

# **COMMUNICATING TO SUCCESS: DECISION-MAKING IN CROWDFUNDING**

**Naomi Moy**

**BBus (International Business) with Distinction**

Primary Supervisor: Professor Benno Torgler

Associate Supervisor: Dr Markus Schaffner

Submitted in fulfilment of the requirements for the degree of  
Master of Business (Research)

School of Economics and Finance

Faculty of Business

Queensland University of Technology

2016



# Keywords

Attention · Behavioural Intentions · Communication · Complexity · Confidence ·  
Crowdfunding · Funding · Information · Information signals · Money · Priming ·  
Readability · Success · Text length.

# Abstract

This thesis investigates how the way project creators communicate to contributors, affects crowdfunding success, using 60,000-80,000 Kickstarter projects. Four studies are presented in this thesis.

The first study explores the information volume (word count) of a text on the number of contributors and the amount they contribute. By comparing over 80,000 types of text with different lengths, it is possible to observe if an optimal text length exists. In all categories, I find a consistent inverted U-shaped relationship between text length and overall success. In the case of Kickstarter projects, an optimal text length does indeed exist.

Within the second study, I explore the complexity of the campaign description on the number of backers and amounts raised. As reading ability differs by individual, the readability of a text matters when writing for a broad audience. To measure complexity I focus on the Flesch Reading Ease readability measure, which generates a score and grade level indicator for complexity. For robustness I extend the analysis with 20 other readability measures. In all cases, funds raised and number of contributors, exhibit a curvilinear relationship between readability and overall success.

In addition to complexity, the third study explores how self-confidence shapes success. To do this, confidence is proxied using two novel behavioural variables: a project duration ratio, which considers the project duration chosen in comparison the duration forgone, and a certainty ratio using the linguistic expressions exhibited in the campaign text. The findings provide strong evidence of an inverted U-shaped relationship between self-confidence and success.

Finally, the fourth study explores how the ultimate goal of receiving funds is discussed within the text, and how discussing money affects the number of participants and the overall amount contributed to a crowdfunding project. The emphasis on money is measured using the Linguistic Inquiry Word Count (2007) dictionary *money*. In doing so, I provide evidence that an increased emphasis on

money will negatively affect success as represented by the significantly negative linear relationship.

# Table of Contents

Keywords .....	i
Abstract .....	ii
Table of Contents .....	iv
List of Figures .....	vi
List of Tables .....	vii
List of Abbreviations .....	viii
Statement of Original Authorship .....	x
Acknowledgements .....	xi
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Thesis Outline .....	7
<b>Chapter 2: Length matters .....</b>	<b>9</b>
2.1 Introduction .....	11
2.1.1 Text Length and Information .....	12
2.1.2 Crowdfunding – Kickstarter .....	13
2.2 Data and Methodology .....	14
2.2.1 Project description length .....	15
2.2.2 Description edits and potential endogeneity bias .....	16
2.3 Results .....	18
2.4 Discussion .....	23
<b>Chapter 3: Don't keep it (too) simple .....</b>	<b>25</b>
3.1 Introduction .....	26
3.2 Data & Methodology .....	29
3.2.1 Kickstarter data .....	29
3.2.2 Method .....	29
3.3 Results .....	31
3.3.1 Robustness checks using other readability indices .....	34
3.4 Discussion .....	35
<b>Chapter 4: Don't talk about money .....</b>	<b>39</b>
4.1 Introduction .....	40
4.2 Data and Methodology .....	43
4.3 Results .....	45
4.4 Conclusion .....	47
<b>Chapter 5: Confidence is good; Too much not so much .....</b>	<b>51</b>
5.1 Introduction .....	52
5.2 Data and Methodology .....	54
5.2.1 Data54 .....	

5.2.2 Proxy for confidence: Project Duration .....	54
5.2.3 Proxy for confidence: Linguistic certainty .....	56
5.3 Results .....	57
5.4 Discussion.....	66
<b>Chapter 6: Concluding Remarks.....</b>	<b>67</b>
6.1 Key Findings.....	67
6.2 Limitations & Future perspectives .....	69
<b>References .....</b>	<b>77</b>
<b>Appendices .....</b>	<b>89</b>
Appendix A : <b>Tables</b> .....	89
Appendix B : <b>Figures</b> .....	94
Appendix C.....	98

# List of Figures

<b>Figure 2.1</b> Distribution of the total word count across projects (a) and by category (b). The right hand tail of each distribution is truncated for the purposes of visualisation. ....	15
<b>Fig. 2.2</b> Marginal effects of word count on $\ln(\text{raised})$ .....	19
<b>Fig. 2.3</b> Marginal effects of word count on $\ln(\text{backers})$ .....	20
<b>Figure 3.1</b> Distribution of the (reversed) Flesch Reading Ease score (a) and the corresponding grade level (b) of Kickstarter project description. ....	30
<b>Figure 3.2</b> Fitted models of readability (Flesch reading-ease) with crowdfunding success: Fund raised (a) and Number of backers (b) .....	34
<b>Figure 3.3</b> Fitted models of fund raised (a) and number of backers (b) vs. readability indices .....	35
<b>Figure 5.1</b> Duration ratio prediction on $\ln(\text{raised})$ (a) and $\ln(\text{backers})$ (b).....	59
<b>Figure 5.2</b> Certainty ratio prediction on $\ln(\text{raised})$ (a) and $\ln(\text{backers})$ (b).....	59



## List of Tables

<b>Table 2.1</b>	Summary statistics of non-censored word count. ....	16
<b>Table 2.2</b>	Summary statistics of identified edits.....	18
<b>Table 2.3</b>	Multivariate analysis of description length (word count censored at 99%):.....	21
<b>Table 2.4</b>	Multivariate analysis of description length (word count censored at 99%) on $\ln(\text{Raised})$ .....	22
<b>Table 2.5</b>	Multivariate analysis of description length (word count censored at 99%) on $\ln(\text{Backers})$ .....	22
<b>Table 3.1</b>	Summary Statistics .....	31
<b>Table 3.2</b>	Multivariate analysis using the raw FRES Score.....	33
<b>Table 3.3</b>	Multivariate analysis of the Grade level based on FRES .....	34
<b>Table 4.1</b>	Descriptive statistics per project .....	44
<b>Table 4.2</b>	Simple regression focused on money terms .....	45
<b>Table 4.3</b>	Multivariate analysis of money terms' influence on the funds of $\ln(\text{raised})$ (\$).....	48
<b>Table 4.4</b>	Multivariate analysis of money terms' influence the number of $\ln(\text{backers})$ .....	49
<b>Table 5.1</b>	Multivariate analysis of confidence (duration ratio): funds raised and number of backers.....	60
<b>Table 5.2</b>	Multivariate analysis of confidence (duration ratio) for the three separate periods: funds raised and number of backers .....	61
<b>Table 5.3</b>	Probit model of success probability using duration ratio .....	62
<b>Table 5.4</b>	Multivariate analysis of confidence (certainty ratio): funds raised.....	63
<b>Table 5.5</b>	Multivariate analysis of confidence (certainty ratio): number of backers .....	64
<b>Table 5.6</b>	Probit model of success probability: certainty ratio.....	65

# List of Abbreviations

ARI	Automated Readability Index
FAQ	Frequently Asked Question
FRES	Flesch's Readability Ease Score
LIWC	Linguistic Inquiry Word Count
OLS	Ordinary least squares
US	United States

*For my family.*

# Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: QUT Verified Signature

Date: October 2016

# Acknowledgements

Firstly I would like to thank my supervisors, Professor Benno Torgler and Dr Markus Schaffner. Thank you for the continuous encouragement and support, without it I would not be the researcher that I am today. I would particularly like to thank Benno for being so understanding and patient with my research abilities, and for pushing me to work harder than I ever have.

I am incredibly grateful to my colleague, co-author and friend, Ben. Thank you for always staying positive and nudging me in the right direction.

I would also like to thank the guy's on level 8, for without them I would have gone crazy: Ann-Kathrin, Dave, Juliana, Justin, Marco, Poli, Sam, Steve, Suzanne, Tony, Uwe, Yang, and the E&F honours class of 2015. Thank you for the many coffee sessions, long lunches and early Fridays for drinks, the encouragement and support, the snacks, the company and for everything else I have not mentioned. In addition I would like to thank Annastiina, Jason and Louisa, for letting me have to opportunity to teach. I would also like to thank Di and Lee, for the emotional support.

Lastly, I would like to thank my friends and family. Even though most of you had no idea what I was doing, you always believed in me and for that, I cannot thank you enough. A special note of thanks to my housemates, Jared and Merv, for putting up with my thesis induced states, for constantly feeding me and your unwavering support.



# Chapter 1: Introduction

---

Striving for success is a quintessentially human trait that knows no borders, crossing disciplinary institutions and cultural boundaries. Defined as the accomplishment of an aim, success can occur in many scenarios and is pursued not only by individuals but also by teams and entities. The aim in question can be achieving an award, reaching a milestone, winning a competition or academic achievement. Additionally success can be achieving entrepreneurial or career success. In order to attain success we make a series of decisions that we believe will lead to our intended outcome. However, more often than not, success is dependent on the decisions of others, e.g. an employee can work hard to distinguish themselves from others, but cannot grant themselves a promotion. A prominent example of the dependence on others is seen through crowdfunding, where success is defined as achieving the fundraising goal. Yet to successfully raise funds, you need to be successful in motivating the crowd for support and contributions. Therefore there are two aspects to measuring fundraising success through communication: the motivation of individuals for support and achieving a large quantity of contributions from these individuals. This thesis will study success within crowdfunding, using field data from a controlled environment. Kickstarter data will be used, as it provides a controlled environment as all things are held equal when creating outcomes. The use of such a setting allows us to examine the effects of signals communicating information, which can be found within the campaign texts and details, on the number of contributors contributing and total funds raised. The novelty of this is that previous research has not approached the communication of information and linked it to success in this way.

## *Crowdfunding and Kickstarter*

Crowdfunding helps a potential innovator raise capital from a large audience usually via the Internet to get an idea off the ground. In this case, the audience pools together, individually contributing smaller sums of capital to raise the amount required to get the project going. This style of investment can take shape in several forms, such as altruistic donations, equity purchases, loans, or rewards which is the receipt of something such as the product being created in return for funding (see

Agrawal, Catalini and Goldfarb, 2011; Belleflamme, Lambert and Schwienbacher, 2013; Bouncken, Komorek and Kraus, 2015). Whilst the concept of creating capital from entities other than formal financial institution is not new, crowdfunding is a relatively new trend of revenue raising models, as it did not exist in this format (online) until the early 2000's<sup>1</sup>. Using the Internet has broken down the borders and early-stage entrepreneurs are no longer geographically constrained (Agrawal et al., 2014).

Just as the Internet has broken down the geographic constraints, it has lowered both the communication and search costs for entrepreneurs. As crowdfunding platforms are consumer-driven networking services, matching funders with creators is now more efficient and effective (Agrawal et al., 2014). Crowdfunding also gives projects and the subsequent products a chance to flourish when they would not have previously been able to do so, due to restrictions from traditional revenue raising models, e.g. financial loans from lending institutions (Agrawal et al., 2014). As a result this funding model is particularly popular among resource-deficient early-stage entrepreneurs, as it enables them to gauge interest and generate capital.

Potential contributors are economic agents who base their contributing behaviour on bounded rationality. Compared to the traditional form of investors, these contributors are likely to have a minimal amount of resources they can allocate to a project. Hence, contributors pool their limited resources together in order to achieve the fundraising goal. However, before pooling resources individuals are required make decisions to contribute money on the basis of the information communicated to them by the creator, demonstrating adverse selection. These signals can be direct, e.g. goal amount or indirect, e.g. implied creator confidence levels.

Whilst potential funders may be heterogeneous in their motivations to contribute, Agrawal et al. (2014) recognises that crowdfunding is attractive to funders for five distinct reasons. The first being that crowdfunding increases their access to investment opportunities, which is dependent on the fundraising model. Additionally, they recognise that funders are incentivised by the community participation, early access to new produces and the ability to provide support in both

---

<sup>1</sup> ArtistShare is purported to be the first crowdfunding platform (fan-funding), created in 2003.  
<http://www.artistshare.com/>



a monetary and non-monetary form, and the formalisation of a contract (Agrawal et al., 2014).

Due to the online setting of crowdfunding platforms, we are able to observe the how the creator communicates as well as the crowdfunding outcome. This presents us with a novel homogenous setting to examine the influence of creator generated signals on contributing behaviour and ultimately, effects on success. Allowing us to fill in the research gaps of this relatively new fundraising concept.

One of the more well-known crowdfunding platforms is Kickstarter, which follows a rewards-based all-or-nothing funding model, where funds are not allocated to the creator unless the funding aim is reached (Belleflamme, Ormani and Peitz, 2015). In other words, if the project does not achieve its goal amount, creators do not receive any of the funds contributed and the contribution is returned to the funder. However if the project does achieve or exceed the full-funding goal, then funders<sup>2</sup> expect to receive the reward they selected when contributing. Yet, there is no requirement for creators to make the delivery of rewards an actuality, as like all business ventures there is an element of risk in the project itself not becoming a reality.

Regardless of this perceived risk, individuals are still willing to support projects based on the details provided by creators. In order to provide as much detail as possible, Kickstarter specifies four separate tabs within the campaign, one each for the overall campaign, project updates, comments and a community tab. This thesis will focus on the overall campaign page, as it contains the most information about the project. The campaign tab contains details such as the campaign description, images relevant to the campaign, the number of FAQs, duration of the project, the projects social network links and the number, type and tiers of rewards. In addition to these, the top of the page includes details of the goal amount, number of backers, total amount raised, and the projects category, sub-category, geographic location, creator identifier and a video if one has been submitted.

Kickstarter features a wide variety of projects across fifteen different categories and ninety-four subcategories<sup>3</sup>. Every project on Kickstarter has a definitive aim to

---

<sup>2</sup> Kickstarter refers to contributors as *backers*.

<sup>3</sup> There were originally 13 categories and 49 subcategories.

create something, and cannot be a fundraiser for charities or causes. Therefore each project has the same overall aim of creating a product using the funds raised. This aim is detailed by the creator(s) in the descriptive section of the project campaign page. This section is unrestricted, in that there is no capped word count and no requirements on what the creator can and cannot say. In other words, the creator has the final say on everything expressed within the description.

As a requirement of the Kickstarter funding model, each project lists a series of rewards that are offered to contributors in exchange for their resources. Kickstarter restricts the kind of rewards that can be offered by projects, so that rewards are specific to the project, nonfinancial in nature, are not the resale of another product or an illegal item<sup>4</sup>. In most cases, creators will at least offer initial prototypes or the final product at a heavily discounted rate. Other kinds of rewards can be but are not excluded to, acknowledgements, project paraphernalia or incorporating the backers into the project, i.e. an author could name a character in their novel after a backer. Each reward has an estimated delivery date if the project is successful in raising funds. These rewards can be restricted by the number offered and the number available to contributors, as well as the type of reward listed, all of which are at the discretion of the creator.

Websites are constantly evolving to suite the nature of their consumers, in this aspect Kickstarter is no different. So it should come as no surprise, that in its 7 years of operations some changes have been made. The first major change occurred in June 2011, when Kickstarter adjusted its maximum duration period from 90 to 60 days (Strickler, 2011). In October the following year, Kickstarter went ‘live’ in the United Kingdom, in the sense that creators from the UK are now able to launch a project on Kickstarter (Strickler, Chen and Adler, 2012). Originally the project creation side of the crowdfunding platform was only available to residents of the United States, whilst potential backers of projects could come from any part of the globe. The globalisation of project creators continued, with projects now ‘live’ in Australia, Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, Luxembourg, Netherlands, New Zealand, Norway, Spain, Sweden and Switzerland (*Kickstarter in Canada!*, 2013; McGregor, 2013; McGregor, 2014a,c; *Welcoming Austria, Belgium,*

---

<sup>4</sup> <https://www.kickstarter.com/rules/prohibited>

*Italy, Luxembourg, and Switzerland to the Kickstarter Community!*, 2015; *¡Kickstarter ya esta disponible en España!*, 2015; Strickler, 2015).

In addition, Kickstarter introduced a section on the campaign description page called ‘*risks and challenges*’, to reinforce to contributors that projects are in development and that Kickstarter is not a store (Strickler, Chen and Adler, 2012). And additional adjustments were made to the processes of Kickstarter, such as updating the *Terms of Use*, refining their community guidelines and rules, and a ‘*launch now*’ feature that uses an algorithm to assess the readiness of projects (Strickler, 2014a,b). If the project meets the algorithms criteria, it does not need to be reviewed by a Kickstarter staff member. Additionally, two categories were added in 2014, *Journalism* and *Crafts*, along with the Creator Handbook (Abebe, 2014; McGregor, 2014b). Unmistakably Kickstarter caters to a wide variety of ideas within a controlled setting.

Given the emphasis being placed on innovative activities by a great number of governments, exploring the interactions of information being exchanged through salient signals holds value in two ways: first, it demonstrates how people are affected by effort, complexity, confidence and monetary signals in the decision making process, and second, it explores the effect of such salient signals on success. In order to investigate such effects, data from the crowdfunding website Kickstarter will be used. Within this setting we can investigate the following research questions: Does the amount of information provided by a creator of the descriptive text influence the funding outcome? Will the readability of a text influence contributing behaviour? Will portraying extreme levels of confidence elicit the appropriate response needed to achieve success? Given the ultimate goal is to receive funds, how will over emphasising the purpose of the project (raising funds) affect an individual’s decision-making behaviour? Overall I want to understand how communication affects decision-making behaviour.

Creators of a project are given free rein on the amount of information they detail about the idea, and in some cases creators run free with their words whilst others may keep it very brief. Whilst this results in a large variation of volumes of information, potential contributors still rely on a limited set of information indicators to make a decision, exhibiting bounded rational behaviour. If the creator is not communicating effectively, it could have an unintended consequence on fundraising

success. For example, if the creator details too much information within the description regardless of its quality, it may be perceived by a potential contributor as requiring too much effort to read and an overload of information (Shenk, 2003; Eppler and Mengis, 2004). Since individuals are less likely to engage in behaviour that requires more effort, a funder is less likely to read an excessively lengthy body of text (Song and Schwarz, 2008). On the other hand, detailing too little information may not portray the necessary information nor fulfil the potential contributor's expectations or requirements. Thus sending an unintentionally negative signal to contributors.

Whilst one would expect the quantity of a text to affect a contributors perceptions, the complexity of information contained within the text is a significant factor. As potential contributors come from a variety of backgrounds, their understanding of the project will have an effect on their decision-making. Song and Schwarz (2008) observe that if a text is perceived as being more difficult to read, the less willing an individual was to engage in the activity. A prime example of this is consumers' lack of interest in or motivation to read legal contracts or website privacy policies, despite knowing that these documents contain important information (Jensen, Potts and Jensen, 2005; Stark and Choplin, 2009; Luth, 2010). If the text is too complex, potential contributors may ignore the information and alter the outcome of their decision. On the other hand, simplifying the text to make it more comprehensive improves the clarity of those who are non-experts (Masson and Waldron, 1994; Campbell, 1999; van Boom, Desmet and van Dam, 2016). A caution to the simplification of texts, is that it is possible to oversimplify writing to the point that it may no longer result in the desired outcome and projects may not raise the required amount of funds.

Given that money is the unit of exchange, it often plays an important role in decision-making processes. Notwithstanding its physical importance, the association of money has moved beyond being just an instrument for exchange and is now associated with beliefs, hopes and fears (Trachtman, 1999). Such associations can cause strong behavioural changes when an individual is reminded of money (Jiang, Chen and Wyer, 2014). As the underlying aim of a project is to generate funds, it is likely that project creators will discuss money. If a creator places too much of an

emphasis on money, through the actual word or other associated terms it may deter contributors and the amount contributed.

Similarly, creators would need to be confident in their idea before launching it on a public platform. Yet when a creator describes their idea, it is possible to exhibit too much confidence when providing information (Weinberg, 2009). Whilst being confident in one's own abilities is often regarded as an asset, there are still situations in which being over confident can have a negative effect on outcomes, such as the survival chances of a business (Koellinger, Minniti and Schade, 2007; Tonwsend, Busenitz and Arthurs, 2010). However, this does not necessarily mean creators need to express low levels of confidence, rather it may be more appropriate if moderate levels of confidence are exhibited (Mobius and Rosenblat, 2006). It is possible for us to observe levels of confidence both within a text, using linguistic analysis, and from behavioural variables. Thus I can examine whether there is an optimal level of confidence when raising funds.

A more detailed discussion on the literature for each topic is provided within each chapter.

## **1.1 THESIS OUTLINE**

The structure of this thesis is as follows. In Chapter 2, I present the first study on the effects of the quantity of information on the behaviour of contributors. I use the text length as a proxy for information quantity as it is an easily identifiable cue. In this study, titled "Length matters", I examine whether the text length of a crowdfunding description alters the decisions of backers to contribute to a project. This study analyses over 80,000 project descriptions and the resulting outcomes of amount raised and number of contributors. In Chapter 3, I present a study entitled "Don't keep it simple" which extends upon Chapter 2 by examining the complexity and readability levels in crowdfunding description texts using linguistic analysis. In this study I use the same dataset to observe whether different levels of complexity and readability affect the behaviour of contributors and the amount that is contributed. As a result of how creators portray information, I take the analysis a step further and examine specific wording in Chapter 4 using a smaller dataset of approximately

60,000 projects. In this study entitled “Don’t talk about money”, I present the salient effects of using ‘money’ terminology within text, by using linguistic analysis to identify monetary terms within the descriptive text of the dataset. Finally, in Chapter 5, I focus on a specific characteristic of the project creator by analysing the level of confidence detected within the project. In the study titled “Confidence is good; Too much not so much”, I use two proxies for confidence. The first proxy is based on the deadline chosen by the creator in comparison with the days forgone, and the second proxy focuses on the terminology used by the creator when communicating their idea. Chapter 6 concludes the findings of all four studies and provides suggestions for future research.

