as advertising promotion content focused on fun or interest; a “future-oriented topic” as critical content pointing out problems, and a “neutral topic” as content of a neutral attitude toward companies. This study uses a regression analysis model to perform an empirical analysis by setting a company’s YouTube-related variable as an independent variable and setting a company’s current value and future value-related variable as a dependent variable.

The composition of this paper is as follows. In Section 2, hypotheses are derived through related previous research surveys. Section 3 explains data collection and research models. Section 4 presents the results of empirical analysis, and finally, Section 5 concludes.

2. Prior researches and Development of Hypotheses

2.1. Social Media and Current Company Value

To improve a company’s sustainability in the area of business and corporate value, a company must increase its ability to quickly discover and adapt to rapidly changing business environments.

As a result of examining previous studies on the relationship between social media and current corporate value, we find that the relationship between corporate social media and current corporate value was ambiguous. Previous studies have reached conflicting conclusions on the contribution of social media activities to present corporate value [9,10].

On the positive side, some previous studies have shown that social media-based indicators are important leading indicators of current corporate value [11,12]. Social media can help companies improve their images at low marketing costs [13]. In other words, if a company actively communicates with consumers through social media and manages public opinion well, it helps to improve consumer satisfaction and loyalty [14].

On the other hand, social media can negatively affect corporate value in the short term [15,16]. Borah and Tellis (2015) report that negative reactions from customers on social media can lower corporate value [17]. Colicev et al. (2018) explains that if negative information about a company is spread through social media, it may result in a decrease in corporate value [18].

In addition, there are exceptional cases. Noise marketing can have a positive impact on present corporate value. Bigdellou et al. (2022) found that negative publicity can increase product awareness through noise marketing, which can increase purchase and sales. This phenomenon has also been confirmed in other studies [19]. For example, unrelated posts can act as catalysts, exponentially increasing the number of subscribers to social media [20]. That is, the results of emotional analysis or content analysis using social media data may show mixed reactions [21].

Some preceding studies report that social media can have a positive effect on current corporate value [22,23]. Yet, other prior literatures report that social media has a negative impact on current firm value [24,25]. Although previous studies report conflicting research results between social media and current corporate value, our researchers expect that social media is more likely to have a negative impact on current corporate value in the short term. Since the creation of social media, there have been more negative comments than positive ones. There are also some that are not sincere because some positive responses have advertising characteristics. Therefore, we make the following hypothesis.

Hypothesis 1: *A company’s social media operation will have a negative impact on its current corporate value in the short term.*

2.2. Social Media and Future Corporate Value

In modern society, companies that quickly accept consumers' opinions due to the spread of social media are competitive and sustainable. Now, in order for companies to increase their future corporate value, they must use social media well to reflect consumers' complaints and listen to their opinions. In this study, objective results for this are presented through empirical analysis.

Existing studies explain that social media contributes to the development of companies [26,27]. Since social media is a representative modern IT tool used by many companies, it is expected that a positive connection can be found between social media activities and future corporate value. In other words, consistent with the results of previous studies [28–30], we expect that companies which actively use social media will have higher future corporate value than those that do not. Yet, it may not be easy for the effect of social media to increase corporate value in a short period of time.

Consumer reactions to companies can be positive or negative. It is the comments on social media that can immediately and clearly grasp these customers' reactions [31]. Public opinion about companies through comments on social media can be positive or negative. If it is a positive public opinion, it may not actually lead to the purchase of products or services produced by companies as it is simply interest-oriented gossip [32]. On the contrary, if a company humbly accepts negative comments and improves the problems, these problems may improve in the future and ultimately develop the future value of the company.

Although all these conflicting cases can occur, the future value of companies actively communicating with customers through social media is expected to increase in the long term, as companies can sufficiently cope with consumer reactions and correct and modify for problems, unlike in the short term. Therefore, we develop the following hypothesis.

Hypothesis 2: *Corporate social media operations will have a positive impact on future corporate values in the long run.*

2.3. The Impact of Social Media Contents on the Present and Future Corporate Value

In order to conduct a more detailed analysis of the relationship between social media and corporate value, this study classifies comments from corporate YouTube channels using topic modeling techniques. As a result of classifying the contents of YouTube comments using the topic modeling technique in this study, it was possible to classify them into three contents as follows. First, it is simply the content of advertising promotion focused on fun or interest; second, it points out the problems of the company, and finally, it is neutral, not positive or negative.

If the contents of the YouTube comments are simply advertising promotion content focused on fun or interest, this may have a positive effect on the current corporate value [33]. However, in the long run, it is not expected to have a positive effect on future corporate value [34]. This is because if consumers pursue only the current fun or interest without solving the fundamental problems of the company, there is a limit to fundamentally promoting corporate value.

It may be able to take a long time to improve corporate value. If consumers point out corporate problems through YouTube comments, corporate value may decrease in the short term due to negative public opinion, but if they understand and supplement the problems pointed out by customers, the company's value will improve in the long run [35]. In other words, if the content of the YouTube comment is critical of pointing out the problems of the company, the corporate value may temporarily decrease at this point. However, if the firm responds appropriately to this, actually accepts consumer comments, and corrects and supplements the company's problems, there is a possibility that its value could increase in the long run [36]. Therefore, in this study, we anticipate that the present and future corporate values may vary depending on the specific content of corporate YouTube comments. Hence, we derive specific hypotheses as follows.

Hypothesis 3: *Current and future corporate values will differ depending on the specific content of corporate YouTube comments.*

2.4. The Role of Social Media in Sustainable Business

For the sustainable development and growth of a company, social media can play an important part in improving relationships with customers [37]. If a company operates and manages social media well on its own, the company will be able to develop further by improving the relationship between the company and customers [38]. For companies, active use of social media as a new marketing tool in modern society is essential, and if there are companies that still do not make good use of it, social media should be utilized quickly for sustainable growth.

In particular, if a customer is dissatisfied with a company and the corporate immediately accepts it and solves it, the customer's loyalty for the firm will increase [39]. Maintaining such a good relationship with customers is necessary for continuous growth from the perspective of the company [40]. On the other hand, for customers who only talk positively about the company, it is rather doubtful whether the customer's story is sincere [41]. Customers who give cold and objective advice to the company to grow are more valuable than customers who only say positive things such as flattering them [42]. Namely, the future corporate value is expected to be affected by the content of social media communicating with companies and customers on YouTube channels. Considering these things, this study derives the following detailed hypotheses.

Hypothesis 3-1: *If the content of the corporate YouTube comment is related to advertising promotion that focuses on fun or interest, the current corporate value will increase, but the future corporate value will decrease.*

Hypothesis 3-2: *If the content of the corporate YouTube comment is related to critical of pointing out the problem, the current corporate value will decrease and the future corporate value will increase.*

Hypothesis 3-3: *If the content of the corporate YouTube comment is related to the content of a neutral attitude toward the company, it will not have a significant effect on the current and future corporate values.*

3. Research Design

3.1. Sample Selection

3.1.1. Financial Data

In this paper, the financial data of the company collected to perform empirical analysis were downloaded from Kis-value and TS-2000. The sample target of this study is South Korea's KOPI-listed companies, and the sample period is 2020. The reason why the sample target was selected as a KOSPI-listed firm in South Korea is that many people actively use social media as an IT powerhouse, especially YouTube recently. In particular, South Korean corporates operate many YouTube channels and respond quickly to consumer reactions. In other words, to meet the subject of this study, many people used YouTube to actively record comments, and South Korea, which responds quickly to these consumers' reactions, was selected as a sample target.

