

Louvain School of Management

How is the final amount built in a crowdfunding process ?

Impact of messages sent during the campaign by the project holder

Author : Tanguy De Bels
Promotor : Pr. Caroline Ducarroz
Academic year 2018-2019
Master [60] in management sciences (shift schedule)

Acknowledgments

To my supervisor, the professor Caroline Ducarroz who accepted my late enrollment in this project and the various arrangements to accommodate my situation. Many Thanks!

My partner in crime, wary reviewer, fierce perfectionist and brother, Thomas: our interesting and long-lasting chats were a great source of inspiration for this work. Your interest, encouragements and contributions led to a complementary master thesis and a common theoretical part. I was really pleased to finally work with you and I am eager to repeat this kind of collaborative work.

To the rest of my family: you were always keen to know what I was doing and how I was proceeding, although it is likely that you never grasped what it was all about!

I am also grateful to my friends who have supported me along the way even though I had many complaints regarding my schedule this year.

Special mention to my coworkers and patrons which accommodated with my early arrivals and accepted my numerous moments of frustration during this year. It was fantastic to have the opportunity to deliver my work whenever I had the time even during the weekend. Thank you so much for this flexibility.

Thanks for all your encouragements.

As usual, for every computer project, big thanks to OpenClassrooms which supported me during my whole studies and later during my work.

Abstract

Crowdfunding is a mechanism for collecting the financial contributions of a large number of donors to finance a project launched by a project leader, usually all over several weeks, via an online platform. The donations make it possible as time goes by to reach the aimed amount of a campaign.

The project holder can send information and encourage donations during the period of the crowdfunding. This news system is implemented by messages posted on the project web page.

This paper will study how a campaign global amount is built from the individual donations of donors, and will analyze more specifically the impact on this final amount of the information given by the project holder as crowdfunding progresses.

Moreover, interesting similarities can be found between this crowdfunding mechanism and the English auction mechanism. We will transpose, in the case of crowdfunding, the econometric model elaborated in the article written by Ducarroz C., Yang S. and Greenleaf E. in 2016 in *Journal of Marketing* (Understanding the impact of auction promotional messages), to see if the results obtained in crowdfunding converge with those described in the article.

The results converged on some points and more specifically showed that in crowdfunding on the aggregate level the number of jump-donations and more simply just donations had an impact on the final amount collected. On the micro level, we observed that with the increase in the number of donations the project holder is less likely to issue messages. However, persuasive messages have a positive impact on jump donation.

Contents

Abstract	iii
List of Abbreviations and Definitions	vii
List of Tables	viii
Liste of Figures	ix
Liste of Listings	x
 Introduction	 1
 Theoretical Part	 3
1 Crowdfunding Overview	4
1.1 Origin	4
1.1.1 In the United States	5
1.1.2 In France	5
1.2 Definition	6
1.2.1 Creating Links	6
1.2.2 Transparency and Trust	7
1.2.3 Collaboration and Shared Knowledge	7
1.2.4 Reduction of Entrepreneurial Uncertainty	7
1.3 Crowdfunding Types	8
1.3.1 Donation-Based Crowdfunding	8
1.3.2 Reward-Based Crowdfunding	8
1.3.3 Crowdlending	9
1.3.4 Equity-Based Crowdfunding	9
1.3.5 Typology Summary	10
1.4 Crowdfunding Application	11
1.4.1 Crowdfunding Workflow	11
1.4.2 Integration in Funding	12
1.5 Communication Strategy	14

2	Related work	15
2.1	Communication as a Success Factor	15
2.2	Appeal Modes - Message Content	16
2.2.1	Guilt Appeal	16
2.2.2	Self-Benefit and Others-Benefit Appeals	16
2.2.3	Nostalgia Appeals	16
2.3	Parallel Between Crowdfunding and Auctions	17
2.3.1	Jump Bidding	17
2.3.2	Late Bidding	17
2.3.3	Message Frames	18
	 Empirical Part	 19
3	Context	20
3.1	Objectives	20
3.2	The KissKissBankBank Platform	21
3.2.1	Process to Launch a Campaign	22
4	Model	25
4.1	News Types	25
4.1.1	Description of the campaign (pre-campaign)	25
4.1.2	Persuasive	26
4.1.3	Informative	27
4.1.4	Out of Time Scope (post-campaign)	28
4.2	Modelisation	29
4.2.1	Conceptual Framework	29
4.2.2	Hypotheses	32
4.2.3	Econometric Modeling	36
5	Data Extraction	39
5.1	Tools	39
5.1.1	Basics of the Web	39
5.1.2	Scraping	42
5.1.3	Classifier	44
5.1.4	Implementation Choices	45
5.2	Data Extraction Approach	46
6	Empirical Results	47
6.1	Aggregate Level	47
6.2	Micro Level	49
6.2.1	Contributor Decisions	49
6.2.2	Project Holder Decisions	50

Conclusion	51
7 Discussion	52
7.1 Limitations	52
7.1.1 News Types	52
7.1.2 Missing Information	52
7.1.3 Donation Types	53
7.2 Improvements	54
7.2.1 Additional Indicators	54
7.2.2 Decision Interval	54
7.2.3 Background Information	55
7.2.4 News Types	55
8 Conclusion	56
References	57
Appendices	60
A Web Scraping	61
A.1 Security Issues	61
A.2 JavaScript Execution	62
A.3 Data Storage	63
A.3.1 Pickle over Json	63
A.3.2 Pickle over Marshal	63
A.4 Libraries	64
B Source Code	65
B.1 Objects Code	65
B.2 Script Code	67

List of Abbreviations and Definitions

Abbreviations

KKBB: *KissKissBankBank*

VSE: *very small enterprises*

SME: *small and medium-sized enterprises*

I.e.: *id est* (from Latin “that is”).

E.g.: *exempli grātiā* (from Latin “for the sake of an example”).

Definitions

Campaign: is a connected series of operations designed to bring about the funding of a project by raising many small amounts of money from a large number of people, typically via the Internet. The campaign is led by the project holder and constrained by a set goal and an ending date. More information can be communicated by the project holder through the campaign webpage.

Project holder: is the person proposing the project to the platform and administrating the campaign. He is responsible for bringing the project to completion.

Crowdfunders: are the people who contributed to a campaign by giving an amount of money to see the project reach completion.

Also called *donators*, *donors* or *contributors* in this work.

Donation: is the amount of money given by a crowdfunder to contribute to a project.

In-process messages: represent efforts by founders to reach out to current and potential funders, and to inform interested backers about developments in a project by adding new information during the campaign period.

Also called *news*, *updates* or *actualities* in this work.

Crowd: is the large group of individuals addressed to collect money. With the use of internet for crowdfunding, the crowd is the whole world for some campaigns.

List of Tables

1.1	Platforms from the USA	5
1.2	Platforms from France	5
1.3	Comparison of Crowdfunding Typology	10
4.1	Example of Description of a Project	25
4.2	Example of Persuasive News	26
4.3	Example of Informative News	27
4.4	Example of Informative News Regarding the Field	27
4.5	Example of Out of Time Scope News	28
4.6	Variable Definitions and Summary Statistics of the Data	38
5.1	HTML Identification of an Element	43
5.2	Data Extraction Steps Approach	46
6.1	Regression Aggregate Level Results	47
6.2	Regression Aggregate Level Results - Project Holders Decisions	48
6.3	Regression Aggregate Level Results - Contributors Decisions	48
6.4	Regression Micro Level Results	49

List of Figures

1	Example of Campaign	1
1.1	The Different Financing Phases of a SME's Life Cycle	13
1.2	News Intensities Depending on the Different Phases of a Campaign	14
3.1	KissKissBankBank 10 Years Anniversary Key Numbers	22
3.2	Example of Donation with Reward	23
3.3	Example of Donation without Reward	23
4.1	The Conceptual Framework	31
5.1	HTML Code for one Button	40
5.2	UML Class Diagram of the Campaigns Objects	45
A.1	Rendering of Donations on a Campaign Page	62

Code Listings

5.1	HTML Code Demonstrating the Inclusion of Tags	39
B.1	Campaign Objects Python Code	65
B.2	Code of Scraping Script	67

Introduction

Participatory financing, or crowdfunding, is a mechanism that collects financial contributions from a large number of individuals through an Internet platform to finance a project. The purpose of this research is to show the impact of information given by a project holder during a crowdfunding campaign on the final amount collected. The information are relayed through news, text delivered after the start of the campaign at the discretion of the project holder, appearing on the webpage of a campaign. This webpage contains a section specifically dedicated to these messages besides the other information present regarding a campaign as we can see in the figure below.

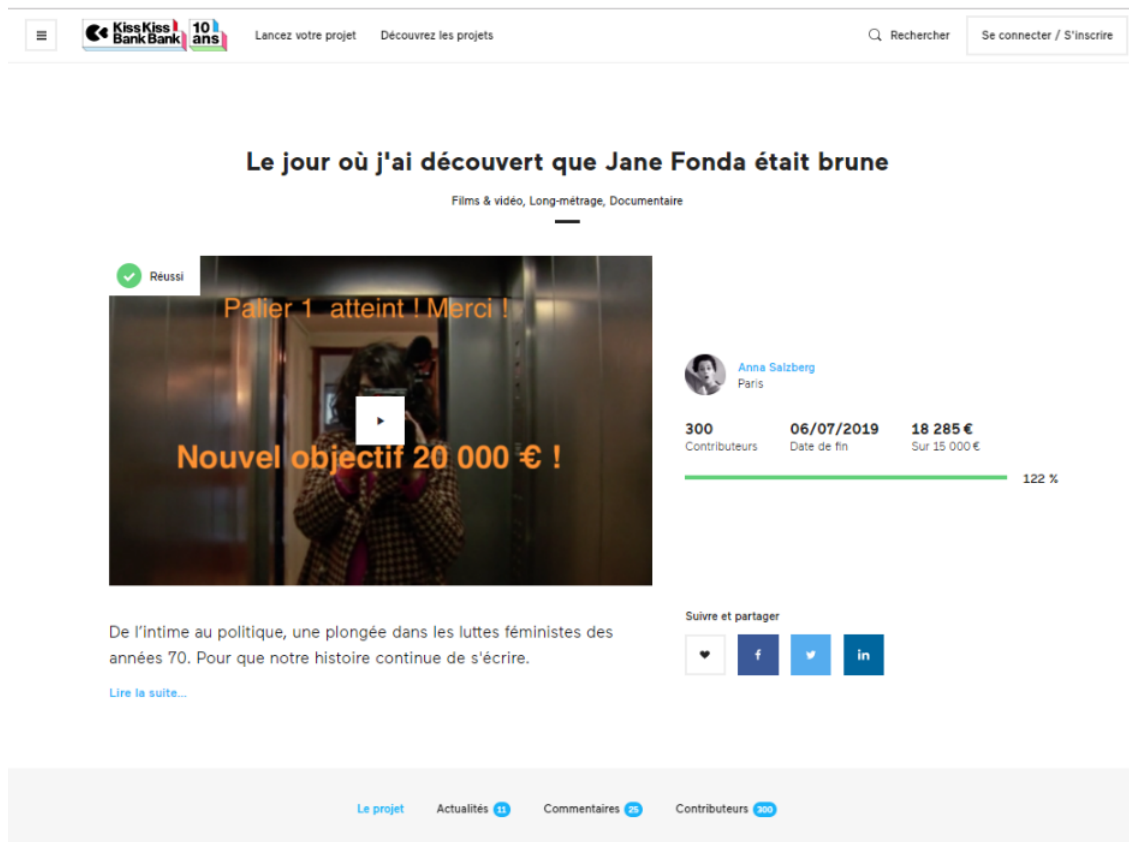


Figure 1: Example of Campaign¹

¹Screen from this campaign: <https://www.kisskissbankbank.com/fr/projects/le-jour-ou-j-ai-decouvert-que-jane-fonda-etait-brune-3>

To do so, this work is aiming to transpose the study made of the impact of in-process promotional messages sent during an auction by Ducarroz, Yang and Greenleaf (2016) to crowdfunding. The above-mentioned paper shed light on the impact of two kind of promotional messages, informative and persuasive messages, during an auction: both would attract more bidders with a superior effect to the later. It would also show that the final amount is directly depending from the number of bidders. Moreover with the growth in the number of bidders, the amounts bid would less often be above the minimal limit.

Several elements of the model shall be altered to adapt to the restrictions of this other kind of money collecting platform. Even though it will still focus its analysis on two levels, a micro and an aggregate level, in our case the donation and the campaign level. All the adaptations will be discussed further in the modelisation section. This transposition shall indicate if the mechanisms work in the same way and if the results are similar.

To achieve this analysis, a literature review regarding the mechanism of crowdfunding and the impact and types of messages sent during a campaign will be made. This review will also be focused on aspect exposed in the article related to the auction by Ducarroz, Yang and Greenleaf (2016). Especially regarding the kind of messages sent during a campaign.

After this review, the hypotheses of the above-mentioned article will be adapted to the crowdfunding field to elaborate a similar econometric model. This model shall allow us deduce the validity of the hypotheses.

To be able to work on the hypotheses and apply a model, a sample of several campaigns information is needed. To do so the website of the chosen platform, *KissKissBankBank*² will be scrapped. These extracted data will be stored and skimmed.

The extraction from the website shall concern:

- Basic information available regarding a campaign
- The contributions made by the donors for the project
- The news published by the project holder during the campaign which may have impacted the donations

It will also be needed to define types of news to refine the analyze of their impact on the campaigns. For this work, the types used in the referential article will be used with slight adaptations. Then a classification of each news of the campaigns meeting the criteria of selection for the sample will be done. The relevance of the news needs to be tested to remove irrelevant ones from the sample.

After applying the model on the data obtained an analysis of the results will determine the validity of the hypotheses. Afterwards a discussion of the limitations and possible improvements of the model will follow.

²<https://www.kisskissbankbank.com>

Theoretical Part

Chapter 1

Crowdfunding Overview

1.1 Origin

Although it was not originally called Crowdfunding, this concept has been around since the 1700s when the first concepts of participatory financing were developed.

Indeed, in the 1700s, some artists, such as Mozart or Alexander Pope, offered written recognition in their artworks in return for financial assistance to enable them to finance their works.

In addition, the crowdfunding principle can be observed through a famous example, the financing of the pedestal of the Statue of Liberty in New York, inaugurated in October 1886. It was financed during two years by a very significant number of micro donations following the appeal launched by the American journalist Joseph Pulitzer (Cazemajour, 2013).

Nowadays, the word "Crowdfunding" has become widely used with the fast-growing development of social media. Indeed, the broad public is now aware of participative financing on online platforms, making it possible to raise funds to finance personal, professional or charitable projects.

These so-called crowdfunding platforms have made it possible to finance other types of projects, such as the financing of the election campaign of Obama in 2008 and 2012 (Ricordeau, 2013).

It is also important to note that the first online platform for musical purposes, ArtistShare, was launched in 2014 and still allows young artists to have their musical projects funded by anyone who would like to take part in the "adventure" (Maalaoui and Conreux, 2014).

Given the emergence of a very large number of crowdfunding platforms in recent years, it is essential to mention a few major companies (Table 1.1 and Table 1.2).

1.1.1 In the United States

Name	Creation Date	Target
ArtistShare	Octobre 2003	• Musical projects
IndieGoGo	January 2008	• Generalist without pre-selection
Kickstarter	April 2009	• Artistic • Technological projects
Rocket Hub	January 2010	• Generalist without pre-selection

Table 1.1: Platforms from the USA³

1.1.2 In France

Name	Creation Date	Target
My Major Company	December 2007	• Artistic projects
KissKissBankBank	September 2009	• Generalist with pre-selection
Mutuzz	July 2010	• Digital creation
Ulule	October 2010	• Creative and innovative projects

Table 1.2: Platforms from France³

³CAZEMAJOUR, L. (2013), « La place du crowdfunding philanthropique », Centre d'étude et de recherche sur la philanthropie, 30 pages.

1.2 Definition

Crowdfunding literally means "financing by the crowd". Other names are used to describe this phenomenon, such as participatory financing, 2.0 financing or socio-financing (Iizuka, 2014).

The European Commission defines participative financing as "a reliable and complementary source of financing, in addition to the general banking sector, and recognises that it is one of the new economic models that contribute to building a pluralistic and resilient social society".

In the broader sense of the term, participatory financing is explained as follows: entrepreneurs try to obtain funds from a very wide audience. Each individual provides a very small amount, rather than raising funds from a very small group of professional investors with large amounts (Cieply and Le Nadant, 2016).

