International Journal of Forecasting 34 (2018) 30–44

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Using past contribution patterns to forecast fundraising outcomes in crowdfunding

Onochie Fan-Osuala [a,*], Daniel Zantedeschi [b], Wolfgang Jank [c]

[a] University of Wisconsin Whitewater, Information Technology and Supply Chain Management, 800 W. Main St. Whitewater, WI 53190, United States
[b] The Ohio State University, Department of Marketing and Logistics, 2100 Neil Avenue, Columbus, OH 43210, United States
[c] University of South Florida, Information Systems and Decision Sciences, 4202 E Fowler Avenue, Tampa, FL 33620-9951, United States

ABSTRACT

The crowdfunding mechanism has proven to be a practical way of raising funds, especially with the widespread use of the Internet. However, one limitation of current crowdfunding platforms is that it is hard for creators and backers to anticipate the success of a campaign. This paper tackles this limitation. We take a two-pronged approach to building our forecasting model. First, we explore the nature and heterogeneity of contribution dynamics in crowdfunding campaigns and compare them across two natural groups (successful and unsuccessful campaigns). We then use insights generated from our exploratory analysis and draw upon the general laws of motion for stochastic processes in order to introduce a new dynamic model for predicting crowdfunding outcomes. Our model incorporates the history and dynamics of both the focal crowdfunding campaign and other campaigns for predicting outcomes. We compare our model to other parametric and semi-parametric benchmark models, and show substantial improvements.

1. Introduction

In the recent years, crowdfunding (individuals collectively contributing money to back different goals and projects through the internet) has proven to be a viable alternative for raising funds (Kuppuswamy & Bayus, 2014). Businesses, entrepreneurs, and individuals have all used this alternative fundraising mechanism to raise the finances to support their businesses, creative projects, and personal goals. There is evidence on both very successful fundraising campaigns (e.g., Pebble smart watches on Kickstarter.com) and unsuccessful fundraising campaigns (e.g., Meatballs LLC in Ahlers, Cumming, Günther, & Schweizer, 2015). Various different factors have been claimed to drive fundraising campaign outcomes, such as goal size, fundraising duration, creator's network size, and signals of quality (Mollick, 2014). However, despite the growing importance and popularity of the crowdfunding mechanism, most platforms still do not provide any analytics tools for creators and backers beyond simple aggregates. For instance, the major crowdfunding platform provider Kickstarter does not provide analytics tools for tracking projects (Wired.com, 2012). Reports from Huffington Post (2013) and Wired.com (2012) show that both creators and backers are interested in tools that can help them to forecast fundraising outcomes and make more informed decisions.

To the best of our knowledge, there are no known or proposed approaches for forecasting crowdfunding outcomes using granular day-by-data data, which makes the approach described here a first attempt.

This paper (1) explores the dynamics of contributions during crowdfunding campaigns, drawing upon the laws of motion; (2) compares the nature and heterogeneity of the contribution dynamics in successful and unsuccessful crowdfunding campaigns; (3) proposes a novel method of forecasting crowdfunding outcomes based on the contribution dynamics of the focal fundraising campaign derived from our exploratory analyses; and (4) compares our method with other parametric and semi-parametric forecasting methods. Our study provides several insights into the crowdfunding phenomenon and offers various

* Correspondence to: Information Technology & Supply Chain Management, College of Business and Economics, University of Wisconsin, Whitewater 809 W. Starin Road Whitewater, WI 53190, United States
E-mail address: fanosuao@uww.edu (O. Fan-Osuala).

http://dx.doi.org/10.1016/j.ijforecast.2017.07.003

practical advantages, while also contributing to the growing body of literature on crowdfunding and forecasting. Theoretically, we present insights into the nature of backers' contribution patterns and the differences between these patterns in successful and unsuccessful crowdfunding campaigns. Practically, our approach offers advantages to several stakeholders in crowdfunding campaigns. First, it benefits crowdfunding platforms by presenting a forecasting model that can be used to provide users with up-to-date analytics. Second, it provides creators with a method for forecasting fundraising outcomes before the end of their campaign. This allows creators to know in time whether or not they will need to shore up their promotion efforts by aggressively seeking donors outside their networks or persuasively pitching to and closing on investors in order to meet their funding goals. A knowledge of probable outcomes can also help creators in planning. For instance, creators who need to deliver short-term rewards to backers can use their forecasts to plan. Finally, backers who are looking for a better way of cherry-picking winners or making better contribution decisions can benefit from our method too.

The rest of this paper is organized as follows. Section 2 presents a short introduction to crowdfunding and its mechanisms, while Section 3 discusses our study context and data. Section 4 presents our methodology, before we round off with the discussion and conclusion.

2. Crowdfunding

Crowdfunding refers to a fundraising campaign where a creator issues an open call to the public through the internet to raise funds either in the form of donations or in exchange for some reward, equity or voting rights to support initiatives for specific purposes (Belleflamme, Lambert, & Schwienbacher, 2014). It has proven to be a practical way for fledgling entrepreneurs to seek early stage funding (Kuppuswamy & Bayus, 2014), and has also generated a lot of interest in academic circles (Agrawal, Catalini, & Goldfarb, 2011; Mollick, 2014). The number of platforms supporting crowdfunding has grown over the past few years, and crowdfunding platforms like Kickstarter, Gofundme, and IndieGoGo now handle millions of dollars' worth of fundraising transactions. It is estimated that over a million projects were funded successfully by crowdfunding platforms in 2012, raising about $2.7 billion (Massolution, 2013). Recently, crowdfunding has drawn the attention of policy makers and regulators, as can be seen from the Jumpstart Our Business Startups Act (JOBS Act) that has been signed into law in the United States. Further, with local governments and non-profits turning to crowdfunding to finance civic projects and programs designed for the common good (Lindsay, 2015), crowdfunding analytics tools could also provide insights that could benefit their fundraising.

2.1. Factors affecting crowdfunding success

Prior studies have identified a number of factors that tend to be associated with successful crowdfunding campaigns. These factors include the project or campaign goal (Kuppuswamy & Bayus, 2014; Mollick, 2014), the crowdfunding campaign duration (Mollick, 2014), the project creator's network size (Kuppuswamy & Bayus, 2014; Zheng, Li, Wu, & Xu, 2014), the contribution frequency (Burtch, Ghose, & Wattal, 2013), and the amount outstanding to the campaign goal (Burtch et al., 2013; Kuppuswamy & Bayus, 2014). Mollick (2014) and Kuppuswamy and Bayus (2014) found that projects with higher campaign goals had a lower probability of being successful, and that the average goal for unsuccessful campaigns is five times that of successful ones. Hou, Wang, and Ge (2015) suggest that this goal size effect may be a result of relatively large funding goals requiring relatively large amounts of contributions from backers to meet the set target. Further, individuals may use a campaign's goal as a proxy for the project's complexity and feasibility, and decide whether or not to fund the campaign based on the goal (Frydrych, Bock, Kinder, & Koeck, 2014; Koch & Siering, 2015).

Researchers have established that individuals' social networks play a significant role in their fundraising success (Shane & Cable, 2002; Shane & Stuart, 2002; Zheng et al., 2014). Not only does the project creator's network serve as an early pool of backers for the project campaign (Mollick, 2014), they also provide endorsements which can serve as quality cues and lead to more external backers (Shane & Cable, 2002). Hence, a project creator's network should have a positive impact on his chances of success.

The crowdfunding literature has also documented that the duration of the fundraising can impact a campaign's success (Cordova, Dolci, & Gianfrate, 2015; Zvilichovsky, Inbar, & Barzilay, 2015). Cordova et al. (2015) show that a project's fundraising duration has a positive impact on its chances of success in a crowdfunding campaign. This may be because fundraising campaigns that run for longer periods of time are more likely to be exposed to higher numbers of potential funders, and as such are more likely to reach their goals eventually.

Burtch et al. (2013) showed that contribution dynamics, which they conceptualized as the *contribution frequency*,[1] are important for predicting crowdfunding outcomes. Likewise, Kuppuswamy and Bayus (2014) and Burtch et al. (2013) showed that the amount required to reach the funding goal also predicts outcomes. Table 1 provides a summary of the factors that we use as predictors and their sources. Our work focuses on using the contribution dynamics of the focal crowdfunding campaign to forecast future values or outcomes. Although this paper is similar to that of Burtch et al. (2013) in that we include dynamic properties of crowdfunding campaigns, we conceptualize these dynamics in a different and more nuanced way. For instance, unlike Burtch et al. (2013), our conceptualization avoids the use of aggregates, instead relying on the natural information flow that happens during a crowdfunding campaign, albeit discretized at the daily level.

[1] This measure was operationalized by Burtch et al. (2013) as the total number of contributions standardized by the number of days over which the campaign was conducted.

Table 1
Summary of the factors that affect crowdfunding success and their sources.

