Volume 31 (1) April 2024 ISSN 0216-423X (Print)
ISSN 2622-2667 (Online)

JABM
JOURNAL of ACCOUNTING, BUSINESS and MANAGEMENT

Index Premium Trends Resulting from Composition Changes to the S&P 500 and Its Implications for Market Efficiency
Dongfang Nie **1-10**

Influences of Gender, Age and Income Differences on Consumers' Purchasing Behavior
Matiur Rahman, Lonnie Turpin and Md. Al Emran **11-21**

The Benefits of Revenue Diversification on Bank Profitability and Stability: An Empirical Study on Indonesian Commercial Banks
Robertus Setiadi and Dwi Nastiti Danarsari **22-37**

Metaheuristic-Driven Optimization for Complex Multidimensional Decision-Making: A Case Study on Prioritizing Airport Locations
Nazila Razi, Rouhollah Bagheri and Hamed Pourabbas **38-55**

Differences in Perceived Value of Team Projects and Learning Styles of Accounting and Marketing Students
Vivek Madupu and Konrad Gunderson **56-64**

Predicting Equity Crowdfunding Success: An Examination of United States Offerings using Sentiment Analysis
Sarah Borchers, Matt Bjornsen, Bree Dority and Suzanne Hayes **65-79**

Comparation of Earnings Quality Measures at Industries in the National Stock Exchange of India
Shikhil Munjal and Gurcharan Singh **80-92**

Self-Control Factor Analysis, Financial Anxiety, and Financial Stress on Financial Satisfaction as an Indication Financial Sustainability: Study of Accounting Students Who Own a Business
Yopy Junianto and Wirawan Endro Dwi Radianto **93-104**

We Learn from History: Earnings Management and Business Scandals in the Early of 2000s
Gerui (Grace) Kang **105-119**

Leadership and Next Generation Unmanned System Integration
Amy T. Clemens and Leslie Huffman **120-132**

Journal of Accounting, Business and Management (JABM) vol. 31 no. 1 (2024) 65-79

Predicting Equity Crowdfunding Success: An Examination of United States Offerings using Sentiment Analysis

Sarah Borchers*
Matt Bjornsen†
Bree Dority‡
Suzanne Hayes§

Abstract

Equity crowdfunding has grown exponentially in the United States since the passage of the JOBS Act in 2013, yet it continues to be a research area that is relatively unexplored in the United States due to the limited availability of data. U.S equity crowdfunding campaigns are notoriously unsuccessful, and this paper develops a predictive model for equity crowdfunding success to determine whether the positivity of the language used, and the length of the campaign description influences an investor's decision to invest. A model is developed on a balanced training set and applied to a test set, and the overall results are evaluated using a confusion matrix to determine the accuracy, precision, and recall of the model. Our overall results indicate that both the tone of sentiment and the length of the campaign description are predictive of an equity crowdfunding campaign's success. Specifically, the potential investors appear to be attracted to positive campaign descriptions that are written with concise language.

Keywords: equity crowdfunding, predictive, sentiment, modeling, financing.

I. INTRODUCTION

Transforming ideas into new business ventures can be both timely and costly for new entrepreneurs. In the past, entrepreneurs have had to rely on angel investors or venture capitalists for funding. However, it can be difficult to attract such individuals. As a means of combating this funding shortage, an emerging form of financing called crowdfunding has gained in popularity over the past decade. Crowdfunding has been defined as "an open call, essentially through the internet, for the provision of financial resources either in form of donation or in exchange for some form of reward and/or

* Associate professor. Department of Accounting, Finance and Economics. College of Business and Technology, University of Nebraska-Kearney. West Center Building, 1917 W 24th St, Kearney, NE 68849, 308-865-8108. E-mail: borcherssj@unk.edu. Web page: https://www.unk.edu/academics/accounting-finance/faculty-staff/sarah-borchers.php.

† Corresponding author. Associate professor. Department of Accounting, Finance and Economics. College of Business and Technology, University of Nebraska-Kearney. West Center Building, 1917 W 24th St, Kearney, NE 68849, 308-865-8107. E-mail: bjornsenmm@unk.edu. Web page: https://www.unk.edu/academics/accounting-finance/faculty-staff/matt-bjornsen.php.

‡ Professor. Department of Accounting, Finance and Economics. College of Business and Technology, University of Nebraska-Kearney. West Center Building, 1917 W 24th St, Kearney, NE 68849, 308-865-8343. E-mail: doritybl@unk.edu. Web page: https://www.unk.edu/academics/accounting-finance/faculty-staff/bree_dority.php.

§ Professor. Department of Accounting, Finance and Economics. College of Business and Technology, University of Nebraska-Kearney. West Center Building, 1917 W 24th St, Kearney, NE 68849, 308-865-8170. E-mail: hayessk@unk.edu. Web page: https://www.unk.edu/academics/accounting-finance/faculty-staff/suzanne-hayes.php.

voting rights in order to support initiatives for specific purposes" (Schwienbacher & Larralde, 2010).