When crowdfunding is highlighted, it is very important to mention *crowdsourcing*. The latter was born in the 1990s through IT and more precisely through open source. Some researchers limit this expression to the field of the Internet and computer science (Good BM. and Su AI., 2013), while the second stream of the literature has focused on semantics by associating all forms of collaboration (Howe, 2006; Kittur, 2010).

Indeed, crowdsourcing represents the core of the crowdfunding principle since it uses the crowd to gather ideas, get feedback, but also to find solutions in order to develop business activities (Doan et al., 2011).

The various principles, developed below, on which the crowdfunding concept is based, namely transparency and trust, make it an innovative alternative means of financing.

1.2.1 Creating Links

The crowdfunding principle is based on two types of links. Indeed, there is first a purely financial link. The latter is manifested by the contribution of funds from crowdfunders to project holders. Secondly, there is an extra-financial link. It is, in contrast, focused on sharing and participation (Rifkin, 2014).

Individuals are motivated by a real motivation, which results in a sense of help and contribution, as well as a real willingness of consumers to become "consumer actors" (Cieply and Le Nadant, 2016).

However, this type of emotion-based one-click decisions is also a source of intrinsic fragility in this funding model. Indeed, since decisions may be guided more by impulse than by reason, the approach according to which the relationship between risk and profitability is positive could then be reversed (Slovic et al., 2004).

1.2.2 Transparency and Trust

The emergence of social networks allows crowdfunding to develop on a very large scale and at a quite impressive speed. Crowdfunding takes advantage of the potential of new Internet technologies to bring people together and increase interaction between different parties within a large community of investors and promoters of innovative projects. Indeed, renewed *trust* in the Internet makes it possible to take advantage of the potential of new technologies while simplifying interactions between the various parties.

In addition, it should be noted that in this kind of financing, crowdfunders are regularly informed of the progress of projects (Cieply and Le Nadant, 2016).

1.2.3 Collaboration and Shared Knowledge

One of the key factors for the success of crowdfunding is that it relies on collective intelligence. Indeed, this ability of a community to converge intelligence and knowledge to progress towards a common goal refers directly to crowdsourcing which, as said earlier, is based on the crowd in order to obtain ideas, solutions, or the search for feedback, whether positive or negative (Cieply and Le Nadant, 2016).

Social evidence theory (Cialdini, 1984) also plays a role. This theory represents the automatic and unreflected adhesion of individuals in a position statement. Indeed, this idea is based on the fact that, if many people believe in something, then it is more reasonable to conform to their thinking, because it is almost impossible for so many individuals to make mistakes (Cialdini, 1984).

According to studies by Kim and Viswanathan (2014), it can also be noted that the decisions of the crowd are strongly influenced by experts, which further reinforces our principle of shared knowledge. Indeed, for individuals, the positive opinion of an expert, who is considered as an erudite in his field, can lead a lambda person to become a potential contributor.

1.2.4 Reduction of Entrepreneurial Uncertainty

This last pillar is a very important indicator for assessing the health of the current project. Indeed, based on an initial analysis of the crowd, it will make it possible to limit market risks given that the larger the number of contributors, the healthier and more viable the project will be. (Belleflamme et al., 2013).

Raising funds from professional investors does not offer the same informational feedback since investors do not make decisions based on consumption, but on their assessment of the overall profitability of the investment opportunity. Crowdfunding would therefore make it possible to limit the potential market risk since projects that have not generated sufficient crowdfunding participation can be dropped relatively quickly (Cieply and Le Nadant, 2016).

1.3 Crowdfunding Types

There are 4 distinct types of crowdfunding, each with their own characteristics:

- Donation-Based
- Reward-Based
- Crowdlending
- Equity-Based

Before giving a detailed explanation of each category, it is important to mention that the two on the left are based on dynamics social and/or solidarity-based, while the last two are more related to finance (Iizuka, 2014).

1.3.1 Donation-Based Crowdfunding

The concept of charitable crowdfunding is mentioned when the literature review refers to platforms for non-compensatory donations. This type of participatory financing is therefore the least suitable for financing companies. It will rather concern NGOs or charities that need an online platform to centralize all their donations.

Today, there are three types of no-reward donation platforms. First, there are donations to charities, associations, political parties or charities. Secondly, there are donations to individuals, particularly in the United States for the payment of medical bills. Thirdly, donations for portfolios or the electronic jackpot can be mentioned. Nowadays, these three segments together represent 20% of crowdfunding platforms (Iizuka, 2014).

These platforms give people the opportunity to invest a certain amount to support the funding of a project, without any tangible compensation. There is therefore no system based on obtaining a reward. One of the motivations that may drive individuals to invest is recognition (Iizuka, 2014).

One of the major characteristics of the so-called "donation-based" platforms is the fact that the project holder preserves his intellectual property (Dehorter, 2012).

1.3.2 Reward-Based Crowdfunding

The principle of this kind of platform is pretty similar to donation. Nevertheless, there are two major differences. First, studies will no longer talk about donations but rather about *income* for the project holder. Second, Second, the donator, considered as an investor, will receive a compensation, which will not be a financial compensation though (Iizuka, 2014). In some cases, it is possible that contributors may receive a symbolic counterpart, such as a thank-you email or a postcard (Maalaoui and Conreaux, 2014). But more frequently, the donator will receive the object or service produced by the funded company itself.

Maalaoui and Conreaux also noted, in their 2014 study, that the larger the capital contribution, the greater the reward.

In this type of funding, each reward is based on the type of transaction performed. Indeed, these platforms offer two possibilities of transactions. First, the simple donation, which will act in the same way as for non-compensatory donation platforms, i.e. the contributor may receive a symbolic counterpart. Second, pre-sales may be available. Pre-sales consists of pre-ordering the final product by providing funds to start production (Ricordeau, 2013; Belleflamme and al., 2013).

This type of crowdfunding presents the advantage of reducing entrepreneurial uncertainty (Belleflamme et al., 2013), since it allows the project owner to gain a first clientele in pre-sales and therefore to ensure that the market responds positively to its product.

It is also important to mention that according to Iizuka (2014), "these are the most represented platforms in the crowdfunding world with more than 40% of sites".

1.3.3 Crowdlending

The primary purpose of equity lending platforms is to connect project promoters (borrowers) with a set of investors willing to invest their savings (Laplume, 2013).

Nowadays, there are two types of lending platforms. On the one hand, there are interest-free loans. These are similar to microfinance because they are, in general, microloans provided to project promoters (Ricordeau, 2013; Hendrickx, 2014). On the other hand, there are loans with interest that require monthly repayment (Hendrickx, 2014).

The interest rate can also be variable depending on the platform. Indeed, in some cases, they classify borrowers according to the level of risk (Laplume, 2013).

According to *L'Echo*⁴, in recent years, the various investors could expect a return on capital of between 3 and 12% spread over a period of 3 to 4 years for the various equity loan platforms.

1.3.4 Equity-Based Crowdfunding

The equity-based crowdfunding principle is based on participation in a company, mainly start-ups and SMEs (Iizuka, 2014). In this case, the investor will be remunerated with shares of the company, which will allow him to receive dividends, or capital gains in case of resale.

It is also important to mention that this type of investment is generally used in start-ups or SMEs in the launching phase (Iizuka, 2014). In this system, the amount of the investment is limited to one million euros, which is perfectly suited to ambitious SMEs and start-ups.

However, it should be noted that this type of crowdfunding also involves certain risks (Iizuka, 2014): if the start-up/SME is in default of payment or if it never reaches its break-even point, investors run the risk of never recovering their capital.

⁴<https://www.lesechos.fr/2016/11/crowdfunding-les-nouvelles-regles-pour-investir-227571>

1.3.5 Typology Summary

Table 1.3 represents a summary of the different forms of crowdfunding and their respective characteristics.

Types	Donation-Based	Reward-Based	Crowdlending	Equity-Based
Concept	Philanthropic funding	- Philanthropic funding - Funding for arts, cultural, ...	Loans between individuals or to SMEs	- Investments - Acquisition of shares in the company
Funded projects	- Humanitarian and associative projects - Personal projects	Cultural and entrepreneurial projects	- Individuals - SMEs	- Young companies in the start-up phase - SMEs, VSEs
Remuneration	None	- Royalties - Counterparties	Interest (optional)	Dividends
Capital requirements	Between 2 K€ and 20 K€	Between 2 K€ and 20 K€	Between 2 K€ and 25 K€	Between 50 K€ and 1 M€
Actors (Examples)	- Helloasso - The Give - KKBB	- Kickstarter - Ulule - KKBB	- Friendsclear - Spear - Union Loan	- WiSeed - Anaxago - Smartangels

Table 1.3: Comparison of Crowdfunding Typology⁵

⁵IIZUKA, M. (2014), «Le crowdfunding: les rouages du financement participatif», EduBanque Editions, Guyancourt (France), 197 pages.

1.4 Crowdfunding Application

1.4.1 Crowdfunding Workflow

Although the notion of Crowdfunding may look simple at first glance, it actually hides a much more complex reality. Indeed, a crowdfunding campaign represents time and involvement from the project owner and the platform holder, which are key variables in the success of a project. A Crowdfunding campaign is generally divided into three distinct phases: pre-campaign, campaign and post-campaign.

Pre-Campaign

The pre-campaign consists of different steps.

The first represents a considerable period of time since it covers the preparation, submission and selection of the project by the platform (Bessière and Stéphaney, 2017).

Then, the project owner will have to expose a showcase to platform holder, which will be composed of a folder that the entrepreneur will have previously prepared. In order to be retained, this application must meet given criteria, which will help the project be the most credible and attractive in the eyes of potential contributors. Concerning the implementation of the showcase, for example, there may be videos that will allow donators to learn more about the project presented when the project gets launched (Bessière and Stéphaney, 2017).

Campaign

The project owner, during the campaign, needs be active on the platform and on social networks to motivate potential contributors to fund the initiative (Meyskens and Bird, 2015). Indeed, project owners will use a news system allowing contributors to stay informed of the project's progress. These various messages will aim to influence the behaviour of contributors or potential contributors, in order to achieve the objective of the crowdfunding campaign.

Bessière and Stéphaney (2017) describe this phase as crucial and intense. Indeed, it is at this moment that it will be necessary to convince contributors, who are part of the crowd, to participate.

Post-Campaign

When fundraising ends, there are two scenarios. The first one is to achieve or exceed the objective. In this case, funds are transferred to the project holder. The second scenario is an unachieved objective. In most platforms, funds will then be reimbursed to each contributor. This method is based on the principle of "All or Nothing" (Bessière and Stéphaney, 2017).

However, other platforms prefer to practice the "keep it all" principle which, as its name implies, allows project holders to keep all the funds collected even in the case of failure (Bessière and Stéphaney, 2017).

1.4.2 Integration in Funding

It is difficult to find funding to create and develop a business. Throughout its cycle, the start-up / SME will use different sources of funding depending on the level of risk incurred by the company. It is therefore important to mention and briefly explain the various phases, as well as the means of financing associated with them.

Indeed, crowdfunding projects are generally set up by start-ups or SMEs, given their high capital requirements due to the difficulties encountered when they are launched. In order to maximize the chances of success of the project, regular monitoring of the project will be necessary. This can be done through news or messages sent to the crowd, namely potential contributors.

Phase 1: Incubation and Priming

In this stage, the company's main concern will be to prepare for its launch, in particular by producing certain prototypes, but also by investing heavily in research and development. Obviously, this step occurs before any product or service is commercialized (FEB, 2013).

Throughout this period, the company will use *love money*, more commonly known as the 3Fs (Friends, Family, Fools), to raise funds. Regarding the amounts collected, they will not be significant and will only ensure the funding of the beginning of the project (Zimmern, 2012).

It is also during this phase that crowdfunding will be the most used, as it will act as a fund booster to address the problem of lack of funding that any company can encounter in its early days. This fundraising event will allow the company to ensure its development, but also to reach phase 2, which will allow them to capture the attention of Business Angels or other venture capital (FEB, 2013).

Phase 2: Launch

During this phase, the first commercialization of the product or service through market analysis will be the major concern. This step will be distinguished by its sales volume and relatively modest turnover (FEB, 2013).

It is also important to mention that it is at this phase that the company, having reached a new stage of development, will become interesting for *Business Angels*⁶ and other venture capital funds (FEB, 2013).

⁶The business angel is an investor who decides to financially support a company whose project seems to him to be innovative. In addition to part of his assets, this person makes available to the company manager his experience and all his skills in the entrepreneurial field.

Phase 3: Growth

This will be referred to both development and commercialization on a large scale. Indeed, in order to ensure the company's sustainability and to have a solid foundation, the company will use venture capital or bank loans to finance itself (FEB, 2013).

Phase 4: Maturity

The company, in its most advanced phase, will tend to use bank loans and/or stock exchange listing, which remains the most important means of financing in terms of funds raised.

SME's Life Cycle

According to FEB (2013), each phase of the company's development corresponds to a very specific and different means of financing. The more the company grows, the less the company will be considered risky.

Figure 1.1 presents the means of financing as well as the type of financial institution capable of providing adequate funding at each phase of the life cycle of a start-up or SME.

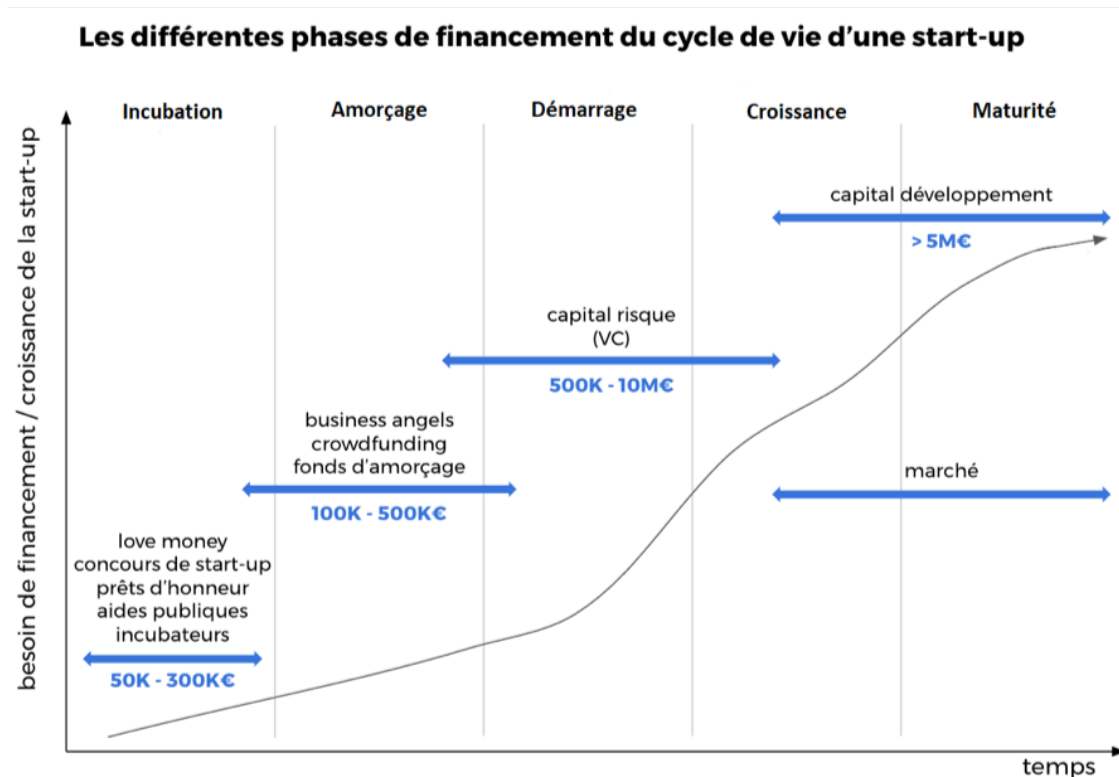


Figure 1.1: The Different Financing Phases of a SME's Life Cycle⁷

⁷EKELAND, M., LANDIER, A. et TIROLE, J. (2016), «Renforcer le capital-risque français», Conseil d'analyse économique, n°33, 12 pages. (modifié par Thomas De Bels)

1.5 Communication Strategy

To make their project as viable as possible, it is recommended that the project owner develop a communication strategy to reach the crowd. Depending on the platform chosen, the entrepreneur will have access to the platform's user network. Indeed, crowdfunding platforms have and make available to the project owner their communication service which will allow the crowd to stay informed of the progress of the project. However, it is important for project owners to have their own independent network, i.e. an advertising campaign targeting people who are not present on the platform (Bessière and Stéphan, 2017).

