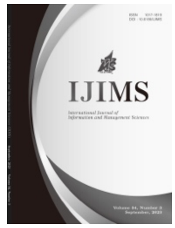
International Journal of
Information
and
Management
Sciences

http://ijims.ms.tku.edu.tw/main.php

International Journal of
Information and Management Sciences



34 (2023), 179-191. DOI:10.6186/IJIMS.202309_34(3).0001

Predicting the Success Rate of Reward-Based Crowdfunding Campaigns: Evidence from Machine Learning

Chia-Ling Chang[1] and *Yen-Sheng Lee*[2*]
[1]Research Division VI, Taiwan Institute of Economic Research, Taiwan
[2]Davis School of Business, Colorado Mesa University, USA

Keywords

Machine learning, crowdfunding success factors, venture capital (VC), entrepreneurial experience

Abstract

Prior research argues that the characteristics of crowdfunding campaigns affect their success rate. We examine this further to understand whether success in funding projects can be predicted by associated project characteristics. We apply machine learning to classify reward-based projects in the crowdfunding market. Specifically, we construct three classification tree-based models and provide evidence that the proposed machine learning models have a strong out-of-sample predictive power over the probability of fundraising success. In addition, the robustness check through both logistic regression and propensity score matching approaches confirms that project characteristics, except for those linked to venture capital, are among the factors behind crowdfunding campaign success. This study can assist entrepreneurs in understanding the impact of project characteristics on the crowdfunding success rate.

1. Introduction

Crowdfunding has broadened access to funding and changed the way capital is raised in the current times. Compared with traditional forms of finance, crowdfunding is still a young and evolving form of finance where entrepreneurs or start-up companies finance new ideas, products, or services by soliciting funds from a large number of crowdfunders via the Internet. Prior research contends that the characteristics of crowdfunding campaigns affect their success rate. Unlike many studies that used regular econometric techniques to identify factors behind success in securing funds in the crowdfunding market, our study adopts a machine learning (ML) approach to distinguish factors behind successful fundraising from those behind failed fundraising. With the advance of computer technology, ML algorithms have been applied recently to forecast asset returns in financial markets. We implement three ML algorithms, namely, decision tree, random forest, and gradient boosted tree, which enable computers to recognize patterns, that is, crowdfunding factors contributing to a positive fundraising outcome

* corresponding author

such that the fundraising proceeds exceed the crowdfunding target set by an entrepreneur. The empirical results of application of the algorithms point out that an entrepreneur's experience is the most important factor determining the success of a crowdfunding campaign. Put differently, the more experienced entrepreneurs are in convincing potential crowdfunders to finance a campaign, the better chance the campaign has that it will thrive. As a robustness check, a logistic regression model is deployed to investigate the relationship of crowdfunding attributes with the odds of succeeding in the crowdfunding market. We discover that majority of crowdfunding attributes are statistically significant and connected to the success rate of a crowdfunding event, evidence in agreement with the results of our ML models.

Additionally, we scrutinize the role of venture capital (VC) in crowdfunding campaigns. In respect of entrepreneurs and VC firms, crowdfunding is not only a financing method but also a means of expanding prospective online contributors. The literature on entrepreneurial finance generally supports the endorsement effect that traditional VC-funded seed ventures signal to the outsiders the superiority (i.e., superior quality) of the VC-backed over non-VC-backed start-ups. There have been few studies, however, exploring the endorsement effect of venture capitalists on the success rate of a crowdfunding campaign. Drawing on propensity score matching method, this study also aims at filling this gap within the literature and analyzes, as research question, the impact of VC endorsement on the chances of a crowdfunding campaign success. It must be noted that, if the use of VC eliminates the information asymmetry inherent in start-ups that raise funds through crowdfunding, VC-backed start-ups viewed as high-quality investments will become more appealing to crowdfunders than their non-VC-backed counterparts viewed as low-quality investments. In other words, VC-backed crowdfunding campaigns may be more successful than those not funded by VC. The results of propensity score matching model imply that VC support does not raise the success probability in the crowdfunding market, contrary to the literature on entrepreneurial finance. In sum, our models with their findings contribute to the existing literature on crowdfunding success rate.

The rest of the paper is organized as follows. Section 2 reviews the related literature, which is followed by the hypothesis and an elaboration of the data and methodology in Section 3. Subsequently, Section 4 discusses the results of the ML algorithms. Section 5 presents the findings of logistic regression and propensity score matching. Section 6 concludes the study.