Crowdfunding has grown exponentially in recent years as a new financing mechanism for start-up companies. In 2017, over $17.2 billion was raised through equity crowdfunding platforms in North America (Crawford, 2021). This accounts for over half of the $28.8 billion raised globally in the same time frame (Szmigiera, 2019). Until recently however, due to regulations in the United States blocking the general solicitation of funding, entrepreneurs were not able to gain funding through internet-based equity crowdfunding (Mamonov & Malaga, 2017). As a result, most of the crowdfunding research to date has focused on countries other than the U.S., particularly the United Kingdoms (Vismara, 2016 and 2018; Horvat et al., 2018; Barbi & Mattioli, 2019; Kleinert & Volkmann, 2019; Kleinert et al., 2020; Du et al., 2021; Coakley et al., 2022; Estrin et al., 2022; and Miglo, 2022) and to a lesser extent Germany (Angerer et al., 2017; Block et al., 2018; and Reichenbach & Walther, 2021), France (Andrieu et al., 2021; Le Pendeven & Schwienbacher 2021), and Italy (Piva & Rossi-Lamastra, 2018; Troise et al., 2021). This has left research on equity crowdfunding in the U.S. relatively unexplored. Although other forms of crowdfunding such as rewards-based, debt-based, or donation-based have been more widely studied, only a handful of studies have focused specifically on equity-based crowdfunding in the U.S. (Agrawal et al., 2014; Mamonov & Malaga, 2017, 2018 and 2019; Hayes et al., 2020; and Dority et al., 2021). Further, given the limited amount of information investors are presented with when choosing which campaigns to invest in, it begs the question of what motivates these investors and helps to predict success. Predicting a campaign's success is arguably the most important issue for an entrepreneur, yet it is an area widely unexplored in the equity crowdfunding arena. In addition, equity crowdfunding campaigns are drastically underfunded or unsuccessful, leaving this an area ripe for additional exploration. The unbalanced nature of the success of equity crowdfunding campaigns poses further challenges to creating a model for predicting success. We use a balanced dataset approach to our model and show an improvement of nearly 45% in the recall of our predictive model, meaning our model was able to identify successful offerings correctly much more frequently.

After reviewing the literature on success factors of equity crowdfunding, Dority et al. (2021) identified nine broad areas shown to influence equity crowdfunding success, including founder-specific factors, firm-specific factors, campaign-specific factors, the financial information disclosed, the product type and stage of development, investors, social capital, updates and discussions, and textual analysis. Regarding the last category – research examining the impact of the investor's feelings about, or sentiment towards, equity crowdfunding offerings – very few studies exist (Block et al., 2018; Horvat et al., 2018; and Dority et al., 2021). Due to the sparse amount of information available to a potential investor, both the way the campaign description is written, and the sentiment of the language used, may play an important role in the decision to invest. Further, when scrolling through several pages of potential campaigns, the length of the campaign description may also attract or deter a potential investor. In addition, the content presented in, as well as the length of, the campaign description is one area in which the entrepreneur has complete control. Therefore, in this paper we develop a simple model for predicting equity campaign success that centers around the length of the campaign description and the overall tone in U.S. equity crowdfunding offerings.

The remainder of this paper is organized as follows. Section II provides the background and hypothesis development. In section III, the data and model are discussed. The step-by-step analysis is detailed in section IV. Our results are discussed in section V, and section VI concludes.

II. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

2.1. Overview

Equity crowdfunding in the U.S. became a viable option for start-up companies within the last seven years. While venture capital, angel investing, and rewards-based crowdfunding were legal, equity crowdfunding was prohibited by the Securities and Exchange Act of 1933 and 1934 until September 2013. In response to the 2007-2008 financial crisis, which made it harder for new ventures to raise capital, Title II of the Jumpstart Our Business Startups (JOBS) Act was passed in September 2013, relaxing the rules for public solicitation of new ventures. The JOBS Act aimed to stimulate economic growth by improving access to public capital markets and eliminating listing requirements for emerging growth companies (SEC, 2013; Colombo et al., 2016). This passage made it easier for start-up companies to acquire funds from accredited investors. An accredited investor is someone who has an annual income exceeding $200,000 or who has assets in excess of $1 million, excluding their primary residence (SEC, 2013; Mamonov & Malaga, 2017). It is estimated that $1.4 billion has been raised by entrepreneurs from accredited investors via equity crowdfunding platforms since the passage of the JOBS Act in 2013 (Mamonov & Malaga, 2017). Title III of the JOBS Act was passed in 2016 and opened equity crowdfunding opportunities to non-accredited investors as well. Due to the difference in regulations surrounding Title II and Title III offerings, our research focuses on Title II equity crowdfunding offerings of the JOBS Act.

Although equity crowdfunding literature in the U.S. is lacking, other forms of crowdfunding are more widely researched and better understood. Most of the research to date has focused on donation-based, rewards-based, or debt-based crowdfunding. Although all forms have similar characteristics, each is unique in terms of the motivators or personal benefits received (Lukkarinen et al., 2016; Mamonov & Malaga, 2017). Crowdfunding transactions are conducted through an internet platform and can be completed from virtually anywhere. Most start-ups of this nature start with a fundraiser initiating a request for funding. Potential investors can then browse the offers available and, if interested, invest money in return for a personal benefit (Ahlers et al., 2015). Through equity-based crowdfunding, investors receive equity in the start-up company in which they are investing (Ahlers et al., 2015; Mamonov & Malaga, 2017). As equity crowdfunding is a relatively new phenomenon in the U.S., this is an area ripe with potential for research. In order to better understand the success factors of an equity crowdfunding campaign and to develop a successful predictive model, we draw on a variety of studies based on US data and non-US data to better understand equity crowdfunding altogether.