Factor	Effect	Source
Project goal	Negative	Beier and Wagner (2015), Cordova et al. (2015), Hou et al. (2015), Koch (2016), Koch and Siering (2015), Mollick (2014) and Zvilichovsky et al. (2015)
Project duration	Positive	Cordova et al. (2015), Koch (2016), Mollick (2014) and Zvilichovsky et al. (2015)
Number of funders	Positive	Colombo, Franzoni, and Rossi-Lamastra (2015), Cordova et al. (2015), Etter, Grossglauser, and Thiran (2013) and Greenberg, Pardo, Hariharan, and Gerber (2013)
Network size	Positive	Koch (2016), Mollick (2014) and Zvilichovsky et al. (2015)
Average contribution/average daily contribution	Positive	Burtch et al. (2013) and Cordova et al. (2015)

3. Study context and data

This section briefly describes the context and structure of the data used in this study. Our data were collected from Kickstarter.com, one of the oldest, largest, and most popular crowdfunding platforms (Kuppuswamy & Bayus, 2014). Since a lot of the empirical works on crowdfunding have used data from Kickstarter (Kuppuswamy & Bayus, 2014; Mollick, 2014), we likewise use data from the platform when building our forecasting model. However, our approach can be adapted readily to other platforms.

3.1. Kickstarter

Kickstarter prides itself on helping to "bring creative projects to life". The platform has successfully provided funding for various projects that eventually led to thriving companies, including Pebble[2] Technology and Ouya.[3] By September 2014, Kickstarter had raised more than $1.3 billion for 69,530 projects (Kickstarter, 2014). Fig. 1 shows a typical Kickstarter crowdfunding campaign homepage. Projects on Kickstarter are grouped into various broad categories (see Fig. 2).

To participate on Kickstarter, individuals must join the community through free registration. Members can create projects for funding, contribute to projects financially, and comment on projects. There are no geographic restrictions on membership, though project creators can only be from certain countries. Backers can pledge a maximum of US$10,000 or its equivalent. Projects on Kickstarter typically run for between 1 and 60, days with 30 days being the recommended time frame.

Kickstarter has two main factors that set it apart from other crowdfunding platforms. First, it operates on an "all-or-nothing" fundraising model. This means that a project must meet its funding goal fully within its fundraising period, otherwise the funds are returned to the backers. Second, contributors on Kickstarter do not receive equity in the projects that they fund, but may receive modest "rewards" (e.g., a thank you note).

3.2. Data

We extracted information for about 2000 projects that were posted on Kickstarter between April 1st, 2014, and May 2nd, 2014. We visited the website using automatic web agents and extracted key information on the campaigns over their fundraising cycle. After collecting the data, we conducted data cleaning and preprocessing. During the data cleaning process, we found that some creators canceled their campaigns during the funding cycle, meaning that such campaigns did not run their full course. We dropped such campaigns from our sample. Because we are interested in predicting fundraising campaigns that seek to raise a sizeable amount of money, we restricted our sample to campaigns that were seeking at least $5000. This led to a further reduction in our sample size. Since we are interested in modeling and predicting the dynamic process of ongoing crowdfunding campaigns, we also restricted our analysis to projects with at least one backer[4] (projects with zero backers will not contribute any information to our model). We likewise restrict our model building to projects with fundraising campaigns that run for at least ten days (as this allows for dynamics, which is a key feature of our model). Typically, project creators will want to observe a fundraising campaign for a short time after its launch before deciding what actions to take to improve its outcome.

After completing the data cleaning and preprocessing, we are left with 618 usable projects.[5]

Table 2 provides descriptive statistics of our sample, while Fig. 2 shows the distribution of campaigns across different categories in Kickstarter.

4. Methodology and results

This section presents our method and results. Fig. 3 displays the flow of our study, with the implementation of our technique. We start by introducing the functional data analysis (FDA) (Silverman & Ramsay, 2005) techniques that we use to develop our forecasting model, then discuss the underlying contribution patterns of the crowdfunding campaigns in our dataset. We apply functional principal components analysis (fPCA) and continue with an

2 www.getpebble.com.

3 www.ouya.tv.

4 Although we are not interested in predicting whether or not a crowdfunding campaign will receive any backers, we used the inverse probability weighting (IPW) technique (Haneuse et al., 2009; Pan & Schaubel, 2008) to check for the performance consistency of our proposed dynamics-based forecasting model relative to other models when we account for campaigns with zero backers. There was no significant difference in performance consistency.

5 The Appendix A also provides detailed instructions as to how to collect and process Kickstarter data.

6 https://www.kickstarter.com/projects/paolaprestini/original-music-workshop-start-up-funds.

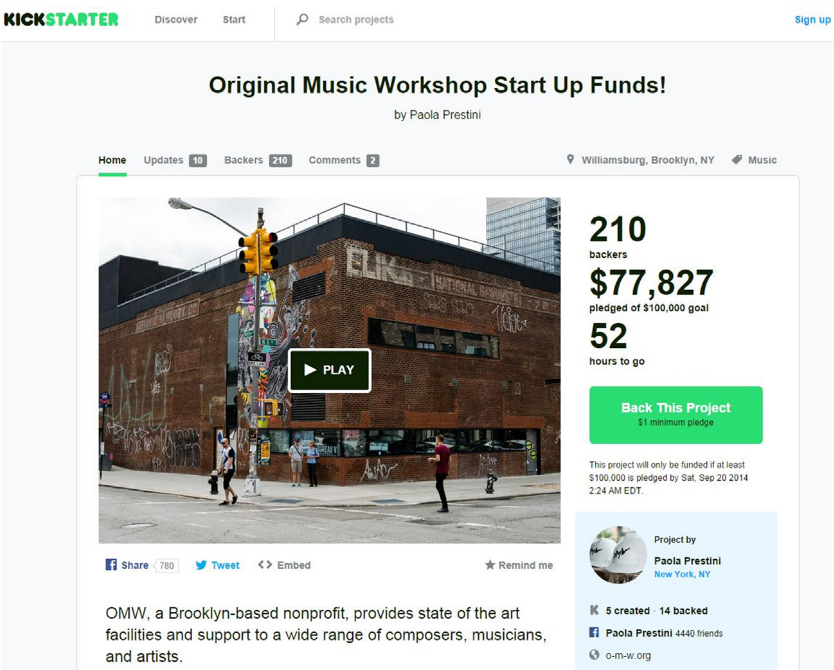


Fig. 1. Kickstarter home page for a crowdfunding campaign (Original music workshop start up funds![6]) that was funded successfully.

Table 2
Descriptive statistics of our sample ($N = 618$).

	Mean	Median	Std. Dev.	Min	Max	25th quantile	75th quantile
Campaign goal ($)	23 655	10 500	47 791.27	5000	580 000	7000	25 000
Amount raised ($)	19 672	5742	141 325.3	1	3 401 361	699	12 853.3
Number of backers	198.7	58.5	664.4	1	11,855	11	154.75
Average contribution per backer ($)	96.53	66.91	114.19	1	1013.41	37.6	114.9
Percent of campaign goal raised (%)	91.52	55.8	297.74	0	6802.72	5.3	108.9
Number of facebook friends	833.2	506.5	920.23	0	4861	252.8	1031.3
Funding duration (days)	29.84	30	4.3	14	45	30	30

exploratory analysis of the patterns. Finally, we incorporate dynamic features into the model and compare it with relevant benchmarks.

4.1. Functional data analysis

Functional data analysis is a relatively new statistical technique that has been applied to the study and understanding of phenomena that exhibit various types of dynamics by capitalizing on a phenomenon's dynamic attributes. FDA's key feature is its ability to examine curves and shapes, or, more generally, functional observations (Silverman & Ramsay, 2005). These functional observations can be functions of time and/or space, etc. FDA has been applied in such diverse domains as medicine and human biology (Ramsay & Silverman, 2002; Zhou et al., 2010), finance (Hays, Shen, & Huang, 2012), and marketing (Foutz & Jank, 2010; Sood, James, & Tellis, 2009), among others. Hays et al. (2012) applied FDA to the modeling of yield curves, which are used as an instrument for portfolio management and for pricing securities. Sood et al. (2009) used FDA to model the diffusion and market penetration of new products and demonstrated that it outperformed other models such as the Bass diffusion model (Bass, 1969). Foutz and Jank (2010) proposed an FDA-based shape approach to the forecasting of movie box-office performances using the price patterns of movies in a virtual stock market. Similarly, Hui, Meyvis, and Assael (2014) developed a Bayesian functional model to help in understanding consumer judgments of TV show pilots using consumer moment-to-moment response data. Consistent across these studies is the exploitation of the underlying dynamics or pattern of the process generating the phenomenon under study. It contributed to or improved the overall explanatory or predictive power of the results, respectively.