According to Bessière and Stéphan (2017), the communication strategy is based on three main axes:

- The *rhythm* of communication, which will have to be adapted according to the duration of the fundraising. In general, campaigns follow a U-shaped curve, i.e. they benefit from strong interaction at the beginning and end of the campaign and a trough in the middle of it.
- The *target* of communication, which is divided into three groups: family and friends, communication relays⁸ and the crowd.
- *Communication relays*, such as Twitter, Facebook or YouTube, which can help project holders make their campaign viral.

Entrepreneurs can also benefit from some coaching from the platform. It will often be able to assist them in their funding and communication strategy, as well as in the presentation of their project and the constitution of possible counterparts (Bessière and Stéphan, 2017). Figure 1.2 shows the importance of interactions, i.e. news and messages, between the crowd and the project owner according to the different phases of the campaign.

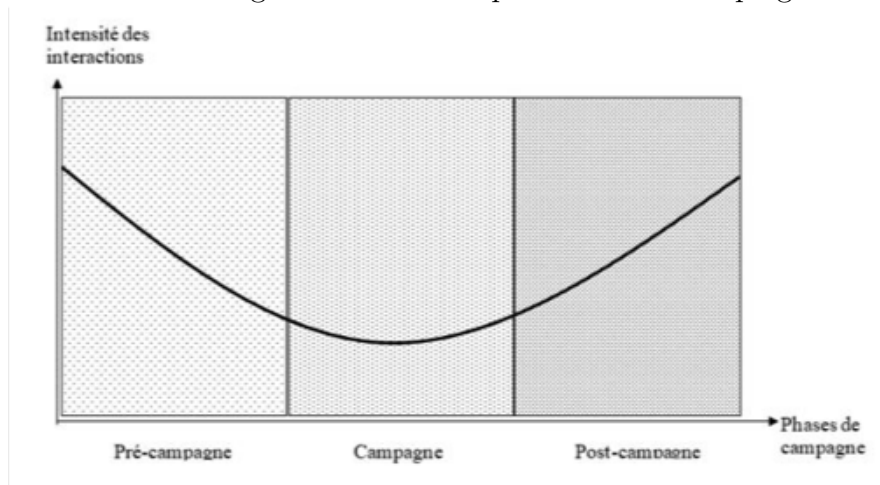


Figure 1.2: News Intensities Depending on the Different Phases of a Campaign⁹

⁸A communication relay is a system, organization or network that links two or more agents sharing the same objective.

⁹Bessière, V., & Stéphan, É. (2017). Le crowdfunding : fondements et pratiques. De Boeck Supérieur, 176 pages.

Chapter 2

Related work

2.1 Communication as a Success Factor

Previous research can help us assess whether communication in crowdfunding campaigns can be considered as a success factor or not.

According to Mollick et al (2014), the presence of strong communication through the regular publication of news, published by the project owner, is a key factor for the success of a fundraising campaign.

Regarding the publication of information and the project holder's means of communication, there is two options in order to keep the crowd informed while conducting the campaign for a project holder.

On one hand the updates. These are unilateral communication tools. Indeed, *updates*, in contrast to *comments*, do not allow project owner to respond directly to a message left by a participant. However, this tool can be used to provide additional information about the project to the crowd (Mollick, 2014). This option is mainly used to provide information on the progress of the project. Indeed, updates can be considered as signals on the quality of the current project (Mollick, 2014). In addition, these messages, without discrimination on the nature of the message, have a positive impact on the crowd's participation in terms of equity financing (Block, Hornuf, and Moritz, 2017).

On the other hand, comments allow the initiator of the project to interact directly with the crowd by taking into account opinions, whether positive or negative, in order to show his involvement and receptiveness to criticism. (Clauss et al., 2017). Therefore, this option aims at making the interactions between the project owner and the crowd as persuasive as possible, in order to collect as much funding as possible to reach the campaign objective as quickly as possible (Clauss et al., 2017).

2.2 Appeal Modes - Message Content

In this kind of practice, donations are encouraged through marketing communication, using appeal processes, in order to reach as many people as possible within the crowd. Three of them have received particular attention in the literature: guilt appeals, self- and others-benefit appeals, and nostalgia appeals (Chen et al., 2016). The content of these different types of messages will aim to convince the crowd, recognizing themselves through one of them, to invest in the crowdfunding campaign, and therefore to achieve the expected objective.

2.2.1 Guilt Appeal

This type of appeal aims at affecting the potential contributor's sensitivity. The issue of responsibility will be indeed highlighted, in order to raise awareness among the crowd and, consequently, to have them contribute by donating, in order to remove this feeling of guilt (Hibbert et al., 2007).

Some project holders use marketing communication practices that raises donors' interest, highlighting information or photos that exacerbate this sense of guilt, in order to achieve their objectives (Burt and Strongman, 2005; Huhmann and Brotherton, 1997). This type of message is frequently used in communications related to charities or health projects (Burt and Strongman, 2005; Huhmann and Brotherton, 1997).

2.2.2 Self-Benefit and Others-Benefit Appeals

Entrepreneurs also often use messages related to *self-accomplishment* to raise as much money as possible. These messages are intended to make potential contributors believe that funding this campaign will be the most rewarding thing they have done so far (White and Pelozo, 2009).

2.2.3 Nostalgia Appeals

Previous research shows that appeals for "nostalgia" also have a very strong impact on the crowd, the purpose of this type of practice is to make the donor contribute by bringing up images of the past that will make people regret the time spent (Ford and Merchant, 2010; Merchant et al., 2011; Zhou et al., 2012).

2.3 Parallel Between Crowdfunding and Auctions

This part will not aim to provide a complete description of the auction principle, however, after an in-depth analysis of the characteristics of crowdfunding through literature, this section will be dedicated to highlighting the links that exist between the crowdfunding and auction processes. Several parallels between the two processes are easily achievable. These include the behavior of bidders and donors, namely jump bidding and latent bidders, and the message frames.

2.3.1 Jump Bidding

Studies by Rothkopf and Harstad (1994) showed that there were several types of bidder behavior in the auction process. One of the most widely used in their studies is jump bidding.

On the auction side, "jump bidding" is defined as a much bigger bid than the minimum increment required. This behavior can be explained by several factors perceived as motivating the bidder. The three most obvious motivations would be the desire to discourage other potential bidders, the willingness to show the value given to the property or even impatience (Rothkopf and Harstad, 1994).

In comparison with jump bidding, the term jump donation will be used for crowdfunding. Some contributors who would like to show their support for a project that is really important to them will make a much bigger donation than the average (Leclercq 2016). The functioning of this behavior is therefore essentially the same from one sales process to another regarding this aspect. However, another characteristic of crowdfunding could be the desire for a more valuable reward. Indeed, some donors may engage in jump donation behavior in order to obtain a bigger reward.

2.3.2 Late Bidding

The term late bidding is defined as bidding that happens in the last moments of an auction. This mechanism can induce two types of behaviors. First, also called sniping, this behavior may refer to a bidder wanting to wait until the last moments of the auction, in order to make a bid to which other potential bidders will not have time to react, and therefore allowing him to win the auction. This type of behavior is mainly used in time-limited auctions. Second, late bidding also refers to bidding behaviour consisting of entering the auction relatively late given the need for more information to be convinced (Backus et al., 2015).

This second meaning will highlight a link with the crowdfunding principle. Donors with late donation behavior, similar to late bidding, will tend to expect a certain quantity of informational or persuasive messages throughout the campaign in order to convince them. They will therefore be likely, as in the case of auctions, to make their first donation when the project has already reached a certain achievement rate. However, studies have shown that in the case of crowdfunding, no sniping behavior has been highlighted (Hornuf and Neuenkirch, 2015).

2.3.3 Message Frames

The literature, covering different sales processes (e.g.: auctions and crowdfunding campaigns) differentiates both "rational" and "emotional" message frameworks (Chandy et al., 2001; Zhang and Gelb, 1996). These two types of messages will therefore have a relatively different impact on purchase intentions in the case of auctions, or in the rate of crowd participation in a crowdfunding campaign.

On one hand there is a "rational" message framework which, through a message, explains how the use of a service or product can anticipate or solve a problem. In general, it can be said that this type of message highlights the utilitarian ¹⁰ aspects of a product.

On the other hand, there is an "emotional" message framework which, in contrast to the previous one, emphasizes the achievement of a positive well-being. We therefore find the hedonic ¹¹ aspects of the product.

Regarding the effectiveness of message frames, previous studies, regarding sales processes, have shown that message frames are more effective when they correspond to the use conditions of the product. Indeed, we can see that utilitarian products will be much better represented by rational messages, while emotional messages will be better for hedonistic products (Johar and Sirgy, 1991; Zhang and Gelb, 1996).

Another factor for the effectiveness of message frameworks is participation. Indeed, other studies have highlighted that in cases of high participation in sales processes, e.g. in auctions or crowdfunding campaigns, rational messages are more effective, while emotional messages are much more appropriate for lower participation (Dens and Pelsmacker, 2010).

In addition, the various studies conducted on the subject have shown that rational message frameworks are more effective in new markets, while emotional message frameworks work better in older markets (Chandy et al., 2001).

¹⁰The adjective utilitarian describes something that is useful or functional.

¹¹Hedonic behaviour values pleasure and well-being above all else.

Empirical Part

Chapter 3

Context

3.1 Objectives

The purpose of this research is to show the impact of information given by a project holder during a crowdfunding campaign on the final amount collected. The information are relayed through news, text delivered after the start of the campaign at the discretion of the project holder, appearing on the webpage of a campaign.

To do so, this work will transpose the study made of the impact of in-process promotional messages sent during an auction by Ducarroz, Yang and Greenleaf (2016) to crowdfunding. The above-mentioned paper shed light on the impact of two kind of promotional messages, informative and persuasive messages, during an auction: both would attract more bidders with a superior effect to the later. It would also show that the final amount is directly depending from the number of bidders.

In the crowdfunding field we will work with campaigns instead of auctions. The final amount reached will be constituted through the contributions of donators, the goal becomes the achievement of the campaign and not the acquisition of a good. The emitter of the messages becomes a person directly involved in the project and concerned by its achievement. It also induces that the communication will be less governed by guidelines since each project holder will be responsible of his own publications.

As explained above, several elements of the model shall be altered to adapt to the restrictions of this other kind of money collecting platform. Even though it will still be possible to observe the impact of messages on the number of donators, and the amount contributed by each of them. These limitations and adaptations will be discussed in a further section.

3.2 The KissKissBankBank Platform

KissKissBankBank & Co. is a European leader in crowdfunding. This French company was launched in September 2009 by Omblin Le Lasseur, Vincent Ricordeau and Adrien Aumont. It advocates a committed vision of the peer-to-peer economy, embodied by three complementary platforms:

- KissKissBankBank¹²: "a pioneer of donation for donation dedicated to creation, innovation and solidarity"¹³.
- LENDOPOLIS: which enables small and medium-sized businesses to obtain professional credits directly from individuals.
- Goodeed: which supports NGOs through free donations.

KissKissBankBank & Co. is the only player in Europe to offer its services in three segments of crowdfunding, offering a solution tailored to a multitude of personal and professional projects. They started their activity with *KKBB*, then launched *LENDOPOLIS* by the end of 2014 and finally *Goodeed* was founded in 2015. The company was purchased in 2017 by *La Banque postale* (KissKissBankBank, 2019).

Since 2009, KissKissBankBank & Co. group sites have brought together nearly 1.3 million members. Together, they helped finance about 27.000 projects of all kinds, for a total of more than 80 million euros, including 70 million on *KKBB* and 11 million on *LENDOPOLIS* (KissKissBankBank, 2019).

This platform has been the incubator for some notable projects such as the documentary "Demain" on solutions (ecological, societal, etc.) existing around the world. This project, led by Mélanie Laurent and Cyril Dion, raised 444.390€ in 2014. Many other projects were launched through this mean and it also allowed people to launch their career or business.

The chosen platform used to collect data for this work is *KKBB*¹⁴. This website is based on the *Kickstarter* model, another well known crowdfunding platform that has a more international impact (KissKissBankBank, 2019; Cazemajour, 2013).

This platform was chosen in particular because of its transparency, as it is one of the only platforms to disclose donors, contribution amounts and dates on a campaign's web page. These features have made *KKBB* the only viable solution for obtaining data without the need to obtain it directly from the platform by access request.

¹²From now on, we will refer to the KissKissBankBank by the abbreviation *KKBBB*

¹³Platform welcomekit webpage: <https://kisskissbankbank.welcomekit.co/>

¹⁴Platform website: <https://www.kisskissbankbank.com/>

For their ten years anniversary, they published key numbers (see Figure 3.1):

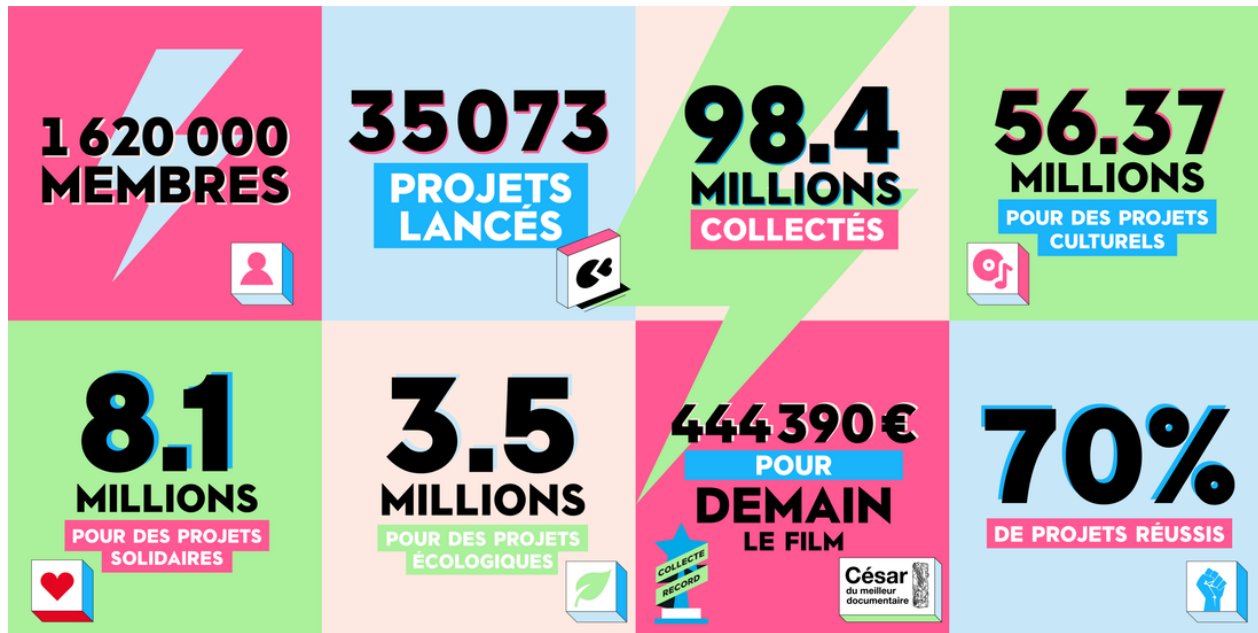


Figure 3.1: KissKissBankBank 10 Years Anniversary Key Numbers¹⁵

This figure reveals some interesting information:

- Only the successful or ongoing campaigns are displayed on the website. The failed ones are not displayed and we can not get access to the data of these campaigns. It is obvious since the figure announces 35.073 projects launched while the website only displays 20.870 projects (funded or live).
- The company still have access to data besides the ones available to the users. The failed campaigns were once live ones and we can then assume that the data are still in possession of *KKBB*.
- Most of the funded projects are linked to culture (almost 60%, 56.37 millions out of 98.4)

3.2.1 Process to Launch a Campaign

In order to propose a campaign, a project holder needs to fill a description of his/her project. *KKBB* team proceeds with a first screening of the projects presented to it, and accepts on average one project out of four. According to Vincent Ricordeau, one of the founders of the company, the rejected projects are not "creative enough" or are in a too "embryonic" state (Les Échos, 2013).