2. Literature Review

This section reviews the literature on key factors behind crowdfunding success. This is followed by the review of literature on the capabilities of venture capitalists in sorting and shepherding start-ups.

2.1 Determinants of Crowdfunding Success

The crowdfunding literature tends to support the notion that the market demand, information transparency, confidence, and crowdfunding experience are associated with the outcomes of a crowdfunding campaign. The attributes of reward proposals may trigger market demand, which, in turn, is critical for success of crowdfunding projects (Wei & Lin, 2016). Backers' willingness to pre-acquire crowdfunding rewards hinges on the information transparency of a crowdfunding project (Oo et al., 2019; Lagazio & Querci, 2018). Several studies (Mollick, 2014; Arrow, 1962) validate the viewpoint that the confidence and experience of entrepreneurs are positively related to the probability of crowdfunding success.

2.2 Screening Process of Venture Capitalists

The literature on entrepreneurial finance agrees on the importance of the role of VC in financial markets owing to VC's screening activities for start-ups. The role of venture capitalists is different from that of individual investors, also known as backers in the context of reward-based crowdfunding. The body of literature on the role of VC in the entrepreneurial setting verifies the standpoint that venture capitalists scout for and coach promising start-ups to boost the overall return on their portfolio company investments (Baum & Silverman, 2004; Baum et al., 2000).

Both high uncertainty and strong information asymmetry lead to adverse selection and financing constraints for individual investors and entrepreneurs, respectively. However, venture capitalists draw on screening activities to resolve the information asymmetry in financial markets. Hence, the literature suggests that the dominance of venture capitalists over start-up investments helps start-ups succeed (Gompers et al., 2008; Zacharakis & Meyer, 2000; Amit et al., 1998; Fried & Hisrich, 1994).

Venture capitalists' assessment of new venture survival aims to reduce the risk of adverse selection. They consider the deal selection process the most important practice contributing to value creation (Gompers et al., 2020; Shepherd, 1999; Tyebjee & Bruno, 1984). Typical investment criteria in screening a deal are market awareness (e.g., size, growth, and access to customers), product differentiation (e.g., patents and technical edge), managerial capabilities (e.g., marketing and management skills and the references of entrepreneurs), and the cash-out potential (e.g., mergers, acquisitions, and public offerings).

Although venture capitalists expend a great deal of effort on collecting information to eliminate low-quality proposals, during a business pitch, they identify signals that would determine whether the proposal will be successful (Kirsch et al., 2009; Rosch, 1975). The research shows that the VC selection criteria comprise interorganizational endorsements (Stuart et al., 1999), top-managers' talent (Zarutskie, 2010; Franke et al., 2008; Beckman et al., 2007; Eisenhardt & Schoonhoven, 1990; Becker, 1962), and founders' human capital (Chatterji, 2008; Colombo et al., 2004; Almus & Nerlinger, 1999; Stuart et al., 1999). These selection criteria are positively correlated with the possibility of securing VC funds.

As most of the related literature examines the role of VC in equity-based crowdfunding (Mamonov & Malaga, 2018; Mamonov et al., 2017; Lukkarinen et al., 2016; Ahlers et al., 2015), we deal with reward-based crowdfunding and investigate whether the endorsement impact brought by VC investment increases the likelihood of entrepreneurial firms to prosper in the subsequent crowdfunding.

3. Hypothesis and Data

Following Chang et al. (2021)'s study, we divide the attributes of a crowdfunding project into five categories: the mastery of market demand, transparency of project information, confidence of proponents, crowdfunding experience of proponents, and support of venture capital (VC). If crowdfunding attributes come into play in the decision-making of both entrepreneurs and backers, we would expect the subsequent outcome (either success or failure) of a crowdfunding campaign to be susceptible to these attributes. Therefore, our crowdfunding success hypothesis is articulated as follows:

H1: Crowdfunding attributes are positively correlated with the likelihood of being successful in the crowdfunding market.