Additionally, the reason why the sample period is a single year, 2020, is that a standard time for collecting YouTube comment data was set, and text data for YouTube comments at this time were collected. The number of YouTube contents owned by a company alone is very large, and the number of comments accordingly is countless. Due to cost and time constraints, it is not possible to collect all this vast amount of data, so we set a standard point and collected YouTube comment text data at that point.

As of the end of 2020, there were 631 KOSPI-listed companies in South Korea. Among them, companies that do not directly operate YouTube channels, firms belonging to the

financial industry, and corporates that do not have the settlement date of the end of the December were excluded from the sample. In addition, companies that could not use the necessary financial variables from the database were also excluded from the sample. After going through this sample selection procedure, the final available sample was 168 firm-year. Table 1 below summarizes the sample selection procedure used in this study.

Table 1. Sample selection criteria.

Criteria	N
South Korea's KOSPI-listed companies as of the end of 2020	631
Less: Companies without YouTube channels, firms in the financial sector, and corporates where the end of December is not the settlement date	(461)
Less: Unavailable financial data required from the database	(2)
Final observations	168

3.1.2. Text Data

The analysis of unstructured greeting data used 'R' (version 3.6.3), an open-source data analysis tool, and the data were collected as of 31 December 2020. In order to extract meaningful keywords from the collected greeting text data, only Korean, which removes stop-words such as promotional text and special characters (speech points), was used as analysis data. In the process of nounizing words, it was based on the Sejong dictionary of the 'KoNLP' package specializing in Korean text mining provided by 'R', and a dictionary that was added and produced directly to suit data analysis was used. In addition, the level of refinement of words was increased by selecting words with more than two syllables among the nounized words and integrating words such as synonyms and synonyms into one.

Term Frequency-Inverse Document Frequency (TF-IDF) was used to calculate the importance of keywords in documents from text data. TF-IDF is a weighted value calculation of how much a particular keyword in a document represents a characteristic of that document, using an analysis method that incorporates both the frequency of a particular keyword appearing in a document and the frequency of common appearance in another associated document [43]. The TF-IDF calculation method is calculated by multiplying the relative frequency of a particular word by the value associated with the inverse document frequency of the proportion of documents in which a particular word appears in the entire document. The higher the TF-IDF value, the higher the value that is significantly repeated in the document. The calculated TF-IDF values were organized in descending order of average to identify keywords with high importance among all documents, and in this study, only values with a TF-IDF value of 1.5 or more were extracted and used for analysis.

Topic modeling refers to a statistical model for discovering abstract themes (topics) of a set of documents, and representative methods include Latent Semantic Analysis (LSA) and Latent Dirichlet Allocation (LDA). First of all, LSA uses a method of reducing the dimension of Document-Term Matrix (DTM) to group close words into topics, and latent Dirichlet allocation (LDA) extracts topics by estimating the combined probability of a specific topic in a document and the probability that a word exists in a specific topic [44].

This study adopted the LDA technique for word topic among the two methods. When words are gathered, they become topics, and when topics are gathered, they become documents. Each word may represent a topic, and a set of topics becomes a document. Here, it can be said that LDA analysis is the way to find out which topic words belong to. In LDA analysis, also called topic model, topic is the same concept as segment in cluster analysis, and the biggest feature of LDA analysis is that analysis does not need to define the distance required by cluster analysis methods based on probability models,

and multiple topics can be included in one document. The algorithms of topic modeling are divided into sampling-based algorithms and variational inference, and this study adopts Gibbs Sampling, the most widely used sampling-based algorithm in LDA analysis. For LDA analysis, the LDA library in the ‘topic models’ package provided by ‘R’ was used.

The results of topic modeling depend on the number of topics and it is important to determine the appropriate number of topics in the entire document set [45], and to determine the optimal number of topics through trial and error, but to avoid overlapping between topics and ensure the level of interpretation of the topics [46]. It is also said that a small number of topics avoid overlapping between topics and ensure structural validity [47]. In this study, since many overlapping keywords began to appear when the number of groups was 3 or more, the number of topic groups was designated as 3 and the number of repetitions was designated as 1000 through hyper-parameters that can be adjusted by the user for ease of interpretation.

As a result of finally classifying the contents of YouTube comments by topic modeling, Topic 1 is ‘fun or interest-oriented advertising promotion content’, Topic 2 is ‘critical content pointing out problems’, and the last, Topic 3, is ‘content of a neutral attitude toward companies’. There are 47 final keywords for Topic 1, 43 final keywords for Topic 2, and 78 final keywords for Topic 3. The summary of the contents and number of keywords by topic is presented in Table 2.

Table 2. Number of topics classified by topic modeling.

Topic	Topic_1	Topic_2	Topic_3
Topic Contents	Advertising promotion contents focused on fun or interest	Critical contents pointing out a problem	The contents of a neutral attitude toward a company
Keyword examples	Advertising	Question	Memories
	Fun	Worry	Utilization
	Love	Development	Search
	Cheering	Future	Information
	Price	experience	Algorithms
	promotion	Planning	Country
	Model	Idea	World
	Actor	Expectation	Staff
	Broadcasting	Reality	Effort
	Purchase	Success	Reality
N	47	43	78

3.2. Descriptive Statistics and Correlation Analysis

Table 3 describes the definitions of variables used in the empirical analysis of this study. Table 4 presents the descriptive statistics of the variables used in the empirical analysis in this study. Channel is a variable that means 1 if you have a company’s YouTube channel or 0 if you do not. The average value of the channel is 0.270 and the standard deviation is 0.444. Currently, the average values of ROA and ROE, which mean corporate value, are 0.006 and -0.064, respectively, and the standard deviations are 0.089 and 0.466, respectively. In addition, the average values of TQ and MB, which mean future corporate values, are 1.690 and 2.135, respectively, and the standard deviations are 2.168 and 3.137, respectively.

Table 3. Definition of variables.

Variables	Definition
Channel	Dummy variable: 1 if the company has a YouTube channel, otherwise 0
Topic_1	TF-IDF average of keywords classified as Topic_1
Topic_2	TF-IDF average of keywords classified as Topic_2
Topic_3	TF-IDF average of keywords classified as Topic_3
ROA	Measures for current company value 1: current net income is divided by total assets
ROE	Measures for current company value 2: current net income is divided by total equity
TQ	Measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets
MB	Measures for future company value 2: market value of equity to book value of equity ratio
FORN	Foreign investors' shareholder ratio
LARG	The largest shareholder's equity ratio
SIZE	Natural logarithm of total assets
LEV	Total liabilities to total assets
BIG4	Dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0
CFO	Operation cash flow is divide by total assets
LOSS	Dummy variable: 1 if the company reports loss, otherwise 0
GRW	(This year-end sales – lagged year-end sales) ÷ lagged year-end sales

Table 4. Descriptive statistics.

Variables	N	Means	SD	Min	Q1	Medium	Q3	Max
Channel	168	0.270	0.444	0.000	0.000	0.000	1.000	1.000
ROA	168	0.006	0.089	−0.293	−0.012	0.018	0.057	0.164
ROE	168	−0.064	0.466	−3.490	−0.019	0.031	0.094	0.221
TQ	168	1.690	2.168	0.389	0.819	1.042	1.601	14.390
MB	168	2.135	3.137	0.265	0.627	1.115	2.186	18.246
FOREIGN	168	0.144	0.166	0.000	0.022	0.069	0.222	0.662
LARGE	168	0.300	0.156	0.091	0.180	0.258	0.406	0.664
SIZE	168	27.746	1.690	24.699	26.613	27.458	28.842	32.116
LEV	168	0.425	0.210	0.076	0.258	0.421	0.569	0.937
BIG4	168	0.649	0.479	0.000	0.000	1.000	1.000	1.000
CFO	168	0.054	0.087	−0.248	0.014	0.063	0.105	0.219
LOSS	168	0.298	0.459	0.000	0.000	0.000	1.000	1.000
GRW	168	−0.038	0.237	−0.728	−0.129	−0.038	0.056	0.682

Table 5 presents the correlation between the variables used in the empirical analysis of this study. Channel has a significant relationship with MB at 0.094, under 10%, but not a statistically significant relationship with the rest of ROA, ROE, and TQ. The correlation coefficient between ROA and ROE is 0.883, and a significant value is derived statistically below 1%. The correlation coefficient between MBs other than TQ is 0.866, and statistically, a significant value is derived below 1%. In addition, there is no statistically significant relationship between ROA and TQ, and the correlation coefficient between ROA and MB is −0.193, indicating a statistically significant relationship under 1%. There is no statistically significant relationship between ROE and TQ, and the correlation coefficient between ROE and MB is −0.307, which is found to have a statistically significant relationship below 1%.