Chapters 4 and 5 have smaller datasets, as there were some learning experiences while working with the data. In the first two studies, a portion of the data was dropped due to inconsistencies. However, in the two empirical studies conducted in the last few weeks (Chapter 2 and 3), the issue was resolved by correcting for the ‘tab’ and ‘new line’ spaces in the raw data file. It would have been relatively time consuming to revise the other two chapters, which would lead to submission delays. Nonetheless, adjustments will be made in preparation of submitting the papers for publication in academic journals. Additionally, it should be noted that the investigation of such a large sample, over 60,000 observations, is likely to cause the variables to be statistically significant regardless of even if the explanatory variable is not economically significant (McCloskey and Ziliak, 1996). Therefore, when explaining the results I use figures to assist in interpreting the economic significance.

Chapters 2 and 3 are in collaboration with my principal supervisor Professor Benno Torgler and Ho Fai Chan. Chapters 4 and 5 are in collaboration with my principal supervisor Professor Benno Torgler, Dr Markus Schaffner and Ho Fai Chan.

## Chapter 2: Length matters

---

### Statement of Contribution of Co-Authors for Thesis by Published Paper


The authors listed below have certified\* that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the Australasian Research Online database consistent with any limitations set by publisher requirements.

In the case of this chapter:

#### Length Matters

July 2016 – In preparation for submission

Contributor	Statement of Contribution
Naomi Moy	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (50%)
	
08/08/2016	
Ho Fai Chan	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (20%)
Benno Torgler	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (30%)

#### Principal Supervisor Confirmation

I have sighted the email or other correspondence from all Co-authors confirming their certifying authorship.

Benno Torgler



08/08/2016

Name

Signature

Date

“... in an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.”

Herbert A. Simon (1971, pp. 40-41)

“It has often been said,  
There’s so much to be read,  
You never can cram  
All those words in your head.

So the writer who breeds  
More words than he needs  
Is making a chore  
For the reader who reads.

That’s why my belief is  
The briefer the brief is,  
The greater the sigh  
Of the reader’s relief is.

And that’s why your books  
Have such power and strength.  
You publish with shorth!  
(Shorth is better than length.)”

Theodor Seuss Geisel (Dr. Seuss)



## 2.1 INTRODUCTION

Gathering information is a crucial part of the decision-making or problem-solving process. For example, we check for approaching traffic before crossing a road, or assess the reviews of hotels before making a booking. It could be expected that decisions made after collecting too little information could lead to a worse outcome, e.g. causing severe traffic accident or receiving an inferior hotel room. On the other hand, in the presence of an overabundance of information, individuals are unable to process every piece of information available in order to form an appropriate or adequate judgement despite the fact that processing all the information would be more likely to lead to a better outcome (increase in the quality of decisions). According to the “economics of attention”, our attention is a scarce resource, and we are only able to allocate our attention to a finite number of items (Huberman and Wu, 2008). So when we are confronted with a surplus of information, such as lengthy texts, we find it difficult to distinguish between relevant and unnecessary pieces of information. Instead, we tend to rely on heuristics or use reasoning to take shortcuts in order to make decisions, as the required costs (attention and cognitive effort to acquire all the relevant information) outweighs the potential benefits (Malhotra, 1982). Given the level of effort required to process information, it is not surprising that a bounded rational individual would choose not to evaluate all the information available.

This paper seeks to examine the relationship between the amount of information provided and decision-making in a crowdfunding setting, where innovative projects and ideas are proposed to the general public for financial support. Since potential contributors rely on the information portrayed within the project description to inform their pledging decision of whether or not fund a project, it is important to understand the role of information quantity in investment behaviour. Hence, this study addresses the following research question: Does the amount of information provided by the creator influence the funding outcome? Using a large empirical dataset from Kickstarter.com and looking at close to 80,000 projects, we observe an inverted U-shape relation between the length of the project description and the number of funders or the amount raised.

### 2.1.1 Text Length and Information

Information overload occurs when an individual receives too much information and thus reaches a point where they are no longer able to process information (Malhotra, 1982; Eppler and Mengis, 2004). In effect, too much information is considered to be too much of a good thing (Shenk, 2003). This is evidenced by the performance of individuals when making decisions, which steadily increases when information is added and then declines once the information becomes too much, resulting in information overload (see Eppler and Mengis, 2004 for an overview). Therefore, providing too much information within a text, thereby increasing its length, could (unintentionally) have a negative effect on an individual's decision. It can lead to feelings of stress, confusion, pressure, anxiety, or low motivation (Eppler and Mengis, p. 328, 2004). Naturally, the propensity to ignore lengthy informative texts is closely related to the attention span of the individual. The problems associated with attention spans can be further exacerbated if the comprehensibility of parts of the texts are too dependent on clear understanding of preceding sections. Readers may gauge the level of effort required to digest information through cues such as the number of pages or thickness of a document.

A common social phenomenon in online discussion forums is the commentary of 'too long didn't read' for articles with large amounts of text. Academic journal articles are often around 20 pages long, many non-fiction books are no longer than 300 pages. A purpose of *Reader's Digest* was to condense otherwise lengthy material (e.g., *Reader's Digest Condensed Books*) (for a discussion see Baron, 2015). Online news articles are also becoming shorter (Baron, 2015)

The effect of changing the writing style, in terms of length of the text, has been investigated in various disciplines. For example, in the informetrics and scientometrics literature, various studies report a significant positive relationship between scientific article length and citation outcome (see Torgler and Piatti, 2013; Falagas et al., 2013; Chan et al., 2015; Weinberger et al., 2015)<sup>5</sup>. In contrast, short and succinct abstracts are more likely to increase citations rather than longer abstracts (Letchford et al., 2016). A large body of literature is also devoted to studying the link between response rates and questionnaire lengths in survey

---

<sup>5</sup> When examining the number of citations, Stremersch et al. (2007) find that the article length has a positive linear relationship with citations. Non-linearity does not increase the model fit.

research. Shorter questionnaires have been observed to increase the quality of responses and the response rate (Burchell and Marsh, 1992; Galesic and Bosnjak, 2009; Herzog and Bachman, 1981; for an overview see Rolstad et al., 2011). On the other hand, Lund and Gram (1998) observe a distinct U-shape relationship between response rate and questionnaire length: shorter questionnaires elicit the greatest response rate yet the longest questionnaire did not necessarily have the lowest response rate.

In a consumer setting such as exhibit labels in a museum, or health claims for foods, the length of a text has been shown to affect an individual's attention. The longer the article of text, the less likely it is to be read or comprehended (see Bitgood and Patterson, 1993; Cota and Bitgood, 1993; Williams, 2005; Bitgood, Dukes and Abbey, 2006). In contrast, purchasing behaviour based on online reviews is positively correlated with the length of a comment,<sup>6</sup> increasing sales and perceived helpfulness (Chevalier and Mayzlin, 2006; Kampouris and Evans, 2013) assuming that there is no information overload leading to diminished attentiveness and decision quality (Malhotra, 1982; Lee and Lee, 2004; Lurie, 2004; Chen, Shang and Kao, 2009; Krasnova, Kolesnikova and Guenther, 2009). In sum, the length of a body of text, perceived as the quantity of words or the number of pages, has an effect on different forms of success such as the sale of an item, the quality of the decisions, or the citation of an article.

### **2.1.2 Crowdfunding – Kickstarter**

Extending upon the current literature, this paper will examine the effects of text length on achieving success in an entrepreneurial setting, by examining data from Kickstarter, the crowdfunding website. Kickstarter is a platform that links innovators with individuals who are willing to contribute funds in exchange for physical (product) and non-physical (a thank you) rewards. In order to convince individuals to contribute to their campaign, creators of the innovation must pitch their idea using text, which may be supported by images and videos. The amount of information a creator is able to portray within each medium is important given the all-or-nothing

---

<sup>6</sup> Chevalier and Mayzlin (2006) use the average length of all reviews on the websites of Barnes and Noble, and Amazon. Kampouris and Evans (2013) examine comments that were under 1000 words in length.

funding model used by Kickstarter, i.e. project creator will not receive any money if funding goal is not reached. In using this setting we can observe how the number of words can deter or increase the number of monetary contributions into a project, as both the descriptive text and outcome are discernible. Furthermore, Kickstarter campaigns provide us with a controlled setting, or a real-world laboratory by holding most things equal when creating outcomes (same goal, platform, possibilities and restrictions). Thus, Kickstarter descriptors are the ideal texts to study. We expect that an overabundance of information, quantified by the number of words, will decrease the amount raised and the number of contributors to a project (inverted U-shaped relationship between project text length and funding success). Such non-linearity has not been explored in former studies (Evers, 2012; Geva, Barzilay and Oestreicher-Singer, 2016; Crosetto and Regner, 2014; Du et al., 2015).

## **2.2 DATA AND METHODOLOGY**

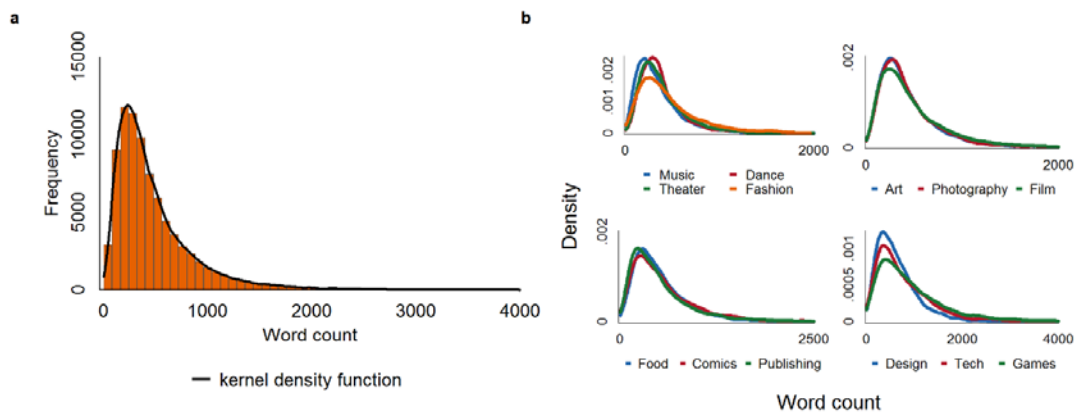
This study uses a dataset comprising detailed information on all Kickstarter campaigns during April 2009 and May 2013 (Neight-Allen, 2013). During this time period, 87,265 projects were created across all the available categories, however 3,851 were removed as these projects were still ongoing (finished after the scrape date). A number of projects were found to be duplicates, and consequently 203 projects were dropped. A further 1,319 observations were excluded as the projects were cancelled or suspended, or the descriptions contained less than three words. The resulting dataset contains 81,892 projects<sup>7</sup> and spans from 21<sup>st</sup> April 2009 to the 29 April 2013. Each observation contains information related to the project, such as the campaign's full description text, project and creator identifier, category, funding goal, project geographic location, and the launch date. The outcome of the fund raising campaign, namely, overall amount raised (in US dollars) and the number of backers who contributed funds to the project, are also recorded. Of the projects within this dataset, 37,962 (46.36%) succeeded in raising the goal amount, whilst 43,930 (53.64%) failed to achieve their funding goal. Table A1.1 provides the descriptive statistics of the variables used in the following analysis.

---

<sup>7</sup> Mollick (2014) excluded projects with more than 2.5 million or lower than 10 cent fundraising goals, as they were considered non-serious. We retain these projects within the data set as removing them does not alter the results.

### 2.2.1 Project description length

In order to quantify the amount of information provided by the project creator, we use the total word count in the project description section. On average, the length of a typical Kickstarter project is approximately 500 words ( $SD = 465.61$ ). As one would expect, the distribution of word count varies across different project categories, e.g., a project in Technology has, on average, more detailed information than a project in Art. The summary statistics of project description word count are presented in Table 2.1. Project description length of Games, Technology and Design are the longest amongst all categories, ranging from an average of 639 to 954 words, while Music, Dance and Theater has the least amount of words, with 353 to 400 words on average. Moreover, the distribution of project description lengths is right skewed (see Figure 2.1), which is similar to the lognormal distribution observed in other forms of Internet-mediated text such as, emails, Internet discussion forums and Wikipedia articles (Sobkowicz et al., 2013). Noticeably, two projects in the Publishing category have more than ten thousand words as the creators provide sample chapters for potential backers. Although these outliers are genuine, they might heavily affect the estimates of the non-linear relationship between text length and funding outcomes of interest. Therefore, we decided to censor (winsorize) the word count variable at the top 99th percentile for each category for later analysis (see Table 2.1). Specifically, the value of the top 1% longest projects in each category is replaced with the value of the 99th percentile of the sample.



**Figure 2.1** Distribution of the total word count across projects (a) and by category (b). The right hand tail of each distribution is truncated for the purposes of visualisation.

**Table 2.1** Summary statistics of non-censored word count.

<i>Category</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	99 <sup>th</sup> percentile	<i>Max.</i>
Music	18909	352.78	276.66	4	1353	4078
Dance	1146	393.84	297.79	4	1424	4112
Theater	4003	400.10	285.71	5	1567	2572
Art	7518	447.46	354.18	4	1761	5535
Photography	2565	451.38	338.94	6	1766	4082
Fashion	2597	463.11	355.57	4	1633	3275
Film & Video	21538	498.98	418.53	4	2084	5361
Food	2973	531.99	400.60	5	1976	4975
Publishing	9237	547.88	666.95	4	2528	32135
Comics	2190	569.24	486.02	9	2402	5137
Design	3317	638.84	443.23	11	2167	4469
Technology	1762	780.99	610.68	8	2907	5062
Games	4137	954.48	784.61	5	3986	5370
Total	81892	494.90	465.61	4	2230	32135

*Note:* We censored the word count variable at the 99<sup>th</sup> percentile for each category.

### 2.2.2 Description edits and potential endogeneity bias

Throughout the fund raising campaign, project creators are able to edit some details of the project, including the description (text, images and videos) or add additional rewards. Goal and project deadline are not editable after launch. Furthermore, information cannot be modified once the campaign has ended. Kickstarter also provides creators a separate ‘Project updates’ tab to communicate with (potential) backers, providing updated information. Nevertheless, there have been occasions where the creator has edited the content of the description during the campaign. Unfortunately, the number of such edits of the description text were not documented. Therefore, potential endogeneity bias might arise due to causality issues, for example, creator urging potential backers to contribute as the deadline approaches or overfunded projects might include additional rewards to attract more backers.

Venugopal and Bagadia (2015) examines the dynamic effect of edits within Kickstarter campaign texts and the impact that edits may have on success. Based on a smaller sample of 19,299 Kickstarter projects, they observe that 64% of projects had no edits throughout the entirety of the project. The authors report that a vast majority of projects with edits were only edited once or twice and were more likely to occur in the first few days of the campaign. They also note an increase in the number of edits at the end of projects’ duration. The authors believe that these edits could be caused

by creators showing appreciation for backers as the goal has been reached or urging more contributions as they are close to achieving the goal. However, no statistical evidence are provided for these claims. Additionally, the authors provide a metric to measure the significance (size) of the edit but find no statistical significance between the extent of edits and project success.

A second study by Xu, Yang and Rao (2014) examines the updates provided in the *update tab* by the creator *before* the outcome of the campaign is determined. Of the 8,529 projects they examined, 58.6% had at least one update and an update was significant in increasing the chances of success. Using the Latent Dirichlet Allocation (LDA) the authors classify the updates into seven different categories. From most frequent to least frequent type they are: *Social Promotion*, *Progress Report*, *New Content*, *Reminder*, *Answer Question*, *New Reward*, and *Appreciation*. The authors find that *Reminder*, *Progress Report* and *New Reward* updates are the most influential amongst the categories in terms of predicting project success. In order to account for the potential endogeneity bias, we use cues similar to those identified by Xu et al. (2014) in an attempt to identify projects in which edits may have happened. For example, if the description contains words or phrases such as ‘\*ve reached’, we identify as progress as it picks up achievement of goal being fully or partially reached (Progress Report) and ‘days to go’ as a *Reminder*. The full list of search terms are listed in Appendix A1.2. In addition, if text contains the word ‘Update’ or ‘UPDATE’ (case sensitive) but none of the search terms for *Progress Report* or *Reminder*, we classify them as general edits, which might provide information on *New Content* or *New Reward*. Thus, we code our edit indicator as *no edits*, or *general*, *reminder* and *progress* edits.

Whilst we have a significant portion of projects without an identified edit, 6,478 (7.91%) of the projects we examine are identified as containing an edit. Given that Venugopal and Bagadia (2015) determine 36.26 % of their observations contain edits, with a large portion of them being minor edits (e.g. correcting typos), an edit identification of 7.91% seems reasonable. We find that 72.95% of creators who edited their description were successful in obtaining funding, and just over 80% of the projects with a reminder or progress edit achieved funding (see Table 2.2). This is similar to the findings of Xu et al. (2014) where *Reminder* has a stronger effect on

achieving success, followed by *Progress Report* and general updates (*New Content* or *Reward*).

**Table 2.2** Summary statistics of identified edits

	Total	Successful	Unsuccessful	Percentage successful
No edit	75414	33236	42178	44.07
General	3733	2461	1272	65.93
Reminder	592	481	111	81.25
Progress	2153	1784	369	82.86
Total (edit)	6478	4726	1752	72.95

*Notes:* 220 progress projects had been identified with a reminder *and* progress edit. They are coded as a progress given the achievement of the goal would occur after the reminder.

## 2.3 RESULTS

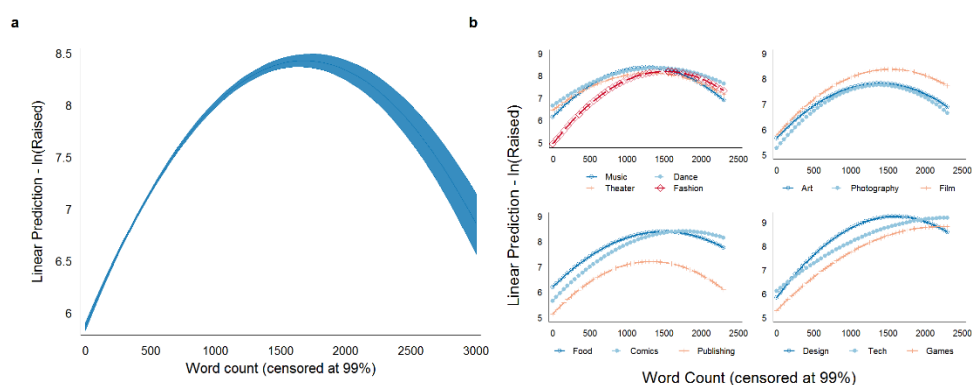
To observe the effects of a creator’s description of the project and the subsequent length on the fundraising success of the project, we conduct a multivariate analysis using OLS regressions on the natural logarithm of the funds raised and number of backers contributing. Standard errors are clustered over project creator.

Under the assumption that an excessive amount of text can negatively affect funding success, the analysis of the text word count controls for non-linearity by including a squared term for the word count. Using a simple regression model we attempt to predict the effects of text length and funding success (Table 2.3). Whilst not accounting for other factors (see specifications (1) and (4)), the coefficient for the linear (quadratic) term of word count is positive (negative) significant at 1% level. For every one word increase in the description, we expect a  $0.0031 - 2*(8.7e-07)*wordcount$  percentage change in the amount raised, and a  $0.0023 - 2*(6.0e-07)*wordcount$  percentage change in the number of contributors. This suggests that having more words in the project description increases the overall amount raised and the number of contributors but with a diminishing or even negative effect when the text gets too long. Whilst it is possible to examine the confidence interval to predict the effect of the word count on the number of backers or funding raised, these confidence intervals should be treated with caution. As exemplified in Figure 2.2 and Figure 2.3, the confidence interval range is different in the lower word count to the upper word count. In total 2% of the projects went beyond the turning point of optimal number of words. The highest rates are observed in the categories *Games* and *Publishing* (6-7%).



In all further specifications we control for the category of the projects, project size (*Goal*), campaign length (*Duration* and *Duration*<sup>2</sup>), project location (*Latitude* and *Longitude*), number of existing or previous projects by the same creator (*Project #*), and average number of projects in the same *sub-category* during the project campaign (*Avg. # project in sub-cat*). Furthermore, in specifications (3) and (6), we controlled for potential description edit (identifiers *Edits*) and exclude projects that are not located in the USA (approximately 7% of the projects are based outside the USA)<sup>8</sup>. Holding other factors constant, the linear term of word count is positive and the squared term is negative while the coefficients are statistically significant at the 1% level. This demonstrates a robust inverted U-shaped relationship between text length and overall funding success, as evidenced in Figure. 2.2 and Figure. 2.3 (panel a).