Our sample consists of 274,836 reward-based crowdfunding projects collected from both Crunchbase website (https://www.crunchbase.com/home) and the FINDIT database (https://findit.org.tw/English/index.aspx). Market demand is proxied by the actual proceeds (*Funding*), the numbers of reward proposals ($N_{proposal}$) and backers (N_{backer}) in a crowdfunding project. Information transparency is measured by the number of comments ($N_{comment}$) posted by backers and numbers of updates (N_{update}), pictures ($N_{picture}$), and videos (N_{video}) uploaded by proponents. Confidence is gauged by the duration of a crowdfunding campaign (*Duration*) and the natural logarithm of the target amount of funding (*Lngoal*) set by a proponent. Experience is quantified by the numbers of past successful ($N_{success}$) and failed ($N_{failure}$) campaigns, and the number of crowdfunding projects ($N_{backing}$) with which a proponent helps other proponents. VC support is proxied by the amount of a VC's investment (*VCmoney*) and a dummy variable (VC_d) that takes on the value of one if a VC supports a crowdfunding start-up, and zero otherwise. Whether a crowdfunding project succeeds or fails is indicated by a dummy variable ($Success_d$) that equals one if actual crowdfunding proceeds are higher than the fundraising goal, and zero otherwise.

Table 1 lists all the crowdfunding variables and summarizes the statistics of each variable under investigation. The *Funding*, *VCmoney,* VC_d, and $Success_d$ statistics are notable. The average of crowdfunding proceeds is only \$11,895 and a relatively small deal, compared with other large transactions in financial markets (e.g., initial public offerings). The averages of *VCmoney* and VC_d are \$1,044 and 0.0003, respectively, suggesting that a VC does not participate as much in the crowdfunding market as in the VC market. The average success rate of a crowdfunding campaign ($Success_d$) is merely 0.28 and indicates a low probability of a successful crowdfunding project.

Table 1: *Summary and statistics of variables*

Category	Variable	Symbol	Description	Average	Std. dev.
Market demand	*Proceeds*	*Funding*	A project's actual proceeds in U.S. dollars secured by an entrepreneurial start-up	\$11,895	\$116,095
	Reward proposals	N_{proposal}	The number of reward proposals in a crowdfunding campaign	7.91	6.01
	Backers	N_{backer}	The number of backers (sponsors) in a crowdfunding campaign	135	1,078
Transparency	*Comments*	$N_{comment}$	The number of comments posted by backers in a crowdfunding campaign	52	1,298
	Updates	N_{update}	The number of updates entered by a proponent	3.84	9.16
	Pictures	$N_{picture}$	The number of pictures posted by a proponent	7.66	11.64
	Videos	N_{video}	The number of videos uploaded by a proponent	0.78	0.82
Confidence	*Duration*	*Duration*	The duration of a crowdfunding campaign set by a proponent in number of days	33.16	10.86
	Goal	*Lngoal*	The natural logarithm of a proponent's target fundraising amount	8.77	1.76
Experience	*Success*	$N_{success}$	The number of past successful crowdfunding projects of a proponent	0.92	3.26
	Failure	$N_{failure}$	The number of past failed crowdfunding projects of a proponent	1.03	1.11
	Backing	$N_{backing}$	The number of crowdfunding projects with which a proponent helps other proponents	6.37	24.95
VC support	*VCmoney*	*VCmoney*	The amount of a VC's investment in U.S. dollars before launching a crowdfunding campaign	\$1,044	\$201,632
	VC dummy	VC_d	A binary dummy variable. Coded 1 (VC-supported), if a VC supports a start-up before launching a crowdfunding campaign, and 0 (non-VC-supported) otherwise	0.0003	0.0183
Crowdfunding outcome	*Success dummy*	$Success_d$	A binary dummy variable. Coded 1 (successful), if a project's actual proceeds are equal to or greater than the fundraising target, and 0 (unsuccessful) otherwise	0.28	0.45

4. Methodology and Results

To forecast the outcome (classes or labels in the context of ML) of a crowdfunding project, we rely on ML classification approach where classification results are labeled as $Success_d = 1$ or 0, and review decision tree learning models (Kuhn & Johnson, 2013) that are among the most popular machine learning algorithms given their simplicity. We briefly outline how to appraise the effectiveness of decision tree algorithms, namely, predictive accuracy and feature importance score. This is followed by the empirical estimations from three decision tree learning models.