Table 5. Correlation analysis results.

Variables	ROA	ROE	TQ	MB	FOREIGN	LARGE	SIZE	LEV	BIG4	CFO	LOSS	GRW	Channel
ROA	1.000 ***												
ROE	0.883 ***	1.000 ***											
TQ	0.031	−0.033	1.000 ***										
MB	−0.193 ***	−0.307 ***	0.866 ***	1.000 ***									
FORN	0.171 ***	0.152 ***	0.130 **	0.101 *	1.000 ***								
LARG	0.060	0.043	−0.087 *	−0.091 *	−0.013	1.000 ***							
SIZE	0.218 ***	0.174 ***	−0.027	−0.086 *	0.449 ***	0.111 **	1.000 ***						
LEV	−0.326 ***	−0.350 ***	−0.028	0.141 ***	−0.181 ***	0.028	0.117 **	1.000 ***					
BIG4	0.125 **	0.075	0.066	0.005	0.258 ***	0.218 ***	0.510 ***	0.023	1.000 ***				
CFO	0.483 ***	0.444 ***	0.062	−0.001	0.186 ***	0.090 *	0.155 ***	−0.081 *	0.107 **	1.000 ***			
LOSS	−0.660 ***	−0.547 ***	−0.035	0.098 *	−0.144 ***	−0.043	−0.190 ***	0.216 ***	−0.123 **	−0.348 ***	1.000 ***		
GRW	0.327 ***	0.313 ***	0.163 ***	0.088 *	0.039	−0.058	0.009	−0.098 *	0.014	0.280 ***	−0.321 ***	1.000 ***	
Channel	0.015	−0.033	0.072	0.094 *	0.227 ***	−0.029	0.313 ***	0.072	0.165 ***	0.041	−0.029	−0.025	1.000 ***

* < 0.10, ** < 0.05, *** < 0.01.

3.3. Research Model

This paper uses a regression analysis in order to test hypotheses (Figure 1). Our researchers conduct 2SLS analysis to control the endogeneity that can occur in the research model. The 2SLS analysis results expect to show more objective and robust results.

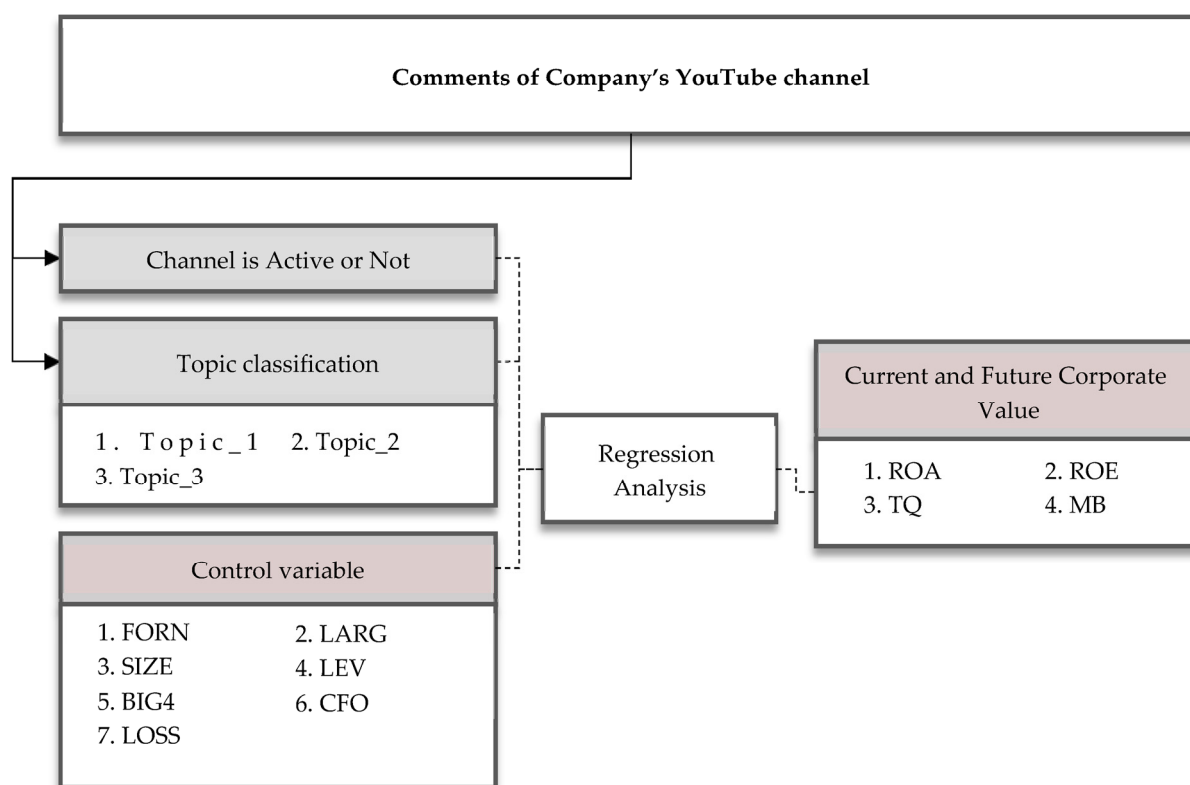


Figure 1. Research framework.

The 1st stage model of the 2SLS study model of this study is as follows:

[1st stage model]

$$\text{Channel} = \beta_0 + \beta_1 \text{FORN} + \beta_2 \text{LARG} + \beta_3 \text{SIZE} + \beta_4 \text{LEV} + \beta_5 \text{BIG4} + \beta_6 \text{LOSS} + \beta_7 \text{GRW} + \beta_8 \text{ROA} + \beta_9 \text{IND Effect} + \varepsilon \quad (1)$$

where: Channel: Dummy variable 1 if the company has a YouTube channel, otherwise 0; FORN: Foreign investors' shareholder ratio; LARG: The largest shareholder's equity ratio; SIZE: Natural logarithm of total assets; LEV: Total liabilities to total assets; BIG4: Dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: Operation cash flow is divided by total sales; LOSS: Dummy variable: 1 if the company reports loss, otherwise 0; GRW: (This year-end sales - lagged year-end sales) ÷ lagged year-end sales; ROA: Measures for current company value 1: current net income is divided by total assets; IND Effect: Dummy variables by industry to control differences in the industry to which a company belongs.

In Equation (1), Channel, dummy variable 1 if the company has a YouTube channel, otherwise 0, are employed as dependent variables. In addition, FORN [48], LARG [49], SIZE [50], LEV [51], BIG4 [52], LOSS [53], GRW [54], and ROA [55] are established as variables to control the effect of the characteristics of a company's social media channel.