Furthermore, as the main effect might vary (to different degrees) among project categories, we extended our analysis by examining each category individually using the same structure of specifications (see (3) and (6)). As seen in Table 2.4 and Table 2.5, the main result (inverted U-shape) is evident in each category (statistically significant at a 1% level), but the turning point varies over categories as shown in Figures 2b and 3b. Additionally, the same pattern can also be found within each sub-category<sup>9</sup> (see Figures B1.1 and B1.2).

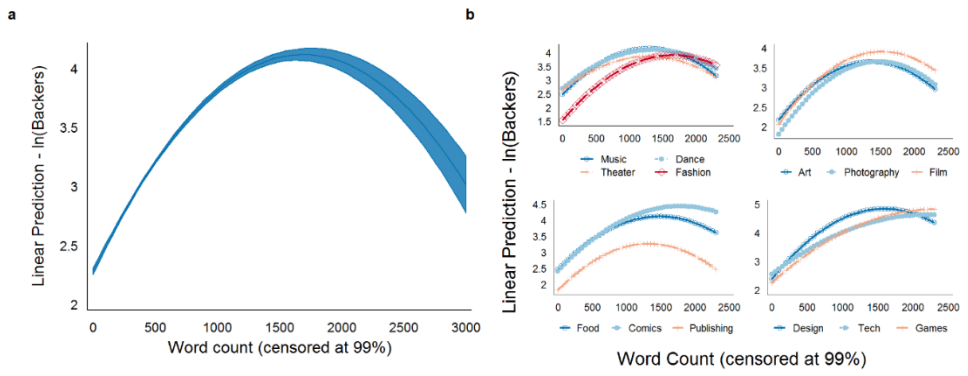


**Fig. 2.2** Marginal effects of word count on  $\ln(\text{raised})$ .

Panel (a) visualises the results of specification (3) in Table 2.3; Turning point is 1681.77. Panel (b) uses the same specification by categories.

<sup>8</sup> Kickstarter officially opened to the UK market on 31<sup>st</sup> October 2012, approximately 6 months before the sample period.

<sup>9</sup> The only exception is *Electronic Music*, where the coefficient of the quadratic term is not statistically significant.



**Fig. 2.3** Marginal effects of word count on  $\ln(\text{backers})$ . Panel (a) visualizes the results of specification (6) in Table 2.3; Turning point is 1689.28. Panel (b) uses the same specification by categories.

**Table 2.3** Multivariate analysis of description length (word count censored at 99%):

	$\ln(\text{Raised})$	$\ln(\text{Raised})$	$\ln(\text{Raised})$	$\ln(\text{Backers})$	$\ln(\text{Backers})$	$\ln(\text{Backers})$
	(1)	(2)	(3) <sup>^</sup>	(4)	(5)	(6) <sup>^</sup>
Word count	.0031*** (55.12)	.0032*** (58.42)	.003*** (52.72)	.0023*** (50.81)	.0024*** (53.95)	.0022*** (48.03)
Word count <sup>2</sup>	-8.7e-07*** (-28.25)	-9.1e-07*** (-29.87)	-9.0e-07*** (-28.77)	-6.0e-07*** (-23.47)	-6.4e-07*** (-26.14)	-6.4e-07*** (-25.38)
Category						
Comics		.023 (0.44)	-.025 (-0.45)		.33*** (7.58)	.26*** (5.98)
Dance		.74*** (14.74)	.81*** (15.81)		.48*** (12.57)	.52*** (13.23)
Design		.65*** (14.23)	.67*** (13.94)		.61*** (15.95)	.6*** (15.18)
Fashion		-.36*** (-7.34)	-.3*** (-5.99)		-.38*** (-10.00)	-.35*** (-9.04)
Film & Video		.13*** (4.01)	.18*** (5.31)		-.043* (-1.81)	-.023 (-0.91)
Food		.36*** (8.11)	.41*** (9.03)		.26*** (7.60)	.29*** (8.18)
Games		-.14*** (-2.72)	-.17*** (-3.15)		.28*** (6.85)	.21*** (5.02)
Music		.53*** (18.30)	.55*** (18.23)		.42*** (18.95)	.42*** (18.22)
Photography		-.24*** (-4.88)	-.24*** (-4.57)		-.22*** (-5.97)	-.22*** (-5.70)
Publishing		-.62*** (-18.34)	-.6*** (-16.81)		-.4*** (-15.26)	-.39*** (-14.32)
Technology		.42*** (7.00)	.42*** (6.81)		.38*** (8.08)	.36*** (7.49)
Theater		.53*** (14.29)	.57*** (14.77)		.35*** (12.11)	.36*** (12.28)
Goal (\$)		1.3e-07 (1.64)	6.2e-08 (0.87)		4.9e-08 (1.10)	1.3e-09 (0.04)
Duration (days)		.014*** (7.56)	.013*** (6.76)		.0061*** (4.24)	.0053*** (3.61)
Duration <sup>2</sup>		-.00018*** (-9.67)	-.00017*** (-8.70)		-.00012*** (-8.33)	-.00011*** (-7.43)
Latitude		.0032*** (3.95)	.014*** (12.24)		.006*** (9.56)	.014*** (15.89)
Longitude		.00032 (1.53)	.0011*** (3.16)		.0008*** (4.83)	.0016*** (5.67)
Project #		-.017* (-1.91)	-.019** (-2.17)		-.01 (-0.67)	-.012 (-0.82)
Avg. # project in sub-cat		.00063*** (6.14)	.00052*** (4.93)		.00038*** (4.52)	.00032*** (3.68)
Edits						
General edit			.89*** (26.12)			.85*** (29.74)
Reminder			1.2*** (17.92)			1.1*** (18.97)
Progress			1*** (26.78)			.98*** (28.96)
N (Obs.)	74665	74665	69754	74665	74665	69754
N (Cluster)	66999	66999	62527	66999	66999	62527
R <sup>2</sup>	0.100	0.138	0.154	0.100	0.141	0.163

Notes: *t*-statistics in parentheses. The reference group for *Project category* and *Edits* are *Art*, and *No edit*, respectively. <sup>^</sup>Projects with U.S. location only. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

**Table 2.4** Multivariate analysis of description length (word count censored at 99%) on  $\ln(\text{Raised})$ 

	Art (7)	Comics (8)	Dance (9)	Design (10)	Fashion (11)	Film & Video (12)	Food (13)	Games (14)	Music (15)	Photography (16)	Publishing (17)	Technology (18)	Theater (19)
Word count	.0031*** (17.30)	.0031*** (10.07)	.0024*** (5.40)	.0043*** (15.40)	.0043*** (13.29)	.0034*** (29.40)	.0029*** (10.47)	.0032*** (13.44)	.0035*** (26.38)	.0036*** (10.05)	.0031*** (18.34)	.0026*** (7.16)	.0026*** (9.76)
Word count <sup>2</sup>	-1.1e-06*** (-11.13)	-8.8e-07*** (-5.60)	-8.7e-07*** (-3.28)	-1.3e-06*** (-9.81)	-1.4e-06*** (-8.16)	-1.1e-06*** (-17.60)	-9.8e-07*** (-6.59)	-7.2e-07*** (-7.46)	-1.4e-06*** (-14.24)	-1.3e-06*** (-6.25)	-1.2e-06*** (-13.52)	-5.6e-07*** (-3.48)	-9.8e-07*** (-5.68)
N (Obs.)	6283	1994	1025	2924	2180	17678	2770	3596	16736	1951	7490	1527	3600
N (Cluster)	5785	1663	915	2623	2054	16186	2612	2982	15580	1832	6967	1428	3218
R <sup>2</sup>	0.112	0.181	0.088	0.232	0.191	0.133	0.119	0.305	0.105	0.137	0.111	0.222	0.081

Notes: t-statistics in parentheses. Marginal effects in italics. In all specifications, we control for *Goal*, *Duration*, *Duration*<sup>2</sup>, *Latitude*, *Longitude*, *Project number*, *Average number of project in the same sub-category*, and *Edits*. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

**Table 2.5** Multivariate analysis of description length (word count censored at 99%) on  $\ln(\text{Backers})$ 

	Art (20)	Comics (21)	Dance (22)	Design (23)	Fashion (24)	Film & Video (25)	Food (26)	Games (27)	Music (28)	Photography (29)	Publishing (30)	Technolog y (31)	Theater (32)
Word count	.0022*** (15.96)	.0023*** (8.85)	.0021*** (5.98)	.0031*** (13.08)	.0029*** (11.56)	.0024*** (28.12)	.0022*** (10.30)	.0022*** (11.75)	.0026*** (24.75)	.0025*** (8.92)	.0022*** (16.81)	.0019*** (6.62)	.0019*** (8.87)
Word count <sup>2</sup>	-8.0e-07*** (-10.32)	-6.4e-07*** (-4.82)	-7.7e-07*** (-3.52)	-9.7e-07*** (-8.10)	-8.7e-07*** (-6.20)	-8.0e-07*** (-16.56)	-7.6e-07*** (-6.29)	-4.9e-07*** (-6.17)	-1.0e-06*** (-12.87)	-8.5e-07*** (-4.98)	-8.3e-07*** (-12.08)	-4.3e-07*** (-3.31)	-7.4e-07*** (-4.99)
N (Obs.)	6283	1994	1025	2924	2180	17678	2770	3596	16736	1951	7490	1527	3600
N (Cluster)	5785	1663	915	2623	2054	16186	2612	2982	15580	1832	6967	1428	3218
R <sup>2</sup>	0.106	0.186	0.102	0.178	0.170	0.141	0.129	0.297	0.113	0.127	0.123	0.201	0.083

Notes: t-statistics in parentheses. Marginal effects in italics. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively. In all specifications, we control for *Goal*, *Duration*, *Duration*<sup>2</sup>, *Latitude*, *Longitude*, *Project number*, *Average number of project in the same sub-category*, and *Edits*. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

## 2.4 DISCUSSION

We examine the influence of information quantity on funding success. Success is measured in terms of funds contributed and the number of backers contributing, and amount of information is measured by the word count of the project description. Using these variables, we are able to demonstrate that a non-linear relation exists between information quantity and funding outcome. In the lower word count range, increasing the number of words tends to have a positive effect on funding. Yet, in all product categories on Kickstarter, there is an optimal number of words, beyond which being too verbose would reduce the project creator's ability to attract contributions and contributors (backers), as evidenced by the inverted U-shape presented in Fig. 2.2 and Fig. 2.3. Using (too) many words (going beyond the turning point which happened in 2% of the cases), regardless of whether the creator is emphasising certain points or providing extra detail, will have a negative impact on the overall success of the project by deterring backers and their funds. This effect could be caused by a perception that extra effort is required to read all the text, or by experiencing an information overload. On the other hand, much shorter texts may not have an appropriate amount of detail to convince the potential contributor that the project is of high quality, or even legitimate, as projects with less words are more likely to be a scam (Shafqat, Malik and Kim, 2016).

Whilst we have attempted to consider the influences of text length cues on decision-making behaviour within a limited cognitive capacity (bounded rational individuals) and increasing fundraising success, our examination is not without limitations. It is highly likely that we have an omitted variable bias. For example, social network sizes, the presence and scope of images and videos, the frequency and timing of updates, as well as spelling errors within text have all been found to significantly affect success (Evers, 2012; Mollick, 2014; Cumming, Leboeuf and Schwienbacher, 2015; Gao and Lin, 2015). Nor have we been able to control for the distinctive personal characteristics of backers, such as age or category interests, which may influence their decisions to fund due to the potential of shaping their information capacity or interest. Although we consider the quantitative aspect of the text, we do not consider the qualitative features contained in the text, which is an aspect that could be extended upon in the future. Finally, we have attempted to

account for a suspected endogeneity bias. However, future studies could work with more precise data monitoring all the changes over time.

## Chapter 3: *Don't keep it (too) simple*

---

### Statement of Contribution of Co-Authors for Thesis by Published Paper


The authors listed below have certified\* that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the Australasian Research Online database consistent with any limitations set by publisher requirements.

In the case of this chapter:

#### Don't keep it (too) simple

July 2016 – In preparation for submission

Contributor	Statement of Contribution
Naomi Moy	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (50%)
	
08/08/2016	
Ho Fai Chan	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (20%)
Benno Torgler	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (30%)

#### Principal Supervisor Confirmation

I have sighted the email or other correspondence from all Co-authors confirming their certifying authorship.

Benno Torgler



08/08/2016

\_\_\_\_\_  
Name

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Date

“It is likewise to be observed, that this society has a peculiar cant and jargon of their own, that no other mortal can understand...”

Jonathan Swift (1841), *Gulliver’s Travels*, p. 68.

### 3.1 INTRODUCTION

Every day we convey and receive a large body of information through various physical and non-physical forms, each with varying levels of complexity. We utilise this information to make decisions that can be integral to our daily lives. Yet it is possible that receiving an overabundance of information can be considered too much of a good thing (Shenk, 2003). This information overload is not uncommon, and is a consequence of the various channels, both physical and virtual, through which the recipient may be bombarded. Information overload is affected by the way we communicate. Specifically, if a body of text is perceived as difficult to read, an individual is less willing to engage in the activity (Song and Schwarz, 2008). Such circumstances can be problematic when the outcome of a situation depends heavily on complete comprehension of a text. A prominent example of this is crowdfunding, where the entirety of the text of the campaign description portrays information about the innovative idea, with the purpose of convincing the reader to contribute funds. If potential contributors give up on reading or fail to understand the text because it is too difficult to comprehend, then it is unlikely the project will achieve success.

For more than 90 years, researchers have been investigating the readability of texts to better understand how people communicate and comprehend information. Research has been conducted across a broad series of textual forms and topics (for an overview see Fry, 1987 and DuBay, 2007). One of the more notable works on readability is an assessment of educational texts used by educators to match learning materials with an individual’s reading ability (Fry, 2002). From a policy perspective, the use of various readability measures to categorise educational resources should improve learning environments by decreasing the level of complexity of learning texts (Allington, 2005; Lockwood, 2005; Kasule, 2011; Burton, 2014; Fatoba, 2015). However one study has found that enhanced learning outcomes are not significantly related to textbook readability (Peng, 2015). While teachers use these measures to improve learning within a classroom setting, health professionals use readability indices to align informative health materials with patients’ reading levels (Elliott and



Shneker, 2009; Wilson, 2009; McInnes and Haglund, 2011; Risoldi Cochran, Gregory and Wilson, 2012). Increasing a patient's comprehension of medical materials resulted in better health outcomes (Baker et al., 2002), and consequently, health care agencies recommend targeting reading levels of a 6th grade equivalent<sup>10</sup>.

Since the late 1940's, readability indices have been used to improve the performance of newspaper articles (Gunning, 1969). In many cases, lowering the complexity levels of newspaper articles increased the number of readers (DuBay, 2007). Readability also matters for manufacturer liability, consumer protection, or even court cases (Wright, 1981; Fry, 1987). If product instructions are misunderstood it can produce unintended and serious consequences such as accidents, or even chemical explosions resulting from incorrect mixing of materials (Fry, 1987).

Misunderstandings can also arise from the complex requirements and terminology used within legal documents. Due to the complicated nature of such forms, it is unsurprising that consumers regularly fail to comprehend legal documents, such as online privacy agreements and financial or insurance contracts (Jensen, Potts and Jensen, 2005). It follows that various legislative bodies are implementing policies around minimum readability requirements (Friman, 1994).

The use of readability formulas to communicate information more effectively has not gone unnoticed within the research areas of advertising, marketing and consumer behaviour. For example, several studies of online reviews demonstrate that products with easier-to-read reviews experience higher sales (Liu and Park, 2015; Fang, Ye, Kucukusta and Law, 2016). In addition, as a review's level of complexity increases, its helpfulness decreases (Wu, 2013). Within text research in advertising (advertising copy) observes a threshold for readability, whereby consumers prefer moderate levels of syntactic complexity (Bradley and Meeds, 2002; Chebat, Gelinas-Chebat, Hombourger, and Woodside, 2003).

---

<sup>10</sup>

[http://www.dhhs.tas.gov.au/publichealth/about\\_us/health\\_literacy/health\\_literacy\\_toolkit/assessing\\_readability](http://www.dhhs.tas.gov.au/publichealth/about_us/health_literacy/health_literacy_toolkit/assessing_readability)

Readability has also been examined within the context of crowdfunding, where conflicting responses with respect to the effect of text complexity and success have been reported. Gao and Lin (2015) find that debt-based crowdfunding projects with simpler descriptions are less likely to default on their loans, and hence are more successful. In line with this result, Evers et al. (2012) finds that more complex campaign descriptions are significantly linked to the project's success. On the other hand, Xu et al. (2014) observe that the overall effect of readability on success is decreasing, and backers appreciate more sophisticated descriptions. In two different studies by Cumming et al. (2015) and Gorbatai and Nelson (2015), the readability levels of the description did not appear to have a significant effect on the outcome of success; however, projects that fail appear to be easier to read. Furthermore, whilst Greenberg et al. (2013), Chung and Lee (2015), and Zhou et al. (2015) incorporate readability indices into their analysis, they do not report specific results.

We present an analysis of text readability and project success using data collected from the popular crowdfunding website Kickstarter. This platform uses an all-or-nothing funding model, where creators do not receive funding unless the goal is reached. Given the entirety of the project is communicated over the Internet, a significant proportion of the information provided is through text based cues. Such a setting is highly controlled and approximates a real-world laboratory as project creators have similar motivations or the same goal (get funding success) and act under the same restrictions (same platform and rules). Thus, we have a relatively clean environment and a large number of projects (more than 70,000) to analyse whether and how readability affects the level of funds raised and the number of contributors (backers).

It is possible that we may observe a non-linear effect. If the text is too simple, the idea may not seem like such a novelty, and if it is too complicated it may be too difficult to comprehend. As a crowdfunding project is reliant on the general public for financial support and subsequent success, creators really need to be able to communicate appropriately in order to attract potential contributors. These potential contributors rely on the readability of the project text in order to form their pledging decisions. Thus, providing empirical insights is valuable to better understand how to communicate more efficiently and appropriately to maximise the chances of success.

## 3.2 DATA & METHODOLOGY

### 3.2.1 Kickstarter data

In May 2013, detailed information was collected from all 87,265 Kickstarter campaigns that occurred from April 2009 to May 2013 (Neight-Allen, 2013). There were 3,851 live projects upon collection, and 1,522 projects were either duplicates, suspended, or cancelled. As a result these projects were dropped from the dataset. The final dataset spans from the 21st of April 2009 to the 29th of April 2013 and contains 81,892 individual observations. Project information was also collected, including the full descriptive text of the campaign, identifiers for the creator and project, the launch date, funding goal, the project's geographic location, and category. The final outcome of the amount raised and the number of backers are also recorded. In total, of the projects in our dataset over \$466,000,000 was raised by successful projects from approximately 6,412,600 contributors, with each project raising on average \$6,423.21 from roughly 89 backers (see Table 3.1 for descriptive statistics).

### 3.2.2 Method

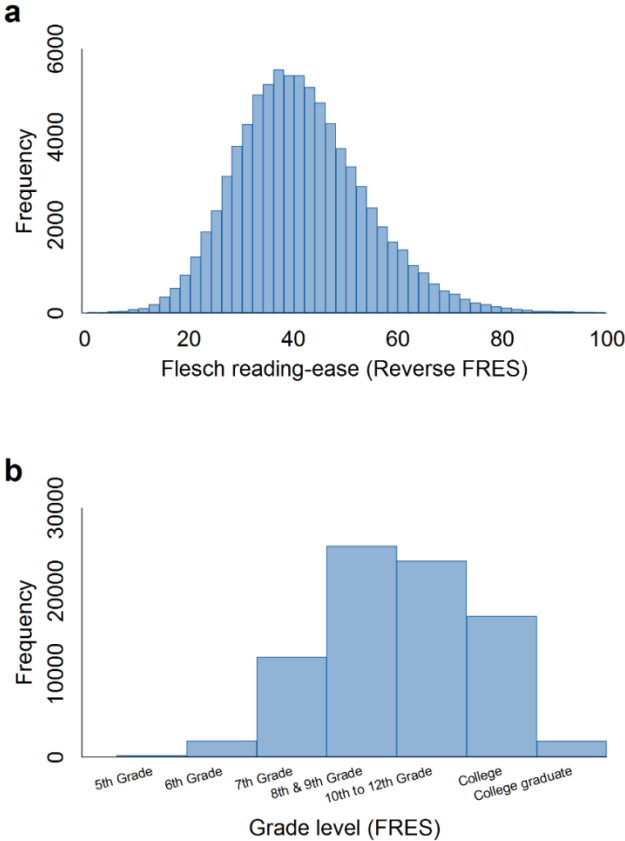
To measure reading ease of the written content in Kickstarter projects, we first use the Flesch Reading Ease Scale (FRES) to calculate text complexity scores in the project description section (Flesch, 1948). The FRES is the most cited measure according to Google Scholar (see Table A3.1), and takes into account the number of syllables per word and average sentence length of a written passage and is formulated as follows:

$$FRES = 206.835 - 1.015 * \left( \frac{Word}{Sentence} \right) - 84.6 * \left( \frac{Syllable}{Word} \right) \quad (3.1)$$

where *Word*, *Sentence*, and *Syllable* are equal to the total number of words, sentences, and syllables in a piece of text, respectively. For most text samples, the resultant score would range from 0 to 100, where 0 being very difficult to read and 100 being easiest to understand. Flesch (1979) propose a rough translation of the FRES into the school grades which the text can be easily comprehend. For the calculation of readability of all Kickstarter project description text we use the R

package *quanteda* (Benoit and Nulty, 2013). In our analysis, we reverse the original scale such that higher values indicate that the text is more difficult to read<sup>11</sup>. In Figure 1 we present the distribution of the *FRES* and the corresponding grade level of all Kickstarter project descriptions.