4.1 Decision Tree Learning

The goal of the decision tree learning approach is to create a model that predicts the value of a target (dependent) variable based on input (independent) variables. Decision trees where the target variable takes discrete and continuous values are called classification and regression trees, respectively. A decision tree classification method builds a decision tree where a dataset is split based on dependent variables. Specifically, the decision tree starts from a root node that recursively partitions the data space into subsets that contain instances with similar outcomes (e.g., successful or failed crowdfunding projects). In an iterative process, we can repeat this splitting procedure until the decision tree yields the largest information gain (conceptually akin to identifying the minimum of the loss function) such that successful or failed projects are grouped together. Unlike a simple decision tree model, ensemble methods (for example, random decision forests and gradient boosted trees) use more than one decision tree. Random decision forests construct multiple decision trees by repeatedly resampling the subsets of the data with replacement. Gradient boosted trees allow optimization along with stochastic gradient boosting algorithm where, at each iteration, a subsample of the data is drawn at random without replacement.

4.2 Effectiveness of Decision Tree Algorithms

As decision tree algorithms are constructed to obtain final classification results, the effectiveness of ML models can be measured by either predictive accuracy (ACC) or feature importance score (FIS). In all our ML models, we stipulate the training sample size as 70% of the total sample size. The remaining 30% is utilized as the out-of-sample testing dataset, and the ACC of the out-of-sample dataset is obtained as follows:

$$ACC = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions\ (out\text{-}of\text{-}sample\ size)} \tag{4.1}$$

Gini Index is used to estimate the node impurity and feature importance is the measurement of a reduction in the node impurity weighted by the number of observations that reach a particular node from the total number of observations. The following equation calculates Gini Index (GI_m):

$$GI_m = \sum_{k=1}^{|K|} \sum_{k' \neq k} P_{mk} P_{mk'} = 1 - \sum_{k=1}^{|K|} P_{mk}^2 \tag{4.2}$$

where K is the number of classes (K=2 because of two classes of $Success_d$ being 1 or 0), and P_{mk} is the proportion of class k at node m.

As Gini Index (GI_m) attains the importance of a node m, and a single feature (independent variable) can appear in several branches of a decision tree, we compute the *FIS* of feature i as shown below:

$$FIS_i = \frac{\sum_{m:node\ m\ slplits\ on\ feature\ i} GI_m}{\sum_{m \in all\ nodes} GI_m} \tag{4.3}$$

Note that the FIS_i are normalized against the sum of all feature importance scores across all features. Therefore, the sum of all FIS_i is equal to one.

4.3 Results of Decision Tree Algorithms

Decision tree, random forest, and gradient boosted tree algorithms are implemented in Python 3.8. The predictive accuracies of the three models fall in the range of 97%–99%, with *ACC* of 97%, 99%, and 99% for decision tree, random forest, and gradient boosted tree algorithms, respectively. Not only do the precise predictions of the out-of-sample dataset hint that our ML models are effective, but they also are in favor of our hypothesis (*H1*), that is, crowdfunding features are corelated with the odds of a successful crowdfunding campaign.

Table 2: *Results of the feature importance scores (FIS)*

Decision Tree		**Random Forest**		**Gradient Boosted Tree**	
Variable	*FIS*[a]	Variable	*FIS*	Variable	*FIS*
Success	0.8525	*Success*	0.3059	*Success*	0.8232
Failure	0.0963	*Failure*	0.2462	*Failure*	0.1223
Backers	0.0338	*Backers*	0.1225	*Backers*	0.0157
Goal	0.0100	*Proceeds*	0.1039	*Comments*	0.0128
Pictures	0.0072	*Goal*	0.0552	*Pictures*	0.0116
VCmoney	0.0000	*Pictures*	0.0487	*Updates*	0.0056
VC dummy	0.0000	*Updates*	0.0474	*Proceeds*	0.0044
Backing	0.0000	*Comments*	0.0444	*Goal*	0.0036
Duration	0.0000	*Backing*	0.0143	*Videos*	0.0003
Videos	0.0000	*Reward proposals*	0.0052	*Backing*	0.0001
Updates	0.0000	*Duration*	0.0038	*Duration*	0.0000
Comments	0.0000	*Videos*	0.0024	*Reward proposals*	0.0000
Reward proposals	0.0000	*VCmoney*	0.0000	*VCmoney*	0.0000
Proceeds	0.0000	*VC dummy*	0.0000	*VC dummy*	0.0000
Sum of *FIS*	1.0000		1.0000		1.0000

Note: [a] The values of the feature importance score (*FIS*) are normalized so that the sum of the feature importance scores is equal to one.