To control the impact of corporate governance on social media, FORN and LARG is established in the model. SIZE is used because it can affect social media operated according to the size of the company, and LEV is used because there may be a difference in corporate social media depending on the corporate debt ratio. BIG4 is added as a control variable to the model to control the impact of accounting auditing. LOSS uses it because

it is concerned that there would be a difference in social media policies or contents from companies that reported losses, and GRW is used as a control variable in anticipation that corporate growth can affect social media activation. In addition, our researchers add the industrial dummy variable since the industries to which the companies operating social media belong are different. The reason for using ROA in the 1st stage model is to control the effect of the current corporate value on the future corporate value when the dependent variables are TQ and MB, which mean the future value of the company, in the 2nd stage. In other words, in the 2nd stage research model, when the dependent variables are ROA and ROE, which mean the present value of the company, ROA is not used as the control variable of the 2nd stage model.

In the empirical analysis model of the 2nd stage of this study, in order to test Hypothesis 1, Channel is established as an independent variable; ROA and ROE, which mean the present value of a company, are set as dependent variables. Control variables of the model are established by prior literatures [48–57]. The empirical analysis model of the 2nd stage to test Hypothesis 1 is as follows.

[2nd stage model for testing Hypothesis 1]

$$\text{ROA/ROE} = \beta_0 + \beta_1 \text{Channel} + \beta_2 \text{FORN} + \beta_3 \text{LARG} + \beta_4 \text{SIZE} + \beta_5 \text{LEV} + \beta_6 \text{BIG4} + \beta_7 \text{CFO} + \beta_8 \text{LOSS} + \beta_9 \text{GRW} + \text{IND Effect} + \varepsilon \quad (2)$$

where: ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs.

To verify Hypothesis 2, TQ and MB, which mean the future value of a company, are set as dependent variables. Additionally, Channel is established as an independent variable. The empirical model of the 2nd stage to test Hypothesis 2 is as follows.

[2nd stage model for testing Hypothesis 2]

$$\text{TQ/MB} = \beta_0 + \beta_1 \text{Channel} + \beta_2 \text{FORN} + \beta_3 \text{LARG} + \beta_4 \text{SIZE} + \beta_5 \text{LEV} + \beta_6 \text{BIG4} + \beta_7 \text{CFO} + \beta_8 \text{LOSS} + \beta_9 \text{GRW} + \text{IND Effect} + \varepsilon \quad (3)$$

where: ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷

lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs.

For the test of Hypothesis 3, Topic_1, 2, and 3 are set instead of Channel as independent variables.

Topic_1 means ‘fun or interest-oriented advertising promotion content’, Topic_2 images ‘critical content pointing out problems’, and the last Topic_3 indicators ‘content of a negative company’. As the dependent variables, ROA and ROE, which mean the present value of the company, and TQ and MB, which mean the future value of the company, were used, respectively. The empirical analysis model of the 2nd stage to verify Hypothesis 3 is as follows.

[2nd stage model for testing hypothesis 3]

$$ROA/ROE/TQ/MB = \beta_0 + \beta_1 \text{Topic}_1, 2, 3 + \beta_2 \text{FORN} + \beta_3 \text{LARG} + \beta_4 \text{SIZE} + \beta_5 \text{LEV} + \beta_6 \text{BIG4} + \beta_7 \text{CFO} + \beta_8 \text{LOSS} + \beta_9 \text{GRW} + \text{IND Effect} + \varepsilon \quad (4)$$

where: ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors’ shareholder ratio; LARG: the largest shareholder’s equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs.

4. Empirical Results

4.1. The Results of Hypotheses 1 and 2

Table 6 presents the results of empirical analysis for Hypothesis 1 and Hypothesis 2 of this paper. First of all, looking at the results of Hypothesis 1, when the dependent variable is ROA, β value of the independent variable Channel is −1.953 and the t-value is −5.377, which is statistically significant at the 1% level. When the dependent variable is ROE, β value of the independent variable Channel is −0.719 and the t-value is −4.715, which is statistically significant at the 1% level. These results, in support of Hypothesis 1, imply that firms which directly operate and manage YouTube accounts now have a lower corporate value than those that do not. Comments on the YouTube account express many negative opinions and complaints about the company. The results of this study suggest that these factors can reduce the current corporate value from a short-term perspective.

Table 6. 2SLS results (Topic_1, dependent variables ROA/ROE/TQ/MB, n = 168).

Variables	1st Stage		2nd Stage (ROA)		2nd Stage (ROE)		2nd Stage (TQ)		2nd Stage (MB)	
	β	t Value	β	t Value	β	t Value	β	t Value	β	t Value
Intercept	−2.123	−4.724	−1.657	−5.734 ***	−4.621	−5.163 ***	25.633	4.225 ***	37.056	4.027 ***
Channel			−0.719	−5.377 ***	−1.953	−4.715 ***	10.684	3.818 ***	14.909	3.512 ***
FORN	0.370	2.314 **	0.244	4.453 ***	0.655	3.857 ***	−2.421	−2.127 **	−2.341	−1.356
LARG	−0.312	−2.831 ***	−0.227	−4.965 ***	−0.593	−4.195 ***	2.96	3.11 ***	4.218	2.922 ***
SIZE	0.082	5.301 ***	0.067	6.026 ***	0.187	5.416 ***	−0.968	−4.135 ***	−1.433	−4.035 ***
LEV	−0.013	−0.130	−0.097	−6.985 ***	−0.338	−7.877 ***	0.587	1.994 **	2.119	4.746 ***
BIG4	−0.006	−0.143	−0.006	−1.073	−0.035	−1.867 *	0.204	1.658 *	0.181	0.968
CFO			0.35	9.181 ***	0.873	7.401 ***	0.426	0.511	2.49	1.97 **
LOSS	0.023	−0.079	−0.075	−11.186 ***	−0.144	−6.936 ***	−0.17	−1.127	−0.3	−1.307
GRW	0.081	1.095								
ROA	−0.098	−0.375					1.268	1.484	−3.717	−2.867 ***
IND Effect	Included		Included		Included		Included		Included	
Adj_Rsq	18.71%		57.75%		50.78%		19.10%		24.34%	
F value	3.496 ***		15.825 ***		12.187 ***		3.516 ***		4.429 ***	

ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales − lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs. * < 0.10, ** < 0.05, *** < 0.01.

Next, the results for Hypothesis 2 are as follows. When the dependent variable is TQ, β value of the independent variable Channel is 10.684 and the t-value is 3.818 which is statistically significant at the 1% level. When the dependent variable is MB, β value of the independent variable Channel is 14.909 and the t-value is 3.512, which is statistically significant at the 1% level. This is the result of supporting Hypothesis 2. In other words, it means that companies which directly operate and manage YouTube accounts have higher future corporate value than those that do not. In the short term, a company may lose its value due to customer comments or complaints, but in the long term, this study finds that if it interacts with customers through social media channels and solves their complaints, a trust relationship between the company and customers can eventually improve its future corporate value.

In addition, this study shows the results after controlling the endogeneity using the 2SLS model, and presents more objective and reliable analysis results by using multiple proxy measures, current corporate value, and future corporate value. In addition, the adjusted R-square value, which means explanatory power between variables, is also derived as an appropriate value, and the F-value, which means the validity of the research model, is also statistically significant.

4.2. The Results of Hypothesis 3

Table 7 shows the results of empirical analysis for Hypotheses 3-1. In this study, YouTube comment contents were classified into three groups by topic modeling techniques. Among them, Topic 1 is related to advertising promotion contents focused on fun or interest.

Table 7. 2SLS results (Topic_1, dependent variables ROA/ROE/TQ/MB, n = 168).