The average readability level of a Kickstarter project using FRES was 41.58 which is the equivalent of an 8<sup>th</sup>-9<sup>th</sup> year grade level (see Figure 3.1 and Table 3.1). There is a large variation in the FRES values across different projects, demonstrating thus a large variation in the writing style of creators with projects ranging between a minimum FRES score of 6.38 (very easy) and a maximum of 103.44 (very difficult).



**Figure 3.1** Distribution of the (reversed) Flesch Reading Ease score (a) and the corresponding grade level (b) of Kickstarter project description.

<sup>11</sup> A total of 21 (and 69) projects have a Flesch Reading Ease scores below 0 (and above 100), therefore, we replace the top and bottom 0.1% observations with the values of the 0.1<sup>th</sup> (6.38) and 99.9<sup>th</sup> (103.44) percentile.

**Table 3.1** Summary Statistics

	N	Mean	SD	Min	Max
Backers		88.57	843.28	0	91585
Raised (\$)		6423.31	68773.82	0	10300000
Flesch Reading Ease Score ( <i>FRES</i> )		41.58	12.69	6.38	103.44
Grade level ( <i>FRES</i> )		4.57	1.12	1	7
Word count		494.90	465.61	4	32135
Goal (\$)		15276.92	221083.8	.01	21500000
Duration (days)		37.42	16.03	1	91.95
Avg. # project in subcategory		120.58	85.46	1	467.63
Funded					
No	43930	.54			
Yes	37962	.46			
U.S. dummy					
No	5729	.07			
Yes	76613	.93			

### 3.3 RESULTS

In this section, we perform a multiple regression analysis to estimate the effect of readability of the Kickstarter project description on its subsequent funding outcome. The dependent variables of our OLS regressions are the natural logarithm of the funds raised (see specifications (1) to (3) in Table 2 and (7) to (9) in Table 3) and the number of backers contributing (specifications (4) to (6) in Table 2 and (10) to (12) in Table 3). The main independent variables of interest are the raw Flesch Reading Ease scores (Table 3.2) and the corresponding grade levels (Table 3.3). We employ as controls the total word count of the description text to take into account the *quantity* of information potential backers are exposed to (specifications (2), (5), (8) and (11)). We also control for project specific variables such as project category, funding goal, and project duration and location (geographic coordinates). To take into account prior experience on Kickstarter, we create a variable that captures the number of previous Kickstarter projects of a creator. Furthermore, to account for the potential project competition or a field specific interest (e.g., level of general innovation, attention or interest), we use the average number of Kickstarter projects in the *same project sub-category* that were running during the entire duration of each project.

The results show that both funds raised and number of contributors significantly increase with the complexity of the project description. When

considering just the complexity of the campaign description (specifications (1) and (4)), an additional increase in the FRES score is expected to raise a further  $0.0054 - 2 \times (4.6e04) \times \text{FRES}$  percent of funds or increase the number of backers by  $0.0034 - 2 \times (3.1e-04) \times \text{FRES}$  percent. The negative coefficients of FRES squared term in all specifications suggest a curvilinear relationship, i.e. dependent variable increases at a diminishing rate (significant at 1% level, see Figure 3.2). However, of all Kickstarter projects, only 508 (406) projects lie beyond the turning point of 78.92 (80.73) which suggests that the negative effect of high complexity in description text on funds raised (number of backers) is small. This is supported by the results reported in Table 3.3, where the FRES is translated into grade level. All else being equal, projects with description deemed to be easily understood by college students (or higher) have, on average, 17 to 22 (5 to 13) percent more funds raised (backers) compared to those with reading ease level of 8th and 9th grade. On the other hand, projects with readability of 6th grade level have 30 percent less funds raised and 19 less backers compared to the reference group.

**Table 3.2** Multivariate analysis using the raw FRES Score

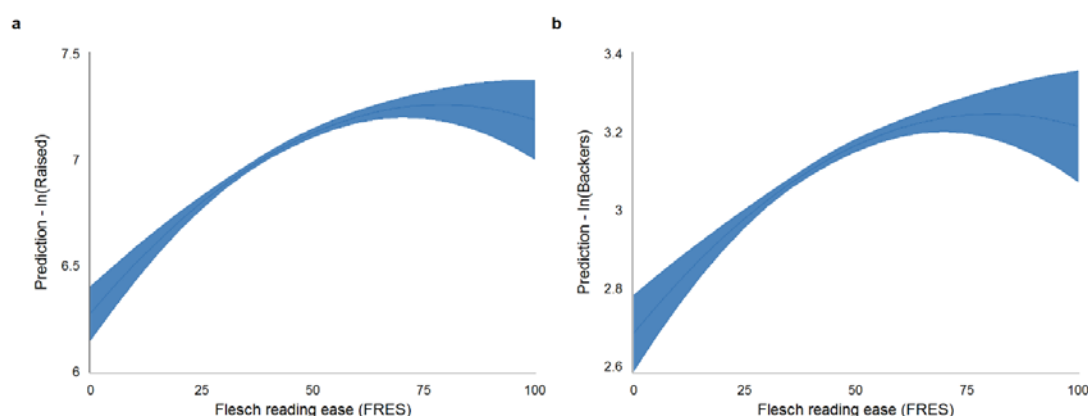
	$\ln(\text{Raised})$ (1)	$\ln(\text{Raised})$ (2)	$\ln(\text{Raised})$ (3) <sup>^</sup>	$\ln(\text{Backers})$ (4)	$\ln(\text{Backers})$ (5)	$\ln(\text{Backers})$ (6) <sup>^</sup>
FRES	0.054*** (18.73)	0.022*** (7.60)	0.025*** (8.59)	0.034*** (15.18)	9.2e-03*** (4.14)	0.014*** (6.28)
FRES <sup>2</sup>	-4.6e-04*** (-15.09)	-1.4e-04*** (-4.55)	-1.6e-04*** (-5.07)	-3.1e-04*** (-13.11)	-6.3e-05*** (-2.63)	-8.6e-05*** (-3.64)
Word count		3.0e-03*** (53.53)	3.2e-03*** (55.31)		2.2e-03*** (49.79)	2.4e-03*** (51.33)
Word count <sup>2</sup>		-8.4e-07*** (-27.28)	-9.2e-07*** (-28.35)		-5.9e-07*** (-22.94)	-6.6e-07*** (-25.18)
Category						
Comics			0.11* (1.94)			0.37*** (8.13)
Dance			0.7*** (13.71)			0.45*** (11.48)
Design			0.71*** (14.83)			0.62*** (15.66)
Fashion			-0.25*** (-5.02)			-0.32*** (-8.35)
Film & Video			0.19*** (5.62)			-0.015 (-0.58)
Food			0.42*** (9.15)			0.29*** (8.07)
Games			-0.032 (-0.61)			0.32*** (7.33)
Music			0.65*** (21.24)			0.48*** (20.64)
Photography			-0.24*** (-4.67)			-0.22*** (-5.81)
Publishing			-0.57*** (-16.24)			-0.38*** (-13.94)
Technology			0.41*** (6.58)			0.35*** (7.22)
Theater			0.52*** (13.55)			0.33*** (11.18)
Goal (\$)			6.0e-08 (0.82)			-1.2e-09 (-0.03)
Duration (days)			0.012*** (6.20)			4.6e-03*** (3.13)
Duration <sup>2</sup>			-1.7e-04*** (-8.46)			-1.1e-04*** (-7.26)
Latitude			0.013*** (11.71)			0.014*** (15.34)
Longitude			8.6e-04*** (2.44)			1.4e-03*** (4.98)
Project #			-0.015* (-1.75)			-9.1e-03 (-0.61)
Avg. # project in sub-cat			5.3e-04*** (4.94)			3.1e-04*** (3.57)
N (Obs.)	74665	74665	69754	74665	74665	69754
N (Cluster)	66999	66999	62527	66999	66999	62527
R <sup>2</sup>	0.008	0.104	0.142	0.004	0.101	0.142

Notes: *t*-statistics in parentheses. The reference group for *Project category* is *Art*. <sup>^</sup>Projects with U.S. location only. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

**Table 3.3** Multivariate analysis of the Grade level based on FRES

	$\ln(\text{Raised})$ (7)	$\ln(\text{Raised})$ (8)	$\ln(\text{Raised})$ ^(9)	$\ln(\text{Backers})$ (10)	$\ln(\text{Backers})$ (11)	$\ln(\text{Backers})$ ^(12)
Grade level (FRES)						
5th Grade	-1.33*** (-7.58)	-0.44** (-2.43)	-0.43** (-2.50)	-0.84*** (-6.49)	-0.16 (-1.23)	-0.18 (-1.38)
6th Grade	-0.66*** (-12.20)	-0.24*** (-4.46)	-0.30*** (-5.70)	-0.45*** (-10.82)	-0.13*** (-3.13)	-0.19*** (-4.71)
7th Grade	-0.29*** (-11.48)	-0.14*** (-5.74)	-0.16*** (-6.70)	-0.17*** (-8.24)	-0.05*** (-2.59)	-0.08*** (-4.11)
10th to 12th Grade	0.10*** (5.09)	0.08*** (4.07)	0.10*** (5.13)	0.05*** (2.99)	0.03* (1.86)	0.06*** (4.20)
College	0.17*** (7.75)	0.20*** (9.64)	0.22*** (10.16)	0.05*** (2.86)	0.07*** (4.48)	0.13*** (7.69)
College graduate	0.11** (2.21)	0.27*** (5.48)	0.33*** (6.22)	0.01 (0.24)	0.13*** (3.49)	0.22*** (5.43)
Word count		3.0e-03*** (53.66)	3.2e-03*** (55.40)		2.2e-03*** (49.81)	2.4e-03*** (51.37)
Word count <sup>2</sup>		-8.5e- 07*** (-27.37)	-9.2e- 07*** (-28.42)		-5.9e- 07*** (-22.96)	-6.7e- 07*** (-25.23)
Controls	No	No	Yes	No	No	Yes
N (Obs.)	74665	74665	69754	74665	74665	69754
N (Cluster)	66999	66999	62527	66999	66999	62527
R <sup>2</sup>	0.008	0.103	0.142	0.004	0.101	0.142

Notes: *t*-statistics in parentheses. The reference group for *Project category* and *Grade level* are *Art*, and 8<sup>th</sup> & 9<sup>th</sup> *Grade*, respectively. ^Projects with U.S. location only. The symbols \*, \*\*, \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.



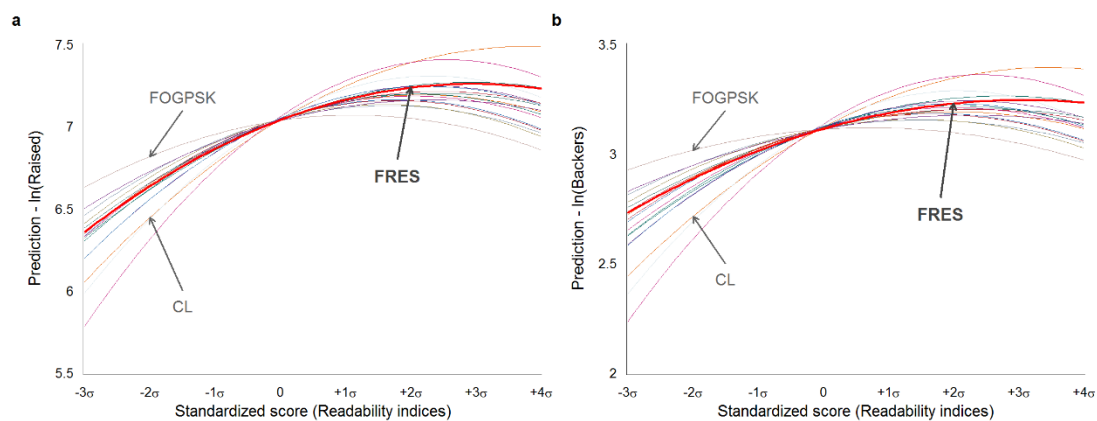
**Figure 3.2** Fitted models of readability (Flesch reading-ease) with crowdfunding success: Fund raised (a) and Number of backers (b)

### 3.3.1 Robustness checks using other readability indices

In addition to the Flesch Reading Ease Scale, we test the robustness of our results by replacing *FRES* with other readability measures which use different or additional



linguistic features to estimate how easily written text can be understood. In Table A3.1 we provide a summary of the readability measures used for robustness checks<sup>12</sup>. All readability measures calculated using the Kickstarter sample are highly correlated with the Flesch Reading Ease scores (see Figure A3.1). The Pearson's correlations between the FRES and other readability measures report ranges from 0.73 ( $DB_2$ ) to 0.99 ( $FRES_{PSK}$ ), with an average of 0.86. To compare all models, we standardise each readability scores with a mean equal to 0 and standard deviation of 1. The distribution shape of the readability scores are very similar (see Figure A3.2). We retest the model in specifications (3) and (6) with each standardized readability score. The coefficients for the linear (quadratic) component of each readability score are positive (negative) and statistically significant at 1% level. Centering (standardization) the readability measures allows us to compare the estimated effects on funding outcome at the mean level; we observe consistent non-linear effects for all the measurements, although the degree of non-linearity varies (Figure 3.3). The number of projects beyond the turning point (optimal level of complexity) varies between 0.17 (CL) and 10.28% ( $FOG_{PSK}$ ) for the amount raised and between 0.27 and 18.08% for the number of backers (CL and  $FOG_{PSK}$  again, see Table A3.2).



**Figure 3.3** Fitted models of fund raised (a) and number of backers (b) vs. readability indices

### 3.4 DISCUSSION

Readability measures are quite often used as checks to demonstrate that a given piece of writing affects the audience in the way intended by the author. In this study we use

---

<sup>12</sup> Dubay (2007) provides an excellent review on the construction, development and application of different readability measures.

readability measures to determine whether levels of complexity affect funding success and the level of participation in a crowdfunding setting. Einstein once pointed out in a much-quoted sentence: “everything should be as simple as possible, but not simpler”. What we observe here is that the level of complexity in a campaign description does indeed affect the outcome in terms of number of backers contributing and the funds raised. However, if complexity increases, then the likelihood of attracting more backers and funds increases, but not at a continuous rate – as evidenced by the non-linearity shown in Figure 2 and 3. After a certain level of complexity, readers might struggle to comprehend the information which may reduce the willingness to fund and participate. However, very few projects go beyond the optimal level of complexity. This could indicate that potential contributors put effort into understanding Kickstarter projects. It could also be that complexity is seen as indicator of quality or innovation (“if it is too simple anyone can do it!”).

Whilst we have endeavoured to consider the effects of readability on communicating information insofar as it affects decision-making processes and specific project success, our analysis is not without limitations. The *FRES* levels and grade level equivalents should be treated with caution and used as rough guides rather than the rule (Klare, 1974) although we should note that other proxies used in the robustness checks show similar results. Moreover, the responses of readers to complexity levels are dependent on the readers’ preferences and knowledge. If the material is within the readers’ competence and interest, their threshold of handling and accepting complexity is higher (in our figures this lies more to the right-hand side). Thus, complexity becomes less important (Smith and Senter, 1967; Putrevu, Tan and Lord, 2004) or, at least, less of a problem. This could explain why very few projects surpass the turning point of optimal complexity. Future research could examine the preferences, interests, knowledge and skillset of the readership, collecting background characteristics of backers or more detailed, dynamic information such as scrolling times and number of unique pages viewed.





## Chapter 4: *Don't talk about money*

---

### Statement of Contribution of Co-Authors for Thesis by Published Paper


The authors listed below have certified\* that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the Australasian Research Online database consistent with any limitations set by publisher requirements.

In the case of this chapter:

#### **Don't talk about the money**

July 2016 – Submitted manuscript currently under review

<b>Contributor</b>	<b>Statement of Contribution</b>
Naomi Moy 	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (50%)
08/08/2016	
Ho Fai Chan	Performed the data analysis and interpretation, and provided critical revisions. (30%)
Markus Schaffner	Developed the study concept, collected the data, and provided critical revisions. (10%)
Benno Torgler	Developed the study concept, performed the data analysis and interpretation, and drafted the manuscript. (10%)

#### Principal Supervisor Confirmation

I have sighted the email or other correspondence from all Co-authors confirming their certifying authorship.

Benno Torgler



08/08/2016

Name

Signature

Date

“We can be emotionally engaged by many kinds of texts, and therefore motivated to read them.”

Lea and Webley (2006, p. 166).

#### 4.1 INTRODUCTION

Because both its physical and non-physical transfer can trigger strong emotive reactions and behavioural responses (Sanfey, Rilling, Aronson, Nystrom and Cohen, 2003; Jiang, Chen and Wyer, 2014), money has become more than just an instrument for obtaining valued goods<sup>13</sup>. In fact, its association with beliefs, hopes and fears (Trachtman, 1999) has generated a growing body of literature on the psychology of money (Lea, 2008). Nevertheless, although money is a key instrument in exchanges, it is less efficient in other contexts, including one-way transfers or grants (Boulding, 1981) and activities like gift giving and blood donation. In the latter, the introduction of monetary compensation can even crowd out the supply of donated blood (Titmus, 1970) by reducing the intrinsic motivation to perform voluntary acts or civic duties (Frey, 1997), whereas the inclusion of non-monetary incentives increase the propensity of blood donation (Lacetera, Macis and Slonim, 2013). It remains unclear, however, whether an over-emphasis on money can also crowd out cooperation in the business environment; for example, by sending the buyer an unintendedly unpleasant signal that leads to a crowding out of contributions and support. Yet little existing research examines the link between money saliency and behaviour, focusing instead on the associations between money and emotions. For example, to assess whether increased earnings increase happiness, some studies examine behaviours like reciprocity or utilitarianism as a result of money priming (see Blanchflower and Oswald, 2004; Tong, Zheng and Zhao, 2013; Malcman, Rosenboim and Shavit, 2015). Only Jiang et al. (2014) and Vohs, Mead, and Goode (2006) use an experimental laboratory setting to examine the effects of money priming on decision makers, which Vohs et al. (2006) identify as to enhance individualism and to reduce collectivism. In all such research, however, it is difficult to find reliable data source that documents both behaviour and the saliency of money.

We overcome this difficulty by taking advantage of the increasing practice of crowdfunding in which interested parties (whether individual or corporate) jointly

---

<sup>13</sup> Money is generally defined as “a medium that can be exchanged for goods and services and is used as a measure of their values on the market” (<http://www.thefreedictionary.com/money>).

participate in an online community to contribute funds for an idea (Mollick, 2014; Profatillov, Bykova and Olkhovskaya, 2015). These contributions are often given on an individual basis enabling large amounts to be raised from numerous small investments (Profatillov et al., 2015). This non-traditional revenue-raising model unlocks the investor market in such a way that project creators can raise capital they might otherwise have been unable to access. It thus helps new creations to market, which can spur further novel production because many result from or enable cumulative innovation (Breedon, 2012; Bruton, Khavul, Siegel and Wright, 2015). In some instances, even established companies use crowdfunding platforms to initiate capital for new products or test the market for product viability (Belleflamme, Lambert, and Schwienbacher, 2014; Belleflamme, Omrani, and Peitz, 2015).

One unique aspect of the crowdfunding revenue-raising model is that it does not necessarily perceive rewards and donations as traditional buyer-seller relationships, particularly when the contributor is buying into a product that does not yet exist (Belleflamme, Lambert, and Schwienbacher, 2013) with the expectation that it will materialize later. In some cases, if participants do not cognitively process idea-related information and engage in deliberate evaluation, donations may even be based simply on empathetic feelings (Dickert, Sagara and Slovic, 2011). Other participants, however, may consider the costs and benefits rather than focusing solely on social interactions or emotions like empathy (Gasiorowska, Zaleskiewicz and Wygrab, 2012). Hence, in the intrinsically social crowdfunding milieu, where benefits accrue from collective decision-making (Agrawal, Catalini and Goldfarb, 2014; Bruton et al., 2015), creators must consider the context in which contributors are active. For example, small contributors might be more attracted to the entrepreneurial idea itself than to the intention to make money, particularly as there is no guarantee that the project will succeed beyond the fundraising stage. In such an environment, an overemphasis on money may backfire. Likewise, many projects are created by new inventors, who are proposing completely new innovations with unidentifiable risks and/or may lack the expertise to set up a successful business venture.

As a result, when creators are priming potential contributors by describing their planned projects, the wording they use is especially important for triggering the desired behavioural responses. In particular, if they are to attract backers, project

creators must avoid such pitfalls as over-emphasizing money, which may signal self-interest and reduce public generosity (Kuang, Weber and Dana, 2007; Xiao and Houser, 2009; Reutner and Wänke, 2012<sup>14</sup>). We leverage the importance of these project descriptions to identify how the prevalence of money terms affects subsequent behaviour and the success of the fund-raising effort. To do so, we use an innovative data set from the crowdfunding website Kickstarter, a pioneer in fixed-goal fundraising through non-equitable investment (Gerber and Hui, 2013; Voelker and McGlashan, 2013; Belleflamme et al., 2014). First implemented in the U.S. in April 2009 and still one of the most successful crowdfunding platforms available, as of February 2016, Kickstarter had launched 283,769 projects, of which 100,585 were successfully funded<sup>15</sup>. Any project created through Kickstarter must have a finite goal and raise capital mainly through the pre-purchasing rewards model, in which the rewards are restricted, directly project related, and subject to a specified expected delivery date.