Table 2 reports the statistics of FIS on each crowdfunding feature. The top two features are $N_{success}$ and $N_{failure}$ as the sum of their FIS is well above 50%. The estimated values of FIS of past successful (failed) crowdfunding campaigns are 0.8525, 0.3059, and 0.8232 (0.0963, 0.2462, and 0.1223) in decision tree, random forest, and gradient boosted tree models, respectively. The third most important feature is Backers, regardless of the three ML estimations. Compared with those of $N_{success}$ and $N_{failure}$, the magnitudes of FIS of Backers are much small and about 12% in the random forest model and in the neighborhood of 1.5% to 3.3% in the decision tree and gradient boosted tree models. Given that the remainder of FIS are trivial, our findings imply that entrepreneurs' prior experience, in either successful or failed crowdfunding projects, has a predictive power over the outcomes of subsequent crowdfunding campaigns.

5. Robustness Check

The main purpose of the robustness check is to see whether the effects of crowdfunding variables on campaign success obtained in the ML tests are sensitive to conventional econometric methodologies, two of which are logistic regression and propensity score matching.

5.1 Logistic Regression

This study uses logistic regression to explore the impact of crowdfunding traits on the probability of project funding success. Similar to the crowdfunding success dummy variable ($Success_d$) described in Section 3, L_i is defined as an observable variable to register the binary outcomes (success or failure) for any crowdfunding project i. We define a project as a success if the project reaches its funding threshold; hence, a successful project takes binary value one (L_i=1). A project is defined as a failure if the proceeds fall short of the funding goal; hence, the binary value of a failed project is zero (L_i=0). Let L_i^* denote the latent (unobservable) dependent variable, which is determined by the logistic regression model below:

$$L_i^* = \sum_{k=1}^{K} \beta_k x_{k,i} + \epsilon_i \tag{5.1}$$

where $x_{k,i}$ is the kth explanatory variable, β_k is the slope coefficient associated with the kth regressor, and ε_i is the error term. L_i^* has the following mapping function with the observable dependent variable L_i:

$$L_i = \begin{cases} 0, \text{if } L_i^* \leq \delta \\ 1, \text{if } L_i^* > \delta \end{cases} \tag{5.2}$$

where δ is the unknown threshold to be estimated along with the slope parameters, β_k (k=1, 2, ..., K) in Equation (5.1). If L_i^* strictly exceeds a certain threshold value δ, then the project is successful (L_i=1); otherwise, the project has failed (L_i=0). Under the assumption of the error term ε_i being normally distributed, the corresponding marginal probability with respect to $x_{k,i}$ (the kth explanatory variable) can be derived as follows:

$$\frac{\partial \Pr(L_i = m)}{\partial x_{k,i}} = \begin{cases} -\phi\left(\delta - \sum_{k=1}^{K} \beta_k x_{k,i}\right)\beta_k, \text{if } m = 0 \\ \phi\left(\delta - \sum_{k=1}^{K} \beta_k x_{k,i}\right)\beta_k, \text{if } m = 1 \end{cases} \tag{5.3}$$

where $\phi(.)$ is the operator of a normally distributed probability density function (PDF) and is always positive. Therefore, the sign of the estimated value of β_k decides the positive/negative sign of the estimated marginal probability of a project's success (L_i=1) or failure (L_i=0). Specifically, if the sign of the estimated β_k is positive (negative), this implies the marginal probability of a project being successful increases (decreases) as $x_{k,i}$ increases. Because there are only two types (success and failure) of fundraising outcomes in our study, we avoid a common problem that the sign of the marginal probability of intermediate types of outcomes (i.e., more than two types of crowdfunding project outcomes) cannot be determined by the sign of the estimated β_k. Because computing marginal probabilities becomes extraneous in our logistic regression model, it is sufficient to obtain the estimates of all unknown parameters, including $\{\beta_1, ..., \beta_K, \delta\}$, for the purpose of our model design. These unknown parameters of the above logistic regression model, characterized by Equations (5.1)–(5.3), can be estimated by the maximum likelihood method.