¹⁵Page of the anniversary: <http://welcome.kisskissbankbank.com/10ans/>

Once this selection is made, projects are published on the website, which creates the interface between the creators of the project and their contributors. Projects are funded by Internet users through donations. The project holder can offer different rewards corresponding to different levels of donations to encourage bigger donations from the contributors. The contributor can choose a counterpart of his choice corresponding to the amount of his donation, for example, the project holder behind the publishing of a new book could propose a signed copy. He is also allowed to get in touch with *kisskissbankers* via internal messaging. It is also possible to make a donation without expecting a reward (Les Échos, 2013).

On the platform the communication is totally in the hands of the project holders. It means that they decide of their own guidelines regarding their announcements as stated in section 1.5. Each project holder can have his/her own plan and some even decide to not interact at all during the campaign (through the system of news at the very least). On the other hand, some continue to interact with their donators, with this system, years after the achievement of their campaign.

This platform is using both the *donation based* and *reward based* type of crowdfunding presented in section 1.3. Figures 3.2 and 3.3 respectively show examples of a donation with expectation of a reward and of a donation without any expectation of any return on investment .



Figure 3.2: Example of Donation with Reward¹⁶

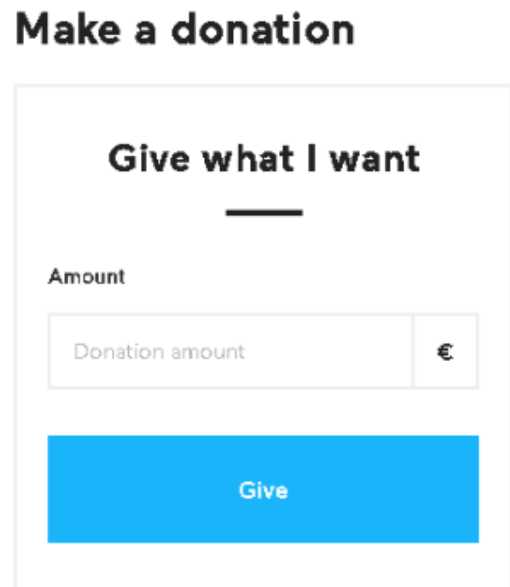


Figure 3.3: Example of Donation without Reward¹⁶

¹⁶Screens of donation options from this campaign: <https://www.kisskissbankbank.com/en/projects/les-amis-toukoul>

Rewards promised by the project holder to potential contributors, following their donation, are supposed to act as incentives to attract contributors to donate. It is also interesting to note that some rewards may be depleted at some point which urges the donor to act as soon as possible to gain his wanted reward. These rewards act as counterparts for the donations made by the contributors and can be chosen during the donation. Some reward can only be obtained by donating at least a minimal amount of money which means that donations can increase with the attractiveness of the rewards.

The projects are described with a budget (amount to reach to be able to launch the project), a time allowed for the collection and the counterparts set by the project holder related to the amounts committed. The financial goal set by the project holder may be exceeded during the collection but the campaign must reach his goal within the allotted time, otherwise the project fails and the contributors are reimbursed. This is the rule of "all or nothing" applied by the company and mentioned in section 1.4.1. If the goal is achieved with time remaining, the campaign continues to accept donations to facilitate the project or allow for further development. Project holders can also set new objectives or make promises to keep contributors attracted. (KissKissBankBank, 2019).

Payments are made by credit card via the PayBox secure payment system. KKBB receives a commission of 5% on the collected funds billed to the creators, plus 3% of secure banking fees (Les Échos, 2013).

Chapter 4

Model

4.1 News Types

We will be analyzing the impact of messages sent by the project holder during a crowd-funding campaign. These messages, called *news*, would be visible from contributors and users, who did not contribute yet, visiting the webpage of the project. We will refine the analysis by defining the kind of messages a project holder can send by observing the nature of the messages sent. Each kind could have a different impact and it will help determine if the effects are similar. We defined the following classes of messages to do so.

We tried to keep ourselves as close as possible from the ones announced by Ducarroz, Yang and Greenleaf (2016) but a few adaptations were necessary especially since the messages are not as easily defined as in auction. This issue will be further discussed in the limitations section (section 7.1).

4.1.1 Description of the campaign (pre-campaign)

The description of a campaign does not belong to any type since it is the first message seen by a potential donator while visiting the webpage. Before anything else the description is the most important and initial information given which will drag contributors into the campaign.

Néo, le jeu des nouveaux métiers	Neo, the game of new jobs
Un jeu de cartes créatif sur les métiers et les compétences de demain!	A creative card game about the trades and skills of tomorrow!
Pour découvrir, connaître, réfléchir, créer et enrichir l'univers des nouveaux métiers!	To discover, know, reflect, create and enrich the world of new professions!
... Et ouvrir le champ des possibles !	... And open up the field of possibilities!

Table 4.1: Example of Description of a Project¹⁷

¹⁷Description from the campaign 'Neo le jeux des nouveaux métiers': <https://www.kisskissbankbank.com/fr/projects/neo-le-jeu-des-nouveaux-metiers/tabs/news>

Since the topic of this study is the messages delivered during the campaign, it seemed clearer to not mix up this initial description of a project and news. Especially because a description could be changed (slightly) by the project holder (e.g. by adding rewards) without sending a notification or a news to the donators.

The description usually contains the information regarding the goal, the deadline, the rewards and an extended summary of the project. Its content, composition and presentation could vary strongly from one project holder to another and eventually miss these information as we can see in the table 4.1.

4.1.2 Persuasive

The persuasive class groups all messages with direct asks from the project holder towards the potential or actual contributors. These asks could be directly or indirectly linked to the objective either by inciting to give money or by asking to diffuse further the campaign to gain more donators. These messages would be classified as "emotional" ones by Zhang and Gelb (1996).

In order to ease the classification we decided that only direct asks would be counted. Of course there is a possibility of implied asks especially since the format of the messages sent during the campaign has no guidelines but accepting those behaviours would prove the classification to be much harder. This is a limitation in our work that will be discussed later with leads to resolution.

75% grâce à vous !	75% thanks to you !
Bonjour à tous les nouveaux soutiens !	Hello to all the new supporters!
75% grâce à vous !	75% thanks to you!
Nous espérons que très bientôt notre jeu va prendre le chemin de la fabrication ! Nous vous mobilisons encore pour atteindre dès que possible notre premier palier, voire même le dépasser !	We hope that very soon our game will go into production! We are still mobilizing you to reach or even exceed our first level as soon as possible! Share and talk about it!
Partagez et parlez en autour de vous !	Following the game of new jobs on our social networks will help you in this process! The game is available on Twitter, Linkedin, Instagram and Facebook!
Suivre le jeu des nouveaux métiers sur nos réseaux sociaux vous accompagnera dans cette démarche ! Le jeu est présent sur Twitter, Linkedin, Instagram et Facebook !	We're going to make it!
Nous allons y arriver !	Thank you! Thank you! Thank you! Thank you! Thank you! Thank you!
Merci ! Merci ! Merci !	

Table 4.2: Example of Persuasive News¹⁸

¹⁸Example of Persuasive news from the campaign 'Neo le jeux des nouveaux métiers': <https://www.kisskissbankbank.com/fr/projects/neo-le-jeu-des-nouveaux-metiers/tabs/news>

4.1.3 Informative

The informative class is a grouping of every news with information regarding the campaign or the field of it. Those messages can be seen as a mean to keep the donators implicated and aware of the project without necessary asking anything more from them. These messages would be qualified as "rational" ones by Zhang and Gelb (1996).

These actualities can also contain information relative to the reward system or the state of achievement of the project. Of course this kind of message could trigger a will to contribute and were most certainly made with that intention but it clearly contains information that can be helpful for potential contributors to form their willingness to donate, we decided to classify it as informative.

Qui sait...?;)	Who knows...?;)
Et oui, sur le CD vous pourrez retrouver toutes sortes de plaintes, chansons à danser, gavottes...! Peut-être bien qu'il y aura cette chanson... Et en entière https://soundcloud.com/jkernomclerivet/cetait-un-petit-navire-jkerno-mclerivet	And yes, on the CD you can find all kinds of complaints, dancing songs, gavottes...! Maybe there will be this song... And in full! ;) https://soundcloud.com/jkernomclerivet/cetait-un-petit-navire-jkerno-mclerivet

Table 4.3: Example of Informative News¹⁹

A news could also be related to the campaign from further by giving information related to the field of the project and not the project itself. It keeps the perks of a classic informative actuality with a possible gain of interest or of contributors by strengthening the sense of community belonging.

Le jour où j'ai découvert que Jane Fonda était brune, partage une information avec vous	The day I discovered that Jane Fonda was brunette, share information with you
Une exposition en hommage à Simone Veil se tient à Reims! Jusqu'au 2 juillet . Courrez-y!	An exhibition in homage to Simone Veil is held in Reims! Until July 2. Run! Run!

Table 4.4: Example of Informative News Regarding the Field²⁰

¹⁹Informative news from the campaign 'Nouveau CD : Jérémy Kerno & Marc Clériveret - A haute voix': <https://www.kisskissbankbank.com/fr/projects/jeremy-kerno-marc-clerivet-premier-album-jeremy-kerno-marc-clerivet/tabs/news>

²⁰Informative news from the campaign 'Le jour où j'ai découvert que Jane Fonda était brune': <https://www.kisskissbankbank.com/fr/projects/le-jour-ou-j-ai-decouvert-que-jane-fonda-etait-brune-3/tabs/news>

4.1.4 Out of Time Scope (post-campaign)

News published after the deadline was reached would not have any interest either for us since it could no longer influence the final amount collected anymore. A class for this kind of message was created.

As we can see below, this kind of news is the opportunity to make a statement (in the example a summary of the campaign and thanks) but also to remind the donators of the product to come and to reinforce the community. It would be out of the scope of this work but this could be seen as a way to reinforce the image of the project holder for future campaigns.

Merci encore et encore	Thank you very much
Vous avez été 300 à nous soutenir, et nous avons atteint grâce à vous la somme de 18 285€, c'est énorme! Génial! Je suis immensément touchée par votre générosité et implication, merci du fond du cœur!!! Évidemment cela m'engage d'autant plus à faire un beau film pour vous toutes qui l'attendez. Pour que nos regards fassent langue, et que l'histoire des femmes continuent de s'écrire. Je me remets au travail, portée par tous vos messages d'encouragement et me sens plus forte grâce à votre soutien. Nous allons donc pouvoir terminer le montage du film, et déposer les prochaines aides pour les finitions dès octobre. Nous vous donnerons des nouvelles régulièrement. Je me rends aussi compte que j'ai maintenant un public, que je vous dois beaucoup, et que j'ai terriblement hâte de vous présenter ce film que vous avez participé à faire naître. Encore merci, vive les femmes, vive nos luttes ! Bisous	You were 300 to support us, and we reached thanks to you the sum of 18 285€, it's enormous! Awesome! I am immensely touched by your generosity and involvement, thank you from the bottom of my heart! Obviously this commits me all the more to making a beautiful film for all of you who are waiting for it. So that our eyes can speak with one another, and so that women's history can continue to be written. I am getting back to work, carried by all your messages of encouragement and feel stronger thanks to your support. We will therefore be able to complete the editing of the film, and submit the next aids for finishing as early as October. We will keep you informed regularly. I also realize that I now have an audience, that I owe you a lot, and that I am terribly looking forward to presenting you this film that you helped to create. Thank you again, long live women, long live our struggles! Kisses

Table 4.5: Example of Out of Time Scope News²¹

²¹After deadline news from the campaign 'Le jour où j'ai découvert que Jane Fonda était brune': <https://www.kisskissbankbank.com/fr/projects/le-jour-ou-j-ai-decouvert-que-jane-fonda-etait-brune-3/tabs/news>

4.2 Modelisation

The following model is greatly inspired from the one used by Ducarroz, Yang and Greenleaf (2016) and most of its components and arguments are kept. We transposed it to the crowdfunding issue with the needed adaptations, as the context is different. The main changes is that every donation is regarded as the entry of a new donator and the use of contributions instead of bids which do not send the same competitive message instead of bids. The reasons and explanations will be given in the sections regarding the limitations of this model.

The model is a direct adaptation and may not be as suited as the original was for the system of auctions. Some leads will be mentioned in a later section for a later work related to the subject. It should allow to refine the model and obtain more robust results.

4.2.1 Conceptual Framework

First, the framework works on two axes, which will lead to two level of analysis:

- Single interaction actions and their responses that can be multiple
- The final result out of these single promotional actions and responses

To do so, it includes both the impact of in-process messages on contributors and a speculation of the strategy the project holder uses to issue these messages. The latter is included because through the chosen, project holders that issue news repeatedly during a campaign can easily observe and analyze contributors responses and incorporate the conclusions of this analysis into later news for the same project.

Moreover, the framework includes different kinds of in-process messages as seen in the previous section, because project holders may use different kind of messages as stimuli to the potential crowdfunders and these may differ in their impact and in the strategy for issuing them.

Finally, in-process news can affect not only contributors who have already donated and can donate again, but also those who are still watching the event and have not yet decided on whether to contribute or not but may do so in the future.

Campaign Setting

Campaign events and their messages are watched both by participants who already donated (contributors) and by interested potential contributors who have not yet donated but might do so in the future ("latent contributors" keeping the terminology of the previous model).

We consider a setting in which a project holder is trying to obtain the funding for a project and accept every amount of donation. Any amount higher than the minimal donation of the campaign would be considered a "jump-donation" (to stay close to the terminology of the previous model) and the result of a desire to contribute more to the project.

Due to the imprecision of the events' time, while we will be discussing about donation interval, we actually will not be able to differentiate two donations occurring the same day. So the interval will be between donations with the same timestamp.

Campaign-Level and Donation-Level Focus

Our framework, as we can see in Figure 4.1, studies each in-process message at two levels: an aggregate, campaign-level, one and a "micro", donation-level, one.

On the one hand, we study how the final collected amount of the campaign is influenced by the total number of each kind of message issued during the campaign. Here, our analysis is at the “aggregate” level, in which the unit analyzed and observed is each amount collected for each campaign. The final amount collected during a campaign is examined by the model. Contributors and project holders decisions can be summed within each campaign as the total number of donators, and as the total number of each kind of message. The final amount is the campaign outcome.

The right side of the figure shows the campaign-level in our framework: the final campaign outcomes that the model further examines. Donator and project holder decisions can be summed within each campaign as the total number of donators and jump-donations, and as the total number of each kind of message. The final amount collected is also a campaign outcome. We discuss each element of this framework and eventual relationships among them and potential psychological mechanisms underlying them (represented by dotted boxes in Figure 4.1), when we develop hypotheses.

On the other hand, our “micro” level examination will study how in-process messages can affect new donators entries, as well as how project holders decisions on whether to issue a message depend on information on past campaign activity. Here, the unit of observation and analysis is each donation or donations per unit of time, and each messages, since we are working with dates and lack the precision of hours, we will establish that each message is sent at the start of the day and the donations of the same day are following it. These signs of interest observed in donators behaviors are also of interest because their collective impact in turn affects the final amount collected during a campaign.

The left side of the figure shows the donations-level (micro-level) in our framework: it presents the campaign at the level of donations and in-process messages. It shows that campaign activities prior to a particular donation interval (between donation $j - 1$ and donation j) affects project holder decisions in this donation interval, which are followed by donators decisions (both potential and already contributors). Once another donation is made, the sequence repeats, until no further donations are perceived and the campaign ends.

Campaign Participants and Their Decisions

We analyze the decisions of two sets of participants in the campaign —contributors and the project holder— and how they interact. At each point in the campaign, each contributor or latent one can observe the campaign state through earlier donations and messages and decide whether to make a donation and if a donation is made, whether to do a jump-donation. On the project holder side, at each donation interval, decide whether to issue each kind of message.

Messages Types in Campaigns

As stated above, more than one type of in-process messages can be used during a same campaign. Here, we draw a distinction between two different kinds of in-campaign messages, those were described earlier in the previous section 4.1: informative and persuasive. The messages otherwise labelled will not be included in the data submitted to the model, those labels served to exclude irrelevant data.