Because not all projects succeed past the fundraising stage, a key part of a Kickstarter project is the campaign page, which details the idea or project being created, the plans for its production and delivery, and the available rewards. This campaign page also allows a buyer, whose greatest risk is project failure and loss of capital contribution, to interpret the signals and characteristics of the business and assess whether the return is worth the risk. By documenting all these interactions online, Kickstarter facilitates identification of the links between money salience and funding success. In particular, because it encompasses a large set of like structured and similarly restricted projects (e.g., description length, collection time), the data allows us to directly measure whether money saliency in creator communication affects potential investors' engagement in crowdfunding.

---

<sup>14</sup> Using a coordination game within a laboratory setting, Kuang et al. (2007) demonstrate that the effectiveness of advice is dependent on the perceived motivation of the source of the advice, which Xiao and Houser (2009) link to a reduction in generosity from dictators with *ex ante* monetary requests. Reutner and Wänke (2012) investigate the reminders of physical money (touch money) within a laboratory setting, detecting higher levels of self-interest and a lowering in social behaviour.

<sup>15</sup> See <https://www.kickstarter.com/help/stats?ref=footer>



## 4.2 DATA AND METHODOLOGY

Because crowdfunding via the Internet gives access to a much wider community of potential creators and contributors while instantly relaying project information, it not only allows creators to raise capital within their own social spheres and across the globe (Agrawal et al., 2015; Mendes-Da-Silva et al., 2016) but provides complete documentation of the creator/contributor exchanges in each project. For this analysis, in May 2013, we collected detailed project information on all Kickstarter campaigns launched between April 21, 2009, and May 2013 (Neight-Allen, 2013), a total of 87,265 projects across all available categories during that time period. Subsequently, 17,666 projects were dropped because the projects were either cancelled or contained data inconsistencies, leaving a data set of 69,599 observations (April 2009 to November 2012). The final data set contains the number of funders and total amount raised together with the descriptive text from each project's campaign page. It also includes project start date, duration, geographic location, goal amount, and identifiers for the project creator and project category. Using the project descriptive text, we measure the level of emphasis on money in each project by counting the number of 'money' terms in the Linguistic Inquiry and Word Count 2007 (LIWC) dictionary of money category, which is then normalized by the text length (total word count). Projects with a descriptive text length less than three were excluded from the analysis.

As Table 4.1 shows, 37,308 (53.60%) of the 69,599 Kickstarter projects failed to raise the goal amount and 32,291 (46.40%) succeeded in funding their goal.<sup>16</sup> Successful projects raised a combined amount of more than \$353,100,000 from 4,749,499 backers with a further \$44,077,157 designated for unsuccessful projects. The majority of projects (approximately 93%) originated from the U.S. Over all projects, a total of 5,366,721 people (backers) contributed or were willing to contribute funds to purchase the product. The most frequent project category was film or video, followed closely by music. On average, each project raised \$5706.92 ( $SD = 65,384.29$ ) in capital from 77.11 ( $SD = 742.61$ ) backers. A money term

---

<sup>16</sup> Mollick (2014) argues for the exclusion of non-serious projects; that is, those whose goals are unreasonably low (e.g., the 10 cent minimum) or high (the 2.5 million maximum). We retain these projects in the data set because excluding them does not change the results.

appears on average 6.39 ( $SD = 6.61$ ) times in a project and are mentioned within 2.5 percent of a projects text (normalized money term).

**Table 4.1** Descriptive statistics per project

	<i>N</i>	<i>Mean</i>	<i>SD</i>	Min.	Max.
Money term (count)		6.39	6.61	0	292
Money term (normalized)		.016	0.12	0	.25
Duration (days)		38.41	16.61	1	91.95
Before rule change	21535	46.02	21.45	1	91.95
After rule change <sup>a</sup>	48064	34.99	12.49	1	60.04
Backers		77.11	742.61	0	87142
Goal (\$)		13938.2	198062.6	.01	21500000
Raised (\$)		5706.92	65384.29	0	10300000
Word count		470.10	437.21	4	32135
Funded		.46		0	1
No	37308	.54			
Yes	32291	.46			
Official category					
Comics	1784	.03			
Dance	1013	.01			
Design	2468	.04			
Fashion	2007	.03			
Film & video	18972	.27			
Food	2386	.03			
Games	3102	.04			
Music	16510	.24			
Photography	2268	.03			
Publishing	7713	.11			
Technology	1320	.02			
Theatre	3602	.05			
U.S. dummy					
No	3562	.5			
Yes	66037	.95			

<sup>a</sup>In June 2011, Kickstarter adjusted the maximum project duration from 90 days to 60 days (Strickler, 2011).

Notes: Art is the reference group, with 6,454 observations. *N* is 69,599 unless otherwise stated. Geographic location is excluded as non-informative.

To measure the effect of money saliency on investment behaviour, we conduct OLS regression analysis using two different dependent variables: number of backers and total amount raised. Because the textual content of each project is unrestricted, creators can freely discuss project ideas, as well as associated risks and costs. To identify the money terms in this content, we draw on the LIWC, Pennebaker et al.'s (2007) extensively reviewed linguistic program created to capture the social and psychological states of a writer from a text. This corpus contains 80 sentiment dictionaries linked to different emotional and cognitive categories classified under

different dimensions, including emotions, functions and personal concerns (Pennebaker et al., 2007). Our analysis is based on the dictionary of 136 stemmed and descriptive words associated with money (see Appendix C).

We obtain the frequency count of money terms for each project by word stemming the campaign text after numbers, punctuation and tab spaces were removed. The resulting word count is then normalized by text length. We use this score in our OLS multivariate analysis, in which the standard error is clustered over the project creator. Each analytical model controls for the project category (with art as the reference), geographic location, project duration, overall word count, money raised and goal amount.

### 4.3 RESULTS

Our first attempt to predict the association between money and funding success is a simple linear regression focusing only on our key independent variable, money term (Table 4.2). Taking the natural log has restricted the analysis to projects that have raised funds (63,168). With no other influences factored in, the inclusion of a money term reduces both the overall amount raised and the number of contributors, as indicated by the negative coefficients, which are -4.3 and -2.5, respectively, significant at the 1% level. These coefficients indicate that, for each increase in the normalised money term by one unit (from 0 to 1) the amount of funding (number of backers) is likely to decrease by 4.3 percent (2.5 percent). Thus, moving from the minimum (zero percent) to the maximum (25 percent) money terms, in our data decreases the amount of funding by around 1 percent and the number of backers by 0.63%. If we consider the confidence interval, the effect of a one word increase in money terms in the campaign description is likely to decrease the amount raised (number of contributions) by a percentage between -5.83 and -2.93 (-3.66 and -1.51).

**Table 4.2** Simple regression focused on money terms

	<i>ln</i> (Raised) (1)	<i>ln</i> (Raised) (2)	<i>ln</i> (Backers) (3)	<i>ln</i> (Backers) (4)
Money term	-4.3*** (-6.2)	-4.3*** (-5.9)	-2.5*** (-4.7)	-2.5*** (-4.6)
Constant	7.1*** (515)	7.1*** (473)	3.1*** (295)	3.1*** (274)
Cluster	No	Yes	No	Yes
N (Obs.)	63168	63168	63168	63168

N (Cluster)		57071		57071
R <sup>2</sup>	.0006	.0006	.00035	.00035

*Notes:* The *t*-statistic in parentheses. \*, \*\* and \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

We check the robustness of this result by controlling for additional variables related to backers (Table 4.3) and amount of funds raised (Table 4.4). In the first set of specifications ((5) and (11)), we control for project category, the rule change that reduced the time allowed to obtain full funding ((6) and (12)), the duration of the project with duration squared ((7) and (13)), and the goal amount ((8) and (14)). We include the former because Burtch et al. (2013) and Mendes-Da-Silva et al. (2016) show that project duration has a positive effect on both level and number of contributions. The duration change is included as Kickstarter has recommended in their tips, after the duration change that projects run for 30 days. Goal was included in the specifications to control for the size of the project. In all specifications the money term coefficient remains statistically significant at the 1% level. Furthermore, projects in Fashion, Photography and Publishing raised less funds and attracted less backers compared to the reference group (Art project) and vice versa for projects in other categories. Having a shorter duration increases the chances of raising funds and attracting backers. Over increasing the campaign duration, on the other hand, produces a negative coefficient that is consistently statistically significant for number of backers, even though the quantitative effect is not large. Goal was not statistically significant to increasing a projects success in any of the models. Finally, we control for projects' geographic locations using an indicator variable (U.S. or other country) (specifications (9) and (15)) and the geographic coordinate (latitude and longitude) for Kickstarter projects just in the U.S. as over 90% of the project plan descriptions are from the U.S. (specifications (10) and (16)). The coefficients of money terms remain statistically significant at the 1% level in all specifications and the size of the effect on both fund raised and backers attracted are stable. Interestingly, the west coast performs better than the east coast in terms of both fund raised and number of backers, while northern states perform better than southern states.

#### 4.4 CONCLUSION

Our results indicate that an increased use of money terms does indeed reduce project funding success, as measured by both the amount of funding generated and the number of contributors (backers). This negative finding for an over-emphasis on money and the implied crowding out effect could indicate that backers care about the project creator's motivations. Hence, being too money focused, regardless of the benefits to the contributor, has a negative effect on the ability to raise the funds necessary to bring the idea to fruition. This valuing of creators' ideas or intrinsic motivations above their drive to make money may stem from crowdfunding's collaborative setting or perhaps because, on average, backers provide only small contributions. Whichever the case, it also remains to investigate whether an emphasis on money can trigger risk and loss aversion behaviour. In addition, because textual analysis is a relatively new field, it is still difficult to pick up on social cues like irony, emphasis and multi-use words. Our results are also heavily U.S. driven because of the majority of creators' reside there. Nevertheless, Kickstarter is in the process of expanding, so it might be valuable for future research to investigate whether our results hold in other countries and cultures.

**Table 4.3** Multivariate analysis of money terms' influence on the funds of  $\ln(\text{raised})$  (\$)

	(5)	(6)	(7)	(8)	(9)	(10) U.S. only
Money term	-6.019*** (-8.156)	-5.768*** (-7.788)	-5.729*** (-7.780)	-5.705*** (-7.747)	-5.362*** (-7.286)	-5.156*** (-6.807)
Category						
Comics	.1504** (2.45)	.1464** (2.38)	.1393** (2.267)	.1393** (2.267)	.167*** (2.722)	.1737*** (2.711)
Dance	.6207*** (11.5)	.6215*** (11.5)	.61*** (11.321)	.6102*** (11.324)	.6161*** (11.378)	.6644*** (11.913)
Design	1.006*** (17.9)	.9982*** (17.7)	.9763*** (17.390)	.9742*** (17.352)	.9837*** (17.608)	.9826*** (16.877)
Fashion	-.3536*** (-6.14)	-.3633*** (-6.3)	-.3759*** (-6.533)	-.3762*** (-6.540)	-.3499*** (-6.088)	-.3259*** (-5.523)
Film & video	.3009*** (9.68)	.3047*** (9.79)	.2984*** (9.613)	.297*** (9.566)	.303*** (9.780)	.3288*** (9.972)
Food	.5662*** (11.2)	.5585*** (11.1)	.5446*** (10.804)	.5433*** (10.777)	.5723*** (11.359)	.5788*** (11.154)
Games	.4276*** (7.15)	.417*** (6.97)	.3987*** (6.681)	.3959*** (6.631)	.407*** (6.817)	.4131*** (6.652)
Music	.3832*** (12.6)	.3844*** (12.6)	.3715*** (12.177)	.3714*** (12.175)	.3983*** (13.030)	.4241*** (13.218)
Photograp hy	-.1655*** (-3.1)	-.1601*** (-2.99)	-.1633*** (-3.060)	-.1632*** (-3.058)	-.1942*** (-3.659)	-.2169*** (-3.683)
Publishing	-.4665*** (-12.4)	-.4712*** (-12.5)	-.4785*** (-12.710)	-.4791*** (-12.723)	-.4634*** (-12.330)	-.4849*** (-12.283)
Technolog y	.7842*** (10.5)	.7837*** (10.5)	.7705*** (10.416)	.768*** (10.383)	.7745*** (10.513)	.7965*** (10.265)
Theater	.5146*** (13.3)	.5183*** (13.4)	.5131*** (13.331)	.5131*** (13.331)	.5338*** (13.867)	.5513*** (13.732)
Duration change		.06763** * (3.84)				
Duration			.02456*** (11.776)	.02452*** (11.753)	.02435*** (11.710)	.02613*** (11.932)
Duration <sup>2</sup>			- * (-13.820)	- * (-13.807)	- * (-13.759)	- * (-13.981)
Goal				1.26e-07 (0.999)	1.26e-07 (1.001)	1.13e-07 (0.940)
U.S. dummy					-.5056*** (-13.757)	
Latitude						.02724*** (14.749)
Longitude						-.001728*** (-3.731)
Constant	6.839*** (242)	6.789*** (220)	6.407*** (122.540)	6.407*** (122.543)	6.87*** (111.934)	5.114*** (48.599)
N (Obs.)	63168	63168	63168	63168	63168	58471
N (Cluster)	57071	57071	57071	57071	57071	52957
R <sup>2</sup>	.029	.029	.032	.032	.035	.039

Notes: The *t*-statistic is in parentheses; art is the reference group. \*, \*\* and \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.

**Table 4.4** Multivariate analysis of money terms' influence the number of *ln*(backers)

	(11)	(12)	(13)	(14)	(15)	(16) U.S. only
Money term	-	-	-3.043***	-3.032***	-2.805***	-2.691***
	3.279*** (-6.009)	2.888*** (-5.278)				
Category						
Comics	.3774*** (7.525)	.3712*** (7.405)	.3828*** (7.639)	.3828*** (7.639)	.4012*** (8.018)	.4072*** (7.843)
Dance	.3872*** (9.301)	.3885*** (9.317)	.3789*** (9.123)	.379*** (9.125)	.3829*** (9.175)	.4045*** (9.378)
Design	.8203*** (17.886)	.8081*** (17.612)	.8039*** (17.545)	.803*** (17.524)	.8093*** (17.740)	.8008*** (16.876)
Fashion	-	-	-.4287***	-.4289***	-.4114***	-.3898***
	.4111*** (-9.543)	.4261*** (-9.879)				
Film & video	.04137* (1.776)	.04719** (2.024)	.05002** (2.146)	.04937** (2.119)	.05331** (2.289)	.07557*** (3.052)
Food	.3808*** (9.851)	.3688*** (9.522)	.369*** (9.543)	.3684*** (9.527)	.3877*** (10.019)	.392*** (9.863)
Games	.6584*** (13.555)	.6418*** (13.226)	.6411*** (13.216)	.6398*** (13.196)	.6471*** (13.352)	.6362*** (12.658)
Music	.2771*** (11.634)	.279*** (11.734)	.2775*** (11.609)	.2774*** (11.608)	.2953*** (12.295)	.3145*** (12.565)
Photography	-	-	-.1846***	-.1846***	-.2051***	-.2215***
	.1956*** (-4.928)	.1871*** (-4.704)				
Publishing	-	-	-.307***	-.3073***	-.2969***	-.3169***
	.3043*** (-10.609)	.3116*** (-10.841)				
Technology	.6485*** (11.195)	.6476*** (11.219)	.6464*** (11.240)	.6452*** (11.221)	.6496*** (11.343)	.6517*** (10.962)
Theater	.3208*** (10.826)	.3266*** (11.026)	.3208*** (10.868)	.3208*** (10.867)	.3345*** (11.322)	.3405*** (11.053)
Duration change		.1053*** (7.55)				
Duration			.01433*** (9.242)	.01431*** (9.229)	.01419*** (9.173)	.01527*** (9.441)
Duration <sup>2</sup>			-	-	-	-
			.0002038*** (-13.020)	.0002037*** (-13.012)	.0002025*** (-12.955)	.0002163*** (-13.151)
Goal				5.84e-08 (0.988)	5.80e-08 (0.992)	5.23e-08 (0.945)
U.S. dummy					-.3354*** (-11.587)	
Latitude						.02561*** (18.021)
Longitude						-.0008245** (-2.291)
Constant	2.966*** (141)	2.887*** (124)	2.767*** (71.646)	2.767*** (71.644)	3.074*** (65.963)	1.655*** (20.623)
N (Obs.)	63168	63168	63168	63168	63168	58471
N (Cluster)	57071	57071	57071	57071	57071	52957
R <sup>2</sup>	.034	.034	.039	.039	.041	.047

Notes: The *t*-statistic is in parentheses; art is the reference group. \*, \*\* and \*\*\* represent statistical significance at the 10%, 5% and 1% levels, respectively.





## **Chapter 5: Confidence is good; Too much not so much**

---

### **Statement of Contribution of Co-Authors for Thesis by Published Paper**


The authors listed below have certified\* that:

1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student's thesis and its publication on the Australasian Research Online database consistent with any limitations set by publisher requirements.

In the case of this chapter:


**Confidence is good; Too much, not so much**

July 2016 – Submitted manuscript currently under review in *Psychological Science*

<b>Contributor</b>	<b>Statement of Contribution</b>
Naomi Moy	Developed the study concept, collected the data, performed the data analysis and interpretation, and drafted the manuscript. (45%)
	
08/08/2016	
Ho Fai Chan	Developed the study concept, performed the data analysis and interpretation, and provided critical revisions. (25%)
Markus Schaffner	Collected the data, and provided critical revisions. (5%)
Benno Torgler	Developed the study concept, performed the data analysis and interpretation, and drafted the manuscript. (25%)

Principal Supervisor Confirmation

I have sighted the email or other correspondence from all Co-authors confirming their certifying authorship.

Benno Torgler  08/08/2016

\_\_\_\_\_  
Name Signature Date

## 5.1 INTRODUCTION

Although confidence in one's own abilities can be an asset, a large and consistent body of evidence demonstrates that individuals are often over-confident (see, e.g., Myers, 1998, Weinberg, 2009). In fact, as De Bondt and Thaler (1995) point out, over-confidence is “perhaps the most robust finding in the psychology of judgment” (p. 389). One commonly cited illustration of this tendency is that around 70–80 percent of drivers claim to drive more safely than the median driver (Svenson, 1981). Nevertheless, many research findings show that moderate levels of overconfidence could in fact be beneficial and advantageous (Mobius and Rosenblat, 2006 and for an overview Weinberg, 2009). For example, the link between confidence and performance has been explored within the forecasting and trading market (Biais, Hilton, Mazurier and Pouget, 2005; Holmes and Silverstone, 2010). In particular, Koellinger, Minniti and Schade (2007) use survey data to provide evidence that entrepreneurs' over-confidence in their skills and abilities are more likely to lead to market entry. They also observe a negative correlation between confidence and perceived business survival chances. Another closely related study by Herz, Schunk and Zehnder (2014) uses an experimental setting complimented by questionnaires with numerically based answers to explore confidence with innovative activity. The authors observe that different forms and levels of confidence can have contrasting effects on innovative activities, such as overestimating the precision of information leading to lower innovative activities, whilst overestimating success will increase innovative activities. Our paper extends this literature beyond the experimental setting by using real-world data<sup>17</sup> from the crowdfunding web site Kickstarter to examine how confidence affects project success. Because of its extensive documentation of interactions and exchanges, this web site provides easy access to the communication between a wide community of crowd funders willing to both create and contribute to new ideas and projects (Agrawal, Catalini and Goldfarb, 2015; Mendes-Da-Silva, Rossoni, Conte, Gattaz and Francisco, 2016).

Nevertheless, whilst it is felt and experienced on an individual level, confidence is often challenging to measure. In general, it can be operationalized

---

<sup>17</sup> Park and Santos-Pinto (2010), for example, use field surveys to study chess and poker players. Whilst Holmes and Silverstone (2010) use the quarterly business survey data from the New Zealand Institute of Economic Research.

when the frequency of correct judgements can be measured, such as by the proportion of correct answers to a set of questions (Belianin, 2015). On the other hand, some scholars argue that most evidence is based on verbal statements of subjective estimates similar to those reported in Svenson (1981) rather than on the observation of actual choices among alternatives (Hoelzl and Rustichini, 2005). Survey data are especially prone to ambiguous interpretation (Hoelzl and Rustichini, 2005; Herz et al., 2014). It is therefore not surprising that many results are derived in experimental settings such as interaction games (e.g., Mayseless and Kruglanski 1987; Camerer and Lovallo, 1999; Kirchler and Maciejovsky, 2002; Moore and Kim, 2003; Mobius and Rosenblat 2006) or the administration of survey questionnaires with numerically based answers (e.g., Fischhoff, Slovic and Lichtenstein, 1977; Soll and Klayman, 2004; Moore and Healy, 2008; Robinson and Marino, 2015).