To investigate the determinants of the probability of project funding success (L_i), we regress L_i, viewed as the dependent variable, on an array of independent variables in Equation (5.4) as follows:

$$L_i = \beta_0 + \beta_1 N_{proposal} + \beta_2 N_{backer} + \beta_3 N_{comment} + \beta_4 N_{update} + \beta_5 N_{picture} + \beta_6 N_{video} + \beta_7 Duration + \beta_8 Lngoal + \beta_9 N_{success} + \beta_{10} N_{failure} + \beta_{11} N_{backing} + \beta_{12} VCmoney + \beta_{13} VC_d \tag{5.4}$$

Table 3 presents the estimated coefficients of the logistic regression in which most of the results are consistent with the literature on crowdfunding. The estimated slope coefficients with respect to market demand, information transparency, and crowdfunding experience apart from $N_{failure}$ are positive and statistically significant at the 1% significance level. Except for the project duration, funding goal, and the number of failed projects, all these explanatory variables positively contribute to the success rate of crowdfunding campaigns.

Referring to the literature on VC financing, we look at both *VCmoney* and VC_d variables because they indicate whether the presence of VC firms and injection of VC seeding money increase the success probability of crowdfunding. Given that the estimated slope coefficients of *VCmoney* and VC_d are not statistically significant, the evidence concerning both coefficients is clear—VC endorsement does not lead to a statistical increase in the probability of crowdfunding success.

Table 3: *Results of logistic regression on crowdfunding success (L_i)*

Category	Variable	Symbol	Coefficient[a]
-	*Constant*	-	4.95***
Market demand	*Reward proposal*	N_{proposal}	0.05***
	Backer	N_{backer}	0.01***
Transparency	*Comment*	$N_{comment}$	0.00***
	Update	N_{update}	0.15***
	Picture	$N_{picture}$	0.04***
	Video	N_{video}	0.30***
Confidence	*Duration*	*Duration*	–0.02***
	Goal	*Lngoal*	–0.68***
Experience	*Success*	$N_{success}$	1.20***
	Failure	$N_{failure}$	–3.24***
	Backing	$N_{backing}$	0.00***
VC support	*VCmoney*	*VCmoney*	0.00
	VC dummy	VC_d	0.48

Note: [a] *, **, and *** represent 10%, 5%, and 1% significance levels, respectively.

5.2 Propensity Score Matching

In the context of adverse selection problem mitigated by VC endorsement, we investigate whether the involvement of VC could help crowdfunders to differentiate high-quality projects (VC-backed) from the low-quality (non-VC-backed) ones. To examine the influence of VC endorsement, if any, on crowdfunding success rate, we work with the framework of Rosenbaum and Rubin (1983), that is, propensity score matching (PSM). As is standard in potential outcome models, PSM takes on predicted probabilities of groups, for example, VC-backed versus non-VC-backed projects, based on an observed predictor (VC_d). Being a statistical matching technique, PSM reduces the bias due to confounding variables that could be found in estimating probabilities among groups. Following the PSM model of Zhao and Vinig (2017), we consider the equation below:

$$Success_{d,i} = \alpha_0 + \alpha_1 VC_{d,i} + \alpha_1 X_i + \epsilon_i \tag{5.5}$$

where $Success_{d,i}$ is a dichotomous indicator of the potential outcome for project i. $Success_{d,i}$ =1 is observed when the project wins the campaign and $Success_{d,i}$ =0 otherwise. Let VC_d denote an indicator variable for VC endorsement. Confounding variable, X_i is a vector of characteristics of the i^{th} project and includes all the variables listed in Table 1, except for VC_d and $Success_d$.

In traditional PSM estimation, each VC-backed project is paired with a few non-VC-backed projects, under the assumptions of conditional independence and common support. Controlling for confounders (X) that could affect the odds of securing VC investment as well as crowdfunding outcomes, the average treatment effect (VC endorsement effect in our study) can be illustrated by the mean within-match difference in the result variable ($Success_d$) between VC-backed and non-VC-backed campaigns. Mathematically, the average treatment effect (ATE) for VC-backed projects can be expressed as follows:

$$ATE = Prob.(Success_d, VC_d = 1|X) - Prob.(Success_d, VC_d = 0|X) \tag{5.6}$$

where $Prob.\ (Success_d,\ VC_d{=}1|X)$ and $Prob.\ (Success_d,\ VC_d{=}0|X)$ are the probabilities of being successful in VC-backed and non-VC-backed projects, respectively, conditional on confounding effect of X.