Crowdfunding exhibits different forms of dynamics (e.g., contributions, backers, online buzz), and these can have implications for the outcome. The dynamics of contributions in a crowdfunding campaign have peaks and troughs over time, with periods of high and low activity (Kuppuswamy & Bayus, 2014). Hence, the crowdfunded amount may be related to contribution c at time t not only through the average contribution $\bar{c}(t)$, but also through the dynamic variations in contributions c during the crowdfunding process. FDA allows us to model some of the dynamic properties that define these variations (e.g., velocity

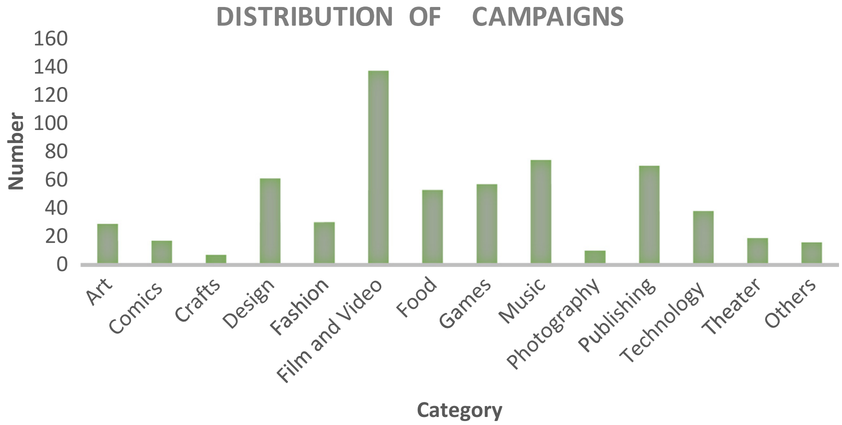


Fig. 2. Distribution of campaigns across categories.

and acceleration) through the use of derivatives. Similarly, FDA is useful for reducing the dimensionality of the high-frequency and dimension data that the crowdfunding process can generate, without losing much information. Following these arguments, and consistent with prior studies, we adopt FDA when developing our forecasting model, on the basis that the dynamics of contributions in the process can lead to better forecasts.

FDA treats each function (curves, shapes, etc.) as a unit of observation, which clearly distinguishes it from the traditional statistical models that use data vectors as input and output variables. As a non-parametric approach, it assumes only smoothness and permits as much flexibility as is required by the data (Xiong & Bharadwaj, 2014). It also has the advantage that it does not assume or fix the number parameters (such as those related to the mean and variance), nor does it specify fixed parameters upfront (Silverman & Ramsay, 2005). Moreover, it is best suited for the analysis of non-stationary time series because it does not require a *stationarity* assumption (Ramsay & Dalzell, 1991). A typical FDA process involves (1) recovering the functional objects (known as smoothing, e.g. using penalized splines); (2) computing the dynamics and/or recovering the functional principal components (fPC); and (3) modeling via functional regression analysis, for example. Fig. 3 provides an overview of the steps taken in this analysis.

4.1.1. Recovering contribution evolution curves

Since our focus is on the contribution dynamics during a campaign, we start by estimating the underlying functional process as a continuous curve $S_i(t)$ that characterizes each campaign. Once these curves have been estimated, their dynamics are then explored. A variety of methods exist for smoothing data and recovering the underlying curves; however, we focus on the penalized smoothing spline because it offers great flexibility for one-dimensional smoothing problems, such as that presented here, as well as convenience in estimating curve rates of change (derivatives). The literature has shown the effectiveness of this technique for recovering the underlying curves of processes from discrete observations (Reddy & Dass, 2006; Silverman & Ramsay, 2005), even in the presence of noise in the raw data. If t_i represents the ith time period in a campaign and C_{ji} represents the amount contributed to project j in period i, the aim of penalized smoothing splines is to identify a continuous smooth function f_j that minimizes a penalized residual sum of squares (similar to least squares minimization in regressions) based on the equation

$$PENSS_j = \sum_{i=1}^{n} \left(C_{ji} - f_{ji}\right)^2 + \lambda \left(PEN_m\right), \tag{1}$$

where $\left(C_{ji} - f_{ji}\right)^2$ measures the fit of the function, PEN_m measures the roughness (variability) of the function f based on its mth derivative, and the smoothing parameter λ provides the tradeoff between fit and roughness. Fig. 4 shows plots of the functional curves recovered for the project campaigns (both successful and unsuccessful). While the functional curves in Fig. 5 show the contribution trajectories of different crowdfunding campaigns, we also calculate the first and second derivatives of the contribution trajectories in order to obtain estimates of their "velocities" and "accelerations", where *velocity* refers to the rate at which the contribution amount is changing and *acceleration* to the rate at which the contribution velocity is changing. Section 4.2 uses these derivatives to describe further the nature and heterogeneity in the contribution dynamics of successful and unsuccessful campaigns.

Both *velocity* and *acceleration* help to determine how quickly a campaign's current amount (amount raised so far in the campaign) is moving towards its final outcome. To obtain a better understanding of velocity in the fundraising context, assume that individual i contributes x dollars at interval Δt after the previous contribution; then, the velocity of contribution within the interval Δt is the amount contributed per unit time $\dot{x}$ in that interval and is given by $\frac{x}{\Delta t}$. This velocity $\dot{x}$ is subject to change, since different time intervals elapse between contributions (similar to the findings presented by Kuppuswamy & Bayus, 2014), and different individuals contribute different amounts even for the same intervals between contributions. These variations are bound to affect the rate at which a campaign approaches its final funding outcome. Potential backers can also use these variations in rates as signals to help them decide whether to contribute to a campaign or not. Similarly, for

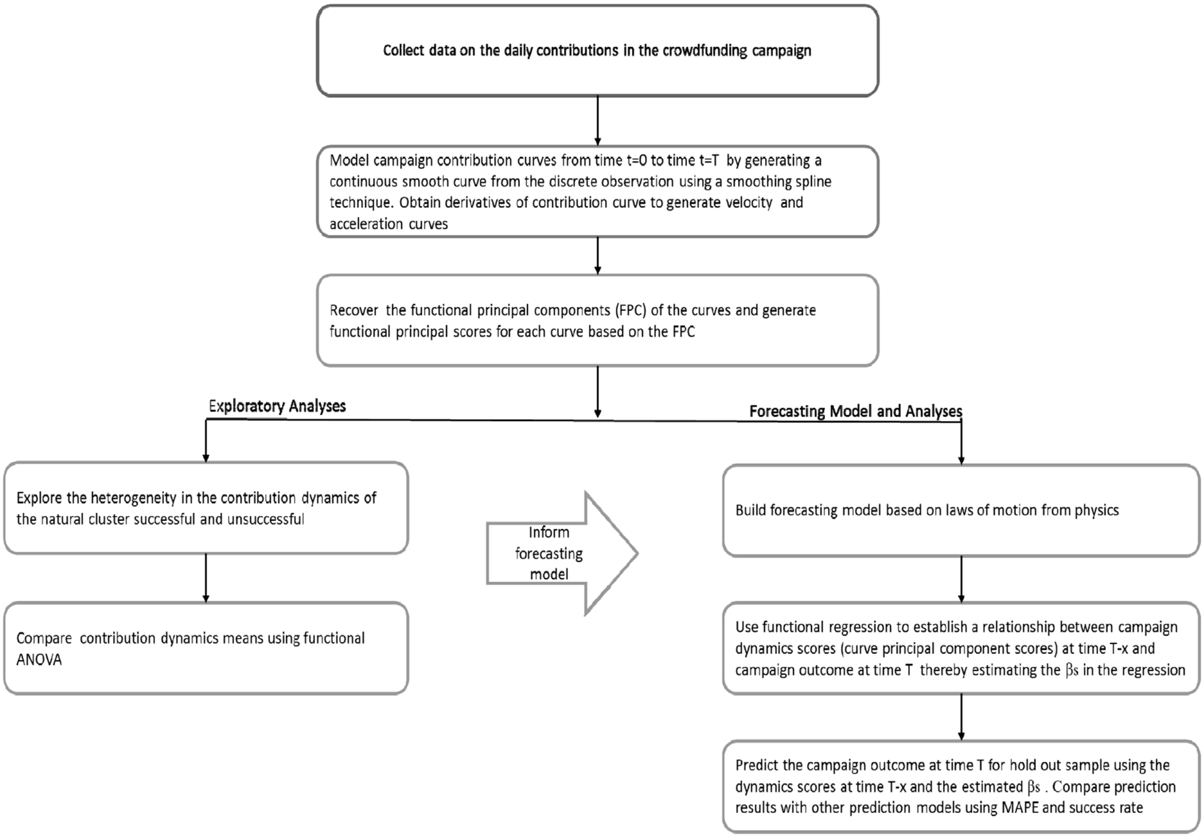


Fig. 3. Flowchart for the exploratory analysis and implementation of the forecasting model.

acceleration, assume more individuals j, k, and l contribute the amounts x_j, x_k and x_l sequentially in decreasing time intervals Δt_j, Δt_k, and Δt_l, and that $\Delta t_j + \Delta t_k + \Delta t_l \leq \Delta t$, resulting in velocities $\dot{x}_j$, $\dot{x}_k$, and $\dot{x}_l$; this means that even if the amounts contributed by the individuals are the same, the velocity will change as a result of the decreasing times, leading to an increase in acceleration. Thus, the acceleration measures how much the velocity is changing over time (which could be either increasing or decreasing) and will reflect the variations in both the numbers of new contributions and the amounts contributed in a given time interval. It will show surges or declines in activity in a campaign over given intervals. Intuitively, a "viral" campaign should see an increasing acceleration during the period in which the campaign experiences "virality", providing evidence of exponential-type growth in contributions throughout the time period.