Variables	1st Stage		2nd Stage (ROA)		2nd Stage (ROE)		2nd Stage (TQ)		2nd Stage (MB)	
	β	t Value	β	t Value	β	t Value	β	t Value	β	t Value
Intercept	0.188	0.941	−0.033	−0.342	−0.977	−1.365	−3.866	−1.015	−1.217	−0.220
Topic_1			0.099	1.200	1.064	1.716 *	−2.444	−0.737	−8.356	−1.737 *
FORN	−0.020	−0.184	0.006	0.185	−0.271	−1.048	1.840	1.337	4.436	2.222 **
LARG	−0.079	−0.890	−0.011	−0.320	0.399	1.526	0.298	0.215	−2.282	−1.132
SIZE	0.066	0.578	0.004	1.028	0.045	1.639	0.180	1.217	0.091	0.425
LEV	0.114	1.251	−0.128	−4.091 ***	−1.152	−4.891 ***	0.731	0.548	4.920	2.541 **
BIG4	0.055	0.515	0.004	0.332	−0.052	−0.530	−0.675	−1.299	−0.372	−0.493
CFO			0.328	5.425 ***	0.471	1.037	1.690	0.629	6.133	1.574
LOSS	0.045	0.357	−0.098	−9.438 ***	−0.362	−4.614 ***	0.271	0.496	0.156	0.197
GRW	0.167	1.875 **								
ROA	0.329	0.404					3.760	1.048	−8.784	−1.687 *
IND Effect	Included		Included		Included		Included		Included	
Adj_Rsq	20.14%		72.25%		42.99%		25.46%		25.02%	
F value	1.504*		11.114 ***		3.929 ***		2.296 ***		2.266 ***	

ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: Operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs. * < 0.10, ** < 0.05, *** < 0.01.

When the dependent variable is ROA, β value of the independent variable Topic_1 is 0.099 and the t-value is 1.200, which is not statistically significant. When the dependent variable is ROE, β value of the independent variable Topic_1 is 1.064 and the t-value is 1.716, which is statistically significant at the 10% level.

When the dependent variable is TQ, β value of the independent variable Topic_1 is −2.444 and the t-value is −0.737 which is not statistically significant. When the dependent variable is MB, β value of the independent variable Topic_1 is −8.356 and the t-value is −1.737, which is statistically significant at the 10% level. This is the result of supporting partially Hypothesis 3-1.

These results can be interpreted as that if customers simply refer to fun or interesting content or advertising and promotion-related content through corporate YouTube comments, the current corporate value can be improved in the short term, but it has a negative impact on future corporate value in the long run. In other words, the fact that customers only respond to immediate fun or interest, advertisements, or promotions without sincerity means that they are not helpful in the long run for companies.

Table 8 presents the research results for Hypothesis 3-2. Topic 2 is related to customers' critical contents pointing out a problem as to the company.

Table 8. 2SLS results (Topic_2, dependent variables ROA/ROE/TQ/MB, n = 168).

Variables	1st Stage		2nd Stage (ROA)		2nd Stage (ROE)		2nd Stage (TQ)		2nd Stage (MB)	
	β	t Value	β	t Value	β	t Value	β	t Value	β	t Value
Intercept	0.108	0.587	−0.038	−0.408	−1.067	−1.510	−3.851	−1.027	−0.435	−0.080
Topic_2			−0.080	−1.535	−0.704	−1.791 *	2.755	1.310	5.231	1.705 *
FORN	−0.021	−0.210	0.002	0.054	−0.308	−1.184	2.006	1.457	4.701	2.340 **
LARG	0.008	0.101	−0.029	−0.974	0.201	0.911	0.691	0.588	−0.704	−0.411
SIZE	0.030	0.284	0.005	1.464	0.060	2.250 **	0.142	0.999	−0.024	−0.114
LEV	−0.061	−0.725	−0.113	−4.479 ***	−0.973	−5.120 ***	0.469	0.432	3.514	2.219 **
BIG4	−0.039	−0.395	0.008	0.730	−0.001	−0.009	−0.742	−1.620	−0.790	−1.183
CFO			0.309	5.045 ***	0.300	0.650	2.208	0.823	7.332	1.875 *
LOSS	−0.015	−0.126	−0.096	−9.424 ***	−0.333	−4.348 ***	0.238	0.448	−0.046	−0.059
GRW	0.188	1.983 **								
ROA	−0.655	−0.874					4.118	1.149	−8.536	−1.633
IND Effect	Included		Included		Included		Included		Included	
Adj_Rsq	19.89%		72.46%		43.11%		26.16%		24.95%	
F value	1.426 *		11.216 ***		3.943 ***		2.344 ***		2.262 ***	

ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs. * < 0.10, ** < 0.05, *** < 0.01.

When the dependent variable is ROA, β value of the independent variable Topic_2 is −0.080 and the t-value is −1.535, which is not statistically significant. When the dependent variable is ROE, β value of the independent variable Topic_2 is −0.704 and the t-value is −1.791, which is statistically significant at the 10% level.

When the dependent variable is TQ, β value of the independent variable Topic_2 is −2.755 and the t-value is 1.310 which is not statistically significant. When the dependent variable is MB, β value of the independent variable Topic_2 is −5.231 and the t-value is 1.705, which is statistically significant at the 10% level. This is the result of supporting partially Hypothesis 3-2.

The results imply that, among YouTube's comments, criticism of companies or pointing out problems of firms may immediately decrease the current corporate value due to damage to the corporate image. However, the results of this study suggest that from a long-term perspective, if a company humbly accepts and improves these criticisms from customers, the future corporate value may increase.

Table 9 shows the empirical results for Hypothesis 3-3. Topic 3 is related to the contents of a neutral attitude toward a company.

Table 9. 2SLS results (Topic_3, dependent variables ROA/ROE/TQ/MB, n = 168).

Variables	1st Stage		2nd Stage (ROA)		2nd Stage (ROE)		2nd Stage (TQ)		2nd Stage (MB)	
	β	t Value	β	t Value	β	t Value	β	t Value	β	t Value
Intercept	0.114	0.443	−0.127	−1.074	−1.703	−1.906 *	−0.158	−0.034	3.823	0.551
Topic_3			0.072	1.025	0.493	0.928	−3.123	−1.113	−3.253	−0.789
FORN	0.040	0.338	0.004	0.102	−0.285	−1.085	1.990	1.441	4.513	2.224 **
LARG	0.070	0.714	−0.045	−1.446	0.075	0.317	1.370	1.086	0.137	0.074
SIZE	−0.096	−0.753	0.006	1.633	0.065	2.342 **	0.105	0.705	−0.054	−0.249
LEV	−0.053	−0.522	−0.096	−3.591 ***	−0.838	−4.167 ***	−0.250	−0.225	2.478	1.517
BIG4	−0.016	−0.135	0.015	1.269	0.049	0.565	−0.970	−2.119 **	−1.136	−1.689 *
CFO			0.309	4.931 ***	0.334	0.706	2.440	0.895	7.252	1.811 *
LOSS	−0.031	−0.216	−0.094	−9.031 ***	−0.320	−4.069 ***	0.129	0.241	−0.211	−0.270
GRW	0.177	1.901 **								
ROA	0.327	0.360					3.841	1.075	−9.373	−1.786 *
IND Effect	Included		Included		Included		Included		Included	
Adj_Rsq	18.89%		72.17%		42.04%		25.87%		23.57%	
F value	1.480 *		11.070 ***		3.817 ***		2.325 **		2.170 **	

ROA: measures for current company value 1: current net income is divided by total assets; ROE: measures for current company value 2: current net income is divided by total equity; TQ: measures for future company value 1: the sum of the market value of equity and the book value of debt, all divided by the book value of total assets; MB: measures for future company value 2: market value of equity to book value of equity ratio; Channel: dummy variable 1 if the company has a YouTube channel, otherwise 0; Topic_1: TF-IDF average of keywords classified as Topic_1; Topic_2: TF-IDF average of keywords classified as Topic_2; Topic_3: TF-IDF average of keywords classified as Topic_3; FORN: foreign investors' shareholder ratio; LARG: the largest shareholder's equity ratio; SIZE: natural logarithm of total assets; LEV: total liabilities to total assets; BIG4: dummy variable; 1 if a company is audited by a large foreign accounting company called Big4, otherwise 0; CFO: operation cash flow is divided by total sales; LOSS: dummy variable: 1 if the company reports loss, otherwise 0; GRW: (this year-end sales - lagged year-end sales) ÷ lagged year-end sales; IND Effect: dummy variables by industry to control differences in the industry to which a company belongs. * < 0.10, ** < 0.05, *** < 0.01.