The description will still intervene in the econometric model since it would still have an impact on the amount donated and the appeal of the campaign

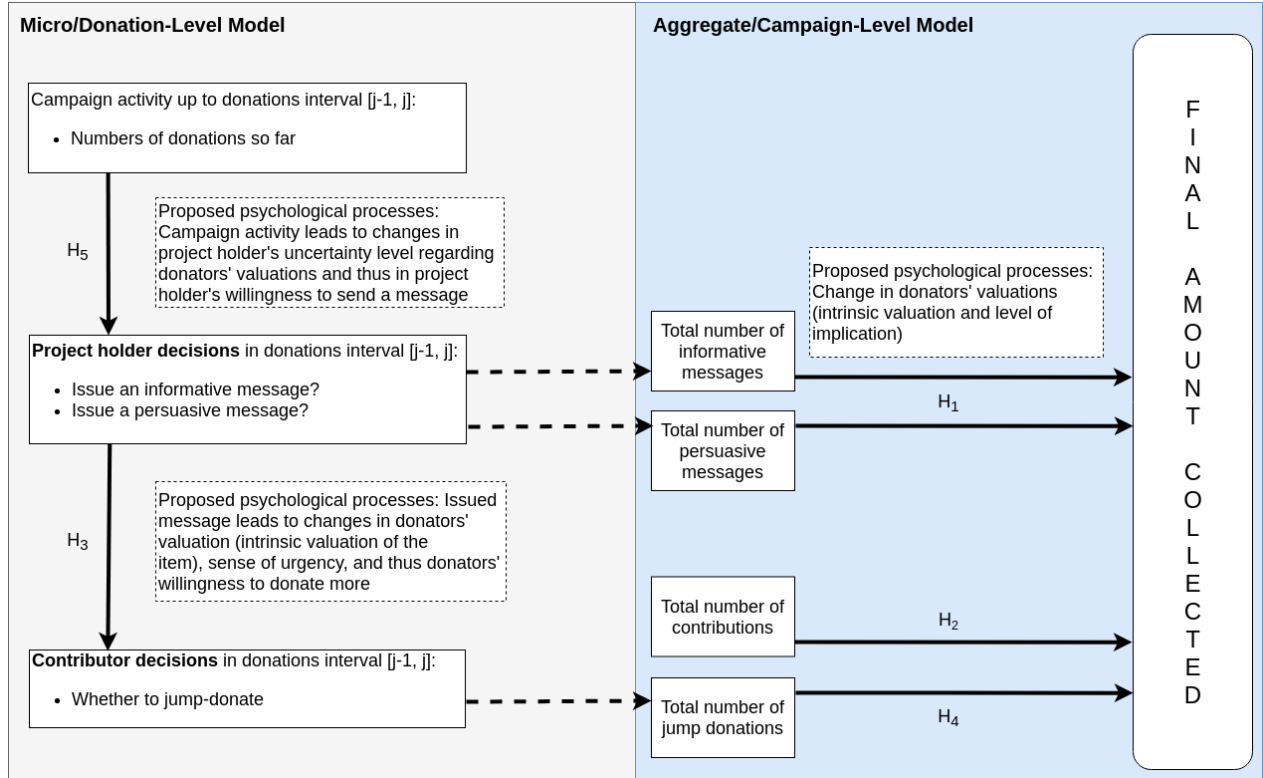


Figure 4.1: The Conceptual Framework

4.2.2 Hypotheses

Aggregate Level: Impact of Total Number of Each Kind of Message on the Total Amount Collected

In a campaign, the final amount collected is determined by the donations and represents the value that contributors give to the project.

Donator valuations of the project are unstable and subject to variations during the campaign, since the preferences of latent donators may not be well-defined but constructed “on the spot”. Thus, as news and more information become available during the campaign, these valuations are updated. The variations are supposedly due to the impact of new value indicators which will impact the preference. We predict that *informative* messages are likely to increase the valuations of the donators because they can provide new information about campaign attributes and levels of these attributes. This would then reduce the uncertainty, for potential donators, on whether or not to contribute to the campaign.

We also predict that persuasive messages have a positive effect on the valuations of donators and the final amount collected, firstly by encouraging donators to donate. The perspective of reaching an objective can increase their willingness to pay. Those messages can lead to an appropriation of the project and an increased will to contribute and realize it. The result could be a donation, but could also be actions leading to more coverage and in the end donators which would still influence positively the amount collected for the campaign.

In summary:

H1a: The final amount collected in a crowdfunding campaign increases as the number of informative messages issued during this campaign increases.

H1b: The final amount collected in a crowdfunding campaign increases as the number of persuasive messages issued during this campaign increases.

Aggregate Level: Impact of Total Number of Donators on the Total Amount Collected

We also expect that the increase of the number of donations increases the expected amount collected. Since more contributions means more money and potentially more contributors and more coverage, we propose that the number of donations has a positive impact on the final amount collected. We will check whether this relationship occurs in our data set, creating the following hypothesis:

H2: The final amount collected increases as the number of donations increases.

Micro Level: Impact of Messages on the Amount Given via Contributor Decisions at the Donation Level

We propose that in-process messages affect the contributor behavior, whether to make a jump-donation, so we develop corresponding hypotheses. We then expect that this donator behavior, when summed at the campaign level, affect the achievement ratio of a campaign.

We propose that in-process messages encourage jump-donations for two reasons. Their desire to donate more than just the minimum required to contribute to the project can be due to the attractiveness of a reward only available to an higher donation, this attractiveness would clearly depend of the donators valuation on the rewards. Since informative and persuasive messages can provide more information regarding the rewards accessible to the donators, we propose that they both encourage jump-donations.

The possibility of missing their chance to fund the project or of losing the expected reward can also encourage jump-donations. It is due to the fact that if the project is not entirely funded they will not be compensated and the project will not be funded, even partially. A *persuasive* message accentuating and highlighting those possibilities will thus give donators a reason to eventually contribute more to be sure of the campaign achievement. However, informative messages may not create this donating urgency since the change made by urgency is not operated on the valuation of the project. Jump-donations probabilities may then increase upon reception of a persuasive message but remain unchanged for informative messages. Therefore, we predict that whereas both types of messages increase the likelihood of jump bidding, persuasive messages have a larger overall impact on jump bidding than informative ones. In summary:

- H3a:* An informative message issued in the previous donation interval increases the likelihood of the current donation being a jump-donation.
- H3b:* A persuasive message issued in the previous donation interval increases the likelihood of the current donation being a jump-donation.
- H3c:* A persuasive message issued in the previous donation interval creates a larger increase in the likelihood of the current donation being a jump-donation than does an informative message.

• Impact of the aggregate outcomes of donation-level donators decisions

Second, to evaluate the impact of messages on the final amount collected through jump-donations, we examine how an increased tendency to jump-donations because of messages, as hypothesized previously, affects the final amount. It seems logical that the higher the donation, the higher the final amount collected would grow. Then we propose that in-process messages affect the final amount by affecting the jump-donations.

H4: The final amount collected increases as the number of jump-donations increases.

Thus, if H3a and/or H3b, along with H4, are supported, in-process messages can also increase final amounts collected, acting through the donation-level mechanism of affecting jump-donating.

The Decision of The Project Holder on Whether to Send a Message During Each donation interval

As discussed earlier, an analysis of in-process messages also requires to understand what motivates project holders to issue these messages during an event. So our conceptual framework also focuses on the strategy of the project holder for issuing messages. The proposition is that the emitter makes dynamic decisions about message sending during the campaign, which are influenced by donators' behaviors. Just as donating behavior can change during the campaign, the propensity to issue a message can also change as the perceptions of the contributors and the achievement of the campaign change.

We regard the donating process as an interaction between the project holders and contributors (and latent ones). The project holder communicates with these parties by issuing messages, and contributors answers by showing their implication through a donation. We propose that the emitter believes that messages can influence the donators, which, in turn, influences the final collected amount.

In order to devise a message-issuing strategy, the project holder needs information on the contributors' valuations, among other campaign variables. Given that the project holder does not have a priori information regarding the valuations of the donators. This information needs to be collected through observing their donations. The emitter is therefore uncertain about the likelihood that individual donators, and donators collectively, will continue to enter the campaign. The project holder can use messages to reduce this uncertainty which can also affect his future message issuing because responses of contributors to messages helps the project leader update his information about donators' valuations.

One indicator could be obtained by the project holder by monitoring how long it takes to get a donations after issuing a message. Thus, we propose that as the uncertainty of the project holder about the valuations of the donators increases, the emitter becomes more likely to issue messages to help resolve that uncertainty.

We first propose that this uncertainty decreases as the number of contributions grows. If a campaign is successful with many donations, it is very likely that many donators have high valuations and are vehemently promoting the project, making the emitter less likely to issue an in-process message. However, on the contrary latent contributors may not have high enough valuations. Or not feeling the emergency or need to contribute. This absence of donations would lead to emitter's uncertainty and make the emitter more likely to issue a message.

In summary:

H5a: As the number of donations increases during the campaign, the project holder becomes less likely to issue an informative message.

H5b: As the number of donations increases during the campaign, the project holder less likely to issue a persuasive message.

4.2.3 Econometric Modeling

Aggregate Level: Impact of Total Number of Each Kind of Message on Amount Collected

The following regression model of the final amount collected is proposed at the campaign level:

$$(1) \quad Amount_i = \beta_{P,0} + \beta_{P,1} \ln(NInfo_i + 1) + \beta_{P,2} \ln(NPers_i + 1) + \beta_{P,3} NDon_i + \beta_{P,4} NJump_i + u_{P,i}$$

For the i th campaign, $NInfo$ and $NPers$ are the number of informative and persuasive messages sent, respectively; $NDon$ is the total number of donations and $NJump$ the number of jump-donations.

The main objective of the equation above is to test H_{1a} and H_{1b} , regarding the aggregate impact of the two kinds of messages on the final amount collected. However, the equation also includes the number of donations and of jump-donations in order to test H_2 and H_4 , the necessary conditions to establish the impact of messages on final amounts via affecting jump donating.

If the four independent variables of main interest in (1) are correlated with the error term $u_{P,i}$, this relationship could lead to endogeneity and to biased estimations. Endogeneity might occur, for example, if donators and the project holder observe factors that affect amounts donated, such as available alternative (concurrent project with same content, ...) that are not observable to the researcher. Even with the knowledge of this fact, this issue was not addressed. Lack of documentation on the characteristics of a campaigns' quantifiable success factors was a brake for the model regarding that aspect, we used the empirical data collected for each campaign for the variables. We report definitions and summary statistics of the variables used in our model in Table 4.6. Potentials instrumental variables or indicators will be mentioned in section 7.2.1 as potential improvements to the model.

Micro Level: Impact of Messages on Given Amount via Contributor Decisions at the Donation Level

• Impact of in-process messages on jump-donation at the donation level on the final amount collected

Let $Y_{JUMP,ij} = 1$ (or $= 0$) if the j th donation in the i th campaign is (is not) a jump-donation. The probability of a jump-donation at donation j can be expressed using the logistic normal as:

$$(2) \quad Prob(Y_{JUMP,ij}) = \frac{\exp(Y_{JUMP,ij}^*)}{1 + \exp(Y_{JUMP,ij}^*)}$$

Where $Y_{JUMP,ij}^*$ is the latent propensity for the j th donation to be a jump donation:

$$(3) \quad Y_{JUMP,i,j}^* = \beta_{JUMP,i0} + \beta_{JUMP,i,j} Info_{i,j-1} + \beta_{JUMP,i,j} Pers_{i,j-1} \\ + \beta_{JUMP,i,j} NDon_{i,j-1} + \epsilon_{JUMP,i,j}$$

The subscript $j - 1$ denotes donating activity at donation $j - 1$ and also denotes message activity in donation interval $[j - 1, j]$. Among the covariates used are whether an informative ($Info_{j-1}$) or a persuasive ($Pers_{j-1}$) message was issued in donations interval $[j-1, j]$. These variables test the impact of messages on jump donating (H3)²².

The Decision of The Project Holder on Whether to Send a Message During Each donation interval

The indicator variable $Y_{INFO,i,j}$ ($Y_{PERS,i,j}$) takes a value of 1 if the campaign holder sent an informative (persuasive) message in donation interval $[j - 1, j]$ and a value of 0 if no such message was sent. The probability that an informative (persuasive) message is issued within the donation interval is assumed to follow the following logistic normal regression specification.

$$(4) \quad Prob(Y_{INFO,i,j}) = \frac{\exp(Y_{INFO,i,j}^*)}{1 + \exp(Y_{INFO,i,j}^*)}$$

$$(5) \quad Prob(Y_{PERS,i,j}) = \frac{\exp(Y_{PERS,i,j}^*)}{1 + \exp(Y_{PERS,i,j}^*)}$$

Where $Y_{INFO,i,j}^*$ ($Y_{PERS,i,j}^*$) is the project holder's unobserved propensity to an informative (persuasive) message.

$$(6) \quad Y_{INFO,i,j}^* = \beta_{INFO,i0} + \beta_{INFO,i,j} Info_{i,j-1} + \beta_{INFO,i,j} Pers_{i,j-1} \\ + \beta_{INFO,i,j} NDon_{i,j-1} + \epsilon_{INFO,i,j}$$

$$(7) \quad Y_{PERS,i,j}^* = \beta_{PERS,i0} + \beta_{PERS,i,j} Info_{i,j-1} + \beta_{PERS,i,j} Pers_{i,j-1} \\ + \beta_{PERS,i,j} NDon_{i,j-1} + \epsilon_{PERS,i,j}$$

(6) and (7) use the same covariates as the jump-donating model. Among these, is $NDon_{j-1}$, the number of donations so far in the campaign. This variable test how donating activity affects the project holder's message-sending decisions (H5)²³.

²²Since donators are not competing with each other as in auctions, a message could have an impact on many of the following donations without the interventions of other campaign actors diminishing its effect. Working in campaign days or in the interval between messages would make it possible to maintain the impact of a message on several donations.

²³Another possible variable would be the ratio of achievement of the campaign which would give the urgency felt by the project holder regarding the achievement of his project. It could be seen as an equivalent of the PercentMktPrice variable of the reference article model. It could also be related to the time left before the end of the campaign.

Variable Name	Variable Definition	Mean	SD	Min	Max
Campaign Level					
Amount (DV)	Final amount collected, in euros	13627.86	28244.426	3037	367172
NInfo (AX1)	Number of intervals in which the auctioneer sends an informative message	11.34	6.533	0	32
NPers (AX2)	Number of intervals in which the auctioneer sends a persuasive message	3.82	4.133	0	22
NDon (AX3)	Total number of donations in the campaign	109.40	97.210	12	705
NJump (AX4)	Total number of jump donations in the campaign	91.05	84.343	4	604
Donation Level					
Y_{JUMP} (DV)	Whether donation j is a jump donation ^a	.83	.374	0	1
Y_{INFO} (DV)	Whether an informative message is sent in current donations <i>interval</i> ^a	.36	.48	0	1
Y_{PERS} (DV)	Whether a persuasive message is sent in current donations <i>interval</i> ^a	.13	.34	0	1
$Info_{j-1}$ (MX1)	Whether an informative message was sent in previous donations <i>interval</i> ^a	.32	.465	0	1
$Pers_{j-1}$ (MX2)	Whether a persuasive message was sent in previous donations <i>interval</i> ^a	.14	.343	0	1
NDon (MX3)	Number of donations so far in the campaign (at donation $j-1$)	98.17	1.5.136	1	705

Table 4.6: Variable Definitions and Summary Statistics of the Data²⁴

²⁴Notes: Data are based on the 21834 donations and the 4893 messages from the 200 campaigns used in the model estimation. DV = dependent variable; AX = aggregate-level variable; MX = micro-level variable.
^a Binary variable.

Chapter 5

Data Extraction

5.1 Tools

5.1.1 Basics of the Web

Before explaining how the data were extracted from KKBB website, it is necessary to have a better grasp of the composition of a webpage. Web programming mostly relies on three languages to design a website when it comes to the composition of a webpage: *HTML*, *CSS* and *JavaScript*. Some substitutes exist but these are the most basic ones and are used on the rendering of the KKBB website pages.

HTML

The Hyper Text Markup Language (in short HTML) is used to define the content of a page and create it. HTML consists of a series of short codes typed into a text-file by the site author, the tags. Later, the browser used (for example Google Chrome) will interpret the tags and render a page containing all the information present in the file.

The tags also allow to organize the content, define sections and add specific rendering effects in the page through a hierarchy. We can for example find in the code of a sentence other tags specific to the words. In the following case, we will format a word in italic.

```
<!DOCTYPE html>
<html>
  <head>
    <title>Page Title</title>
  </head>
  <body>
    <p>This is a <i>HTML</i> text generated with tags.</p>
  </body>
</html>
```

Code Listing 5.1: HTML Code Demonstrating the Inclusion of Tags

²⁵Most of the following information and design of the tools is related to my background as teaching assistant at the polytechnic school of Louvain and my master degree in computer sciences which explains the lack of references.