4.1.2. *Functional principal components analysis (FPCA)*

Having recovered the heterogeneous contribution curves and their derivatives, it is important to identify the distinguishing basic shapes that characterize these curves. Functional principal component analysis (FPCA) helps us identify these basic shapes, which can then be used to create functional principal component scores (PCScore) for the curves, making them more amenable to numerical analysis (Silverman & Ramsay, 2005). FPCA is analogous to ordinary principal components analysis (PCA), in that it can extract key features from repeated measurement data by projecting the original data to a new space of reduced, orthogonal dimensions. Let f_i represent each curve of interest from a sample of j curves measured at p discrete time points; then, all j curves of interest $F^{all} = [f_1, f_2, \ldots, f_j]$ can be denoted by a $[j \times p]$ matrix. Its $[p \times p]$ correlation matrix, $M := \mathrm{Corr}(F^{all})$, can be represented in spectral form as $M = P^T \Lambda P$, where Λ is the diagonal matrix of eigenvalues $[\lambda_1, \lambda_2, \ldots, \lambda_p]$ and $P = [e_1, e_2, \ldots, e_p]$ is the corresponding matrix of eigenvectors. In this example, each e_i is a $[1 \times p]$ vector, which, in reality, will represent a continuous function in time that captures a unique characteristic of the curves of interest (Foutz & Jank, 2010). Since each e_i captures a percentage of the variability of F^{all} like the conventional PCA, the e_i's represent the principal components. In theory, j different principal component curves are required to represent all j curves in each class of dynamics perfectly – trajectory, velocity, acceleration, and so forth. In practice, though, only the first few principal components (i.e. the ones which explain most of the variation) are chosen, which effectively results in a reduction of our high-dimensional feature space. In our case, we use the first two principal components because they account for about 98% of the variability in the curves.[7]

[7] We do not rotate the principal components (PC) generated using the singular value decomposition algorithm because we are not looking to interpret the PCs based on an axis of rotation.

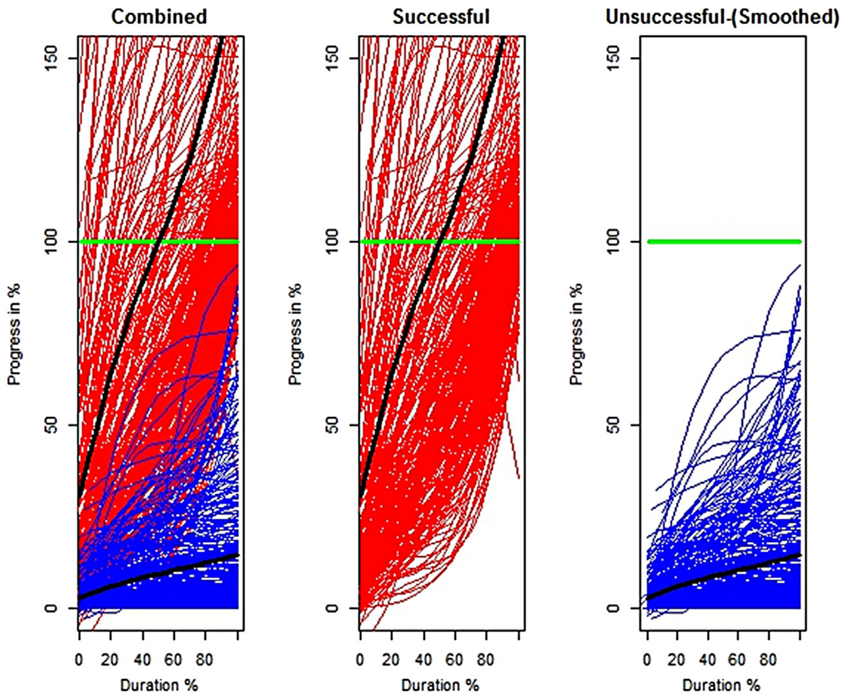


Fig. 4. Normalized contribution dynamics of crowdfunding campaigns. Note: The black lines represent the mean contribution trajectory and the green lines the campaign goal; the red (blue) lines correspond to successful (unsuccessful) campaigns. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These first two principal components are used to generate the PCScore. The PCScore represents how much weight each curve of interest f_i has when it is decomposed into the corresponding principal components. For instance, the PCScore of campaign i' s contribution trajectory f_i will be the result of an inner product of $f_i = [f_{i1}, f_{i2}, \ldots, f_{ip}]$ and the selected principal components matrix $P_{selected} = [p \times 2]$.

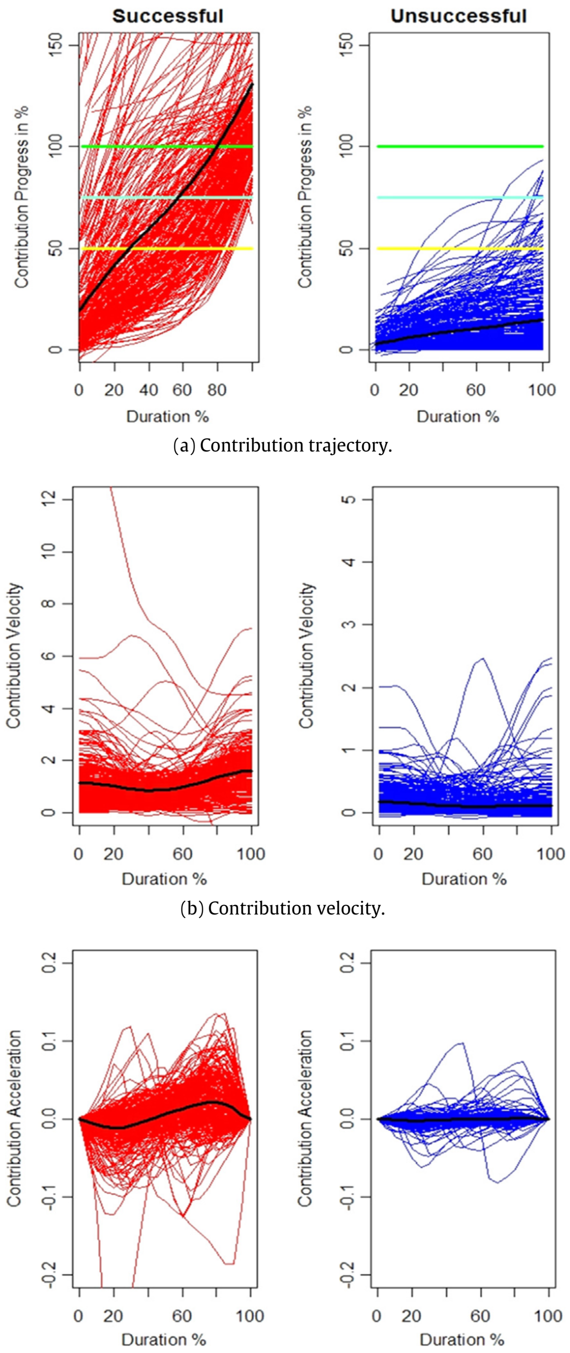
4.2. Exploring the dynamics of successful and unsuccessful fundraising campaigns

After determining the different PCScores for the contribution dynamics (trajectory, velocity, and acceleration), we investigate the natural groups of successful and unsuccessful fundraising campaigns and compare their dynamics. Using both visual and statistical techniques, we compare the contribution dynamics of the groups based on the differences in their trajectories, velocities, and acceleration. Figs. 5 and 6 show plots of these comparisons. Our goals are: (1) to obtain an understanding of the nature of the dynamics exhibited by successful and unsuccessful campaigns, and (2) to determine whether there are significant differences between the two groups. We achieved the first goal by estimating the normalized mean (population average) trajectory of the crowdfunding campaigns in each group (successful and unsuccessful) and extracting the corresponding normalized funding level or time coordinate for each point of interest (funding level or time). We achieved our second goal by comparing the means of the trajectories of the two groups using a functional analysis of the variance (fANOVA) (Silverman & Ramsay, 2005).

The exploratory analysis generated the following insights. On average, the contribution trajectories of successful campaigns start out with higher contribution-to-funding goal ratios than those of unsuccessful campaigns. Similarly, the average successful fundraising campaign reaches its goal by 80% of the way through the campaign period. The average successful campaign reaches 25% of its funding goal at 4.8% of the way into the campaign period, 50% of its funding goal 28.2% of the way into the campaign period, and 75% of its funding goal 57.1% of the way into the campaign period. In contrast, the average unsuccessful campaign raises just 14.8% of its funding goal over the entire period of the campaign. By 25% of the way through the campaign, the average unsuccessful campaign has raised only about 6.65% of its funding goal. By the 50% (75%) mark, the average unsuccessful campaign will have raised just 9.5% (12%) of its funding goal.