When the dependent variable is ROA, β value of the independent variable Topic_3 is −0.072 and the t-value is 1.025, which is not statistically significant. When the dependent variable is ROE, β value of the independent variable Topic_3 is 0.493 and the t-value is −0.928, which is not statistically significant as well.

When the dependent variable is TQ, β value of the independent variable Topic_3 is −3.123 and the t-value is −1.113 which is not statistically significant. When the dependent variable is MB, β value of the independent variable Topic_3 is −3.253 and the t-value is −0.789, which is also not statistically significant. These results support Hypotheses 3-3.

As a result of statistical analysis, as expected by our researchers, we find that the contents of a neutral attitude toward a company of corporate YouTube comments have no significant relationship with current corporate value and future corporate value. Neutral contents which are not just for fun and interest or not for constructive criticism or complaints of the company, are not related to the company's current and future corporate value.

5. Discussion

In this paper, comments on YouTube channels operated by companies were collected using the tax mining technique, and the comments on the corporate YouTube channels were classified into three groups as follows through topic modeling. The contents of Topic_1 are advertising promotion contents focused on fun or interest, the contents of

Topic_2 are critical contents pointing out a problem, and the contents of Topic_3 are the contents of a neutral attitude toward a company.

The focus of this paper is to understand how current and future corporate values change depending on what customers mention a lot through corporate YouTube channels. To investigate this, in this study, the 2SLS model was used to control the endogenous problem of the research model. In addition, for the current corporate value and future corporate value, we tried to present objective and reliable research results as much as possible using multiple proxy variables widely used in previous studies.

The implications of the research results of this paper are as follows. First, the results of the study have shown that companies which directly operate and manage YouTube accounts now have lower corporate values than those that do not. Many of the comments on YouTube accounts express many negative opinions and complaints about the firm. On social media, customers tend to make negative comments about corporates because they are anonymous. This suggests that this can damage the corporate image and reduce the current corporate value in the short term.

Second, this study has found that companies that directly operate and manage YouTube accounts have higher future corporate value than those that do not. Our researchers can explain that in the short term, the current corporate value may be lowered due to customer opinions or complaints, but in the long term, the trust relationship between companies and customers can ultimately improve future corporate value by interacting with customers through social media channels and resolving customer complaints.

Third, the empirical analysis results of this study have shown that if a customer simply mentions interesting content or advertising/promotion-related content through corporate YouTube comments, the current corporate value may improve in the short term, but it negatively affects the future corporate value in the long term. In other words, the fact that customers only respond to immediate fun, interest, advertisements, and promotions without sincerity means that they are not helpful to the firm in the long run.

Fourth, as a result of empirical analysis in this paper, we have found that if customers criticize companies or point out problems of companies through YouTube comments, the current corporate value may decrease due to damage to the firm's image. However, the results of this study suggest that from a long-term perspective, the future corporate value may increase if the firm humbly accepts and improves these points from the customers.

Fifth, the results of this study found that neutral contents, not just for fun and interest and not for constructive criticism or dissatisfaction with the company, have no relation with the company's current and future corporate value.

The summary of the implications of the research results of this paper is as follows. Companies can see that directly operating social media strategically is more advantageous in improving future corporate value from a long-term perspective. Additionally, through YouTube channels, firms can see that implementing policies which respond well to customer complaints or negative comments and continue to manage customers' opinions well are more helpful to firms in the long run.

6. Conclusions

This study conducted empirical analysis on how customers' reactions on social media affect the present and future firm value. This research used Korean KOSPI-listed companies that actually own and operate YouTube channels, and collected data through text mining on comments in YouTube.

Information on companies does not only have quantitative data, but also qualitative data. Yet, owing to the limitations of the research methodology, the accounting and financial fields have mainly studied using only quantitative data and corporate financial data. The research model of this study significantly differs from existing research methodologies in that it quantified unstructured text data called corporate YouTube comments and conducted empirical analysis through a regression analysis model along with corporate financial data.

As a convergence study, this paper combines text mining techniques and regression analysis by quantifying qualitative text data. The results of this study are expected to contribute to further expanding the scope of the study by presenting an innovative methodology that quantifies and variables text data.

In this paper, the TF-IDF value of YouTube comments was calculated, keywords were extracted, and keywords were classified into three groups through topic modeling. The characteristics of three groups of the YouTube comments could be transformed into a “current-oriented topic” as advertising promotion content focused on fun or interest; a “future-oriented topic” as critical content pointing out problems, and a “neutral topic” as content of a neutral attitude toward companies.

This research utilized a regression analysis model to perform an empirical analysis by setting a company’s YouTube-related variable as an independent variable, and setting a company’s current value and future value-related variable as a dependent variable.

The results of this research are as follows. First, this paper found that companies which directly operate and manage YouTube accounts currently have lower corporate value than those that do not. Second, this study also found that companies which directly operate and manage YouTube accounts have higher future corporate value than those that do not. Third, the results showed that if a customer simply mentions interesting content or advertising/promotion-related content through corporate YouTube comments, the current corporate value may be improved in the short term, but in the long term, it has a negative effect on future corporate value. Fourth, the results of this research also presented that if a customer criticizes a company or points out a company’s problems through YouTube comments, the current corporate value decreases due to damage to the company’s image, but it was found that the future corporate value increases. Fifth, this paper found that neutral content, not just for fun and interest, nor for constructive criticism or dissatisfaction with the company, was not related to the company’s current and future corporate value.

The contributions and expected effects of this paper are as follows. First of all, this paper provides useful information through research results that companies are more advantageous in improving future corporate value from a long-term perspective by strategically operating social media directly. In addition, the research results of this study objectively demonstrated through YouTube channels that it is more helpful for companies in the long run to respond well to customer complaints and negative opinions, and to implement policies that continuously manage customer opinions. Finally, the research method used in this paper, that is, the research methodology that conducted empirical analysis through quantification of unstructured tax data, is expected to provide guidelines for many scholars to expand the scope of data available for empirical research in the future.

The results of this study supplies the following implications. From a theoretical point of view, this study presented a new research methodology in that statistical analysis was performed by quantifying qualitative data. From a management perspective, the findings of this paper provide useful information that companies should pay more attention to customer management through social media to achieve sustainable development. Finally, from a practical point of view, the empirical analysis results of this study suggest that it is necessary for companies to actively operate social media to improve corporate value in the future.

The limitation of this study is that due to the vast amount of YouTube comments, only comments from major YouTube contents were used as data for each company. In future studies, richer data will be used to present more objective research results. In addition, this study will be developed to analyze whether there is a significant difference between a company’s YouTube channel comments and a company’s future value by industry.

Author Contributions: Conceptualization, H.J.N.; methodology, H.J.N. and Y.H.K.; software, H.J.J.; validation, Y.H.K. and H.J.N.; formal analysis, H.J.J. and H.J.N.; resources, H.J.J.; data curation, H.J.J. and Y.H.K. writing—original draft preparation, H.J.N.; writing—review and editing, Y.H.K. and H.J.J.; supervision, H.J.N.; project administration, Y.H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by the Semyung University Research Grant of 2021.