2.2. Attracting an Investor

Often, equity crowdfunding research focuses on the types of individuals that raise capital via equity crowdfunding platforms and on the factors that increase the likelihood of a successful equity crowdfunding offering. To be successful in an equity crowdfunding offering, the start-up must attract an interested investor. The prior literature has shown a vast array of reasons investors choose to invest via equity crowdfunding; however, the primary reason for investing in equity crowdfunding is the financial benefit received (Lukkarinen et al., 2016).

2.3. Success Factors

Arguably, the most crucial factor that leads to a successful equity crowdfunding offering is attracting an investor. Typically, a campaign is set up with a target amount of funding. Prior literature shows lower funding goals and/or lower minimum investment requirements are associated with a higher probability of success (Cordova et al., 2015; Lukkarinen et al., 2016; Block et al., 2018; Horvat et al., 2018; and Mamonov & Malaga, 2018); however, more recently, Miglo (2022) finds the effect of the target amount is U-shaped, indicating the target should not be too small or too large. Further, entrepreneurs who sold a smaller fraction of their companies at listing and had more social capital, experienced a higher likelihood of success (Ahlers et al., 2015; Vismara, 2016; and Nitani et al., 2019). Attributes of the product/service, particularly novelty (Horvat et al., 2018) and the perceived innovativeness (Le Pendeven & Schwienbacher, 2021), and the founder's perceived commitment are also documented success factors (Shafi, 2021). Investors carefully evaluate the founder and management team (Piva & Rossi-Lamastra, 2018; Barbi & Mattioli, 2019; Kleinert & Volkmann, 2019; and Reichenbach & Walther, 2021), and Coakley et al. (2022) find that compared to founder teams, solo founders are less likely to succeed in the initial equity crowdfunding offering and are also more likely to fail thereafter.

Campaign success is also linked with the amount of external financing and accelerator attendance (Ralcheva & Roosenboom, 2020). Business accelerator programs provide access to mentors, business networks, and capital. Investors may view participation in these selective programs as a signal of venture quality. As for external financing, a higher number of investors investing in the campaign is important (Vulkan et al., 2016; Horvat et al., 2018; Kleinert & Volkmann, 2019; and Estrin et al., 2022), particularly private investors during the hidden phase (Lukkarinen et al., 2016), early investors (Vismara, 2018), lead investors (Li et al., 2016; Vulkan et al., 2016), investors with public profiles (Vismara, 2018), investors who are geographically concentrated (Horvat et al., 2018), and angel/venture capital investors (Kleinert & Volkmann, 2019; and Mamonov & Malaga, 2018 and 2019).

2.4. Description Sentiment

Limited research has examined the impact of the positivity or negativity of the language used in an entrepreneur's campaign description in attracting an investor (Horvat et al., 2018). However, humans use inferences about text passages to determine if the overall tone of a language is positive or negative. These inferences ultimately impact feelings about a certain text passage and can have a significant impact on the subsequent decisions made. Thus, the language used in the crowdfunding campaign description sets the overall tone for the campaign and can be used as a signal to relay campaign success to potential investors. Text mining tools can be utilized to determine the positive or negative nature of the text, which can then be used to examine its impact on an equity crowdfunding offering (Silge & Robinson, 2018). Even though equity crowdfunding has gained in popularity over the last several years, the research is still somewhat scattered on the factors that ultimately contribute to crowdfunding success. Further, it has not been consistently determined whether sentiment of the campaign description, through both the tone and the length of the description, plays a role in campaign success. Therefore, we seek to answer the following research question:

> Can the sentiment of the language used in an equity crowdfunding description be used to predict the success of an equity crowdfunding offering?

2.5. Theoretical Underpinnings

As with most equity offerings, equity crowdfunding brings about the presence of information asymmetry between the potential investor and the start-up company (Agrawal et al., 2014; Ahlers et al., 2015; and Yan, 2015). These information asymmetry problems may be magnified by the internet-based platforms used in equity crowdfunding that commonly exist in early-stage ventures (Agrawal et al., 2014). Signaling theory is the most widely referenced theory in the equity crowdfunding literature surrounding information asymmetry. Signaling theory arises due to two parties being privy to different information. This information asymmetry leads to the parties using clues or "signals" to make decisions (Connelly et al., 2011). Regarding equity crowdfunding, a variety of solutions have been proposed to help mitigate the risks involved with information asymmetry including the use of clear and appropriate descriptions of the funding opportunity help to reduce information asymmetry (Mamonov et al., 2017).

Building on the importance of the description field in reducing information asymmetry, is the notion that the positivity of the description field may also aid in reducing information asymmetry. A very small stream of research has used text mining or sentiment analysis regarding the product or project description field in equity crowdfunding offerings in order to determine its impact on equity crowdfunding success. In their exploratory study of real estate equity crowdfunding offerings, Mamonov et al. (2017) used text mining of project descriptions of real estate ventures on the Patch of Land real estate crowdfunding platform and found that the descriptions helped to reduce information asymmetry between investors and entrepreneurs which led to an increased probability of campaign investment. Block et al. (2018) use German equity crowdfunding platform data and focus on campaign updates. They find that updates with easier language significantly increases crowd participation during a campaign but not on the amount raised. Horvat et al. (2018) use United Kingdom equity crowdfunding platform data and measure several aspects of the campaign pitch text including: length, writing quality, and stylistic aspects of the language. Dority et al. (2021) use U.S. equity crowdfunding data and focus on campaign descriptions. They find inverse U-shaped relationships exist between information quantity, information quality, and tone and equity crowdfunding campaign success. Our paper expands on this small body of research to examine the content presented and its role in predicting crowdfunding campaign success. Given the potential importance of non-financial information in attracting investors and running a successful campaign, we hypothesize the following:

H_1: the sentiment of the campaign description, as measured by tone and word count, positively predicts U.S. equity crowdfunding campaign success.