An analysis of the groups to determine the difference in means using ANOVA and Welch's t-test shows that there are significant differences between the mean trajectories of successful and unsuccessful campaigns, with the contribution trajectories of successful campaigns typically being higher than those of unsuccessful campaigns. The box plot in Fig. 6(a) shows the differences between the contribution trajectories of the two groups.

The magnitude of the average contribution velocity, as shown in Fig. 5(b), is higher for successful campaigns than for unsuccessful campaigns. An analysis of the groups using both ANOVA and Welch's t-test shows that the mean contribution velocities are significantly different, as can be seen from the boxplot in Fig. 6(b). It is also interesting

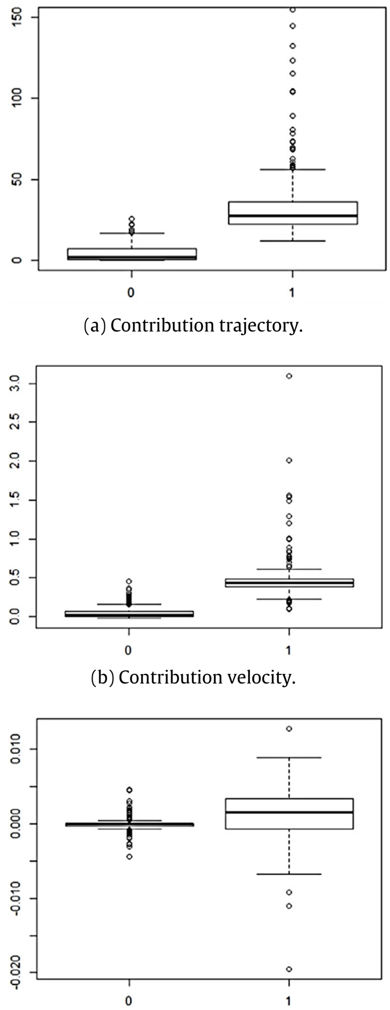


(a) Contribution trajectory.

(b) Contribution velocity.

(c) Contribution acceleration.

Fig. 5. Plots of the (a) contribution trajectory, (b) velocity, and (c) acceleration. Note: The black lines represent the average dynamics in each panel, while the green, aquamarine and yellow lines represent the 100%, 75% and 50% threshold contribution lines respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) Contribution trajectory.

(b) Contribution velocity.

(c) Contribution acceleration.

Fig. 6. Box plots showing comparisons between the means of the dynamics of successful and unsuccessful campaigns. Note: "0" represents unsuccessful campaigns, while "1" is successful campaigns.

to note that the contribution velocity for successful campaigns follows a "U" shape similar to that identified by Kuppuswamy and Bayus (2014) in their study of backers rates. For successful campaigns, the shape of the mean velocity shows that the velocity tends to be higher near the end of the campaign than at the start. The flurry of contributions near the end of the fundraising campaign that this represents can be explained by the classic goal-gradient effect (Hull, 1932) that was recently revamped by Kivetz, Urminsky, and Zheng (2006) in the context of consumer behavior and investigated empirically in the crowdfunding context by Kuppuswamy and Bayus (2017). This suggests that individuals tend to be more energized to attain a goal when it is in sight. On average, successful crowdfunding campaigns show their minimum velocities at around the 40% period mark. On the other hand, unsuccessful fundraising campaigns tend to lose momentum early, then often show a weak last minute burst that is not enough to reach the goal.

Both successful and unsuccessful fundraising campaigns show nearly non-existent accelerations (close to

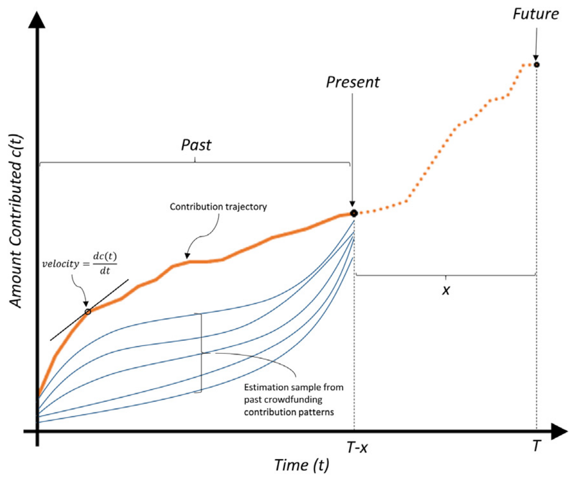


Fig. 7. Illustration of the forecasting scheme used for outcome predictions.

zero for both clusters); see Fig. 5(c). However, an analysis of the clusters using ANOVA and Welch's t-test shows that there is a significant difference between the mean contribution accelerations, as is shown in Fig. 6(c). The plot also suggests that, on average, successful campaigns initially decelerate before accelerating towards the goal, then decelerate again after the goal has been reached. The initial deceleration could be explained as being a natural consequence of the early-stage contribution velocity that the average successful campaign starts up with, but which it cannot maintain throughout the campaign. On the other hand, unsuccessful campaigns maintain a constant low level of acceleration. In other words, unsuccessful campaigns do not exhibit the acceleration patterns that appear to be a distinguishing feature of successful campaigns. From the exploration of crowdfunding projects' contribution dynamics, we can inductively infer the crowding in and crowding out effects that have been documented in the literature (Burtch et al., 2013; Kuppuswamy & Bayus, 2017), leading us to believe that the dynamics provide rich information that may prove useful for forecasting when modeled using more flexible semi-parametric approaches.

4.3. A prediction model for crowdfunding campaigns

Drawing from our exploration of the dynamics of crowdfunding campaigns, we now build a model for forecasting fundraising outcomes. Fig. 7 illustrates our forecast scheme, which combines information from the dynamics of other earlier campaigns with that from the focal campaign in order to predict outcomes.

Our exploratory analysis indicated that the contribution trajectory, velocity, and acceleration are important in the process of forming the outcome of the fundraising campaign over the campaign period. Thus, we base our prediction model on a functional differential equation that captures the relationship. We propose a non-homogenous second order linear differential equation of the form

$$\beta_0(t) f(t) + \beta_1(t) f'(t) + f''(t) + k = 0, \quad (2)$$

where $f(t)$, $f'(t)$, and $f''(t)$ represent the contribution trajectory, velocity and acceleration, respectively, and k represents the funding goal. This is a plausible model for describing the funding outcome of a crowdfunding campaign because there are likely to be forces proportional to the position, velocity, and acceleration that affect each campaign which are similar to those proposed by the laws of motion. In fact, functional differential equation models are believable for many processes that exhibit varying but predictable dynamics (Wang, Jank, Shmueli, & Smith, 2008). In Eq. (2), $\beta_0(t)$ reflects the force(s) acting on the contribution trajectory at time (t) when the campaign has achieved a particular level of contributions. For instance, at a time very close to the funding goal, the goal-gradient effect or a crowding-out effect may become the driving force. Similarly, $\beta_1(t)$ reflects forces that are proportional to the contribution velocity at time t. For instance, a campaign could go "viral" as a result of media mentions or celebrity support triggering a flurry of contributions.

Rearranging Eq. (2) so that k (which is the project's goal or outcome) is on one side of the equation and the dynamic parameters are on the other side, and drawing on insights from our exploratory analysis of the fundraising campaigns, we obtain our final functional forecasting model (FDM):

$$Outcome_T = \beta_0 + \sum_{p=1}^{4} \left(\beta_p PartialDynamics_{T-x,p}\right) + \varepsilon_j, \quad (3)$$

where $PartialDynamics_{T-x,p}$ represents the functional principal component scores (PCScores) that are generated for a project campaign's contribution curves and derivatives up to time $T - x$. Recall from our analysis in Section 4.1.2 that we use functional principal components analysis (FPCA) to generate PCScores, which are analogous to the principal component scores in conventional principal component analysis (PCA). These PCScores allow us to make our curves amenable to numerical analysis. For the partial dynamics, the PCScore for a project campaign is generated not from the project's complete contribution curve and its derivatives between times $t = 0$ and $t = T$, but from the project's partial contribution curves and its derivatives, which go from times $t = 0$ to $t = T - x$. Hence, we can use information on a contribution curve up to time $t = T - x$ to predict the outcome at time $t = T$.

Note that we use only the first two functional principal components for each of the dynamic components of a project campaign – contribution trajectory, contribution-velocity, and contribution acceleration – since they account for about 98% of the variations in the curve. Hence, there are two PCScores for each dynamic component of each curve in our sample (i.e., two PCScores for the contribution trajectory, etc.). However, our model in Eq. (3) drops the acceleration parameters, because, first, our exploratory analysis showed that the acceleration is almost zero for both successful and unsuccessful campaigns, and therefore will not add much to the performance of our forecasting model; and, second, this allows us to obtain a parsimonious model with comparable or higher predictive capabilities, which is often preferred (Shmueli, 2010).

Our FDM model derived from Eq. (2) can be considered an inverse differential equation problem where the dynamic functions $f(t), f'(t)$, and $f''(t)$ are known and we are interested in estimating the β coefficients from the data, and can then use them as predictors of the outcome k. Since our model maintains the linearity in the response equation (Eq. (2)), it can be estimated by least squares methods.