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Conflicts of Interest: The authors declare no conflict of interest.

References

- Correia, P.P.; Medina, I.G. Digital Social Media: An Interactive Technology Incorporated as a Competitive Advantage for Business. *Int. J. Interact. Mob. Technol. (Ijim)* **2014**, *8*, 23. <https://doi.org/10.3991/ijim.v8i2.3576>.
- Weiner, D. Crisis communications: Managing corporate reputation in the court of public opinion. *Ivey business journal* **2006**, *70*, 1–6.
- Spieler, B.; Ballard, D.H.; Mazaheri, P.; Legro, N.; Catanzano, T.; Dey, C.; Prejean, E.; Fontentot, J.; Martin, M.D.; Danrad, R.; et al. Social Media in Radiology: Overview and Usefulness of Online Professional #SoMe Profiles. *Acad. Radiol.* **2021**, *28*, 526–539. <https://doi.org/10.1016/j.acra.2020.03.014>.
- Grégoire, Y.; Salle, A.; Tripp, T.M. Managing social media crises with your customers: The good, the bad, and the ugly. *Bus. Horiz.* **2015**, *58*, 173–182. <https://doi.org/10.1016/j.bushor.2014.11.001>.
- Schweitzer, L. Planning and Social Media: A Case Study of Public Transit and Stigma on Twitter. *J. Am. Plan. Assoc.* **2014**, *80*, 218–238. <https://doi.org/10.1080/01944363.2014.980439>.
- Park, S.O.; Na, H.J.; Kwon, O. Comparative effect of company-driven SNS activity vs. consumer-driven SNS activity on firm value: Evidence from facebook. *Comput. Ind.* **2016**, *82*, 186–195. <https://doi.org/10.1016/j.compind.2016.07.008>.
- Fülöp, M.T.; Breaz, T.O.; He, X.; Ionescu, C.A.; Cordoş, G.S.; Stănescu, S.G. The role of universities' sustainability, teachers' wellbeing, and attitudes toward e-learning during COVID-19. *Front. Public Health* **2022**, *10*, 981593. <https://doi.org/10.3389/fpubh.2022.981593>.
- Akram, U.; Fülöp, M.; Tiron-Tudor, A.; Topor, D.; Căpuşneanu, S. Impact of Digitalization on Customers' Well-Being in the Pandemic Period: Challenges and Opportunities for the Retail Industry. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7533. <https://doi.org/10.3390/ijerph18147533>.
- Chung, S.; Animesh, A.; Han, K.; Pinsonneault, A. Financial Returns to Firms' Communication Actions on Firm-Initiated Social Media: Evidence from Facebook Business Pages. *Inf. Syst. Res.* **2020**, *31*, 258–285. <https://doi.org/10.1287/isre.2019.0884>.
- AlQershi, N. Strategic thinking, strategic planning, strategic innovation and the performance of SMEs: The mediating role of human capital. *Manag. Sci. Lett.* **2021**, *11*, 1003–1012. <https://doi.org/10.5267/j.msl.2020.9.042>.
- Griffith, J.; Najand, M.; Shen, J. Emotions in the stock market. *Journal of Behavioral Finance* **2020**, *21*, 42–56.
- Zhang, H.; Gupta, S.; Sun, W.; Zou, Y. How social-media-enabled co-creation between customers and the firm drives business value? The perspective of organizational learning and social Capital. *Inf. Manag.* **2020**, *57*, 103200. <https://doi.org/10.1016/j.im.2019.103200>.
- Romero, N.L. ROI. Measuring the social media return on investment in a library. *Bottom Line* **2011**, *24*, 145–151. <https://doi.org/10.1108/08880451111169223>.
- Benthous, J.; Risius, M.; Beck, R. Social media management strategies for organizational impression management and their effect on public perception. *J. Strat. Inf. Syst.* **2016**, *25*, 127–139. <https://doi.org/10.1016/j.jsis.2015.12.001>.
- Liu, Y.; Shankar, V.; Yun, W.; Liu, V.S.Y.; Cleeren, K.; Dekimpe, M.G.; van Heerde, H.J. Crisis Management Strategies and the Long-Term Effects of Product Recalls on Firm Value. *J. Mark.* **2017**, *81*, 30–48. <https://doi.org/10.1509/jm.15.0535>.
- Hsu, L.; Lawrence, B. The role of social media and brand equity during a product recall crisis: A shareholder value perspective. *Int. J. Res. Mark.* **2016**, *33*, 59–77. <https://doi.org/10.1016/j.ijresmar.2015.04.004>.
- Borah, A.; Tellis, G.J. Halo Effects in Social Media: Do Product Recalls Hurt or Help Rival Brands. (April 20, 2015). 2015. Available online: <https://ssrn.com/abstract=2596921> (accessed 20 January 2023).
- Colicev, A.; Malshe, A.; Pauwels, K.; O'Connor, P. Improving Consumer Mindset Metrics and Shareholder Value through Social Media: The Different Roles of Owned and Earned Media. *J. Mark.* **2018**, *82*, 37–56. <https://doi.org/10.1509/jm.16.0055>.
- Bigdellou, S.; Aslani, S.; Modarres, M. Optimal promotion planning for a product launch in the presence of word-of-mouth. *J. Retail. Consum. Serv.* **2022**, *64*, 102821. <https://doi.org/10.1016/j.jretconser.2021.102821>.
- Kobayashi, M. Blogging around the globe: Motivations, privacy concerns, and social networking. In *Computational social Networks*; Springer: London, UK, 2012; pp. 55–86.
- Smith, B.G.; Smith, S.B.; Knighton, D. Social media dialogues in a crisis: A mixed-methods approach to identifying publics on social media. *Public Relat. Rev.* **2018**, *44*, 562–573. <https://doi.org/10.1016/j.pubrev.2018.07.005>.