III. RESEARCH METHODOLOGY

3.1. Success

Because equity crowdfunding success is one of the most widely researched areas in equity crowdfunding, success has been defined fairly consistently across the literature. Success is most frequently defined as the venture's ability to attract greater than or equal to the minimum amount of sought funding (Colombo et al., 2015; Yan, 2015; Horvat & Papamarkou, 2017; Malaga et al., 2017; and Mamonov et al., 2017). Other measures of success have included the number of investors attracted, the percentage of the capital campaign raised, or the amount of capital pledged during a campaign on a given day (Colombo et al., 2015; Vismara, 2016; and Block et al., 2018). We define success as the venture's ability to attract greater than or equal to their minimum funding goal.

Our sentiment variables include both the tone of the language used and the length of the description. Tone is quantified using the tone of the language used in the campaign description as measured by the Loughran and McDonald (LM) dictionaries. The Loughran and McDonald Sentiment Word Lists (Loughran & McDonald, 2011) are the most widely used in sentiment research and were created with financial communication in mind. We apply the LM Sentiment Word Lists to the sample of equity crowdfunding campaign descriptions to create the following sentiment-related variables:

1) Word count: the total number of words in the description of the offering after removing stopwords, such as "a," "I," "and," "the," and "of";
2) Positive word count: the number of words from the LM positive word list that appear in a campaign description;
3) Negative cord count: the number of words from the LM negative word list that appear in a campaign description.

Following the method of Courtney et al. (2016) for measuring sentiment in rewards-based crowdfunding, we first estimate an overall sentiment score, as measured by the overall tone of the language used in an equity crowdfunding description, as follows:

$$\textbf{Tone} = \frac{\textbf{Positive Word Count - Negative Word Count}}{\textbf{Word Count}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \textbf{1}$$

Thus, tone is a continuous variable ranging from -1.0 to 1.0 that captures the overall net tone of the campaign description. The second variable of interest, word count, captures the length of the description of the campaign. We measure word count as the total number of words in a campaign description after removing stopwords.

3.2. Data Source

The data for the project is a proprietary data set provided by FinMkt, formerly known as Crowdnetic. FinMkt aggregates project-level data across seventeen U.S. crowdfunding platforms targeting the opportunities created by Title II of the JOBS Act. Data is received by FinMkt directly from the individual crowdfunding platforms. The dataset contains information on 6,615 crowdfunding offerings from a period starting October 1, 2013 and ending September 30, 2016, however, we only utilize data for the equity campaigns. The date range corresponds to the timeframe before Title III of the JOBS Act was passed.

Accordingly, we use cross-sectional data that looks at crowdfunding offerings in the U.S. from 2013-2016. Each offering is firm specific and is listed one time in the dataset. The focus of this paper is on equity crowdfunding offerings; therefore, a subset of the data is used that includes only equity crowdfunding offerings. There are 3,216 closed equity crowdfunding offerings that contain complete data for the analysis. The dataset covers eight sectors and thirteen equity specific platforms. Table 1 provides descriptive statistics for the dataset. The majority of campaigns are written with a slightly positive tone and contain approximately 67 words. Further, at an average asking amount of $2.3 million, it is not surprising that only 2.0% of campaigns are successful.

Insert Table 1 here.

3.3. Estimation Methods

The dependent variable used to test the hypothesis, success, is a binary variable. The independent variables of interest, tone and readability, are continuous variables as calculated above. However, because a vast majority of Title II equity crowdfunding campaigns are unsuccessful, 98% of the campaigns in our sample are unsuccessful, leaving an unbalanced dataset which causes predictive issues with our model. Most machine learning algorithms assume that the data are balanced meaning there are almost equal parts of both outcome variables (Krawczyk, 2016). For equity crowdfunding campaigns, this is not the case, as the majority of these campaigns are unsuccessful. In predictive modeling, the dataset is split into a training set and a test set and a model is developed on the training set. If the training set contains 98% unsuccessful offerings, the model is trained to identify virtually all campaigns as unsuccessful. This creates a problem, especially in research where the stakes are high in predictive modeling, such as identifying a rare type of cancer. Although the stakes may not be as high for entrepreneurs, one could argue that it is also very important for entrepreneurs to know how to market their campaigns to be successful.

Table 1

Descriptive Statistics

Table 1 reports descriptive statistics for the sample of United States equity crowdfunding campaigns from 2013 to 2016. All variables are defined in Appendix A.