One essential strength of working with Kickstarter data is that project success or failure takes place in a very controlled setting, one that approximates a real-world laboratory by holding most things equal when generating outcomes. More specifically, all project creators perform in the same environment under the same conditions or restrictions (e.g., maximum project funding duration) using the same instrument (project description) to achieve the same ultimate goal (funding success). This environment is also highly transparent, and the data produced is reliable because of low variable errors and large data points. As a result, many factors can be held constant, allowing substantial control over the situation, albeit somewhat less than in a laboratory experiment. In contrast to the latter, however, the Kickstarter setting comprises actual processes involving real individuals with strong incentives to be successful (e.g., idea fulfilment, monetary payoff).

We leverage this actuality by using two novel behavioural variables to proxy project creators' confidence at the start of the campaign; namely, the project duration deviation from the maximum allowed time and the level of *certainty* and *tentativeness* found in the campaign description of over 60,000 Kickstarter projects. By doing so, we not only contribute to the growing literature linking self-confidence to actual performance but also extend the somewhat limited evidence on the implications of self-confidence in an entrepreneurial setting. It is currently known, for example, that independent spirits, innovators, or leaders tend to be more self-confident and less likely to imitate peers (Bernardo and Welch, 2001) and that

confidence can promote aggressive strategies (e.g., investment in risky assets) that bring higher than expected profits (Hirshleifer and Luo, 2001). We extend this knowledge by identifying a non-linear inverted U-shaped relation between confidence and crowdfunding success.

## 5.2 DATA AND METHODOLOGY

### 5.2.1 Data

In May 2013, we collected detailed project information on all available Kickstarter campaigns launched between April 21, 2009, and May 2013 (Neight-Allen, 2013) for a total of 87,265 projects across all available categories during that period. Of these, 17,666 projects were dropped because of data inconsistencies, projects which were cancelled and descriptive text sections of less than three words were excluded, which produced a final sample of 69,599 projects (April 2009 to November 2012). The final data set comprises information provided by the project creator, including scheduled duration and overall funding goal, the descriptive text on the campaign page, the project start date, geographic location, and category (field), identifiers for the project creator, and such outcomes as total amount raised, number of backers (funders/contributors), and whether the funding goal was achieved. The average amount raised by each project was \$5706.92 ( $SD = 65,384.29$ ) donated by an average 77.11 ( $SD = 742.61$ ) backers (see Table A3.1). Successful projects raised a combined total of more than \$353,100,000 from 4,749,499 backers, with a further \$44,077,157 pledged to unsuccessful projects.

### 5.2.2 Proxy for confidence: Project Duration

Kickstarter offers interesting ways to derive behavioural confidence proxies, the first of which is linked to project duration. That is, although Kickstarter sets a maximum allowable time for generating the monies required to successfully fund the project, project creators must define their own crowdfunding duration. Hence, deviations from the maximum, which reduce the time available to crowdfund the project, can be seen as a statement of confidence. We therefore compute confidence as

$$Confidence_{DUR,i} = \frac{M_t - x_i}{M_t} \quad (5.1)$$

where  $M$  denotes the maximum allowed duration for period  $t$  and  $x_i$  the duration chosen by the project creator  $i$ . The range for this variable is from 0 to 1, where the higher the value the more confident the project creator is shown to be (distribution shown in Figure B3.1).

The confidence measure using duration is not without problems, as in June 2011, Kickstarter adjusted the maximum project duration from 90 days to 60 days and added a recommended duration of 30 days to the tip manual<sup>18</sup>. The influence of this change on duration is evident in Figure B3.2 in the Appendix, as there is a downward shift in the confidence ratio after the rule change. Prior to these rule changes, however, on July 21, 2010, an external observer demonstrated the need to consider shorter project durations (Mod, 2010). This also resulted in a September 21, 2010 collaborative Kickstarter blog post based on trends in pricing and duration up until that point and the recommendation that projects of 30 days were more likely to succeed<sup>19</sup>. To account for these events, in the analysis section we consider three time periods: before the initial July 2010 blog post (*period 1*), between this blog post and the official June 2011 rule change (*period 2*), and after the rule change on 17<sup>th</sup> June 2011 (*period 3*). What makes this setting interesting is that in *period 2* and *3*, project creators may perceive, *prior* to launching their campaign, projects with a duration of approximately 30 days are more likely to succeed. Thus, this could lead to an anchoring effect in regards to the project length (Tversky and Kahneman, 1974). Prior research shows that an anchoring effect is hard to avoid and difficult to escape from (e.g., Strack and Mussweiler, 1997; Wilson, Houston, Etling and Brekke, 1996). This could be an explanation of the significant increase in the proportion of projects with a 29 to 31 day duration following the initial blog post (see Figure B3.3 in the Appendix). Therefore, a small deviation in the proxy in *period 2* and *3* from the anchoring point (i.e. 30 days) could mean a larger change to actual confidence level compared to the same amount of change in the confidence proxy in *period 1*. Thus, one may observe a more visible effect of the proxy on the fundraised amount or number of backers in the later two periods compared to *period 1*.

Additionally, reducing the maximum duration allowed from 90 days to 60 days could mean that *positive deviations* from the anchor point (30+ days) in *period 3*

---

<sup>18</sup> See <https://www.kickstarter.com/blog/shortening-the-maximum-project-length>.

<sup>19</sup> See <https://www.kickstarter.com/blog/trends-in-pricing-and-duration>

might signal a stronger reduction in confidence compared to that in *period 2*. This is because the maximum range of positive deviation from the anchoring point in *period 2* (60 days) is doubled to that of *period 3* (30 days). Thus, we propose a simple normalization strategy for these two time periods so that confidence ratios are comparable. To make them comparable we multiply the positive deviation in *period 2* from the anchoring point by a factor 2. For example, a project duration length of 33 days in *period 3* is mapped to 36 days and 60 days is mapped to 90 days<sup>20</sup>.

### 5.2.3 Proxy for confidence: Linguistic certainty

Furthermore, we introduce a second proxy relating to a project creators' communication about their projects. Because Kickstarter places no constraints on project description, project creators can freely discuss and explain both the project-related concepts and the associated risks and costs. As such this description could reflect creator's confidence in the idea or project and the project's ability to succeed. To approximate the level of confidence expressed in the text, we draw on Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, and Francis, 2007), an extensively reviewed text analysis program used to capture the social and psychological state of an author within a given text. This corpus includes 80 different dictionaries linked to various emotional and cognitive categories, such as *anger* and *causation* (Tausczik and Pennebaker, 2010). As our second proxy of confidence, we utilize two cognitive mechanism categories from the LIWC2007 dictionaries, namely, *certainty* and *tentativeness* (Tausczik and Pennebaker, 2010), which are represented by a total of 191 associated words and word stems. We took these two measures because certainty can be viewed as an expression of confidence, while tentativeness signals uncertainty (Kahneman and Tversky, 1982). To construct the proxy, we first perform Porter stemming (Porter, 1980) on the description text, with all the numbers and punctuation removed. Then we count the number of occurrences of the word stems of the two dictionaries in the stemmed text for each project. Finally we compute the relative use of words in the certainty dictionary to words in the tentativeness dictionary, which is formulated as follows:

---

<sup>20</sup> We assume a linear scaling factor because of its simplicity. However, one could also consider a non-linear mapping for the positive deviation (e.g. quadratic function).

$$Confidence_{CERT,i} = \frac{C_i}{C_i + T_i} \quad (5.2)$$

where  $C_i$  and  $T_i$  denote the frequency word count of *certain* and *tentative* words in project  $i$ , respectively. Hence, the higher the confidence score the greater the confidence level (see Figure B3.1 for the distribution)<sup>21</sup>. The average confidence rate based on project duration is 0.44 ( $SD = 0.22$ ), while that the self-confidence based on text analysis is 0.44 ( $SD = 0.22$ ).

### 5.3 RESULTS

To measure the effects of creator confidence on project success, we first conduct simple OLS regressions using two dependent variables, namely number of backers and total amount raised (in both cases, using the natural logarithm) in which the standard errors are clustered over the project creator. We complement these regressions applying a probit model that uses project success (target achievement) as the dependent variable. We then analyse each confidence proxy separately. Each analytical model controls for category (with art as the reference), project goal, and whether the project was created in the United States. In additional specifications we extend these controls by introducing the exact geographic location of U.S. projects, and duration change to reflect Kickstarter policy amendments. For the specification with the certainty confidence ratio, we also control for the number of days the project is running and its squared term.

Under the assumption that too much confidence can be problematic, our analysis of duration-based confidence (see Table 5.1 and Table 5.2) controls for non-linearity by including a confidence squared term. The linear term is positive and the squared term is negative demonstrating an inverted u-shaped effect of confidence on success (statistically significant at the 1% level). It also assesses the three time periods separately when  $Confidence_{DUR}$  is used (see Table 5.2). In specifications (1) and (4), we control for project size and/or resource intensity by considering category and project goal (in monetary terms and as set by the project creator). Holding all else constant, an increase in the confidence exhibited by the duration

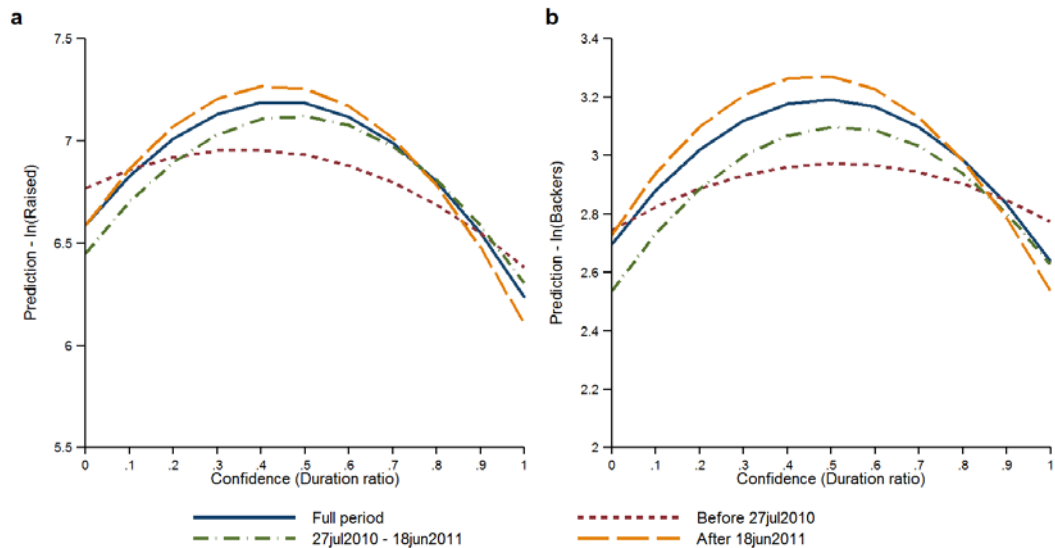
---

<sup>21</sup> 1436 projects do not contain at least one word count in either dictionary; hence, the ratio has not been computed.

ratio is expected to raise (increase backers) by  $2.7 - 2*(3.1)* Confidence_{DUR} (2 - 2*(2.1) * Confidence_{DUR})$  percent. As over 90% of the projects originate in the U.S., we also include a U.S. dummy. In specifications (2) and (5), we add in the duration change dummy, and then in (3) and (6), we focus only on the U.S. by adding the project's geographic coordinates (latitude and longitude). In the subsequent analyses, specifications (7) and (10) address only projects launched before the initial blog post, (8) and (11) are restricted to projects between this initial blog and the duration rule change, and (9) and (12) cover projects that occur after the duration change (see Table 5.2). Nevertheless, because project owners have a specific interest in whether their goal is achieved, we also calculate a probit model on success probability (see Table 5.3) whose marginal effects indicate the strong influence of self-confidence.

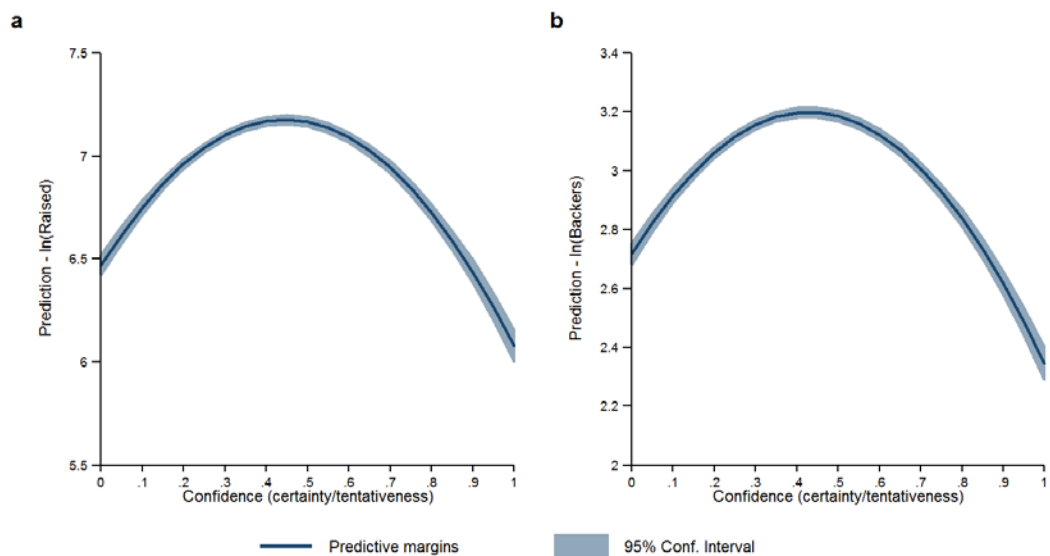
In all specifications, the coefficients for  $Confidence_{CERT}$  and its squared term, used to measure the confidence level evident in the campaign description, are statistically significant at the 1% level (Table 5.4 to Table 5.6). We further control for non-linearity in project duration when the dependent variable is amount raised (Table 5.4) and number of backers (Table 5.5), as these two forms of observations may be positively influenced by increased project duration (Burtch et al. 2013, Mendes-Da-Silva et al. 2016). This analysis is extended to the overall success of the project, using a probit model (Table 5.6), where confidence is indicated to have a significant effect. As illustrated in Figure 5.1 and Figure 5.2, it is revealed that the early turning points occur around confidence ratio of 0.5.





**Figure 5.1** Duration ratio prediction on  $\ln(\text{raised})$  (a) and  $\ln(\text{backers})$  (b).

(a) The turning point (full, period 1, 2, 3) = 0.443, 0.364, 0.475, 0.434 (for period cut off, see Figures B3.2 and B3.3), using Table 5.1, specification (1) and Table 5.2, specification (7 – 9). (b) The turning point is (full, period 1, 2, 3) = 0.486, 0.517, 0.521, 0.463 (for period cut off, see Figures B3.2 and B3.3). Uses Table 5.1, specification (4) and Table 5.2, specification (10 - 12).



**Figure 5.2** Certainty ratio prediction on  $\ln(\text{raised})$  (a) and  $\ln(\text{backers})$  (b).

(a) The turning point = 0.445 and represents Table 5.4, specification (15). (b) the turning point = 0.429 using Table 5.5, and represents specification (19).

**Table 5.1** Multivariate analysis of confidence (duration ratio): funds raised and number of backers

	$\ln(\text{Raised})$	$\ln(\text{Raised})$	$\ln(\text{Raised})$	$\ln(\text{Backers})$	$\ln(\text{Backers})$	$\ln(\text{Backers})$
	(1)	(2)	(3) <sup>^</sup>	(4)	(5)	(6) <sup>^</sup>
Confidence (duration ratio)	2.7*** (22.62)	2.8*** (23.32)	2.9*** (23.19)	2*** (22.49)	2.1*** (23.52)	2.2*** (23.38)
Confidence <sup>2</sup> (duration ratio)	-3.1*** (-21.89)	-3.3*** (-22.72)	-3.4*** (-22.54)	-2.1*** (-19.75)	-2.3*** (-21.00)	-2.3*** (-20.89)
Category						
Comics	.16** (2.57)	.15** (2.40)	.15** (2.39)	.39*** (7.93)	.38*** (7.71)	.39*** (7.54)
Dance	.61*** (11.25)	.61*** (11.27)	.65*** (11.83)	.38*** (9.06)	.38*** (9.06)	.4*** (9.29)
Design	.97*** (17.30)	.95*** (16.98)	.95*** (16.30)	.8*** (17.52)	.78*** (17.10)	.77*** (16.29)
Fashion	-.38*** (-6.63)	-.4*** (-6.93)	-.37*** (-6.34)	-.43*** (-9.94)	-.45*** (-10.36)	-.42*** (-9.61)
Film & video	.29*** (9.25)	.29*** (9.42)	.31*** (9.56)	.041* (1.78)	.047** (2.03)	.068*** (2.73)
Food	.55*** (10.99)	.54*** (10.69)	.54*** (10.52)	.38*** (9.85)	.36*** (9.43)	.37*** (9.30)
Games	.41*** (6.85)	.39*** (6.47)	.4*** (6.36)	.65*** (13.40)	.63*** (12.91)	.62*** (12.27)
Music	.37*** (12.30)	.38*** (12.34)	.4*** (12.52)	.28*** (11.69)	.28*** (11.76)	.3*** (12.01)
Photography	-.21*** (-3.89)	-.2*** (-3.73)	-.22*** (-3.78)	-.22*** (-5.49)	-.21*** (-5.26)	-.23*** (-5.16)
Publishing	-.46*** (-12.27)	-.47*** (-12.55)	-.49*** (-12.48)	-.29*** (-10.29)	-.31*** (-10.69)	-.33*** (-10.87)
Technology	.76*** (10.23)	.75*** (10.22)	.78*** (9.98)	.64*** (11.07)	.63*** (11.06)	.63*** (10.67)
Theater	.52*** (13.62)	.53*** (13.82)	.55*** (13.67)	.33*** (11.10)	.34*** (11.37)	.34*** (11.07)
Goal	1.4e-07 (1.06)	1.3e-07 (1.03)	1.2e-07 (0.97)	6.7e-08 (1.07)	5.9e-08 (1.00)	5.3e-08 (0.95)
U.S. dummy	-.51*** (-13.95)	-.51*** (-13.95)		-.34*** (-11.69)	-.34*** (-11.70)	
Duration Change		.13*** (7.21)	.12*** (6.32)		.14*** (9.69)	.13*** (8.81)
Latitude			.027*** (14.62)			.025*** (17.96)
Longitude			-.0016*** (-3.53)			-.00074** (-2.05)
Constant	6.8*** (142.10)	6.8*** (136.41)	5*** (52.34)	2.9*** (77.16)	2.8*** (72.37)	1.4*** (18.68)
N (Obs.)	63168	63168	58471	63168	63168	58471
N (Cluster)	57071	57071	52957	57071	57071	52957
R <sup>2</sup>	0.039	0.040	0.044	0.044	0.046	0.051

<sup>^</sup> denotes U.S. only projects

Notes: *t*-statistics are in parentheses; art is the reference group. \*, \*\*, and \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively.

**Table 5.2** Multivariate analysis of confidence (duration ratio) for the three separate periods: funds raised and number of backers

	<i>ln</i> (Raised) Period 1 (7)	<i>ln</i> (Raised) Period 2 (8)	<i>ln</i> (Raised) Period 3 (9)	<i>ln</i> (Backers) Period 1 (10)	<i>ln</i> (Backers) Period 2 (11)	<i>ln</i> (Backers) Period 3 (12)
Confidence (duration ratio)	1*** (3.11)	2.8*** (10.94)	3.2*** (21.15)	.87*** (3.40)	2.2*** (11.10)	2.4*** (21.05)
Confidence <sup>2</sup> (duration ratio)	-1.4*** (-3.47)	-3*** (-10.19)	-3.6*** (-20.74)	-.84*** (-2.67)	-2.1*** (-9.36)	-2.5*** (-19.21)
Category	-.011	-.043	.21***	.1	.22**	.46***
Comics	(-0.06) .54**	(-0.33) .47***	(3.00) .66***	(0.69) .4**	(2.15) .27***	(7.90) .41***
Dance	(2.37) .55***	(4.52) .5***	(10.31) 1***	(2.38) .42***	(3.26) .52***	(8.36) .86***
Design	(2.69) -.52**	(3.75) -.26**	(16.44) -.4***	(2.60) -.7***	(4.72) -.37***	(16.59) -.43***
Fashion	(-2.12) .39***	(-1.99) .14**	(-6.22) .31***	(-4.09) .024	(-3.77) -.099**	(-8.77) .091***
Film & video	(4.19)	(2.31)	(8.15)	(0.34)	(-2.23)	(3.18)
Food	.67*** (3.98)	.47*** (4.00)	.54*** (9.43)	.47*** (3.60)	.34*** (3.84)	.37*** (8.40)
Games	-.31 (-1.45)	-.29** (-2.11)	.51*** (7.78)	-.19 (-1.17)	.0026 (0.02)	.76*** (14.31)
Music	.43*** (4.42)	.27*** (4.51)	.4*** (10.95)	.3*** (3.98)	.18*** (3.54)	.31*** (11.14)
Photography	.048 (0.32)	-.14 (-1.49)	-.27*** (-4.01)	-.067 (-0.60)	-.24*** (-3.28)	-.23*** (-4.58)
Publishing	-.066 (-0.55)	-.47*** (-6.04)	-.49*** (-11.12)	.0088 (0.09)	-.36*** (-5.98)	-.31*** (-9.14)
Technology	.089 (0.47)	.17 (1.14)	1*** (11.13)	.066 (0.44)	.16 (1.36)	.84*** (12.15)
Theater	.35*** (2.75)	.41*** (5.62)	.61*** (12.84)	.15* (1.65)	.21*** (3.87)	.4*** (11.07)
Goal	-5.2e- 08*** (-5.06)	8.4e- 06*** (4.38)	1.5e-07 (0.86)	-2.8e- 08*** (-4.31)	3.5e- 06*** (3.03)	6.3e-08 (0.77)
U.S. dummy	-.36*** (-2.58)	-.4*** (-5.94)	-.56*** (-12.34)	-.23** (-2.04)	-.22*** (-4.31)	-.38*** (-10.69)
Constant	6.9*** (42.48)	6.6*** (68.66)	6.9*** (116.56)	2.9*** (22.14)	2.7*** (37.07)	2.9*** (63.68)
N (Obs.)	4865	14350	43953	4865	14350	43953
N (Cluster)	0.022	0.033	0.049	0.020	0.032	0.055
R <sup>2</sup>	-.011	-.043	.21***	.1	.22**	.46***

^ denotes U.S. only projects.