22. Larivière, B.; Joosten, H.; Malthouse, E.C.; Van Birgelen, M.; Aksoy, P.; Kunz, W.H.; & Huang, M.H. Value fusion: The blending of consumer and firm value in the distinct context of mobile technologies and social media. *J. Serv. Manag.* **2013**, *24*, 268–293.
23. Dong, J.Q.; Wu, W. Business value of social media technologies: Evidence from online user innovation communities. *J. Strat. Inf. Syst.* **2015**, *24*, 113–127. <https://doi.org/10.1016/j.jsis.2015.04.003>.
24. Nguyen, H.; Calantone, R.; Krishnan, R. Influence of Social Media Emotional Word of Mouth on Institutional Investors' Decisions and Firm Value. *Manag. Sci.* **2020**, *66*, 887–910. <https://doi.org/10.1287/mnsc.2018.3226>.
25. Luo, X.; Zhang, J. How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm?. *J. Manag. Inf. Syst.* **2013**, *30*, 213–238. <https://doi.org/10.2753/mis0742-1222300208>.
26. Matthews, L. Social media and the evolution of corporate communications. *The Elon Journal of Undergraduate Research in Communications* **2010**, *1*, 17–23.
27. Beckman, T.; Colwell, A.; Cunningham, P.H. The Emergence of Corporate Social Responsibility in Chile: The Importance of Authenticity and Social Networks. *J. Bus. Ethic.* **2009**, *86*, 191–206. <https://doi.org/10.1007/s10551-009-0190-1>.
28. Cornelissen, J.P. *Corporate communication: A guide to theory and practice*. Corporate Communication; Sage Publications: Thousand Oaks, CA, USA, 2020; pp. 1–336.
29. Treem, J.W.; Leonardi, P.M. Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association. *Ann. Int. Commun. Assoc.* **2013**, *36*, 143–189. <https://doi.org/10.1080/23808985.2013.11679130>.
30. Qiu, L.; Kumar, S. Understanding Voluntary Knowledge Provision and Content Contribution Through a Social-Media-Based Prediction Market: A Field Experiment. *Inf. Syst. Res.* **2017**, *28*, 529–546. <https://doi.org/10.1287/isre.2016.0679>.
31. Lambret, C.V.; Barki, E. Social media crisis management: Aligning corporate response strategies with stakeholders' emotions online. *J. Contingencies Crisis Manag.* **2018**, *26*, 295–305. <https://doi.org/10.1111/1468-5973.12198>.
32. Molitorisz, S. More top-down than peer-to-peer: Talking to Australians about their ideal news source. *Media Int. Aust.* **2020**, *175*, 109–123. <https://doi.org/10.1177/1329878x19891586>.
33. Kang, E.; Lee, J.; Kim, K.H.; Yun, Y.H. The popularity of eating broadcast: Content analysis of “mukbang” YouTube videos, media coverage, and the health impact of “mukbang” on public. *Health informatics journal* **2020**, *26*, 2237–2248.
34. Appel, G.; Grewal, L.; Hadi, R.; Stephen, A.T. The future of social media in marketing. *Journal of the Academy of Marketing Science* **2020**, *48*, 79–95.
35. Noam, E. *Overcoming Market Power. Regulating Big Tech: Policy Responses to Digital Dominance*; Oxford University Press: Oxford, UK, 2021; p. 55.
36. Salovaara, A.; University of Helsinki; Lyytinen, K.; Penttinen, E.; Case Western Reserve University. Aalto University High Reliability in Digital Organizing: Mindlessness, the Frame Problem, and Digital Operations. *MIS Q.* **2019**, *43*, 555–578. <https://doi.org/10.25300/misq/2019/14577>.
37. Baird, C.H.; Parasnis, G. From social media to social customer relationship management. *Strat. Leadersh.* **2011**, *39*, 30–37. <https://doi.org/10.1108/10878571111161507>.
38. Agnihotri, R. Social media, customer engagement, and sales organizations: A research agenda. *Ind. Mark. Manag.* **2020**, *90*, 291–299. <https://doi.org/10.1016/j.indmarman.2020.07.017>.
39. Sashi, C.M. Customer engagement, buyer-seller relationships, and social media. *Manag. Decis.* **2012**, *50*, 253–272.
40. Lim, W.M.; Rasul, T. Customer engagement and social media: Revisiting the past to inform the future. *J. Bus. Res.* **2022**, *148*, 325–342. <https://doi.org/10.1016/j.jbusres.2022.04.068>.
41. Gallagher, J.; Ransbotham, S. Social media and customer dialog management at Starbucks. *MIS Quarterly Executive* **2010**, *9*, 197–211.
42. Sashi, C.M.; Brynildsen, G.; Bilgihan, A. Social media, customer engagement and advocacy: An em-pirical investigation using Twitter data for quick service restaurants. *Int. J. Contemp. Hosp. Manag.* **2019**, *31*, 1247–1272. <https://doi.org/10.1108/IJCHM-02-2018-0108>.
43. Zhu, Z.; Liang, J.; Li, D.; Yu, H.; Liu, G. Hot Topic Detection Based on a Refined TF-IDF Algorithm. *IEEE Access* **2019**, *7*, 26996–27007. <https://doi.org/10.1109/access.2019.2893980>.
44. Du, Y.; Yi, Y.; Li, X.; Chen, X.; Fan, Y.; Su, F. Extracting and tracking hot topics of micro-blogs based on improved Latent Dirichlet Allocation. *Eng. Appl. Artif. Intell.* **2020**, *87*, 103279. <https://doi.org/10.1016/j.engappai.2019.103279>.
45. Hadi, M.A.; Fard, F.H. Aobtm: Adaptive online biterm topic modeling for version sensitive short-texts analysis. In Proceedings of the 2020 IEEE international conference on software maintenance and evolution (ICSME) IEEE 2020, Adelaide, SA, Australia, 28 September 2020–2 October 2020; pp. 593–604.
46. Brookes, G.; McEnery, T. The utility of topic modelling for discourse studies: A critical evaluation. *Discourse Stud.* **2019**, *21*, 3–21. <https://doi.org/10.1177/1461445618814032>.
47. Tushev, M.; Ebrahimi, F.; Mahmoud, A. Domain-specific analysis of mobile app reviews using keyword-assisted topic models. In Proceedings of the 44th International Conference on Software Engineering, Pittsburgh, PA, USA, 25–27 May 2022; pp. 762–777. <https://doi.org/10.1145/3510003.3510201>.
48. Heidari, M.; Rafatirad, S. Semantic Convolutional Neural Network model for Safe Business Investment by Using BERT. In Proceedings of the 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS) IEEE, Paris, France, 14–16 December 2020; pp. 1–6. <https://doi.org/10.1109/snams52053.2020.9336575>.
49. McGurk, Z.; Nowak, A.; Hall, J.C. Stock returns and investor sentiment: Textual analysis and social media. *J. Econ. Financ.* **2020**, *44*, 458–485. <https://doi.org/10.1007/s12197-019-09494-4>.

50. Alsartawi, A.M.M. Assessing the Relationship Between Information Transparency Through Social Media Disclosure And Firm Value. *Manag. Account. Rev. (MAR)* **2019**, *18*, 1–20. <https://doi.org/10.24191/mar.v18i2.697>.
51. Putra, F.K.G.; Harymawan, I.; Nasih, M.; Agustia, D. A Quest to Minimize Cost of Debt By Utilizing Human Resources Disclosure. *Pol. J. Manag. Stud.* **2020**, *21*, 342–355. <https://doi.org/10.17512/pjms.2020.21.2.24>.
52. Ferri, L.; Spanò, R.; Ginesti, G.; Theodosopoulos, G. Ascertaining auditors' intentions to use blockchain technology: Evidence from the Big 4 accountancy firms in Italy. *Meditari Account. Res.* **2020**, *29*, 1063–1087. <https://doi.org/10.1108/medar-03-2020-0829>.
53. Hudders, L.; De Jans, S.; De Veirman, M. The Commercialization of Social Media Stars: A Literature Review and Conceptual Framework on the Strategic Use of Social Media Influencers. In *Social Media Influencers in Strategic Communication*; Routledge: New York, NY, USA, 2021; pp. 24–67. <https://doi.org/10.4324/9781003181286-3>.
54. Barrot, J.S. Scientific Mapping of Social Media in Education: A Decade of Exponential Growth. *J. Educ. Comput. Res.* **2020**, *59*, 645–668. <https://doi.org/10.1177/0735633120972010>.
55. Nisar, T.M.; Prabhakar, G.; Strakova, L. Social media information benefits, knowledge management and smart organizations. *J. Bus. Res.* **2019**, *94*, 264–272. <https://doi.org/10.1016/j.jbusres.2018.05.005>.
56. Shen, C.-W.; Chen, M.; Wang, C.-C. Analyzing the trend of O2O commerce by bilingual text mining on social media. *Comput. Hum. Behav.* **2018**, *101*, 474–483. <https://doi.org/10.1016/j.chb.2018.09.031>.
57. Bao, X.; Sun, B.; Han, M.; Lin, H.; Lau, R.Y. Quantifying the impact of CEO social media celebrity status on firm value: Novel measures from digital gatekeeping theory. *Technol. Forecast. Soc. Chang.* **2023**, *189*, 122334. <https://doi.org/10.1016/j.techfore.2023.122334>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.