In the example above, we are defining a sentence and applying a format effect on a word which is part of it. Note that other tags are needed and mandatory to compose a page.

Complete webpage needs far more code as we can see in the code for the button "Start my project" on the home page of the KKBB website.

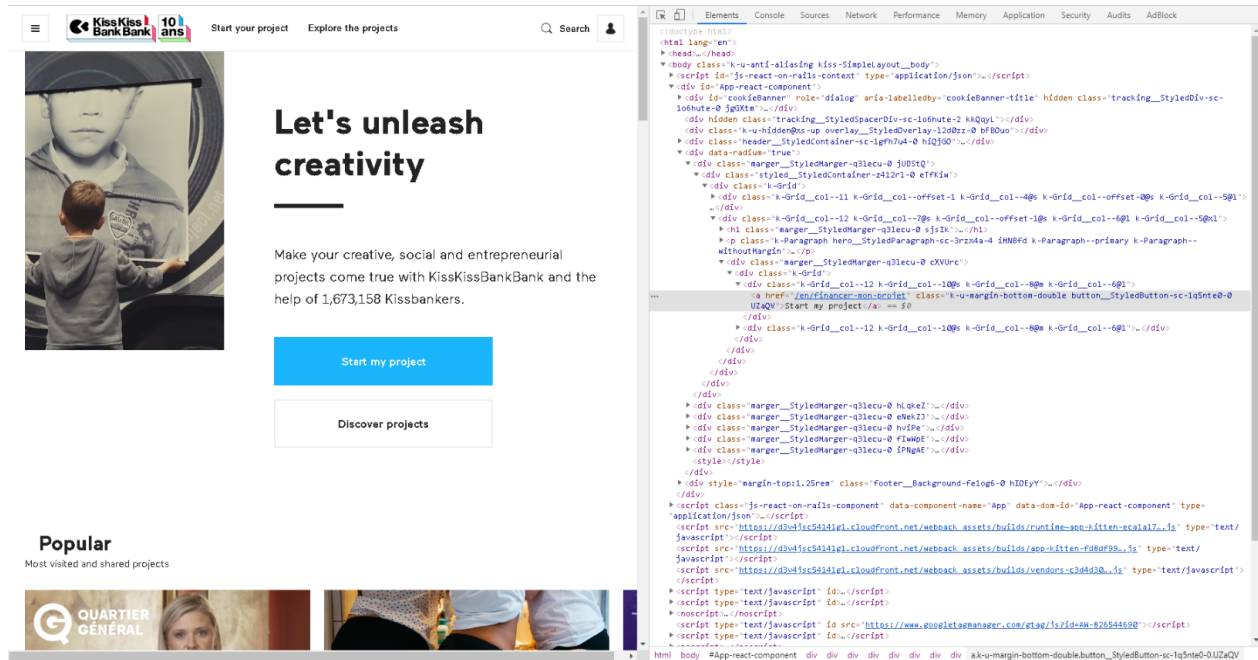


Figure 5.1: HTML Code for one Button²⁶

CSS

The Cascading Style Sheets (in short CSS) is the standard way to define the presentation of an HTML page. It allows to define the fonts, colors and layout of the elements of a page. Its main advantage is to allow to reduce the number of tags needed in the HTML code. This reduction in code also allows to download page much faster due to the reduction in the size of the files.

It is possible through the definition of all the formatting tags as classes in one file. All pages would check this file for formatting before rendering. It also allows to modify all the pages of a website by simply modifying this file and keep the whole site consistent regarding its look.

The CSS was important to differentiate some of the scrapped elements since most of them did not even have an id to facilitate the recovery of the data. But this technology did not have much more impact on the scraping so we will not discuss it further.

²⁶Inspection of the button on the homepage: <https://www.kisskissbankbank.com>

JavaScript

JavaScript (in short JS) is a language made to add all sorts of dynamic elements to a website. Without the use of it, HTML would only be able to output static pages. Once these are loaded, the only way to change the seen content would be by clicking on a link that would change the displayed page. With JS it is possible to add dynamic and interactive elements to websites.

JavaScript is a client-side scripting language, which means the source code is processed by the client's web browser rather than on the web server. This means JavaScript functions can run after a webpage has loaded without communicating with the server. For example, a JavaScript function may check a web form before it is submitted to make sure all the required fields have been filled out. The JavaScript code can produce an error message before any information is actually transmitted to the server (techterms, 2019).

JS scripts are integrated into the browser, which means they have access to the information about the browser and HTML page, and can edit this information, thus changing how elements are displayed on the screen. This access to information grants JS great power to modify the browsing experience. They can also react to events, like a click on a button, or points to a certain page element (yourhtmlsource, 2019).

This last part was essential in the scraping of the KKBB website and will be further discussed later but the JS scripts integrated in the pages were an important challenge to the extraction of the data.

5.1.2 Scraping

Web scraping is a technique of data scraping used for extracting content from websites, via a script or program, in order to transform it to allow its use in another context. It is done through the *Hypertext Transfer protocol* (in short HTTP) or a web browser.

This technique uses both fetching and extracting from web pages.

- Fetching is the downloading of a page (which is done each time a page is displayed in a browser). It is done through *web crawling*.
- Extracting is done by parsing, searching and copying the data.

There is plenty of methods to execute these tasks but we will focus on the HTML parsing one, since it is the one used in our work. KKBB website has a large collection of pages generated dynamically from an underlying database and data fitting in the same category are typically encoded into similar pages by a common template or script. In *data mining*, a program that detects such templates from a particular information source, will extract its content and translate it in something more processable. This is basically what we have been doing to extract the data from KKBB pages.

It is also important to note that web scraping is a grey area ethically speaking. Most of it is used to reuse information gained from other websites. This usually leads to perform referencing, which consists in improving the positioning and visibility of sites in search engine or directory results pages by editing the content of the website. This practice does not have very good reputation among authentic contributors because it can be likened to looting. To encourage webmasters not to use a web scraping method, Google sanctions sites that use them by removing them from its results pages.

On the other hand, in a company context, which is of course the owner of its data, this technology significantly reduces the time and cost of integrating enterprise applications. Although this type of integration has been denigrated in the past for its lack of reliability and performance, there are now professional tools on the market that allow this integration while respecting the security and governance constraints imposed by large companies.

Our work only collects data from the public domain and for non-commercial purposes. All these are directly available online and can be collected by hand. More information regarding the security issues of scraping are provided in the appendix A.1.

The scraping work of the campaigns present on KKBB website was divided in four subtasks for the sake of simplicity and related to the fact that four types of pages were to be scraped.

- Scraping of campaign url addresses
- Scraping of campaign basic descriptions
- Scraping of campaign news
- Scraping of campaign donations

HTML Scraping

Most of the data extraction was made using *Python* and its web scraping referential module *BeautifulSoup*. To do so, a HTTP request was sent to the server of KKBB to obtain the response that would be provided to a browser for a webpage referenced by an url address. This response contained the HTML content associated with the web page.

The extraction was then realised. We started by parsing the HTML code then we proceeded to execute a sequence of searches. These searches were executed on specific selectors with particular attributes. This part was the most tedious as it needed to analyze beforehand the whole structure of the templates. This analyze was made by inspecting the pages on KKBB website with the inspect function of the *Google Chrome Developer Tools*, referencing the useful selectors and attributes (with possibly specific values) combinations.

For example, to scrap the aimed amount of a campaign we would first need to obtain a grouping HTML structure by searching the following combination of selector (*div*), attribute (*class*) and value:

<div	class=	"informations-and-media__StyledInfo-sc-121rta3-3 fzHbGm marger__StyledMarger-q3lecu-0 hAgZUM">
selector	attributes	values

Table 5.1: HTML Identification of an Element

We would then collect all the sub *div* and retrieve the information from the second one present in this sub-selection. This data would be the aimed amount.

Note that the HTML source extracted did not provide all the information needed since some data were generated dynamically. The information regarding the process of extraction is provided in the appendix A.2.

Special cases

While running the scraping, we ran into outliers data that needed to be ruled out of our sample. It was taken into account by defining the expected value domain and format for an element. If an element did not fit those criteria, the campaign would not be taken into account and the data related to it would not be extracted. A few examples are: presales for a concert, handywork, participations by attending or enrollment, ...

These non financial demands could of course be accounted and translated to a monetary evaluation of the donation. Since this would not be done easily with automation and we had a large enough sample, we just did not take those campaigns.

5.1.3 Classifier

In the model presentation (section 4) we presented the way we define the types of messages that the project holder was able to convey to contributors. In order to incorporate this information in the current model and associate it with the message, we made an automatic classifier that would classify a message sent by the project holder in one of our categories.

To do so, a text mining analysis of the content and the title of a message was made and then based on the result a category would be attributed to an actuality. Of course during this analysis some marker words were looked up, upon their presence an automatic attribution would be given. But we went further by integrating the detection of patterns into the sentences. It is important to note that these words and patterns were decided based on a sample and there is still room for improvements in this automatic classification. Moreover, some solutions can also be changed by hand, which will be the case.

It has been proven that the manipulation of words enables more flexibility while interpreting a sentence. *Stemming* consists in reducing inflected (or sometimes derived) words to their word stem form (e.g.: “contribution”, “contributed”, “contribute” reduce to the stem “contribut”). Usually close words reduce to the same stem (INGI2263 Computational Linguistic, 2017). Following this idea, the pattern used were all searched in a stemmed version of our texts.

Besides the subtler technique of stemming, a first screening was made during the classification by outcasting some message in specific class made to store the *out of scope* message (e.g.: messages sent after the end of the campaign). Typically those message would not have had any impact during the crowdfunding process and it was easier to rid them out of the sample by a rule-based approach. A few rules were determined beforehand regarding the domain of attributes.

²⁶A more complex and possible approach was *lemmatisation*. This process involves first determining the part of speech (a category to which a word is assigned in accordance with its syntactic functions) of a word, and apply different normalization rules for each part of speech. First, the part of speech is identified prior to attempting to find the root because for some languages the stemming rules change depending on the part of speech of a word. After this identification, we apply the stemming rules. Since this approach is highly conditional upon obtaining the correct lexical category and was more time costly, it was not chosen.

5.1.4 Implementation Choices

Data Representation

Since it would be easier to treat all the incoming text data upon the creation of an object, it was decided to use the objects used in *Python*. The object also allows to include additional values that would make the rule base classification easier (e.g. deadline reached, aimed amount obtained, ...). It would also ease the relational links between objects such as news and campaigns or donations and campaigns or eventually news and donations.

As seen below in the figure 5.2, the data were stored in three types of data structures: campaign, donation and news. Campaigns contains the data regarding the url address of the project, the title, the description, the name of the project holder, the initially aimed amount, the amount actually obtained and the ending date of the campaign. All these information are available on the webpage of the project but it also contains flag values such as "achieved" or "done" for the reasons stated earlier. Donations contains the information regarding the name of the contributor, the date of the donation and the amount. News contains the information related to a message sent such as the title, the content, the date and the kind which will be attributed after classification (see section 4.1).

The reason of the different data structures is simple: a campaign could possess news or not. The same reasoning can be applied to the donations (even though it is less likely to find a campaign without donation). Moreover, the donations and news can be numerous for a same campaign. Those are stored in lists, the advantages of this solution are expose in the appendix A.3.

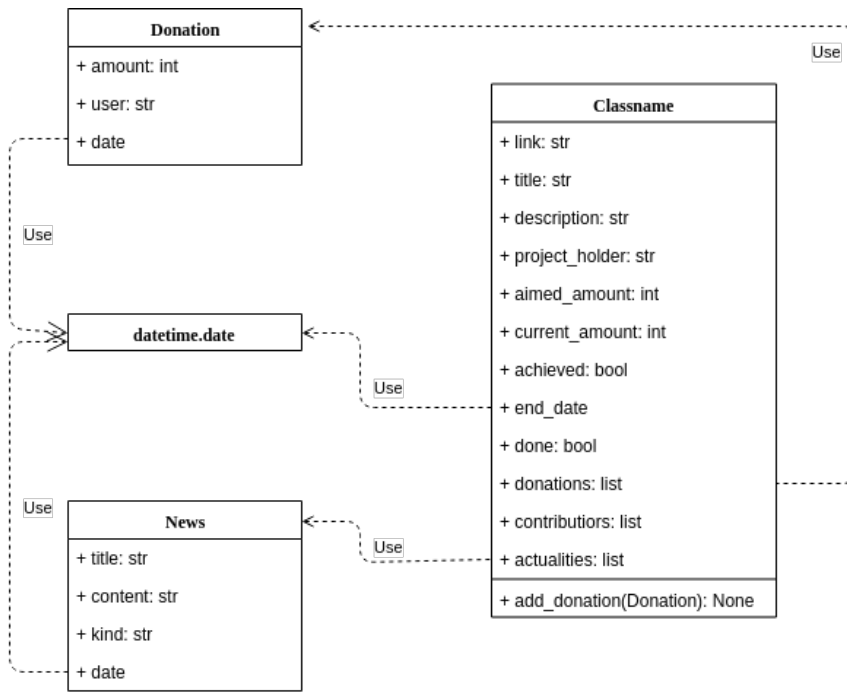


Figure 5.2: UML Class Diagram of the Campaigns Objects

5.2 Data Extraction Approach

In the table 5.2 we can see the steps used to generate and extract a valid sample from the platform *KKBB*. After the extraction, it was possible to work on the data and apply to model to test our hypotheses.

Objectives	Approach
Establishing relevant sample	• Define the needed extraction tools and design of the scripts. • Obtain every campaign URL from the website. • Scrapping over the campaigns web pages to obtain the projects basic information, donations and news. • Skimming²⁷of the data to keep successful campaigns with the most messages.
Messages classification	• Define the characteristics of the messages classes. • Exclude not fitting in messages an out-of-category type. • Classification of the messages of the campaigns with the pre-defined types. • Validation of the classification by an expert, the Pr. P. Scarmure.

Table 5.2: Data Extraction Steps Approach

²⁷The skimming process shall:

- Remove the atypical campaigns which would not fit in the sample due to anomalies in their attributes
- Reduce the number of campaigns to an easily processable and relevant sample

Chapter 6

Empirical Results

6.1 Aggregate Level

In order to evaluate the impact of the four factors of influence chosen for crowdfunding campaigns on the global amount collected, we performed a linear regression on the SPSS software. Among the four identified factors, two significantly explain the amount collected but this is not visible in the model as it is.

As we can see in the Table 6.1, our model is significant even though its variables do not seem so. To better analyze the impact of our variables through the deepening of their impact, we made two restricted models: a first one which includes the result of the decisions of the project holders, and a second one based on the result of the decisions of the contributors.

Model	Variable	β	Sig.
Amount	<i>NInfo</i>	-.026	.683
	<i>NPers</i>	.007	.912
	<i>NDon</i>	.886	.106
	<i>NJump</i>	-.374	.494
Model		Fisher	Sig.
Amount	Regression	17.193	.000

Table 6.1: Regression Aggregate Level Results

Aggregate Impact of Messages on Collected Amount

The estimate for the impact of messages during a campaign are reported in Table 6.2. Results indicate that the impact of informative messages on the final amount collected was positive but not significant, not supporting H1a. The persuasive messages impacted negatively and were also not significant and then not supporting H1b.

An explanation could be that contributors are mostly interested by the description of the project and then the messages should not be determining factors in terms of contributions but this would not be in line with the assumption that contributors are looking for more information to decide. At least as far as informative messages are concerned. Another possibility could be due to a bad match between messages and donations because contributors could read their notifications with a slight delay and contribute in another donation interval than the one associated with the message responsible for it.

Model	Variable	β	Sig.
Amount	<i>NInfo</i>	-.004	.742
	<i>NPers</i>	.024	.960
Model		Fisher	Sig.
Amount	Regression	.057	.945

Table 6.2: Regression Aggregate Level Results - Project Holders Decisions

Impact of Number of Donations and of Jump Donations on Collected Amount

Even though the restricted model is significant, the variables still seem to no be so as we can see in the Table 6.3. We then chose to test them individually.

The impact of the number of donations ($\beta_{Amount, NDon} = .512$) was positive and significant, supporting H2. It was also the case for the impact of the number of jump donations ($\beta_{Amount, NJump} = .504$), supporting H4. These positive coefficients suggest that during a campaign any contribution big or not is perceived as a positive influence when trying to build the amount collected.