Notes: *t*-statistics are in parentheses; art is the reference group. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

**Table 5.3** Probit model of success probability using duration ratio

	Success (1)	Success (2)	Success (3)
Confidence (duration ratio)	1.1*** (14.60) <i>0.40</i>	1.1*** (14.21) <i>0.39</i>	1.1*** (14.65) <i>0.41</i>
Confidence <sup>2</sup> (duration ratio)	-.76*** (-8.36) <i>-0.27</i>	-.72*** (-7.86) <i>-0.26</i>	-.79*** (-8.38) <i>-0.29</i>
Category			
Comics	-.009 (-0.24) <i>-0.00</i>	-.0072 (-0.20) <i>-0.00</i>	.016 (0.42) <i>0.01</i>
Dance	.55*** (11.89) <i>0.20</i>	.55*** (11.88) <i>0.20</i>	.6*** (12.23) <i>0.22</i>
Design	-.018 (-0.55) <i>-0.01</i>	-.016 (-0.49) <i>-0.01</i>	-.018 (-0.54) <i>-0.01</i>
Fashion	-.52*** (-14.71) <i>-0.18</i>	-.52*** (-14.62) <i>-0.18</i>	-.5*** (-13.80) <i>-0.17</i>
Film & video	-.012 (-0.61) <i>-0.00</i>	-.014 (-0.73) <i>-0.01</i>	.019 (0.89) <i>0.01</i>
Food	-.0034 (-0.11) <i>-0.00</i>	-.0022 (-0.07) <i>-0.00</i>	.00011 (0.00) <i>0.00</i>
Games	-.14*** (-4.00) <i>-0.05</i>	-.13*** (-3.91) <i>-0.05</i>	-.11*** (-3.06) <i>-0.04</i>
Music	.21*** (10.06) <i>0.08</i>	.21*** (10.03) <i>0.08</i>	.24*** (10.83) <i>0.09</i>
Photography	-.25*** (-7.70) <i>-0.09</i>	-.25*** (-7.76) <i>-0.09</i>	-.28*** (-7.79) <i>-0.10</i>
Publishing	-.38*** (-16.60) <i>-0.14</i>	-.38*** (-16.51) <i>-0.14</i>	-.39*** (-16.06) <i>-0.14</i>
Technology	-.067 (-1.62) <i>-0.02</i>	-.068* (-1.66) <i>-0.03</i>	-.065 (-1.50) <i>-0.02</i>
Theater	.44*** (15.75) <i>0.16</i>	.44*** (15.69) <i>0.16</i>	.45*** (15.49) <i>0.17</i>
ln(Goal)	-.24*** (-54.13) <i>-0.09</i>	-.24*** (-53.92) <i>-0.09</i>	-.24*** (-51.62) <i>-0.09</i>
U.S. dummy	-.24*** (-10.51) <i>-0.09</i>	-.24*** (-10.50) <i>-0.09</i>	
Duration change		-.026** (-2.14) <i>-0.01</i>	
Latitude			.021*** (18.61) <i>0.01</i>
Longitude			.00028 (0.98) <i>0.00</i>
N (Obs.)	69599	69599	64457
N (Cluster)	62295	62295	57844
Pseudo R <sup>2</sup>	0.082	0.082	0.089
Prob. > chi <sup>2</sup>	0.000	0.000	0.000

Notes: Marginal effects are in italics; *t*-statistics are in parentheses. Art is the reference group. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

**Table 5.4** Multivariate analysis of confidence (certainty ratio): funds raised

	<i>ln</i> (Raised) (13)	<i>ln</i> (Raised) (14)	<i>ln</i> (Raised) (15)	<i>ln</i> (Raised) (16) <sup>^</sup>
Confidence (certainty ratio)	3.2*** (29.06)	3.2*** (28.92)	3.2*** (28.75)	3.2*** (28.22)
Confidence <sup>2</sup> (certainty ratio)	-3.6*** (-31.20)	-3.6*** (-31.09)	-3.6*** (-30.91)	-3.6*** (-30.15)
Category				
Comics	.12* (1.91)	.11* (1.83)	.11* (1.75)	.11* (1.78)
Dance	.64*** (11.54)	.64*** (11.57)	.63*** (11.41)	.68*** (12.09)
Design	.95*** (17.17)	.94*** (17.00)	.92*** (16.70)	.92*** (15.95)
Fashion	-.39*** (-6.72)	-.4*** (-6.90)	-.41*** (-7.08)	-.38*** (-6.50)
Film & video	.25*** (8.15)	.26*** (8.31)	.25*** (8.12)	.28*** (8.40)
Food	.51*** (10.11)	.5*** (9.94)	.49*** (9.74)	.5*** (9.57)
Games	.36*** (6.00)	.34*** (5.78)	.33*** (5.55)	.34*** (5.47)
Music	.38*** (12.50)	.39*** (12.58)	.37*** (12.17)	.4*** (12.42)
Photography	-.22*** (-4.12)	-.21*** (-3.99)	-.22*** (-4.06)	-.24*** (-4.09)
Publishing	-.49*** (-13.01)	-.5*** (-13.15)	-.5*** (-13.29)	-.52*** (-13.20)
Technology	.71*** (9.64)	.71*** (9.65)	.7*** (9.53)	.72*** (9.39)
Theater	.52*** (13.50)	.53*** (13.65)	.52*** (13.53)	.54*** (13.38)
Goal	1.1e-07 (0.92)	1.1e-07 (0.89)	1.1e-07 (0.95)	1.0e-07 (0.88)
U.S. dummy	-.49*** (-13.40)	-.49*** (-13.41)	-.49*** (-13.33)	
Duration change		.082*** (4.68)		
Duration			.024*** (11.57)	.025*** (11.78)
Duration <sup>2</sup>			-.00028*** (-13.53)	-.0003*** (-13.72)
Latitude				.027*** (14.53)
Longitude				-.0017*** (-3.59)
Constant	6.7*** (142.28)	6.7*** (137.11)	6.3*** (100.44)	4.6*** (43.05)
N (Obs.)	62094	62094	62094	57445
N (Cluster)	56171	56171	56171	52098
R <sup>2</sup>	0.046	0.046	0.050	0.053

<sup>^</sup> denotes U.S. only projects.

Notes: *t*-statistics are in parentheses; art is the reference group. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

**Table 5.5** Multivariate analysis of confidence (certainty ratio): number of backers

	<i>ln</i> (Backers) (17)	<i>ln</i> (Backers) (18)	<i>ln</i> (Backers) (19)	<i>ln</i> (Backers) (20)
Confidence (certainty ratio)	2.3*** (27.35)	2.2*** (27.11)	2.2*** (27.12)	2.3*** (26.69)
Confidence <sup>2</sup> (certainty ratio)	-2.6*** (-30.72)	-2.6*** (-30.51)	-2.6*** (-30.47)	-2.6*** (-29.65)
Category				
Comics	.36*** (7.05)	.35*** (6.93)	.36*** (7.17)	.37*** (7.02)
Dance	.41*** (9.69)	.41*** (9.71)	.4*** (9.53)	.43*** (9.86)
Design	.78*** (17.18)	.77*** (16.88)	.76*** (16.85)	.75*** (15.98)
Fashion	-.43*** (-9.97)	-.44*** (-10.30)	-.45*** (-10.38)	-.43*** (-9.66)
Film & video	.013 (0.55)	.02 (0.86)	.022 (0.94)	.043* (1.74)
Food	.35*** (8.93)	.33*** (8.63)	.34*** (8.66)	.34*** (8.52)
Games	.6*** (12.54)	.59*** (12.17)	.59*** (12.20)	.58*** (11.58)
Music	.28*** (11.70)	.29*** (11.88)	.28*** (11.72)	.3*** (12.02)
Photography	-.24*** (-5.93)	-.23*** (-5.70)	-.22*** (-5.64)	-.24*** (-5.47)
Publishing	-.32*** (-11.23)	-.33*** (-11.49)	-.33*** (-11.30)	-.35*** (-11.48)
Technology	.59*** (10.36)	.59*** (10.38)	.59*** (10.40)	.6*** (10.12)
Theater	.34*** (11.29)	.34*** (11.54)	.34*** (11.34)	.34*** (11.01)
Goal	4.3e-08 (0.79)	3.7e-08 (0.72)	5.0e-08 (0.91)	4.4e-08 (0.85)
U.S. dummy	-.32*** (-11.18)	-.32*** (-11.19)	-.32*** (-11.11)	
Duration change		.12*** (8.22)		
Duration			.014*** (8.87)	.015*** (9.10)
Duration <sup>2</sup>			-.0002*** (-12.57)	-.00021*** (-12.70)
Latitude				.025*** (17.84)
Longitude				-.00076** (-2.13)
Constant	2.9*** (79.47)	2.8*** (74.93)	2.7*** (56.04)	1.3*** (15.97)
N (Obs.)	62094	62094	62094	57445
N (Cluster)	56171	56171	56171	52098
R <sup>2</sup>	0.050	0.051	0.055	0.061

^ denotes U.S. only projects.

Notes: *t*-statistics are in parentheses; art is the reference group. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

**Table 5.6** Probit model of success probability: certainty ratio

	Success (1)	Success (2)	Success (3)	Success (4)
Confidence (certainty ratio)	1.3*** (19.54) <i>0.49</i>	1.3*** (19.53) <i>0.49</i>	1.3*** (19.20) <i>0.48</i>	1.4*** (19.14) <i>0.49</i>
Confidence <sup>2</sup> (certainty ratio)	-1.6*** (-22.19) <i>-0.58</i>	-1.6*** (-22.17) <i>-0.58</i>	-1.6*** (-21.82) <i>-0.57</i>	-1.6*** (-21.30) <i>-0.57</i>
Category				
Comics	-.058 (-1.57) <i>-0.02</i>	-.058 (-1.56) <i>-0.02</i>	-.033 (-0.88) <i>-0.01</i>	-.0067 (-0.17) <i>-0.00</i>
Dance	.57*** (11.80) <i>0.21</i>	.57*** (11.80) <i>0.21</i>	.57*** (11.79) <i>0.22</i>	.62*** (12.15) <i>0.22</i>
Design	-.029 (-0.90) <i>-0.01</i>	-.028 (-0.88) <i>-0.01</i>	-.037 (-1.13) <i>-0.01</i>	-.036 (-1.06) <i>-0.01</i>
Fashion	-.52*** (-14.39) <i>-0.18</i>	-.52*** (-14.36) <i>-0.18</i>	-.53*** (-14.73) <i>-0.19</i>	-.51*** (-13.81) <i>-0.18</i>
Film & video	-.029 (-1.47) <i>-0.01</i>	-.03 (-1.51) <i>-0.01</i>	-.017 (-0.84) <i>-0.01</i>	.016 (0.73) <i>0.01</i>
Food	-.01 (-0.33) <i>-0.00</i>	-.01 (-0.31) <i>-0.00</i>	-.019 (-0.61) <i>-0.01</i>	-.015 (-0.45) <i>-0.01</i>
Games	-.16*** (-4.81) <i>-0.06</i>	-.16*** (-4.78) <i>-0.06</i>	-.17*** (-5.02) <i>-0.06</i>	-.14*** (-3.99) <i>-0.05</i>
Music	.2*** (9.47) <i>0.08</i>	.2*** (9.47) <i>0.08</i>	.22*** (10.19) <i>0.08</i>	.25*** (11.04) <i>0.09</i>
Photography	-.27*** (-8.19) <i>-0.10</i>	-.27*** (-8.21) <i>-0.10</i>	-.25*** (-7.65) <i>-0.09</i>	-.28*** (-7.70) <i>-0.10</i>
Publishing	-.4*** (-17.38) <i>-0.15</i>	-.4*** (-17.35) <i>-0.15</i>	-.4*** (-17.05) <i>-0.14</i>	-.4*** (-16.49) <i>-0.14</i>
Technology	-.085** (-2.06) <i>-0.03</i>	-.085** (-2.08) <i>-0.03</i>	-.084** (-2.06) <i>-0.03</i>	-.079* (-1.82) <i>-0.03</i>
Theater	.46*** (15.89) <i>0.17</i>	.46*** (15.87) <i>0.17</i>	.46*** (16.06) <i>0.17</i>	.47*** (15.73) <i>0.17</i>
ln(Goal)	-.27*** (-57.54) <i>-0.10</i>	-.27*** (-57.72) <i>-0.10</i>	-.25*** (-53.96) <i>-0.09</i>	-.25*** (-51.40) <i>-0.09</i>
U.S. dummy	-.24*** (-10.27) <i>-0.09</i>	-.24*** (-10.26) <i>-0.09</i>	-.23*** (-10.03) <i>-0.08</i>	
Duration change		-.0077 (-0.65) <i>-0.00</i>		
Duration (days)			-.0049*** (-3.54) <i>-0.00</i>	-.0046*** (-3.17) <i>-0.00</i>
Duration <sup>2</sup>			-.000019 (-1.37) <i>-0.00</i>	-.000027* (-1.84) <i>-0.00</i>
Latitude				.021*** (18.47) <i>0.01</i>
Longitude				.00032 (1.09)

				0.00
N (Obs.)	68163	68163	68163	63082
N (Cluster)	61123	61123	61123	56719
Pseudo R <sup>2</sup>	0.082	0.082	0.087	0.094
Prob. > chi <sup>2</sup>	0.000	0.000	0.000	0.000

^ denotes U.S. only projects.

Notes: *t*-statistics are in parentheses; art is the reference group. \*, \*\*, and \*\*\* denote significance at 10%, 5%, and 1% levels, respectively.

## 5.4 DISCUSSION

Within this study we examine the salient effects of confidence on the overarching goal of fundraising success. We explore this relationship between confidence and success using two novel behavioural proxies of confidence: the funding success of over 60,000 Kickstarter crowdfunding projects and linguistic expressions of confidence in the descriptive text of the respective campaign pages. By measuring the success in terms of funds contributed, the number of backers, and achievement of the target goal, we demonstrate that confidence does indeed promote success, implying that failure to show enough confidence reduces project creator's ability to attract contributors (backers) and the funds needed for success. Indicating that too much confidence actually hurts the project, as reflected by a consistent and striking non-linear inverted U-shaped association between confidence and success, irrespective of whether it is based on project duration or linguistic expression. Moreover, the effects of the confidence proxy on the duration ratio in *period 2* and *3* are more visible compared to that of *period 1*, as evident in the steepness of the curve (see Figure 5.1a and Figure 5.1b). This result supports the relevance on an anchoring effect, that is, small deviations in the proxy (in *period 2* and *3*) represent larger changes in the level of actual confidence.

Admittedly, our confidence proxies are not free of problems. Project creators could reduce project duration for strategic reasons; for example, to avoid backer procrastination, failure, or forgetfulness in contributing (thereby using duration as an attention device). Nevertheless, the fact that both confidence proxies yield similar results strongly suggests that the choice of duration length is not based entirely on strategic calculation. Hopefully, ongoing developments in the field of text analysis will allow future research to further test the robustness of our results.



## Chapter 6: Concluding Remarks

---

Crowdfunding incentivises innovation by supporting untried ideas from unknown creators. In using the crowd to fund projects from both new and old innovators, not only do the creators get to gauge the interest in their idea but they stimulate economic growth through the new release of information and technology. As a result of the new information and technology being positively received, over time there has been an increase in the number of people willing to create and contribute to projects. Therefore understanding how creators portray their ideas and how potential contributors (investors) respond to this information is important. The advantages of exploring a controlled and detailed field setting allows us to explore how people communicate and receivers are affected by the communication. The key elements of this thesis are: the use of text to analyse information signals from communication and provide empirical evidence of the factors of success using both behavioural and linguistic indicators. To the best of my knowledge, the proxies created for confidence indicators are the first of their type. In addition, despite its simplicity, there has been a very limited empirical investigation into text length, this thesis controls for text length in an entrepreneurial situation. Whilst the empirical analysis of money effects and complexity levels, have not been presented in this form before. Using campaign data from Kickstarter, I find that the pitch's length, complexity and money emphasis, as well as the confidence levels exhibited by the creator, are all significant factors for creating fundraising success.

This chapter concludes the thesis by providing a summary of the findings, a discussion of the limitations and suggestions for future research into decision-making in a crowdfunding context.

### 6.1 KEY FINDINGS

The findings presented in this thesis not only identify individual factors that influence fundraising success but identifies a strong predisposition that backers to will not make a decision when the creator communicates excessively. In the first study the descriptive text length is used to analyse information volume on outcomes,

as text length is a simple cue that readers can identify very quickly. As information is communicated daily through text, it is important to understand at what point people no longer pay attention and find information too complex to understand. In all categories on Kickstarter, there is an optimal number of words that increases the likelihood of success, and going beyond this point by being too long-winded will reduce the project creator's ability to attract contributor support. This non-linear relationship may be instigated by the perceived information overload or the effort to read the text. The second study requires us to examine the text itself, in this case I investigate the readability levels of the text. I identify a similar curvilinear relationship, indicating that readers are willing to read up to a certain level of complexity, before it becomes too complex and they back out. However, there are only a few projects that go beyond the optimal level of complexity. This may indicate several things, that contributors are motivated in understanding ideas, that most contributors have a high level of reading ability or contributors are specifically versed in the area of interest i.e. they're also robotics engineers or musicians. Taking it a step further using specific word terminology, in the third study on money salience and success, I observe a negative linear relationship when monetary terms are emphasised within the campaign text. Such a statistically significant relationship could imply that backers genuinely care about the motivations of the project creator when deciding if they should contribute to a project or not. If the readability of a text can shape decisions it should come as no surprise that signals of confidence also shape the processes of decisions. As confidence influences our everyday decision making, it is important to understand how the confidence of others may affect our own decisions. In the crowdfunding setting, there exists a non-linear (inverted U-shape) relationship between self-confidence and crowdfunding performance based on evidence found using large empirical data, in line with the results found in traditional survey and experimental approaches. Additionally, I use two different confidence proxies, one that represents the creators behaviour when making a decision about the duration of the project, and the other considers the terminology used when the description is written. With both proxies the results remain robust. Thus, supporting previous works. The advantage of the two behavioural proxy measurements for confidence is that they do not suffer from potential self-reported bias commonly seen in survey data. Since both proxies were derived from ex-ante information (i.e. before

realization of success), it allows casual inference to the results with appropriate statistical models used.

Whilst these findings may not necessarily result in specific policy outcomes, this thesis makes useful contributions to the literature of attention, behaviour, text-length, readability, confidence, money priming, communication and success. Using a naturalistic setting allows us to observe how bounded rational decision-makers make decisions when confronted with limited processing capabilities and an abundance of information.

## **6.2 LIMITATIONS & FUTURE PERSPECTIVES**

Our approach to analysing multiple signals that communicate information within the novel setting of crowdfunding has extended our understanding of how individuals communicate to achieve success. Whilst these factors have been found to have a significant effect on increasing the likelihood of success, our models are not without limitations. This thesis analyses some of the more prominent factors of communicating information, however it is still necessary to consider other features contained within the data. The more we know about communication and its effects on decision-making the better we understand factors that lead to success when innovating. Understanding the characteristics of those who innovate and those willing to support innovation, is interesting because it can lead to the release of new information and technology. Nevertheless, while the dataset contains an extensive amount of detail about each project the quantitative analysis is still limited. The following sections note the limitations realised throughout the processes of this thesis and propose future research perspectives.

### ***General analysis limitations***

There are several ways the thesis studies can be limited, other than through a potential omitted variable bias or endogeneity bias. Firstly, linguistic analyses cannot pick up on the connotations or idioms expressed in a text, particularly when the analysis is done on a one-word basis. I was also not able to account for jargon, phrases, internet jargon, and slang that is not captured by the common dictionary. Nor have I been able to control for the individual preferences of readers. Regarding

the confidence of creators, it is possible that the duration is used strategically and may not appropriately reflect a creator's confidence. However, both proxies yield similar results, and as such the project duration may not necessarily be a strategic calculation. Furthermore, while we consider the quantitative effect of the project description, we do not necessarily consider the qualitative features, such as the quality of the innovation being created. We can presume that the quality of the innovation is proxied by the level of success, however, future analysis could consider linking patents to Kickstarter projects.

### ***Project characteristics***

Kickstarter will often promote projects, through classifications such as '*Projects we love*', '*Projects of the day*', '*Recommended for you*' and '*What's popular*', with each option showcasing the project in a different way on the platform, i.e. on the homepage or the project discovery page. To be classified under '*Projects we love*' a project needs to be picked by a Kickstarter staff member as their favourite and is then promoted on the homepage, additionally a button of '*project we love*' is added to the project. Such a features are more likely to direct more potential contributor traffic through to the project, and being featured is observed to have a positive effect on a project's success (Kuppuswamy and Bayus, 2013; Mollick, 2014). To improve the analysis, future research should include indicators for Kickstarter featuring and whether such features are an appropriate proxy for a project's quality.