Variable	N	Mean	Median	Q1	Q3	Std. Dev.
Success	3,216	0.02	0.00	0.00	0.00	0.12
Variables of Interest:						
Tone	3,216	0.01	0.01	-0.01	0.04	0.05
Word Count	3,216	67.31	62.00	37.00	88.00	48.15
Firm Characteristics:						
Amount Requested (in thousands)	3,216	2,318	500.00	150	1,000	39,360
Market Characteristics:						
Monthly VIX Value	3,216	14.34	13.75	13.70	13.95	2.23

We combat the unbalanced data problem by creating a balanced dataset containing 50% successful and 50% unsuccessful offerings. We split this into a training and test set and develop our model on the training set. This allows the model to learn from the successful campaigns and more correctly identify them when applied to the entire model. Because we want to create a simple predictive model to be easily utilized by entrepreneurs, and due to the small number of successful campaigns, we use as few control variables as possible. As previously noted in the literature review, prior research has found that the funding target amount (the funding goal) and the minimum investment amount (Ahlers et al., 2015; Cordova et al., 2015; Lukkarinen et al., 2016; and Block et al., 2018) are related to campaign success. Thus, in all regressions, we control for the minimum investment amount. We also include the Volatility Index (VIX) as a proxy for market conditions in all regressions, as well as quarter-year, platform, sector, and region fixed effects. Our logit regression constructed is as follows:

$$\mathbf{Success(0,1)_i = \beta_0 + \beta_1 Tone_i + \beta_2 Word\ Count_i + \beta' X_i + \varepsilon_i} \quad \ldots\ldots\ldots\ldots \quad \mathbf{2}$$

Success is a binary variable that takes the value of one if the entrepreneur of campaign i successfully raised their stated minimum funding goal amount, and zero otherwise. Tone is a measure of the overall tone of the description as defined previously. Word count is a measure of the length of the description as defined previously as well. We control for the campaign asking amount, the Volatility Index (VIX) as a proxy for market conditions in all regressions, as well as quarter-year, platform, sector, and region fixed effects. All variables used in the study are defined in Appendix A.

IV. RESULTS AND DISCUSSIONS

4.1. Original Training Set versus Test Set

We first develop a training and test set on our entire dataset to develop a predictive model of equity crowdfunding success. The split of the entire dataset can be seen in Table

2 Panel A. The model is developed based on the original training set of 2,122 observations and then applied to the test set of 1,094 observations. The results of the model are presented in Table 2 Panel B. Results indicate that none of the variables modeled have significant predictive power.

To evaluate the effectiveness of the proposed model, we utilize a confusion matrix to calculate the accuracy, precision, and recall of the model. An example confusion matrix as well as associated accuracy, precision, and recall calculations can be found in in Table 2 Panel C and Table 2 Panel D, respectively. The confusion matrix associated with the original model on the original test set is presented in Table 2 Panel E, and the associated accuracy, precision and recall can be found in Table 2 Panel F.

As can be seen in Table 2 Panel F, the original training model performed with approximately 98% accuracy when applied to the original test set. At first, one may think this is an outstanding predictive model. However, in evaluating the dataset further, it becomes apparent why the model is so accurate. The original dataset is extremely unbalanced as to the breakdown of the dependent variable. There are 98% unsuccessful offerings and only 2% successful offerings. The breakdown between successful and unsuccessful offerings is presented in Table 2 Panel G. Because the dataset is so unbalanced, the model was trained to always select the "not successful" option. In doing this, the original model predicted with 98% accuracy. In other words, if the model predicts an unsuccessful outcome every single time, it will be able to predict 98% of the outcomes. For unbalanced models, precision and recall are far better indicators of model success. Precision measures the amount of truly successful offerings out of all of the offerings identified as successful. Recall identifies of all the total successful offerings, what percentage is predicted as positive (Jayaswal, 2020). Table 2 Panel D presents the calculations for accuracy, precision, and recall.

Table 2

Table 2 presents the breakdown of the dataset as well as the results of the original model. Panel A quantifies the breakdown of the dataset into the training and test set. Panel B reports results from a logit regression where the dependent variable is a binary variable indicating whether or not the entrepreneur raised 100% or more of the minimum funding goal amount. All variables are defined in Table 1. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. Panels C - G present the confusion matrix and show the results of the training model when applied to the test set.

Panel A: Initial Training versus Test Set

	Offerings	Percent of Total
Training Set	2,122	66%
Test Set	1,094	34%
Total	3,216	100%

Panel B: Original Test Set Regression Results: Impact of Sentiment on Campaign Success

Variable	Logit Model	
	Coefficient	t-stat
Intercept	-5.616***	-3.580
Variables of Interest:		
Tone	-0.027	-0.007
Word Count	0.002	0.752
Firm Characteristics:		
Amount Requested	-0.00003	-0.632

To be continued Table 2.

Variable	Logit Model	
	Coefficient	t-stat
Market Characteristics:		
Monthly VIX Level	-0.082	-1.274
Platform Controls	Yes	
Quarter Year Controls	Yes	
Region Controls	Yes	
Sector Controls	Yes	
# of Observations	3,216	
AIC	480.464	
Log Likelihood	-209.232	

Panel C: A Confusion Matrix

Actual Class	Predicted Class	
	Failure	Success
Failure	True Negative (TN)	False Positive (FP)
Success	False Negative (FN)	True Positive (TP)

Panel D: Accuracy, Precision, and Recall Calculations

$$\textbf{Accuracy(\%)} = \frac{\textbf{TP+TN}}{\textbf{TP+FP+FN+TN}} \textbf{*100}$$

$$\textbf{Precision(\%)} = \frac{\textbf{TP}}{\textbf{TP+FP}} \textbf{*100}$$

$$\textbf{Recall(\%)} = \frac{\textbf{TP}}{\textbf{TP+FN}} \textbf{*100}$$

Panel E: Confusion Matrix of Original Model on Original Test Set

Actual Class	Predicted Class	
	Failure	Success
Failure	1,073	0
Success	21	0

Panel F: Fit of Original Model on Original Test Set

Measures of Fit	
Accuracy	98.08%
Precision	0.00%
Recall	0.00%

Panel G: Breakdown of the Successful versus Unsuccessful Offerings in the Dataset

Full Dataset	Number of Offerings	Percent of Total
Infrequent (Successful)	49	2%
Frequent (Not Successful)	3,167	98%
Total	3,216	100%

4.2. Creating a Balanced Training Set and Test Set and Developing a Model

To create a better prediction model, we create a balanced dataset to train the model and then run it on an "unbalanced" test set with the same distribution as the original dataset. Finally, we run the model on the entire dataset. Of the sixty-four total successful offerings, approximately 67% were used in the training set and the remainder were used in the test set. A breakdown of the training and test sets can be seen in Table 3 Panel A and Table 3 Panel B, respectively.