4.3.1. Benchmarks: alternative model specifications

We expand on the advantages of our approach by comparing our FDM model with other reference models. Note that, unlike our FDA-based model, FDM, these models use the most recent dynamics information without incorporating information on historical or past dynamics. However, all of these other models are very popular in the context of forecasting (De Gooijer & Hyndman, 2006; Stock & Watson, 1998), especially the normal parametric linear model (Stock & Watson, 1998) and the generalized additive model (GAM; see Hastie & Tibshirani, 1986). The variables that are used as predictors in our models are drawn from the extant literature on crowdfunding, which has identified them as determinants of success, as was discussed in Section 2.1. These include the project goal, the creator's network size, the amount already contributed, and the number of backers (Table 1 shows the sources of these predictors). Further, we allow for a fair comparison between these models and our dynamics-based model by estimating the models using the most recent dynamic information from the variables. However, unlike the FDM model, these models are not able to incorporate information on the historical or past dynamics of the variables. The benchmark models are as follows.

Normal parametric linear model. This is a simple model that fits a line over the estimation sample by minimizing the sum of squared errors, then uses the equation of this line to predict the holdout sample. For example, we predict the outcome of a fundraising campaign at time $T-x$, where T is the period of the campaign and x is the time before the end of the campaign, by using information at time $T-x$ from campaigns in the estimation sample to fit the model. The model estimates can then be used to predict the outcomes of campaigns in the holdout sample. Specifically, the prediction model is given by

$$\begin{aligned} Outcome_T &= \beta_0 + \beta_1 SocNetsize_{T-x} + \beta_2 Goal \\ &\quad + \beta_3 NumCont_{T-x} + \beta_1 AmtCont_{T-x} + \varepsilon_j, \end{aligned} \tag{4}$$

where *SocNetsize* is the size of the creator's network on Facebook, *Goal* is the amount the creator is seeking to raise, $NumCont_{T-x}$ is the number of individuals that have contributed to the campaign by time $T-x$, and $AmtCont_{T-x}$ is the amount that has been raised by time $T-x$.

Generalized additive model. The estimation technique of this model is similar to that of the normal parametric linear model, but its predictors incorporate the nonlinearities in the data. That is, it uses smooth functions to capture the distribution of the predictors, then uses these non-linear forms as predictors. Similarly to the normal parametric model, we fit the estimation sample containing information about the predictors at time $T-x$ and the outcomes of the campaigns, then use the fitted model to predict the outcomes of the holdout sample. Specifically, the prediction model is given by

$$\begin{aligned} Outcome_T &= \beta_0 + \beta_1 SocNetsize_{T-x} + \beta_2 Goal \\ &\quad + f(NumCont_{T-x}) + f(AmtCont_{T-x}) + \varepsilon_j, \end{aligned} \tag{5}$$

where $f(*)$ implies that the parameters in parentheses have been smoothed, and the definitions of the variables are the same as for Eq. (4).

4.3.2. Model estimation

We divide our data (618 campaigns including outliers) into a training set (518 campaigns) for estimation and a holdout set (100 campaigns) for validation. Since our data varied over the time period during which it was collected and we are interested in making predictions for the future, the training set consisted of crowdfunding campaigns that started earlier in the data collection period, while the validation set or holdout sample consisted of campaigns that started later. We chose the holdout sample to consist of campaigns that started on later dates in order to allow for additional exogenous variation with respect to the training set that may be useful when testing competing models (Ebbes, Papies, & Van Heerde, 2011).

Because our dynamic model (FDM) is a functional regression model, we estimated it using least squares estimation techniques that are similar to those in other regression models (Silverman & Ramsay, 2005). Fig. 8 shows fit statistics for the models. FDM consistently fit better than other models. The model fit depends on the length of time x before the end of the crowdfunding campaign: the shorter the time x (i.e., the time before the end of the campaign), the better the R^2 value, the Akaike information criterion (AIC; see Akaike, 1974) and the Bayesian information criterion (BIC; see Schwarz, 1978). The AIC and BIC measure how well a model fits the data; the smaller the AIC and BIC values, the better the model fit. Intuitively, since it should be harder to forecast further into the future, the models should not perform as well when the time interval to the end of the campaign is longer.

4.4. Prediction results and performance

After estimating our model, we check its prediction performance on the holdout set by measuring the prediction performance at five different points during the crowdfunding campaign and comparing it with those of the benchmark models at the same points. The points used are 10%, 30%, 50%, 70%, and 90% of the total campaign duration. We measure the prediction performance using the mean absolute percentage error (MAPE) metric. For each time period $t = T-x$ in the holdout sample, we compute

$$MAPE\,(T-x) = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_{T-x,i} - \hat{Y}_{T-x,i}\right|}{\left|Y_{T-x,i}\right|}.$$

For instance, a prediction model with a forecast error (MAPE) of 0.06 at 30% of the way into the campaign implies that the true funding outcome will lie within ±6% of the forecasted value, on average. In forecasting, the goal is to have prediction models with smaller forecast errors, which

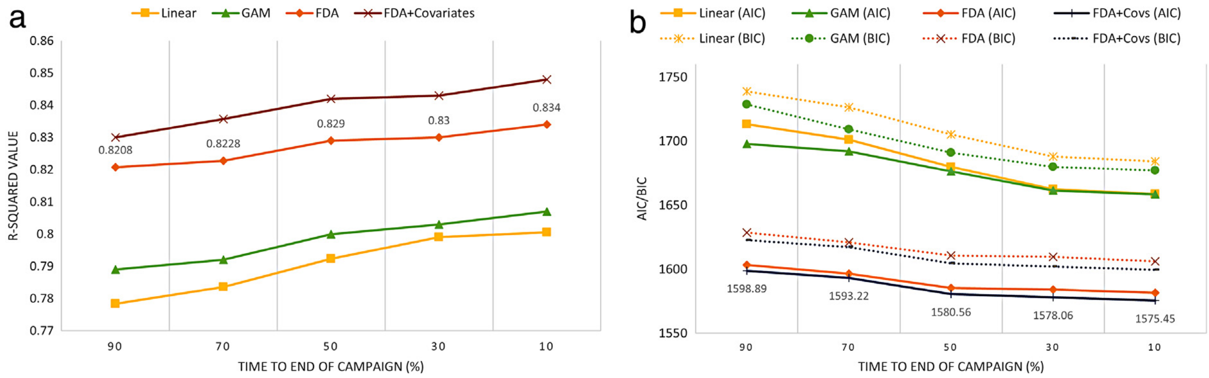


Fig. 8. Model fit depending on the normalized time to the end of the campaign: (a) R^2 value; (b) AIC and BIC.

in turn means better prediction performances. Table 3 and Fig. 9 show the results for all of the models across the different time periods, with Table 3 showing the numerical values of the forecasting error across different stages of the crowdfunding campaign and Fig. 9 illustrating the errors across the various stages. The x-axis in Fig. 9 shows the time since the beginning of the campaign, while the y-axis shows the forecasting error.

Not surprisingly, we see that the predictive performances of all models improve (i.e., MAPE decreases) as we predict closer to the end of the campaign (i.e., as time interval x gets smaller).

The prediction result shows that the dynamics-based model is better than all of the alternative models. At 10% of the way into the fundraising campaign period, the prediction error for the dynamics based model is about 6.29%, while it drops to 1.27% by 90% of the way into the fundraising campaign. The performances of the benchmark models were fairly competitive between themselves, but lagged considerably with respect to our proposed model. For instance, the GAM model was generally better than the linear model, but the latter became almost as good towards the end of the campaign. All of the models generally had prediction errors of less than 12%, indicating that they are good models overall. Interestingly, our model performed very well at the very early stages of the campaign, implying that users of the model can obtain fairly good predictions at the outset of their crowdfunding campaign. One explanation for the superior performance of our model could be that it incorporates information about the dynamics (i.e., contribution trajectory and velocity) of both the focal crowdfunding campaign and other campaigns when predicting outcomes, unlike the other models, which use last state information. Further, as Figs. 8 and 9 show, including covariates in our dynamics model improves the predictive performance further, albeit only marginally.

4.4.1. *Performance quality check*

We probe the quality of the FDM model further by conducting a performance quality check on a subsample of the crowdfunding project campaigns that consists of the 134 crowdfunding project campaigns with the greatest outcome uncertainty; that is, crowdfunding project campaigns with overlapping contribution trajectories (overlapping red and blue lines) but different outcomes (see Fig. 10). This subsample was obtained by including only those campaigns that have raised between 15% and 30% of their target goal by 20% of the way through the campaign. This yields the 134 crowdfunding trajectories in Fig. 10. We then split the subsample such that 104 campaigns are used to train the model and 30 are used for testing. Table 4 shows the performances of all models. We can see that the FDM model consistently performs better than either the GAM or the linear regression model. Furthermore, adding covariates improves the predictive performance of the FDM model. Interestingly, the FDM model performs better on this subsample than on the full sample (Table 3). One possible explanation for this could be that the dynamics of the campaigns in the subsample are more consistent than those of the campaigns that are removed from the sample, leading to better predictions of our functional model.