The estimation of ATE through the above PSM procedures is merely 0.01 and not statistically significant, evidence that VC endorsement has no influence on the outcomes of crowdfunding campaigns.

6. Conclusion

Our paper sheds light on crowdfunding success rate. The findings of our three ML models suggest that the experience (success or failure) of a proponent is a discriminating factor in the success rate of a crowdfunding project. The feature importance score (FIS) of our decision tree-based learning models is calculated as the average decrease in impurity and evaluated by its relative significance. Among the 14 features, the number of past successful crowdfunding campaigns ($N_{success}$) ranks first, followed by the number of past failed crowdfunding campaigns ($N_{failure}$). Empirical evidence from the logistic regression model and propensity score matching is somewhat in line with that from the ML models. Results of the logistic regression prove that crowdfunding characteristics, except for VC support, are statistically significant at the 1% significance level and positively associated with the outcomes of crowdfunding campaigns. We take a further step to quantify the effect of VC endorsement on the probability of crowdfunding success. Results of propensity score matching show that VC endorsement has no impact on the crowdfunding success rate.

Our contributions to the existing literature are twofold. First, to the best of our knowledge, our study is the first to explore crowdfunding results by applying ML algorithms, different from other studies that used conventional econometric methods for the purpose. Our second contribution is that we show that VC-accreditation does not guarantee success of reward-based start-ups, which is conditional on other determinants. All our models, including ML, logistic regression, and propensity score matching consistently point out that VC is not the key to success in a reward-based crowdfunding campaign possibly because of the rewards effect in which the backers' focus on rewards may cause the signaling effect of VC to fail in enhancing their sensitivities to VC support.

As our study responds to a growing interest in the crowdfunding market, the practical implication for decision makers is that entrepreneurs should deepen and advertise their crowdfunding experience to prospective crowdfunders in raising the likelihood of success when initiating a crowdfunding campaign. Although VC plays a vital role in financial markets, it may not be worth the effort for crowdfunding entrepreneurs to pursue VC funding on account of the low impact of VC endorsement on crowdfunding success. A direction for future research is the application of artificial intelligence to predict the outcomes of crowdfunding projects because machine learning is a subset of artificial intelligence. Future studies can address the role of VC in equity crowdfunding campaigns to explore whether VC endorsement affects the decision-making process of equity-based crowdfunders who may be sensitive to the signaling effect of VC.

References

Ahlers, G. K. C., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice*, *39*(4), 955-980. https://doi.org/10.1111/etap.12157

Almus, M. & Nerlinger, E. A. (1999). Growth of new technology-based firms: Which factors matter? *Small Business Economics*, *13(2),* 141-154. https://doi.org/10.1023/A:1008138709724

Amit, R., Brander, J., & Zott, C. (1998). Why do venture capital firms exist? Theory and Canadian evidence. *Journal of Business Venturing*, *13*(6), 441-466. https://doi.org/10.1016/S0883-9026(97)00061-X

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, *29*(3), 155-173. https://doi.org/10.2307/2295952

Baum, J. A. C., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, *21*(3), 267-294. https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<267::AID-SMJ89>3.0.CO;2-8

Baum, J. A. C., & Silverman, B. S. (2004). Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *Journal of Business Venturing*, *19*(3), 411-436. https://doi.org/10.1016/S0883-9026(03)00038-7

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, *70*(5), 9-49. https://doi.org/10.1086/258724

Beckmanc, C. M., Burton, M. D., & O'Reilly, C. (2007). Early teams: The impact of team demography on VC financing and going public. *Journal of Business Venturing*, *22*(2), 147-173. https://doi.org/10.1016/j.jbusvent.2006.02.001

Chang, C., Lee, Y., & Tien, H. (2021). Does venture capital affect crowdfunding performance? *Economic Review: Journal of Economics and Business*, *19*(2), 53-64. https://doi.org/10.51558/2303-680X.2021.19.2.53

Chatterji, A. K. (2009). Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal*, *30*(2), 185-206. https://doi.org/10.1002/smj.729

Colombo, M. G., Delmastro, M., & Grilli, L. (2004). Entrepreneurs' human capital and the start-up size of new technology-based firms. *International Journal of Industrial Organization*, *22*(8-9), 1183-1211. https://doi.org/10.1016/j.ijindorg.2004.06.006

Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures, 1978-1988. *Administrative Science Quarterly*, *35*(3), 504-529. https://doi.org/10.2307/2393315

Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (2008). Venture capitalists' evaluations of start-up teams: Trade-offs, knock-out criteria, and the impact of VC experience. *Entrepreneurship Theory and Practice*, *32*(3), 459-483. https://doi.org/10.1111/j.1540-6520.2008.00236.x

Fried, V. H., & Hisrich, R. D. (1994). Toward a model of venture capital investment decision making. *Financial Management*, *23*(3), 28-37. https://www.jstor.org/stable/3665619

Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, *135*(1),169-190. https://doi.org/10.1016/j.jfineco.2019.06.011

Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2008). Venture capital investment cycles: The impact of public markets. *Journal of Financial Economics*, *87*(1), 1-23. https://doi.org/10.1016/j.jfineco.2006.12.002

Kirsch, D., Goldfarb, B., & Gera, A. (2009). Form or substance: The role of business plans in venture capital decision making. *Strategic Management Journal*, *30*(5), 487-515. https://doi.org/10.1002/smj.751

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling.* New York, NY: Springer. https://link.springer.com/book/10.1007/978-1-4614-6849-3

Lagazio, C., & Querci, F. (2018). Exploring the multi-sided nature of crowdfunding campaign success. *Journal of Business Research*, *90*, 318-324. https://doi.org/10.1016/j.jbusres.2018.05.031

Lukkarinen, A., Teich, J. E., Wallenius, H., & Wallenius, J. (2016). Success drivers of online equity crowdfunding campaigns. *Decision Support Systems*, *87*, 26-38. https://doi.org/10.1016/j.dss.2016.04.006

Mamonov, S., Malaga, R., & Rosenblum, J. (2017). An exploratory analysis of Title II equity crowdfunding success. *Venture Capital*, *19*(3), 239-256. https://doi.org/10.1080/13691066.2017.1302062

Mamonov, S., & Malaga, R. (2018). Success factors in Title III equity crowdfunding in the United States. *Electronic Commerce Research and Applications*, *27*, 65-73. https://doi.org/10.1016/j.elerap.2017.12.001

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, *29*(1), 1-16. https://doi.org/10.1016/j.jbusvent.2013.06.005

Oo, P. P., Allison, T. H., Sahaym, A., & Juasrikul, S. (2019). User entrepreneurs' multiple identities and crowdfunding performance: Effects through product innovativeness, perceived passion, and need similarity. *Journal of Business Venturing*, *34*(5), Article 105895. https://doi.org/10.1016/j.jbusvent.2018.08.005

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192-233. https://doi.org/10.1037/0096-3445.104.3.192

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55. https://doi.org/10.1093/biomet/70.1.41

Shepherd, D. A. (1999). Venture capitalists' assessment of new venture survival. *Management Science*, *45*(5), 621-632. https://doi.org/10.1287/mnsc.45.5.621

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly*, *44*(2), 315-349. https://doi.org/10.2307/2666998

Tyebjee, T. T., & Bruno, A. V. (1984). A model of venture capitalist investment activity. *Management Science*, *30*(9), 1051-1066. https://doi.org/10.1287/mnsc.30.9.1051

Wei, Z., & Lin, M. (2017). Market mechanisms in online peer-to-peer lending. *Management Science*, *63*(12), 4236-4257. https://doi.org/10.1287/mnsc.2016.2531

Zacharakis, A. L., & Meyer, G. D. (2000). The potential of actuarial decision models: Can they improve the venture capital investment decision? *Journal of Business Venturing*, *15*(4), 323-346. https://doi.org/10.1016/S0883-9026(98)00016-0

Zarutskie, R. (2010). The role of top management team human capital in venture capital markets: Evidence from first-time funds. *Journal of Business Venturing*, *25*(1),155-172. https://doi.org/10.1016/j.jbusvent.2008.05.008

Chia-Ling Chang
Research Division VI, Taiwan Institute of Economic Research, Taiwan.
E-mail address: d14500@tier.org.tw
Major area(s): Economics

Yen-Sheng Lee
Davis School of Business, Colorado Mesa University, USA.
E-mail address: ylee@coloradomesa.edu
Major area(s): Finance

(Received January 2023; accepted May 2023)