The logistic regression model was applied to the balanced training set first, and then applied to the redistributed "unbalanced" test set. Results can be seen in Table 3 Panels C, D, and E. The model is performing with approximately 83% accuracy, 8%

precision, an 41% recall. As expected, accuracy decreased from the original model; however, precision and recall were significantly improved.

Table 3

Table 3 presents the process of balancing the dataset and the associated regression results. Panel A shows the breakdown of a balanced dataset showing equal numbers of successful and not successful offerings. Panel B shows the breakdown of the test set based on the balanced training dataset. Panel C reports results from applying the balanced model on the unbalanced test set. All variables are defined in Table 1. ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. Panels D and E present the confusion matrix and show the results of the balanced training model when applied to the unbalanced test set.

Panel A: Balanced Training Dataset		
Training Dataset	**Number of Offerings**	**Percent of Total**
Infrequent (Successful)	32	50%
Frequent (Not Successful)	32	50%
Total	64	100%
Panel B: Test Set Based on Balanced Training Set		
Test Dataset	**Number of Offerings**	**Percent of Total**
Infrequent (Successful)	17	2%
Frequent (Not Successful)	850	98%
Total	867	100%
Panel C: Results of the Balanced Training Model on the "Unbalanced" Test Set		
	Logit Model	
Variable	**Coefficient**	**t-stat**
Intercept	20.876	1.574
Variables of Interest:		
Tone	110.142*	1.763
Word Count	-0.144*	-1.933
Firm Characteristics:		
Amount Requested	0.001	1.422
Market Characteristics:		
Monthly VIX Level	-2.145*	-2.316
Platform Controls	Yes	
Quarter Year Controls	Yes	
Region Controls	Yes	
Sector Controls	Yes	
# of Observations	64	
AIC	79.361	
Log Likelihood	-11.680	
Panel D: Confusion Matrix of Balanced Training Model on "Unbalanced" Test Set		
	Predicted Class	
Actual Class	**Failure**	**Success**
Failure	709	131
Success	16	11
Panel E: Fit of the Balanced Training Model on the "Unbalanced" Test Set		
Measures of Fit		
Accuracy	83.04%	
Precision	7.75%	
Recall	40.74%	

4.3. Applying the Model Developed on the Balanced Training Set to the Entire Dataset

The improved predictive model was then applied to the entire dataset. Regression results are the same as those in Table 3 Panel C as the same model was applied to the "unbalanced" test set; therefore, results of the confusion matrix and fit are the most relevant. Results of the confusion matrix and the fit of the model to the entire dataset can be seen in Table 4 Panel A and Table 4 Panel B, respectively.

Table 4

Panel A: Confusion Matrix of Balanced Training Model on the Entire Dataset		
	Predicted	
Actual	**0**	**1**
0	2,725	442
1	27	22

Panel B: Fit of the Balanced Training Model on the Entire Dataset	
Measures of Fit	
Accuracy	85.42%
Precision	4.74%
Recall	44.90%

4.4. Results

Results of the overall measures of fit for each of the three steps was analyzed in Table 5 Panel A. With the balanced training model, precision and recall are both improved; however, the accuracy of the model decreases. This is due to the model actually predicting some successful cases instead of predicting each offering as unsuccessful. The overall model is 85% accurate, 5% precise, and has a recall percentage of approximately 45%. However, by creating a balanced dataset, we were able to drastically improve our model in order to better help entrepreneurs run a successful equity crowdfunding campaign. The model can be written as follows:

$$\textbf{ln(Odds of Success)= 20.876+110.142(Tone)+0.001(Issue Min. Amount in thousands)-0.144(Word Count)-2.145(VIX)} \quad \textbf{3}$$

As an example, if a campaign was asking for \$1,000,000, had 100 words in their campaign description, during a time when the volatility index was 13.75, wrote their campaign language to include 20% positive wording, then according to our model the campaign would have a 50% chance of success. Interestingly, word count negatively impacts campaign success as lengthier campaign descriptions appear to deter potential investors. In the above scenario, holding all else constant, shortening the campaign description by ten words, increases the probability of success of the campaign to 81%. Relevant information for the above scenarios is presented in Table 5 Panel B. The results of the regression are displayed in Table 5 Panel C.

V. CONCLUSION

Our overall results indicate that both the tone of sentiment and the length of the campaign description are predictive of an equity crowdfunding campaign's success; however, not as we would have originally expected. The potential investors appear to be attracted to positive campaign descriptions that are written with concise language. Overly wordy campaigns negatively impact success. Using our simple model, entrepreneurs can be mindful of the characteristics of a campaign that may increase their likelihood of success. Using a novel approach to the unbalanced data problem, we created a balanced dataset to train our predictive model, which increased the overall precision and accuracy drastically.