We assume that the two variables are too correlated, resulting in the fact that they are not significant when used together. This reasoning is supported by the fact that each jump-donation is obviously a donation and that this first variable is so contained into the second one.

Model	Variable	β	Sig.
Amount	<i>NDon</i>	.886	.106
	<i>NJump</i>	-.374	.494
Model		Fisher	Sig.
Amount	Regression	34.818	.000

Table 6.3: Regression Aggregate Level Results - Contributors Decisions

6.2 Micro Level

In order to evaluate the impact of three factors of influence chosen for crowdfunding events, such as donations and messages, on the events' probability of occurrence, we performed a linear regression on the SPSS software. These variables are the number of donations so far in the campaign and whether a message has been sent in the previous donation interval. We can see the values of our model in Table 6.4.

Model	Variable	β	Sig.
JUMP	$Info_{j-1}$	.008	.231
	$Pers_{j-1}$	.015	.024
	$NDon$	.027	.000
INFO	$Info_{j-1}$	.508	.000
	$Pers_{j-1}$	-.072	.000
	$NDon$	-.068	.000
PERS	$Info_{j-1}$	-.067	.000
	$Pers_{j-1}$	.497	.000
	$NDon$	-.038	.001

Model		Fisher	Sig.
JUMP	Regression	7.612	.000
INFO	Regression	734.204	.000
PERS	Regression	612.603	.000

Table 6.4: Regression Micro Level Results

6.2.1 Contributor Decisions

Impact of in-Process Messages on Jump Donating

An estimate for the impact of previously issuing an informative message on jump donating was not significantly positive, which does not support H3a. Persuasive message do not provide the same effect since the result was significantly positive, supporting H3b. It is also interesting to note that even though both variables are not significant, their different betas ($\beta_{JUMP,Info} = .008$ and $\beta_{JUMP,Pers} = .015$) would seem to support H3c since persuasive messages are more efficient than informative messages to stimulate jump donating.

This could be explained by the fact that plenty of information is already available in the description of the project and the additional information would not have the same impact as an emotional message as described by Zhang and Gelb (1996). It would also be logical that the persuasive and emotional message would lead to more than a simple contribution. Indeed, some donators can show their greatest commitment by making a jump donation.

If H3 is supported, it induces that messages have influence over the jump donating mechanism. We established earlier that by supporting H4 the number of jump-donations has an influence on the final amount collected. So we can deduce that the messages are directly impacting the final amount collected.

6.2.2 Project Holder Decisions

Decision to Send a Message

Regarding the decision to issue messages, it is interesting to note that the number of donations so far has a significant negative impact ($\beta_{INFO, NDon} = -.068$ and $\beta_{PERS, NDon} = -.038$), supporting H5a and H5b. With H2 validated, we can associate the increase in the number of donations with a greater odds of success for the campaign due to the increase of the amount collected which would be closer to the goal. Consequently, the stress factor of the project holder in relation to the completion of his campaign would decrease, and therefore he would not feel the need to increase the number of donations as much as before. It could then reduce its actions and send fewer messages.

We also observe that if the project holder already sent a message in the previous donations interval, the project holder is less likely to send another kind of message as we can see with the negative impact of the variables $Info_{j-1}$ and $Pers_{j-1}$ respectively for PERS and INFO ($\beta_{INFO, Pers} = -.072$ and $\beta_{PERS, Info} = -.067$). Though the opposite would be true and the project holder would tend to send a message of the same type as the previously sent message again ($\beta_{INFO, Info} = -.508$ and $\beta_{PERS, Pers} = -.497$). A possible explanation for these values would be the existence of writing or communication habits of the project holder who would not consciously write his message in a particular and similar style each time.

Conclusion

Chapter 7

Discussion

7.1 Limitations

7.1.1 News Types

The messages addressed during a campaign are not as controlled as in auction where the speaker is a professional with rules on how to express himself regarding a product. Therefore the unprofessionalism of the project holder and the absence of defined regulations regarding the way to address donators creates a wide variety of news (each campaign has its own project holder instead of one auctioneer for a set of auctions). The numerous distinctions between the messages make it much harder to classify properly a news without any guidelines for the project holders.

This lack of guidelines also led to longer messages than in auctions, a campaign is lasting longer than an auction (approximately fifteen minutes) which allows to address the donators with much more information since time is not as spare. Thus the message can contain more information and eventually belong to more than one class of news, it makes the dichotomous classification even harder.

Nevertheless, we decided to keep our news classification as close as possible to the one used in auction by Ducarroz, Yang and Greenleaf (2016) but as stated earlier the news are much more depending of the way the project holder see it fit. This could lead to an impossible classification for a message. What shall we do with a message containing a survey to include a contributor in the choices related to the project for example? Is it supposed to be informative or persuasive? By default with the current classification it would classify as informative but it does seem too broad.

7.1.2 Missing Information

Firstly, as stated earlier, failed campaigns are not displayed on KKBB website. This represents a lost of valuable data since it could also give us hints about behaviours to not adopt while delivering messages during a campaign and eventually demonstrate the negative impact of some some kind of messages.

Secondly, KKBB website does not display any hour related information concerning the donations or the news. It means that when we associate donations to news there is uncertainties concerning the fact that the message was actually already published before the donation. We worked on the assumption that the news was the first thing to happen at the beginning of the day.

Thirdly, another missing information is the external stimulation of the public. We do not know if the project holder is also using other channels to promote his campaign or if he is already popular with a large community and if they are ready to listen to him and eventually support him.

Finally, the distinction between new contributors and old ones is also almost impossible since the means of authentication are numerous and it is not mandatory to keep the one used in the first time, even though it is highly plausible. This particularity means that the detection of new donators can not be fully guaranteed.

7.1.3 Donation Types

Without knowing the expectations of the project owner from its contributors, we may not detect a real response from a contributor if it is not expected or recorded by the platform. If the goal of the project holder behind a message is only to achieve greater media visibility, this would not be recorded by the platform but would still be a kind of gesture by a contributor and a response to the stimulus that is a message from the project leader.

Moreover it was also stated that the currency expected for a campaign was not limited to euros. This irregularity could lead to harder evaluation of a contribution and of how it gets the project holder closer to his aimed amount even as donations are made. Especially since it would prevent an automatic scraping of the data because the information could not be understood by the program with too irregular currencies (e.g. presales).

7.2 Improvements

7.2.1 Additional Indicators

Campaign Phases

Depending on its level of achievement or its deadline proximity, a potential contributor could decide to contribute due to a feeling of emergency. Defining such phases and attribute it to news could reveal some hidden motivators behind contributions.

Tiredness of Impact

A news may not influence a contributor after a set amount of time (which is a phenomenon known in advertising). In this case it would be safe to assume that it is the project that interested the donators (by the description) and that it is not the last news that prompted the donation. It could be interesting to determine the amount of time after which a news lost its effect.

Current Number of Contributors

Moreover the more contributors a campaign gain the more media coverage it would probably gain. It could be interesting to add this value to the model as an influence factor to see if the impact of news is growing or at least maintaining the adhesion of more donators to the campaign. It could be integrated in another model in which the number of contributors would be the dependant variable.

Aggregate Level Dependant Variable

Since the objective of a campaign is to reach an amount and not to reach the highest price possible. Another logical dependant variable could be the achievement ratio which could be higher than one. This could be a more representing value of the influence of messages because we could distinguish campaigns with donations made after reaching the initial goal.

7.2.2 Decision Interval

By selecting campaign days as decision interval instead of donations interval, we could be able to add to the micro-level a donator decision: whether to donate or not. Since the end of the day would be the end of the interval, it would be possible to evaluate if a message led to a donation or not (campaigns with days without donations but with messages has already been observed). This would lead to the possibility to include the impact of these decisions on the aggregate-level by linking it to the number of contributions.

²⁸Those indicators could also be used as instrumental variables to control for the potential endogeneity in our econometric model, using the same method adopted by Ducarroz, Yang and Greenleaf(2016).

7.2.3 Background Information

As explained earlier, some news could be seen as means to entertain the contributors and create a sense of belonging to the community. This kind of messages could grant a sympathy rating to a project holder. If we can find another campaign where he is project leader, it would allow us to observe if the initial contributions are greater than in his previous campaigns due to his popularity and his already existing backers pool.

7.2.4 News Types

Hybrid News

Since the news are harder to classify a solution would be to use hybrid news which could be both persuasive and informative. A way to implement it could be through the use of a Likert scale which could give us a tendency of belonging to a class of news.

Additional Types

Due to the absence of specific guidelines regarding the news redaction we have a wide variety of news with some really atypical ones which do not fit in already existing classes. We could add new types to the classification with probably a more specific effect and a more detailed analysis of this type of message in our models. Some ideas for new types: community management (surveys, polls, ...) or social interaction (thanks, ...).

As stated by Huhmann and Brotherton (1997), other types of message exist such as guilt appeals, self-benefit and others-benefit appeals and nostalgia appeals that could be used and represents more specific effects to the model.

Chapter 8

Conclusion

During a crowdfunding process, the final amount collected is built through donations which can be affected by the messages sent by the project holder throughout the campaign. The goal of this work is to analyze the impact of these messages on two levels. The campaign level by analyzing how the final amounts collected are built through the number of each kind of messages, the number of donations and the number of jump donations. And the donation level through the decisions made by donors and project holders.

To make this analysis, an existing model based on auction was transposed by altering several elements of the model to adapt to the restrictions of this other kind of money collecting platform. The theoretical part highlighted crowdfunding mechanisms and made a parallel between the auction and crowdfunding processes. This work also has limitations due to the relatively strict transposition of the auction model. Suggestions for improvement have been identified that should make the model more robust and more effective in identifying the impact of messages on crowdfunding.

The results showed that on the aggregate level the number of jump-donations and more simply just donations had an impact on the final amount collected which was also the case regarding the auctions. Even though the informative messages do not have the same impact on the final amount collected since the number of any kind of messages were not significant.

On the micro level, we observed that with the increase in the number of donations the project holder is less likely to issue messages as seen in the auction with the increase of variables indicating progress towards the goal. However, contrary to the auctions persuasive messages have a positive impact on jump donation. It is then possible to conclude that even if the messages have no direct impact on the amount collected, the messages still impact the final amount through the mechanism of jump donating since the messages have a positive influence on the contributors' decisions.

References

- Backus, M., Blake, T., Masterov, D. V., & Tadelis, S. (2015). Is Sniping A Problem For Online Auction Markets?
- Belleflamme, P., Lambert, T. & Schwienbacher, A. (2013). Crowdfunding : tapping the right crowd, Ecore Discussion Paper, 38 pages.
- Bessière, V., & Stéphany, É. (2017). Le crowdfunding : fondements et pratiques. De Boeck Supérieur.
- Block, J., Hornuf, L., & Moritz, A. (2017). Which updates during an equity crowdfunding campaign increase crowd participation? *Small Business Economics*, 50(1): 3–27.
- Burt, C., & K. Strongman. (2005). Use of Images in Charity Advertising: Improving Donations and Compliance Rates. *International Journal of Organizational Behavior* 8(8): 571–580.
- Cazemajour, L. (2013). La place du crowdfunding philanthropique, Centre d'étude et de recherche sur la philanthropie, 30 pages.
- Chandy, R. K., G. J. Tellis, D. J. MacInnis, & P. Thaivanich. (2001). What to Say When: Advertising Appeals in Evolving Markets. *Journal of Marketing Research* 38(4): 399–414.
- Chen, S., Thomas, S., & Kohli, C. (2016). What Really Makes a Promotional Campaign Succeed on a Crowdfunding Platform? *Journal of Advertising Research*, 56(1): 81–94.
- Cialdini, R. B. (1984). *Influence: The Psychology of Persuasion*. NY: Quill.
- Cieply, S. & Le Nadant, A. (2016). Le crowdfunding : Modèle alternatif ou généralisation du modèle de marché pour les start-up et les PME ?. *Revue d'économie financière*, 122: 255-272.
- Clauss, T., Breitenacker, R. J., Kraus, S., Brem, A., & Richter, C. (2017). Directing the wisdom of the crowd: the importance of social interaction among founders and the crowd during crowdfunding campaigns. *Economics of Innovation and New Technology*, 1–21.
- Dehorter, N. (2012). *Crowdfunding suivez le guide : panorama et conseils pratiques*, 88 pages.
- Dens, N., & P. D. Pelsmacker. (2010). How Advertising Strategy Affects Brand and USP Recall for New Brands and Extensions. *International Journal of Advertising*, 29(2): 165–194.

- Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the World-Wide Web. *Communications of the ACM*, 54(4).
- Ducarroz, C., Yang, S. & Greenleaf, E. (2016). Understanding the Impact of In-Process Promotional Messages: An Application to Online Auctions. *Journal of Marketing*.
- Ekeland, M., Landier, A. & Tirole, J. (2016). Renforcer le capital-risque français. *Conseil d'analyse économique*, 33, 12 pages.
- FEB (2013), Financement des entreprises : guide pratique pour une stratégie durable, pages 13-15. Retrieved from: http://vbofeb.be/Global/Publicaties/Gratis%20downloads/Brochure%20financiering%20van%20bedrijven_FRlow.pdf
- Ford, J. B., & A. Merchant. (2010). Nostalgia Drives Donations. *Journal of Advertising Research*, 50(4): 450–459.
- Good, BM. & Su, al. (2013). Crowdsourcing for bioinformatics. *Oxford Journals*, 9 pages.
- Hendrickx, M. (2014). Crowdfunding : mode d'emploi : comment trouver de 300euros à 1 million grâce au financement participatif ?. *Edipro*, 122 pages.
- Hibbert, S., A. Smith, A. Davies, & F. Ireland. (2007). Guilt Appeals: Persuasion Knowledge and Charitable Giving. *Psychology & Marketing*, 24(8): 723–742.
- Hornuf, L., & Neuenkirch, M. (2015). Pricing the value of cash flow rights in crowdinvesting: An analysis of innovestment backers. *Research Papers in Economics*, 15(7).
- Hornuf, L., & Schwienbacher, A. (2018). Market mechanisms and funding dynamics in equity crowdfunding. *Journal of Corporate Finance*, 50: 556–574.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14: 1-4.
- Huhmann, B. A., & T. P. Brotherton. (1997). A Content Analysis of Guilt Appeals in Popular Magazine Advertisements. *Journal of Advertising*, 26(2): 35-45.
- Iizuka, M. (2014). Le crowdfunding : les rouages du financement participatif. *EduBanque Editions*, 197 pages.
- Johar, J. S., & M. J. Sirgy. (1991). Value-Expressive Versus Utilitarian Advertising Appeals: When and Why to Use Which Appeal. *Journal of Advertising*, 20(3): 23–33.
- Kim K. & Viswanathan S. (2014). The Experts in the Crowd: the Role of Reputable Investors in a Crowdfunding Market, 49 pages.

- Kittur, A. (2010). Crowdsourcing, collaboration and creativity, *The ACM Magazine for Students*, 17: pages 22-26.
- Laplume, J-F., Bertin, A. & Favre, C. (2013). Le financement participatif, une alternative à lever des fonds traditionnelle. *L'Agence aquitaine du numérique*, 30 pages.
- Leclercq T. (2016). Customer engagement in Value Co-Creation Process: Gamification in a Crowdfunding Platform.
- Maalaoui, A. & Corneaux, P. (2014). Crowdfunding : les clés du financement participatif. Ellipses Editions, Paris, 125 pages.
- Merchant, A., J. B. Ford, & G. Rose. (2011). How Personal Nostalgia Influences Giving to Charity. *Journal of Business Research* 64(6): 610–616.
- Meyskens, M., & Bird, L. (2015). Crowdfunding and value creation. *Entrepreneurship Research Journal*, 5(2): 155-166.
- Mollick, E. (2014). The Dynamics of Crowdfunding: An Exploratory Study. *Journal of Business Venturing*, 29(1): 1-16.
- Ricordeau, V. (2013). Crowdfunding : le financement participatif bouscule l'économie. FYP éditions, Limoges, 96 pages.
- Rifkin J. (2014). La nouvelle société du coût marginal zéro. Les Liens qui libèrent, 510 pages.
- Rothkopf, M., & Harstad, R. (1994). On the Role of Discrete Bid Levels in Oral Auctions. *European Journal of Operational Research*, 74: 572-581.
- Slovic P, Finucane M., Peters E. & MacGregor D. (2004). Risk as analysis and risk as feelings. *Risk Analysis*, 24(2): 311-322.
- White, K., & J. Peloza. (2009). Self-Benefit Versus Other-Benefit Marketing Appeals: Their Effectiveness in Generating Charitable Support. *Journal of Marketing*, 73(4): 109–124.
- Zhang, Y., & B. D. Gelb. (1996). Matching Advertising Appeals to Culture: The Influence of Products' Use Conditions. *Journal of Advertising*, 25(3): 29–46.
- Zhou, X., T. Wildschut, C. Sedikides, K. Shi, & C. Feng. (2012). Nostalgia: The Gift That Keeps On Giving. *Journal of Consumer Research*, 39(1): 39–50.
- Zimmern, B. (2012). Le trou de financement, l'équity gap. Retrieved from: <http://www.irdeme.org/Le-trou-de-financement-l-equity.html>

Appendices

Appendix A

Web Scraping

A.1 Security Issues

Even though our work was well-intended (research purpose), we had to prevent and bypass the usual securities implemented by websites to protect themselves from such practices and attacks (once more, it was not our case). There is currently 20888 projects available on the website on 2320 pages.