5. Discussion

Forecasting the outcomes of crowdfunding campaigns is becoming increasingly important because of the growth in the use of crowdfunding mechanisms by fledgling entrepreneurs, businesses, and individuals. Users have shown interest in having forecasting tools at their disposal (Huffington Post, 2013; Wired.com, 2012), but there has been little work on the development of forecasting models for the crowdfunding context, and most crowdfunding platforms are yet to implement such tools despite having access to large amounts of crowdfunding data. To the best of our knowledge, the model that we demonstrate and assess the merits of here is the first crowdfunding forecasting model to use the contribution histories and contribution dynamics of other comparable crowdfunding campaigns along with those of the focal crowdfunding campaign. FDA is an innovative statistical technique that has gained popularity recently, and allows us to carry out this task of modeling and quantifying dynamics.

Moreover, we explore and analyze the contribution dynamics of crowdfunding campaigns in terms of their contribution trajectories, contribution velocities, and contribution accelerations. Again, FDA proves useful in helping to glean key intuitions on the contribution dynamics of the crowdfunding phenomenon. When we compare the natural clusters of successful and unsuccessful fundraising

Table 3
MAPE results of competing models.

Stage of campaign (in %)	Parametric regression	GAM	FDM	FDM with covariates
10	0.1105	0.0976	0.0629	0.0621
30	0.0766	0.0658	0.0510	0.0502
50	0.0594	0.0507	0.0277	0.0275
70	0.0422	0.0383	0.0221	0.0220
90	0.0211	0.0210	0.0127	0.0126

Table 4
MAPE results of the models.

Stage of campaign (in %)	Parametric regression	GAM	FDM	FDM with covariates
30	0.0407	0.0405	0.0401	0.0379
50	0.0372	0.0370	0.0281	0.0255
70	0.0280	0.0250	0.0168	0.0161
90	0.0178	0.0178	0.0105	0.0101

Note: We start from 30% of the way into the campaign because the sample selection criterion is based on the campaign status at the 20% time point mark.

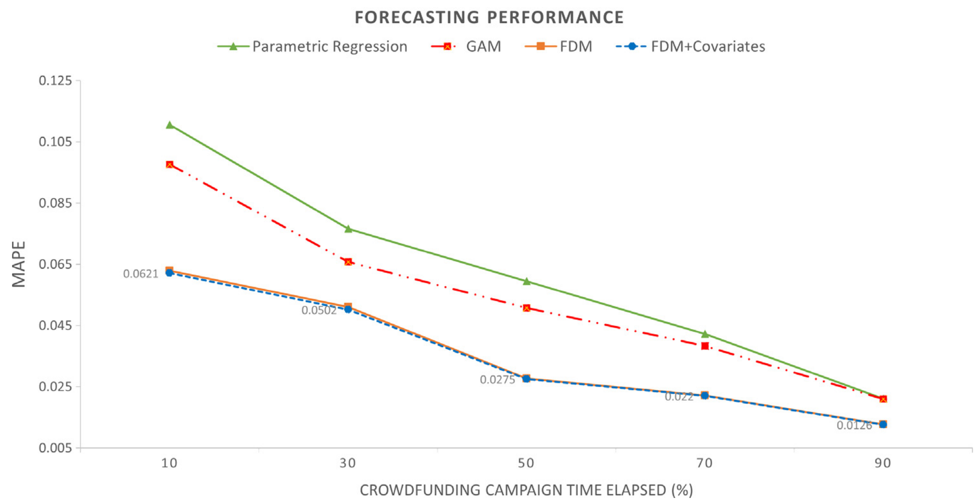


Fig. 9. Prediction performances of competing models, measured in terms of MAPE.

campaigns, our analyses lead to the following important insights. First, in general, successful crowdfunding campaigns reach their target goals long before the campaign is over. The average successful campaign reaches its target goal at about 80% of the way into the campaign, whereas the average unsuccessful campaign will be at about only 12.3% by the same time period. Second, successful campaigns tend to start out with higher contribution velocities than unsuccessful campaigns. This could mean that such creators engage in some offline awareness building prior to the campaign launch that can help their campaigns to take off. For instance, Kuppuswamy and Bayus (2014) suggest that this early awareness is usually among family and friends. The high contribution velocity can also act as a signal of quality for external backers who are not in the creator's social circle but are looking for projects to back. Potential creators should try to build their offline awareness before launching their crowdfunding campaigns, to increase their likelihood of success. Third, both successful and unsuccessful campaigns tend to have little or no contribution acceleration. While this insight seems surprising, it highlights the fact that campaigns "going viral" is not a norm in the crowdfunding environment. Creators should not bank on a "viral effect" to drive them to success when launching campaigns, but instead should work diligently to steer backers to their campaigns to contribute.

On the forecasting front, we show that our dynamics-based model (FDM), developed using FDA techniques, performs better than standard alternative models. Insights generated from the forecasting model include the following.

(1) The use of FDA techniques provides superior forecasting power for dynamic processes that evolve over time. Not only does our model perform better because of its basis in the laws of motion, but FDA also provides a greater flexibility by removing parametric assumptions, which can lead to improvements in predictive power. For instance, the GAM model's relaxation of the parametric assumptions on the normal linear model led to predictive improvements relative to the linear model. FDA's key strength lies in its ability to capture the variety of patterns in the different campaign contribution dynamics without *overfitting* through its use of principal components. Though this increased flexibility

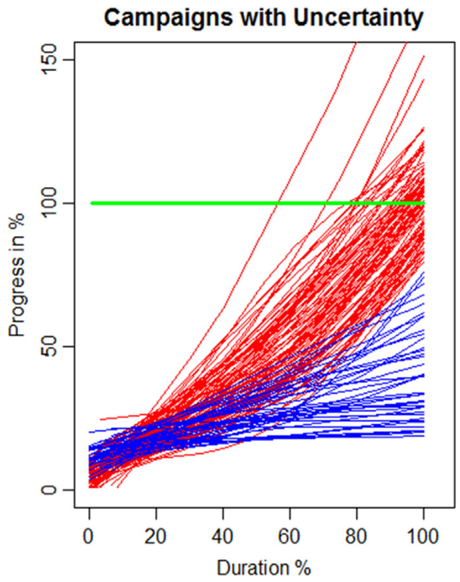


Fig. 10. Overlapping successful and unsuccessful crowdfunding campaigns that have not reached 30% of their goal by 20% of the way through the campaign. Note: Red (blue) curves represent successful (unsuccessful) campaigns, and the green line is the 100% contribution threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

can lead to the disadvantage of variability in the estimates, using a large number of subject observations reduces the problem of variability (Xiong & Bharadwaj, 2014).

(2) The principles of dynamics from the physics literature can also be applied to predictions in the crowdfunding context. This approach is novel because the concept of using the second order equation of motion has not been extended to forecasting in the business literature before. The concept allows us to take advantage of ideas that are well established in the physical sciences literature and apply them to crowdfunding prediction problems.
(3) Incorporating information about the history and evolution of the process that leads to the outcome of the fundraising exercise leads to a better forecasting performance.

From a managerial perspective, our model is very useful because of its ability to forecast funding outcomes accurately from the very early stages of crowdfunding campaigns. This ability to forecast funding outcomes well in advance provides campaign creators with sufficient time allow them to make the necessary adjustments to their crowdfunding campaigns, engage in solicitation activities for more funds, and make post-funding plans (e.g., production plans for crowdfunded products) and arrangements for their crowdfunding campaigns. Furthermore, given that various crowdfunding platforms still do not provide forecasting tools and various third-party platforms are springing up to help campaign creators monitor, track and forecast their campaigns outcomes (e.g., kicktraq.com), our model forms a good forecasting tool which can be implemented easily by these platforms in order to provide more accurate and reliable forecasts.

6. Conclusion

Our explorations and developments contribute to the crowdfunding and forecasting domains in several ways. First, understanding and forecasting outcomes in the crowdfunding context is a non-trivial task that is important to all stakeholders in crowdfunding—the platform owner, creators, and backers. Our method shows the effectiveness of using the contribution dynamics in campaigns as a forecasting measure and provides evidence that it leads to superior forecasts. Second, we provide insights into the nature of contribution dynamics as they relate to both successful and unsuccessful crowdfunding campaigns. This study has at least three clear practical implications:

(1) Our method can be used to make more accurate predictions of crowdfunding outcomes. While this study has forecast outcomes at the ends of campaigns, the model could also be used to forecast outcomes at any future time in a campaign by estimating it with the future time as the focal end time.
(2) Adding relevant covariates to our dynamics model allows for a greater prediction accuracy.
(3) Our technique could also be adapted to other forecasting contexts that exhibit dynamics like the diffusion of products/services and the forecasting of stock prices.

Finally, as crowdfunding continues to grow and offer more fundraising opportunities, we hope that our work will not only provide a forecasting tool, but also raise interest and emphasize the need to build predictive models that will fit the crowdfunding context better.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijforecast.2017.07.003.