Our findings have implications for both start-ups and investors alike. Start-ups should be cognizant of the asking amount when setting their campaign goals. Entrepreneurs appear to be best positioned when using positive campaign wording and concise, easy to understand descriptions. Finally, during times of market volatility, entrepreneurs should avoid launching riskier equity crowdfunding campaigns as the success of the campaign is negatively impacted by times of higher market volatility.

Future research could consider additional tests applied to this concept, including isolating the positivity score, negativity score, and uncertainty score, as well as using the binary variable from the sentiment score to determine if the individual effects impact success more so than the aggregate effect. Further, utilizing additional dictionaries may impact results as well. Although the model does have some predictive power, it has room for improvement. However, developing an easy to understand predictive model for equity crowdfunding success has the ability to aid both entrepreneurs and investors alike, creating room for additional growth in the equity crowdfunding market.

REFERENCES

Agrawal, A., Catalini, C., & Goldfarb, A. (2014). Some simple economics of crowdfunding. *Innovation Policy & the Economy*, *14*(1), 63-97.

Ahlers, G., Cumming, D., Gunther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory & Practice*, *39*(4), 955-980.

Andrieu, G., Le Pendeven, B., & Leboeuf, G. (2021). Equity crowdfunding success for female entrepreneurs: French evidence. *Economics Bulletin*, *41*(2), hal-03227572

Angerer, M., Brem, A., Kraus, S., & Peter, A. (2017). Start-up funding via equity crowdfunding in Germany: A qualitative analysis of success factors. *The Journal of Entrepreneurial Finance (JEF)*, *19*(1), 1-34.

Barbi, M., & Mattioli, S., (2019). Human capital, investor trust, and equity crowdfunding. *Research in International Business & Finance*, *49*, 1-12.

Block, J., Hornuf, L., & Mortiz, A. (2018). Which updates during an equity crowdfunding campaign increase crowd participation? *Small Business Economics*, *50*, 3-27.

Coakley, J., Lazos, A., & Liñares-Zegarra, J. (2022). Equity crowdfunding founder teams: Campaign success and venture failure. *British Journal of Management*, *33*(1), 286-305.

Colombo, M. G., Cumming, D. J., & Vismara, S. (2016). Governmental venture capital for innovative young firms. *Journal of Technology Transfer*, *41*(1), 10-24.

Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal social capital and the attraction of early contributions in crowdfunding. *Entrepreneurship Theory & Practice*, *39*(1), 75-100. https://doi.org/10.1111/etap.12118.

Connelly, S. Certo, T., & Ireland, D. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39-67.

Cordova, A., Dolci, J., & Gianfrate, G. (2015). The determinants of crowdfunding success: Evidence from technology projects. *Procedia Social & Behavioral Sciences*, *181*, 115-124. https://doi.org/10.1016/j.sbspro.2015.04.872.

Courtney, C., Dutta, S., & Li, Y. (2016). Resolving information asymmetry: Signaling, endorsement, and crowdfunding success. *Entrepreneurship Theory & Practice*, *41*(2), 265-290. https://doi.org/10.1111/etap.12267.

Crawford, M. (2021, May 7). *Equity crowdfunding Q1 2021 review.* https://www.startengine.com/blog/equity-crowdfunding-q1-2021-review/

Dority, B., Borchers, S. J., & Hayes, S. K. (2021). Equity crowdfunding: US title II offerings using sentiment analysis. *Studies in Economics & Finance, 38*(4), 807-835.

Du, L., Bartholomae, F., & Stumpfegger, E. (2021). *What factors are relevant for success in UK equity crowdfunding?* https://ssrn.com/abstract=3800747 or http://dx.doi.org/10.2139/ssrn.3800747

Estrin, S., Khavul, S., & Wright, M. (2022). Soft and hard information in equity crowdfunding: Network effects in the digitalization of entrepreneurial finance. *Small Business Economics*, *58*(4), 1761-1781. Doi: 10.1007/s11187-021-00473-w.

Hayes, S., Dority, B., & Borchers, S. (2020). Equity cro wdfunding success: An examination of title II offerings. *Review of Economic & Business Studies, 13*(1), 9-25.

Horvat, E. A., & Papamarkou, T. (2017). *Gender differences in equity crowdfunding*. Presented in Proceedings of the Firth Conference on Human Computation and Crowdsourcing, Quebec City, Canada.

Horvat, E. -A., Wachs, J., Wang, R., & Hannák, A. (2018, July 5-8). *The role of novelty in securing investors for equity crowdfunding campaigns* (pp. 50-59). Presented in The Sixth AAAI Conference on Human Computation and Crowdsourcing (HCOMP), the University of Zurich, Zurich, Switzerland.

Jayaswal, V. (2020, September 14). *Performance metrics: Confusion matrix, precision, recall and F1 score*. Towards Data Science. https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262.

Kleinert, S., & Volkmann, C. (2019). Equity crowdfunding and the role of investor discussion boards. *Venture Capital*, *21*(4), 327-352.

Kleinert, S., Volkmann, C., & Grünhagen, M. (2020). Third-party signals in equity crowdfunding: The role of prior financing. *Small Business Economics*, *54*(1), 341-365.

Krawczyk, B. (2016, April 22). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, *5*, 221-232.

Le Pendeven, B., & Schwienbacher, A. (2021). Equity crowdfunding: The influence of perceived innovativeness on campaign success. *British Journal of Management*, *34*(1), 280-298. Doi: 10.1111/1467-8551.12585.