If we refer to the previous explanations, it means that there was at least 64984 requests addressed to the server of KKBB ²⁹. This kind of traffic in such a short delay from a sole IP address would not seem as usual and could be seen as ill-intended. Several measures exist to prevent websites to be exploited in this way: blocking IP address based on traffic, disabling web services API, CAPTCHA challenges, obfuscation using CSS sprites. Moreover, small variations to the HTML/CSS surrounding important data and navigation elements may render the target website too difficult to scrape due to the diminished ability to automate the scraping process.

Luckily, most of these solutions were not adopted by KKBB but we were still wary of an IP address blocking. To prevent it, randomized delay was add between requests. The campaigns to scrap were also divided into subsamples to allow the scraping to take place from different locations.

²⁹ $20888 * 3 + 2320 = 64984$

A.2 JavaScript Execution

While collecting the data of the campaigns we encountered an issue with the donations collection. Beyond a defined number of donations, the page would not display any more donations but hide it behind a button "En voir plus". Even more upon a click, this button would only display the same number more donations.

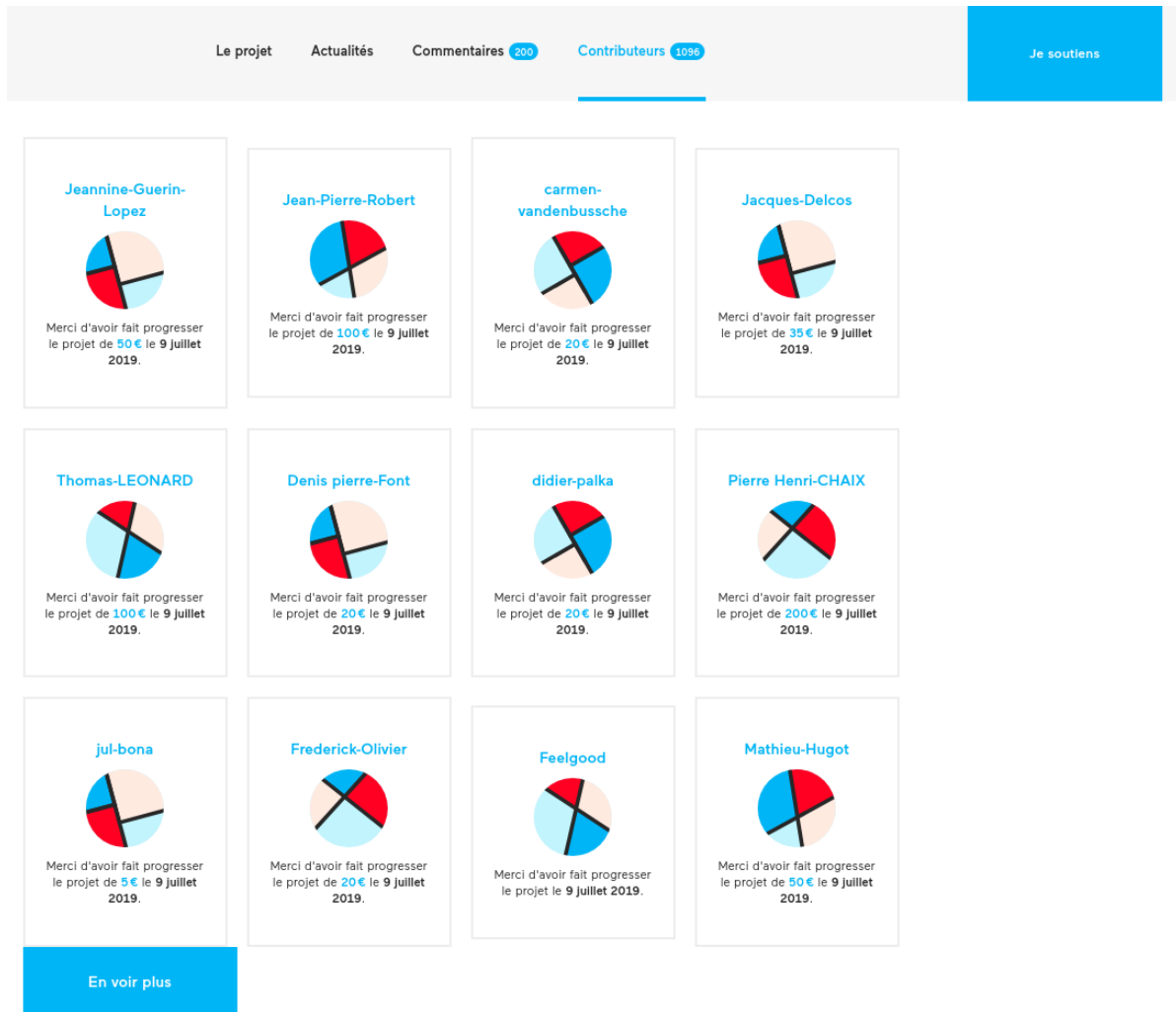


Figure A.1: Rendering of Donations on a Campaign Page³⁰

The HTML code accessible would then only display the information of those donations and no information about the hidden ones. To bypass this mechanism, we simulated the opening of the site in a browser and ran a script to click on the button while it was appearing. We would then proceed normally with HTML scraping on the resulting webpage.

³⁰Donations from the campaign 'QG le media libre': <https://www.kisskissbankbank.com/fr/projects/qg-le-media-libre/tabs/backers>

A.3 Data Storage

Lists of the campaign data structures are stored in files using the pickle module. These lists also allow to easily keep the order of the donations and actualities. This last point is not as relevant as it would seem since the date of the objects would allow us to sort the objects on the fly before any computations but was nevertheless adopted as it was a feature available from *Python* and with the absence of use of a database it would ease belonging relationship storage by attributing a list of donations or news to a campaign.

A.3.1 Pickle over Json

- Pickle is a binary serialization format (instead of text one of json). Which means less storage used.
- JSON, by default, can only represent a subset of the Python built-in types, and no custom classes; pickle can represent an extremely large number of Python types.
- JSON is human-readable, while pickle is not; JSON is also interoperable and widely used outside of the Python ecosystem, while pickle is Python-specific.

A.3.2 Pickle over Marshal

- The marshal serialization format is not guaranteed to be portable across Python versions. Because its primary job in life is to support .pyc files, the Python implementers reserve the right to change the serialization format in non-backwards compatible ways should the need arise. The pickle serialization format is guaranteed to be backwards compatible across Python releases. It is an important point regarding the sustainability of the software.
- The pickle module keeps track of objects it has already serialized, so that later references to the same object will not be serialized again.
- Marshal cannot be used to serialize user-defined classes and their instances. pickle can save and restore class instances transparently.

It was decided to use pickle over the two other solutions mostly for the fact that it would be able to represent custom class such as our objects (Docs.python.3.7, 2019). But it could change, if the data becomes more used or or requires a simpler means of access for other purposes.

Another possible solution was to store it in a comma-separated values (CSV) file which is a delimited text file that uses a comma to separate values. A CSV file stores tabular data (numbers and text) in plain text. CSV files can also be open with *Excel*. This was not done to prevent the need to to translate the objects to a CSV file and inversely which would increase the processing time. If needed, it could be one of the next step for a better representation of the data treated.

A.4 Libraries

The project was developed using *Python* and as such relied a lot on already existing libraries developed by the community to reduce the load of the work. See below for a non-exhaustive list of modules used and their usefulness in order to credit their authors and for future references.

- Pickle³¹: Used to keep our structure intact, reduce the processing time and store our customized objects
- Datetime³²: Used to convert text formatted date to comparable objects and allow to induce a time relation for the elements of our model
- Requests³³: Used to address directly KKBB server and obtain the content sent usually to a webbrowser
- BeautifulSoup³⁴: Used to parse and retrieve information from the HTML code retrieved from the HTTP request
- Selenium³⁵: Used to simulate a local browser through a webdriver and allow to reveal the hidden donations by executing JS scripts

³¹<https://docs.python.org/3/library/pickle.html>

³²<https://docs.python.org/fr/3/library/datetime.html>

³³<https://pypi.org/project/requests/>

³⁴<https://pypi.org/project/beautifulsoup4/>

³⁵<https://selenium-python.readthedocs.io/>

Appendix B

Source Code

All source code can also be found on a public repository: <https://github.com/tdba/KKBB-webscraping>

B.1 Objects Code

```
#!/usr/bin/python3
# -*- coding: utf-8 -*-

import datetime as dt
import dateparser

class Campaign:
    def __init__(self, link, title, description, project_holder,
                 current_amount, aimed_amount, end_date):
        self.link = link
        self.title = title
        self.description = description
        self.project_holder = project_holder
        self.current_amount = current_amount
        self.aimed_amount = aimed_amount
        self.achieved = aimed_amount <= current_amount

    if 'jour' in end_date or 'heure' in end_date or 'minute'
        in end_date:
            self.done = False
            if 'jour' in end_date:
                d = dt.timedelta(days=int(end_date[:2].strip()))
                self.end_date = dt.date.today() + d
            else:
```

```

        self.end_date = dt.date.today()
    else:
        self.end_date = dt.datetime.strptime(end_date, '%A, %B
            %d, %Y').date()
        self.done = self.end_date <= dt.date.today()

    self.actualities = []
    self.donations = []
    self.contributors = []

    def add_donation(self, donation):
        if donation.user not in self.contributors:
            self.contributors.append(donation.user)
        self.donations.append(donation)

class Donation:
    def __init__(self, user, date, amount=0):
        self.user = user
        self.amount = amount
        self.date = dateparser.parse(date).date()

class News:
    def __init__(self, title, content, date, kind=None):
        self.title = title
        self.content = content
        self.date = dt.datetime.strptime(date, '%A, %B %d, %Y').
            date()
        self.kind = kind

```

Code Listing B.1: Campaign Objects Python Code

B.2 Script Code

```
#!/usr/bin/python3
# -*- coding: utf-8 -*-

from bs4 import BeautifulSoup
import requests
from selenium import webdriver
from selenium.common.exceptions import NoSuchElementException

from time import sleep
from random import randint
from math import ceil
from copy import deepcopy
import json
import pickle

import campaign

URL = 'https://www.kisskissbankbank.com'

def max_page():
    url = URL + '/fr/discover?filter=all'
    response = requests.get(url)
    soup = BeautifulSoup(response.content, "html5lib")

    script = json.loads(soup.find('script', class_='js-react-on-rails-component').text)
    nb_projects = script['BROWSING']['projects']['totalCount']

    return ceil(nb_projects/9)

def links(max_p, current_page=1):
    res = []

    while current_page <= max_p:
        url = URL + f'/fr/discover?filter=all&page={current_page}'
        response = requests.get(url)

        sleep(randint(0, 4))
        while response.status_code != 200:
```

```

        response = requests.get(url)
        sleep(randint(0, 4))
        print('Request not treated')

soup = BeautifulSoup(response.content, 'html5lib')

script = json.loads(soup.find('script', class_='js-react-on-rails-component').text)
grid = script['BROWSING']['projects']['edges']

for item in grid:
    res.append(item['node']['publicUrl'])

print(f'Page {current_page} treated.')
current_page += 1

return res

def data_extraction(urls):
    def info_extraction(link):
        response = requests.get(link + '/tabs/description')
        soup = BeautifulSoup(response.content, 'html5lib')

        title = soup.find('head').find('title').text
        desc = soup.find('head').find('meta', attrs={'name': 'description'})['content']
        project_holder = URL + soup.select_one('div.owner-info__StyledOwner-sc-1r3rkij-2.gDZNgQ').a['href']

        prompt_class = 'informations-and-media__StyledInfo-sc-121rta3-3_fzHbGm_marger__StyledMarger-q3lecu-0_hAgZUM'
        prompt = soup.find_all('div', class_=prompt_class)[1:]
        end_date = prompt[0].div.text
        values = prompt[1].find_all('div')
        try:
            actual_value = int(values[0].text[:-1].strip().replace(u'\xa0', u''))
            aimed_value = int(values[1].text[3:-1].strip().replace(u'\xa0', u''))
        except ValueError:
            actual_value, aimed_value = 'No monetary information', 'No monetary information'

    return campaign.Campaign(link, title, desc, project_holder

```

```

        , actual_value , aimed_value , end_date)

def actualities_extraction(p):
    resp = requests.get(p.link + '/tabs/news')
    soup = BeautifulSoup(resp.content , 'html5lib')
    news = soup.select('div.marger__StyledMarger-q3lecu-0.
        denjnR')
    if news[0].section:
        for item in news:
            date = item.select_one('div.marger__StyledMarger-
                q3lecu-0.jgfcQX_>_span').text
            news_title = item.find('h1').text
            content = item.select_one('div.kiss-RichText.
                marger__StyledMarger-q3lecu-0.WSKaG_>_div').text
            actuality = campaign.News(news_title , content ,
                date)
            p.actualities.append(actuality)

def donations_extractions(p, browser):
    browser.get(p.link + '/tabs/backers')
    try:
        button = browser.find_element_by_xpath("//button[text
            ()='En_voir_plus']")
        while button:
            browser.execute_script("arguments[0].click();",
                button)
            sleep(0.2)
            button = browser.find_element_by_xpath("//button[
                text()='En_voir_plus']")
    except NoSuchElementException:
        pass

    inner_html = browser.execute_script("return_document.body.
        innerHTML")
    soup = BeautifulSoup(inner_html , 'html5lib')
    for item in soup.select('div.backers-list__StyledCard-sc
        -13o55gl-1.gKLrNb'):
        user = item.select_one('div_>_span').text
        gift = item.select('div.marger__StyledMarger-q3lecu-0.
            bsSajs_>_p_>_span')

        if len(gift) == 2:
            try:
                amount = int(gift[0].text[:-1].strip().replace
                    (u'\xa0' , u''))

```

```

        except ValueError:
            amount = 'No_monetary_information'

            date = gift[1].text.strip()
            donation = campaign.Donation(user, date, amount)
        else:
            date = gift[0].text.strip()
            donation = campaign.Donation(user, date)
        p.add_donation(donation)

res = []
driver = webdriver.Chrome()
i = 1

while urls:
    try:
        print(f'Project_{i}:_{urls[0]}_scrapping.')
        project = info_extraction(urls[0])
        actualities_extraction(project)
        donations_extractions(project, driver)

        res.append(project)
        i += 1
        sleep(randint(0, 2))
    except:
        pass

    del(urls[0])

driver.close()
return res

if __name__ == '__main__':
    # max_page = max_page()
    # campaign_links = links(max_page)
    # with open('links', 'wb') as f:
    #     pickle.dump(campaign_links, f)

    with open('links', 'rb') as f:
        campaign_links = pickle.load(f)

    while campaign_links:
        print(len(campaign_links))
        number_to_treat = 100

```

```

treated_campaigns = []
campaign_links_to_treat = deepcopy(campaign_links[:
    number_to_treat])
while campaign_links_to_treat:
    campaigns = data_extraction(campaign_links_to_treat)
    treated_campaigns.extend(campaigns)

with open('links', 'wb') as f:
    pickle.dump(campaign_links[number_to_treat:], f)

campaign_links = campaign_links[number_to_treat:]

with open('campaigns', 'rb') as f:
    stored_campaigns = pickle.load(f)
    treated_campaigns.extend(stored_campaigns)
with open('campaigns', 'wb') as f:
    pickle.dump(treated_campaigns, f)

```

Code Listing B.2: Code of Scraping Script

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management

Chaussée de Binche 151, 7000 Mons, Belgique | www.uclouvain.be/lsm