References

Agrawal, A. K., Catalini, C., & Goldfarb, A. (2011). *The geography of crowdfunding*. National Bureau of Economic Research.

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

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, *15*(5), 215–227.

Beier, M., & Wagner, K. (2015). Crowdfunding success: a perspective from social media and e-commerce. In *Proceedings of the International Conference on Information Systems (ICIS)*.

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: tapping the right crowd. *Journal of Business Venturing*, *29*(5), 585–609.

Burtch, G., Ghose, A., & Wattal, S. (2013). An empirical examination of the antecedents and consequences of contribution patterns in crowdfunded markets. *Information Systems Research*, *24*(3), 499–519.

Colombo, M. G., Franzoni, C., & Rossi-Lamastra, C. (2015). Internal social capital and the attraction of early contributions in crowdfunding. *Entrepreneurship Theory and Practice*, *39*(1), 75–100.
Cordova, A., Dolci, J., & Gianfrate, G. (2015). The determinants of crowdfunding success: evidence from technology projects. *Procedia: Social and Behavioral Sciences*, *181*, 115–124.
De Gooijer, J. G., & Hyndman, R. J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, *22*(3), 443–473.
Ebbes, P., Papies, D., & Van Heerde, H. J. (2011). The sense and non-sense of holdout sample validation in the presence of endogeneity. *Marketing Science*, *30*(6), 1115–1122.
Etter, V., Grossglauser, M., & Thiran, P. (2013). Launch hard or go home!: Predicting the success of Kickstarter campaigns. In *Proceedings of the first ACM conference on online social networks* (pp. 177–182). ACM.
Foutz, N. Z., & Jank, W. (2010). Research note-prerelease demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science*, *29*(3), 568–579.
Frydrych, D., Bock, A. J., Kinder, T., & Koeck, B. (2014). Exploring entrepreneurial legitimacy in reward-based crowdfunding. *Venture Capital*, *16*(3), 247–269.
Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013). Crowdfunding support tools: predicting success and failure. In *CHI'13 extended abstracts on human factors in computing systems* (pp. 1815–1820). ACM.
Haneuse, S., Schildcrout, J., Crane, P., Sonnen, J., Breitner, J., & Larson, E. (2009). Adjustment for selection bias in observational studies with application to the analysis of autopsy data. *Neuroepidemiology*, *32*(3), 229–239.
Hastie, T., & Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, *1*, 297–310.
Hays, S., Shen, H., & Huang, J. Z. (2012). Functional dynamic factor models with application to yield curve forecasting. *The Annals of Applied Statistics*, *6*(3), 870–894.
Hou, J., Wang, N., & Ge, S. (2015). Antecedents of crowdfunding project success: an empirical study. In *Proceedings of the 2015 Wuhan International Conference on e-Business*, (52, pp. 610–617).
Huffington Post, (2013). Crowdfunding solutions: 3 tips for using Kickstarter to fund journalism. Retrieved 2014, from http://www.huffingtonpost.com/garrett-goodman/crowdfunding-solutions_b_4085570.html.
Hui, S. K., Meyvis, T., & Assael, H. (2014). Analyzing moment-to-moment data using a Bayesian functional linear model: application to TV show pilot testing. *Marketing Science*, *33*(2), 222–240.
Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review*, *39*(1), 25–43.
Kickstarter, (2014). Kickstarter stats. (Retrieved 11.09.14), from https://www.kickstarter.com/help/stats?ref=footer.
Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research*, *43*(1), 39–58.
Koch, J.-A. (2016). The phenomenon of project overfunding on online crowdfunding platforms–analyzing the drivers of overfunding. In *Proceedings of the 24th European Conference on Information Systems* (ECIS 2016); Istanbul, Turkey 2016.
Koch, J.-A., & Siering, M. (2015). Crowdfunding success factors: the characteristics of successfully funded projects on crowdfunding platforms. In *Proceedings of the 23rd European Conference on Information Systems* (ECIS 2015); Muenster, Germany 2015.
Kuppuswamy, V., & Bayus, B. L. (2014). Crowdfunding creative ideas: the dynamics of project backers in Kickstarter. UNC Kenan-Flagler Research Paper (2013-15).
Kuppuswamy, V., & Bayus, B. L. (2017). Does my contribution to your crowdfunding project matter? *Journal of Business Venturing*, *32*(1), 72–89.
Lindsay, D. (2015). Local governments and nonprofits test crowdfunding for civic projects. The Chronicle of Philanthropy, 7 January 2015. (Retrieved 03.03.17), from https://www.philanthropy.com/article/Local-Governments-and/152005.
Massolution, (2013). 2012CF Crowdfunding industry report. (Retrieved 09.08.14), from http://www.crowdsourcing.org/research.
Mollick, E. (2014). The dynamics of crowdfunding: an exploratory study. *Journal of Business Venturing*, *29*(1), 1–16.
Pan, Q., & Schaubel, D. E. (2008). Proportional hazards models based on biased samples and estimated selection probabilities. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, *36*(1), 111–127.
Ramsay, J. O., & Dalzell, C. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *53*, 539–572.
Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis: methods and case studies*. New York: Springer.
Reddy, S. K., & Dass, M. (2006). Modeling on-line art auction dynamics using functional data analysis. *Statistical Science*, *21*(2), 179–193.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.
Shane, S., & Cable, D. (2002). Network ties, reputation, and the financing of new ventures. *Management Science*, *48*(3), 364–381.
Shane, S., & Stuart, T. (2002). Organizational endowments and the performance of university start-ups. *Management Science*, *48*(1), 154–170.
Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310.
Silverman, B., & Ramsay, J. (2005). *Functional data analysis*. New York: Springer.
Sood, A., James, G. M., & Tellis, G. J. (2009). Functional regression: a new model for predicting market penetration of new products. *Marketing Science*, *28*(1), 36–51.
Stock, J. H., & Watson, M. W. (1998). *A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series*. National Bureau of Economic Research.
Wang, S., Jank, W., Shmueli, G., & Smith, P. (2008). Modeling price dynamics in eBay auctions using differential equations. *Journal of the American Statistical Association*, *103*(483), 1100–1118.
Wired.com (2012). Kicktraq is like Google Analytics for Kickstarter. Retrieved from http://www.wired.com/2012/08/kicktraq-google-analytics-for-kickstarter.
Xiong, G., & Bharadwaj, S. (2014). Prerelease buzz evolution patterns and new product performance. *Marketing Science*, *33*(3), 401–421.
Zheng, H., Li, D., Wu, J., & Xu, Y. (2014). The role of multidimensional social capital in crowdfunding: a comparative study in China and US. *Information & Management*, *51*(4), 488–496.
Zhou, L., Huang, J. Z., Martinez, J. G., Maity, A., Baladandayuthapani, V., & Carroll, R. J. (2010). Reduced rank mixed effects models for spatially correlated hierarchical functional data. *Journal of the American Statistical Association*, *105*(489), 390–400.
Zvilichovsky, D., Inbar, Y., & Barzilay, O. (2015). Playing both sides of the market: success and reciprocity on crowdfunding platforms. Available at SSRN: https://ssrn.com/abstract=2304101.

Onochie Fan-Osuala is an Assistant Professor in IT and Supply Chain Management at the University of Wisconsin Whitewater. He holds a Ph.D. in Business Administration with concentration in Information Systems from the University of South Florida, Tampa. He is interested in using statistics, econometrics and data mining to solve problems in the information systems, operations and marketing domains especially in areas focusing on online marketplaces, social media, crowdsourcing and crowdfunding.

Daniel Zantedeschi is an Assistant Professor in Marketing and a Translational Data Analytics Fellow at The Ohio State University. He was formerly an Assistant Professor at the University of South Florida as well as a Mars Co. Postdoctoral research fellow at the Wharton School of the University of Pennsylvania. He holds a Ph.D. in Information, Risk and Operations Management from the University of Texas at Austin. His research interests lie at the interface between statistical methodology and business analytics.

Wolfgang Jank is the Anderson Professor of Global Management in the Department of Information Systems and Decision Sciences at the College of Business, University of South Florida. Before joining the University of South Florida, Dr. Jank was Associate Professor in the Department of Decisions, Operations & Information Technologies, and Director, Center for Complexity in Business at the Smith School of Business, University of Maryland. He is interested in using ideas from statistics and data mining to solve problems in electronic commerce, marketing, information systems, and operations management. Dr. Jank's research has been published in the statistics, data mining, marketing, information systems, and operations management literature. He has authored over seventy refereed articles, book chapters and conference papers, authored

and edited two books, and presented his work at national and international meetings. Dr. Jank received his MasterâĂŹs degree in Mathematics from the Technical University of Aachen (Germany) and his Ph.D. in Statistics from the University of Florida. After moving to the University of Maryland, he co-initiated a new research area on Statistical Methods in eCommerce and a workshop series on the same name. He has received the Best Information Systems Publication award in 2008. Dr. Jank teaches courses at the undergraduate, graduate and executive levels in the areas of statistics, data mining and quantitative marketing.