Li, X., Tang, Y., Yang, N., Ren, R., Zheng, H., & Zhou, H. (2016). The value of information disclosure and lead investor in equity-based crowdfunding. *Nankai Business Review International*, *7*(3), 301-321.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance, 66*(1), 35-65.

Lukkarinen, A., Teich, J. E., Wallenius, H., & Wallenius, J. (2016). Success drivers of online equity crowdfunding campaigns. *Decision Support Systems*, *87*, 26-38.

Malaga, R., Mamonov, S., & Rosenblum, J. (2017). Gender differences in equity crowdfunding: An exploratory analysis. *International Journal of Gender & Entrepreneurship*, *10*(4), 332-343. https://doi.org/10.1108/IJGE-03-2018-0020.

Mamonov, S., & Malaga, R. (2017). Success factors in title II equity crowdfunding in the United States. Paper presented in Twenty-third American Conference on Information Systems, Boston.

Mamonov, S., & Malaga, R. (2018). Success factors in title III equity crowdfunding in the United States. *Electronic Commerce Research & Applications*, *27*, 65-73.

Mamonov, S., & Malaga, R. (2019). Success factors in title II equity crowdfunding in the United States. *Venture Capital*, *21*(2-3), 223-241.

Mamonov, S., Malaga, R., & Rosenblum, J. (2017). An exploratory analysis of title II equity crowdfunding success. *Venture Capital*, *19*(3), 239-256.

Miglo, A. (2022, January 16). *Comparing crowdfunding theory and practice: The case of technology industry in England.* https://ssrn.com/abstract=3818387 or http://dx.doi.org/10.2139/ssrn.3818387.

Nitani, M., Riding, A., & He, B. (2019). On equity crowdfunding: Investor rationality and success factors. *Venture Capital*, *21*(2-3), 243-272.

Piva, E., & Rossi-Lamastra, C. (2018). Human capital signals and entrepreneurs' success in equity crowdfunding. *Small Business Economics*, *51*(3), 667-686.

Ralcheva, A., & Roosenboom, P. (2020). Forecasting success in equity crowdfunding. *Small Business Economics*, *55*, 39-56.

Reichenbach, F., & Walther, M. (2021). Signals in equity-based crowdfunding and risk of failure. *Financial Innovation*, *7*(1), 1-30.

Schwienbacher, A., & Larralde, B. (2010). Crowdfunding of small entrepreneurial ventures. In D. J. Cumming (ed), *The Oxford handbook of entrepreneurial finance.* Oxford: Oxford University Press.

SEC. (2013). Investor bulletin: Accredited investors. SEC's Office of Investor Education and Advocacy.

Shafi, K. (2021). Investors' evaluation criteria in equity crowdfunding. *Small Business Economics*, *56*(1), 3-37.

Silge, J., & Robinson, D. (2018). *Text mining with R: A tidy approach. Creative Commons Attribution-Non Commercial-Share Alike 3.0,* Retrieved November 19, 2021, from https://www.tidytextmining.com/index.html.

Szmigiera, M. (2019). *Crowdfunding – statistics & facts*. Statista Research Department. https://www.statista.com/topics/1283/crowdfunding/

Troise, C., Matricano, D., Sorrentino, M., & Candelo, E. (2021). Investigating investment decisions in equity crowdfunding: The role of projects' intellectual capital. *European Management Journal*, *40*(3), 406-418. Doi: 10.1016/j.emj.2021.07.006.

Vismara, S. (2016). Equity retention and social network theory in equity crowdfunding. *Small Business Economics*, *46*(4), 579-590.

Vismara, S. (2018). Information cascades among investors in equity crowdfunding. *Entrepreneurship Theory & Practice*, *42*(3), 467-497.

Vulkan, N., Åstebro, T., & Sierra, M. F. (2016). Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights*, *5*, 37-49.

Yan, S. (2015). *Which signaling factors facilitate the success probability of equity crowdfunding?* Bachelor's Thesis, University of Twente.

Appendix A
Variable Definitions

Variable	Description
Success	Binary variable equal to 1 if 100% (or more) of requested amount was raised; 0 if < 100%
Word Count (Readability)	The number of words in the description excluding stopwords
Description Tone (Sentiment)	Loughran and McDonald (2011): percentage of positive words minus percentage of negative words (excluding stopwords)
Issue Amount	Target Offering Amount (in thousands)
Monthly VIX Value	Average monthly Chicago Board Options Exchange Volatility Index (VIX) value

Table 5

Panel A: Comparative Results of the Logistic Regression Model			
	Model developed on unbalanced training set to unbalanced test set	**Model developed on balanced training set to unbalanced test set**	**Model developed on balanced training set to entire dataset**
Accuracy	98.08%	83.04%	85.42%
Precision	0.00%	7.75%	4.74%
Recall	0.00%	40.74%	44.90%
AIC	480.464	79.361	79.361
Log Likelihood	-209.232	-11.680	-11.680

Panel B: Dataset Values for a Potential Campaign			
Tone (percentage of positive words to total words)	**Minimum Asking Amount (in Thousands)**	**Market Volatility**	**Word Count**
0.20	1,000	13.75	100

Ln(Odds of Success)= 20.876+–110.142(.2)+0.001(1,000)-0.144(100)-2.145(13.75)

Panel C: Prediction Values for a Potential Campaign	
Measure	**Value**
Ln(Odds of Success)	0.01065
Odds	1.01071
Probability	0.50300
Prediction	Success