Moreover, a creator cannot rely on being featured by Kickstarter in order to promote their project. In most cases, creators are actively engaging in marketing the campaign through varying forms of media, and continuously updating current and future contributors about the progress of the project. Links to social media, such as Twitter, and a frequent number of updates have been shown to significantly increase funding success (Evers, 2012; Etter et al., 2013; Mollick, 2014; Xu et al., 2014; Cumming et al., 2015). Once creators have managed to attract the prospective backers, they can use visual stimuli, such as images or videos to portray a physical representation of the project or to progress their pitch. The presence of a video and image has been found to increase the chances of success (see Evers, 2012; Mollick, 2014; Mitra & Gilbert, 2014; Cumming et al., 2015), however one study notes no

significance in images increasing success (see Evers, 2012; Xu et al., 2014). It is expected that the number of images will be positively correlated with text length, as well as a field related feature, which in some cases may imply quality e.g. *Photography* and *Journalism*. As such, I expect the inclusion of images to have a similar effect to the length of the text on success.

As previously mentioned, as incentive to contribute creators required to offer a series of rewards specific to the project. As creators are heterogeneous in their nature of creating things, the kind of rewards on offer, the number available and the different price structures, vary across each project and therefore the quality of rewards may vary. Each feature of the rewards has been observed to have a significant impact on funding success as it influences contributing behaviour (Kuppuswamy & Bayus, 2013; Desai et al., 2015; Cumming et al., 2015). Rewards were unable to be controlled for in the analysis as the information was not within the dataset.

### ***Campaign description***

As a result of creators being able to edit the description throughout the duration of the project, there is concern that an endogeneity bias exists within the description of the texts. Due to the dynamic nature of the edits, they are difficult to measure for the whole population of Kickstarter, unless the project was monitored from the day it was launched to the day the campaign ended. Therefore potential edits are controlled for, by identifying key phrases such as ‘*days to go*’ or ‘*GOAL MET*’ within the description based on Xu et al.’s categorisation of updates. The quantitative effect was greater for projects with an identified edit, however the level of significance remains the same. As a result any potential endogeneity bias is expected to have a small effect.

It has been observed that an individual’s behaviour can be strongly influenced by emotion and that emotions play a significant role in the decision-making process (see Lerner, Small and Loewenstein, 2004; Martin and Delgado, 2011). Due to advances in linguistic analysis, it is possible to determine if sentiment exists within a text and estimate the level of sentiment contained within the text. Exploring emotion within texts is not new, with sentiment analysis being examined in newspapers, text messages and the messaging content in online support groups (see Alpers et al.,

2005; Davis, 2011; Holtgraves, 2011). Although some may argue that sentiment analysis in text misses the connotations and idioms assigned to them (Osherenko and André, 2007), sentiment dictionaries and emotive words have been linked with biological processes (see Kissler, Herbert, Winkler, and Junghofer, 2009; Schacht and Sommer, 2009; Martin and Delgado, 2011). Using sentiment detection dictionaries such as the LIWC, it is possible to identify both broad (positive or negative) and discrete (anger or happiness) emotions within a text. An extension on the current analysis would be to determine what level of emotion and type of emotions that exists within crowdfunding content and what effects it has on the overall fundraising success of a project. Which brings us to the next research extension, to explore whether the gender of the creator(s) has an effect on the level of emotional content contained within a project description.

### *Creator characteristics*

It is important to understand the characteristics of a creator, as these features that influence how an individual communicates and ultimately, if they achieve success. Characteristics of the creator such as gender, education levels, and previous experiences and whether or not it was created by a team, are not included in this dataset. Some of these features can be identified on Kickstarter, such as the number of previous projects and the number of team members. In some cases, characteristics such as gender and education, could be identified through Kickstarter or the links to social networks on the creator's profile i.e. Facebook or LinkedIn. Previous research has identified that there exists a gender difference in the expression of emotion within a text, where female readers find concrete words more emotive than male readers (Bauer and Altarriba, 2008). Therefore it is likely that the gender of the creator(s) will have an influence on the amount of emotional content perceived to be in the project description. A further extension would be to consider the team dynamics of projects.

Crowdfunding is an inherently social situation, with large communities already existing on various crowdfunding platforms and social media websites. In line with this, many creators link their social networking and media accounts, i.e. Twitter and Facebook, in an effort to extend their networks and promote their project. As a result the project may receive more contributions. Another factor that may influence the

network and number of potential contributors interested in the project, is the social status of the creator. If a creator is particularly well-known, i.e. a celebrity or a highly successful innovator, then they will have a larger networking base which would influence success. Unfortunately I was unable to control for social networks (i.e. connected to social media accounts or the number of Facebook friends), featured by Kickstarter, and the projects reward and visual features.

### ***Backer characteristics***

In Kickstarter, the success of a project relies on the generosity of contributors, who can be considered as a mix between altruistic donators and traditional investment as they are neither purely donating funds nor investing to seek equitable returns. It has been indicated that backers contribute for early access to new products, community participation, and providing both monetary and non-monetary support (Agrawal et al., 2014). Information on backer characteristics, such as their age, education level, gender, number of projects backed, and the types of projects backed, have not been recorded. For society to function according to the law, the government needs the public to be willing to cooperate; as does a creator when attempting to raise funds from the crowd. Examining factors that characterise an individual, helps us understand the way people cooperate and the kind of individuals who make decisions in situations that are not necessarily reciprocal, and may provide insight into why a person backed a project. For example research has found significant difference in altruistic behaviour and risk preferences across gender (Andreoni and Vesterlund, 2001; Charness and Gneezy, 2012). Future research could examine whether backer preferences, or if intrinsic (extrinsic) motivations, play a role on success and what government agencies could learn from such cooperation.

Several researchers have investigated backer motivations to contribute, gender effects on contributions and potential deterrents to contributions (see Gerber and Hui, 2013; Greenberg and Mollick, 2014; Marom, Robb & Sade, 2016; Trusiak, 2016). To collect this information, some conducted interviews with potential and existing crowdfunding participants (Gerber and Hui, 2013), others used text analysis on the names listed on Kickstarter to determine gender of creators (Marom et al., 2016) or conducted laboratory experiments to determine behaviour (Greenberg and Mollick, 2014). To contribute to the existing literature, future research could

incorporate the additional characteristics of age, gender, education levels, type projects backed and number of projects, to examine preferences on success. In addition, the effects of emotional content within the text should be extended to examine whether creators emotions effect contributions by gender. Following the linguistic route of analysis, it would be interesting to determine whether an emphasis on money causes risk or loss averse behaviour in potential contributors.

### ***Dynamic behaviour***

This thesis is based on static information that is recorded after the outcome of project, and does not include dynamic information about the interactions between the project, creators and backers. In doing so we can investigate the resilience of project creators, whether they have the ability to bounce back again in instance of failure or in the case of success, whether they have learned anything from their previous experience and adjusted accordingly. Project creators have the ability to adjust or edit the descriptions, add videos and images, provide updates and adjust rewards that have no backers or list additional rewards, throughout the duration of the project. As a result it is unknown whether the timing of these adjustments will increase the likelihood of success. A potential research avenue would be to examine whether these observable adjustments influence consumer trust or purchasing behaviour in online entities.

A further extension using dynamic behaviour would be to examine whether shocks of hyped or more successful projects have an impact on the behaviour of creators, i.e. whether other projects mimic patterns of highly successful projects. There are two ways determine a highly successful project, firstly it can estimated by projects that achieve a high percentage of success (relative to the goal) or generating an indicator for projects that “take-off” in the media. The second indicator can be determined by observing the number of mentions of the projects’ title in social media, news articles and google search terms. An initial analysis of the new variables would determine if similar projects, as identified by sub-category and were already ‘live’ when the project took-off, experienced more contributions. This would then be extended to examine the specific communicated characteristics such as text length and readability, and could be extended to key words that were frequently used within the highly successful project. Previous research provides evidence that hyped or successful shocks have a positive impact on the number projects being created and



the amounts being contributed (Geva et al., 2016). Whilst successful and hyped projects impact the success of other projects, another research avenue would be to examine if institutional effects, cultural perceptions and geographic locations have a similar impact to both creator and contributor behaviour.

In addition to this, future research should consider creators who come back to the Kickstarter crowdfunding pool a second or third or fourth time, etcetera. There is no restriction on Kickstarter for the number of projects a creator can make, a creator can launch another a new project or reboot a previously failed campaign. An interesting perspective would be to examine whether creators who try again change their behaviour between the first and second project. An initial assessment between the first two projects would determine whether there is a behavioural change, and if so it could be extended to any subsequent attempts that occur on Kickstarter. To analyse the data, a Heckman Two-step Selection model for example, could be applied to the first and second project. This approach will look at the differences in the length, readability, money terms and confidence in the text, as well as the sentiment contained within the text and if previous success is a factor of any changes.

Future research should incorporate the dynamic information of backer behaviour, such as scrolling times, number of unique pages viewed and contributions as they occur over time. The more details we have about backers the more we are able to understand about their decision-making behaviour.

Crowdfunding has presented a unique setting to study communication and success, and provides a significant amount of information for further exploration of behaviour. Observing how potential contributors make decisions provides insight into factors that influence decision making behaviour when asked to contribute money when the outcome is unknown.



# References

---

- ¡Kickstarter ya esta disponible en España!* (2015, May 19). [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-ya-esta-disponible-en-espana>
- Abebe, N. (2014, June 12). Introducing two new categories: Journalism and crafts [Web log post]. Retrieved from <https://www.kickstarter.com/blog/introducing-two-new-categories-journalism-and-crafts>
- Agrawal, A., Catalini, C., & Goldfarb, A. (2011). *The geography of crowdfunding* (No. w16820). National Bureau of Economic Research
- Agrawal, A., Catalini, C., & Goldfarb, A. (2014). Some simple economics of crowdfunding. *Innovation Policy and the Economy*, 14(1), 63-97. doi:10.1086/674021
- Agrawal, A., Catalini, C., & Goldfarb, A. (2015). Crowdfunding: Geography, social networks, and the timing of investment decisions. *Journal of Economics & Management Strategy*, 24(2), 253-274.
- Allington, R. L. (2005). A special section on reading research-ideology is still trumping evidence. *Phi Delta Kappan*, 86(6), 462.
- Alpers, G. W., Winzelberg, A. J., Classen, C., Roberts, H., Dev, P., Koopman, C., & Taylor, C. B. (2005). Evaluation of computerized text analysis in an Internet breast cancer support group. *Computers in Human Behavior*, 21(2), 361-376.
- Anderson, J. (1983). Lix and Rix: Variations on a little-known readability index. *Journal of Reading*, 26(6), 490-496.
- Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, 293-312.
- Baker, D. W., Gazmararian, J. A., Williams, M. V., Scott, T., Parker, R. M., Green, D., Ren, J. & Peel, J. (2002). Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *American journal of public health*, 92(8), 1278-1283.
- Bamberger, R. & Vanecek, E. (1984). *Lesen–Verstehen–Lernen–Schreiben*. Wien: Jugend und Volk.
- Baron, N. S. (2014). *Words Onscreen: The Fate of Reading in a Digital World*. Oxford University Press, USA.
- Bauer, L. M., & Altarriba, J. (2008). An investigation of sex differences in word ratings across concrete, abstract, and emotion words. *The Psychological Record*, 58(3), 465.
- Belianin, A. (2015). *Overconfidence*. Real-World Decision Making: An Encyclopedia of Behavioral Economics, 305-306.
- Belleflamme, P., Lambert, T., & Schwienbacher, A. (2013). Individual crowdfunding practices. *Venture Capital*, 15(4), 313-333.

- Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, 29(5), 585-609.
- Belleflamme, P., Omrani, N., & Peitz, M. (2015). The economics of crowdfunding platforms. *Information Economics and Policy*, 33, 11-28.
- Benoit, K., & Nulty, P. (2013). *quanteda: Quantitative Analysis of Textual Data. An R library for managing and analyzing text.*
- Bernardo, A. E., & Welch, I. (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, 10(3), 301-330.
- Biais, B., Hilton, D., Mazurier, K., & Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies*, 72(2), 287-312.
- Bitgood, S. C., & Patterson, D. D. (1993). The effects of gallery changes on visitor reading and object viewing time. *Environment and Behavior*, 25(6), 761-781.
- Bitgood, S., Dukes, S., & Abbey, L. (2006). Interest and effort as predictors of reading: A test of the general value principle. *Current Trends in Audience Research*, 19(2), 2.
- Blanchflower, D. G., & Oswald, A. J. (2004). Money, sex and happiness: An empirical study. *Scandinavian Journal of Economics*, 106(3), 393-415.
- Boulding, K. E. (1981). *A preface to grants economics: The economy of Love and Fear*. New York: Praeger Publishers.
- Bouncken, R. B., Komorek, M., & Kraus, S. (2015). Crowdfunding: The current state of research. *The International Business & Economics Research Journal (Online)*, 14(3), 407-415.
- Bradley, S. D., & Meeds, R. (2002). Surface-structure transformations and advertising slogans: The case for moderate syntactic complexity. *Psychology & Marketing*, 19(7-8), 595-619.
- Breedon, T. (2012). *Boosting finance options for business*. London: Department for Business Innovation and Skills.
- Bruton, G., Khavul, S., Siegel, D., & Wright, M. (2015). New financial alternatives in seeding entrepreneurship: Microfinance, crowdfunding, and peer-to-peer innovations. *Entrepreneurship Theory and Practice*, 39(1), 9-26.
- Burchell, B., & Marsh, C. (1992). The effect of questionnaire length on survey response. *Quality and quantity*, 26(3), 233-244.
- Burtch, G., Ghose, A., & Wattal, S. (2013). An empirical examination of the antecedents and consequences of contribution patterns in crowd-funded markets. *Information Systems Research*, 24(3), 499-519.
- Burton, R. S. (2014). Readability, logodiversity, and the effectiveness of college science textbooks. *Bioscene: Journal of College Biology Teaching*, 40(1), 3-10.
- Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review*, 89(1), 306-318.

- Campbell, N. (1999). How New Zealand consumers respond to plain English. *The Journal of Business Communication*, 36 (4), 335-358.
- Caylor, J. S. (1973). Methodologies for determining reading requirements of military occupational specialties. Human Resources Research Organization, Alexandria, Va.
- Chan, H. F., Guillot, M., Page, L., & Torgler, B. (2015). The inner quality of an article: Will time tell?. *Scientometrics*, 104(1), 19-41.
- Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1), 50-58.
- Chebat, J. C., Gelinias-Chebat, C., Hombourger, S., & Woodside, A. G. (2003). Testing consumers' motivation and linguistic ability as moderators of advertising readability. *Psychology & Marketing*, 20(7), 599-624.
- Chen, Y. C., Shang, R. A., & Kao, C. Y. (2009). The effects of information overload on consumers' subjective state towards buying decision in the internet shopping environment. *Electronic Commerce Research and Applications*, 8(1), 48-58.
- Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.
- Chung, J., & Lee, K. (2015). A long-term study of a crowdfunding platform: Predicting project success and fundraising amount. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media* (pp. 211-220). ACM.
- Coleman, E. B. (1965). On understanding prose: Some determiners of its complexity. *NSF Final Report GB -2604. Washington, D.C.: National Science Foundation.*
- Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60(2), 283.
- Cota, A., & Bitgood, S. (1993). Recall of label content: The effects of length and sequence. *Recall*, 8(4), 12.
- Crosetto, P., & Regner, T. (2015). Crowdfunding: Determinants of success and funding dynamics. *Jena Economic Research Papers, 2014*, 035.
- Cumming, D. J., Leboeuf, G., & Schwienbacher, A. (2015). Crowdfunding models: Keep-it-all vs. all-or-nothing. In *Paris December 2014 finance meeting EUROFIDAI-AFFI paper.*
- Dale, E. & Chall, J. S. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 37-54.
- Danielson, W. A., & Bryan, S. D. (1963). Computer automation of two readability formulas. *Journalism & Mass Communication Quarterly*, 40(2), 201-206.
- Davis, R. (2011). A computer-aided affective content analysis of nanotechnology newspaper articles. *NanoEthics*, 5(3), 319-334.
- De Bondt, W. F. M., & Thaler, R. H. (1995). Chapter 13 Financial decision-making in markets and firms: a behavioural perspective. *Handbooks in Operations Research and Management Science*. 9, 385-410.

- Desai, N., Gupta, R., & Truong, K. (2015). Plead or Pitch? The Role of Language in Kickstarter Project Success. Retrieved from <http://nlp.stanford.edu/courses/cs224n/2015/reports/15.pdf>
- Dickert, S., Sagara, N., & Slovic, P. (2011). Affective motivations to help others: A two-stage model of donation decisions. *Journal of Behavioral Decision Making*, 24(4), 361-376.
- DuBay, W. H. (2007). *Smart Language: Readers, Readability, and the Grading of Text*. ERIC.
- Elliott, J. O., & Shneker, B. F. (2009). A health literacy assessment of the epilepsy.com website. *Seizure*, 18(6), 434-439.
- Etter, V., Grossglauser, M., & Thiran, P. (2013, October). Launch hard or go home!: predicting the success of kickstarter campaigns. In *Proceedings of the first ACM conference on Online social networks* (pp. 177-182). ACM.
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The information society*, 20(5), 325-344.
- Evers, M. W., Lourenço, C., & Beije, P. (2012). *Main Drivers of Crowdfunding Success: A Conceptual Framework and Empirical Analysis*. (Master's Thesis, Erasmus Universiteit, Rotterdam, Netherlands).
- Falagas, M. E., Zarkali, A., Karageorgopoulos, D. E., Bardakas, V., & Mavros, M. N. (2013). The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals. *PLoS One*, 8(2), e49476.
- Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, 52, 498-506.
- Fang, I. E. (1967). The 'easy listening formula'. *The Journal of Broadcasting*, 11, 63-68.
- Fatoba, J. O. (2015). Readability of biology textbooks and students' academic performance in senior secondary schools in Ekiti state Nigeria. *European Journal of Research and Reflection in Educational Sciences Vol*, 3(3).
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 552-564,
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Flesch, R. (1979). *How to write in plain English: A book for lawyers and consumers*. New York: Harper.
- French projects are live!* (2015, May 27). [Web log post]. Retrieved from <https://www.kickstarter.com/blog/french-projects-are-live>
- Frey, B. S. (1997). On the relationship between intrinsic and extrinsic work motivation. *International Journal of Industrial Organization*, 15(4), 427-439.

- Friman, M. S. (1994). Plain English statutes-long overdue or underdone. *Loyola Consumer Law Review*, 7(3), 103.
- Fry, E. (2002). Readability versus leveling. *The reading teacher*, 56(3), 286-291.
- Fry, E. B. (1987). The varied uses of readability measurement today. *Journal of Reading*, 30(4), 338-343.
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public opinion quarterly*, 73(2), 349-360.
- Gallagher, J. (2012). The emotion for opening the text message. *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, 22(2), 36-43.
- Gao, Q., & Lin, M. (2015). Lemon or cherry? The value of texts in debt crowdfunding (July 4, 2015) Retrieved from Social Science Research Network [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2446114](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2446114)
- Gasiorowska, A., Zaleskiewicz, T., & Wygrab, S. (2012). *Journal of Economic Psychology*, 33(3), 603-608.
- Gerber, E. M., & Hui, J. (2013). Crowdfunding: Motivations and deterrents for participation. *Journal of ACM Transactions on Computer-Human Interaction*, 20(6), 1-32.
- Geva, H., Barzilay, O., & Oestreicher-Singer, G. (2016). A potato salad with a lemon twist: Using supply-side shocks to study the impact of low-quality actors on crowdfunding platforms (May 9, 2016) Retrieved from Social Science Research Network [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2777474](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2777474)
- Gorbatai, A. D., & Nelson, L. (2015). Gender and the Language of Crowdfunding. In *Academy of Management Proceedings*, (1), 15785.
- Greenberg, J., & Mollick, E. R. (2014). Leaning in or leaning on? Gender, homophily, and activism in crowdfunding. *Social Science Research Network, Working Paper*. Retrieved from <http://ssrn.com/abstract=2462254>
- Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013). Crowdfunding support tools: predicting success & failure. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1815-1820). ACM.
- Gunning, R. (1952). *The technique of clear writing*. New York: McGraw-Hill.
- Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, 6(2), 3-13.
- Simon, H. A. (1971). *Designing organizations for an information-rich world*. Chicago.
- Herz, H., Schunk, D., & Zehnder, C. (2014). How do judgmental overconfidence and overoptimism shape innovative activity? *Games and Economic Behavior*, 83, 1-23.
- Herzog, A. R., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public opinion quarterly*, 45(4), 549-559.
- Hirshleifer, D. & Luo, G. Y. (2001). On the survival of overconfident traders in a competitive securities market. *Journal of Financial Markets*, 4(1), 73-84.

- Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do you put your money on it? *The Economic Journal*, 115(503), 305-318.
- Holmes, M. J., & Silverstone, B. (2010). Business confidence and cyclical turning points: a markov-switching approach. *Applied Economics Letters*, 17(3), 229-233.
- Holtgraves, T. (2011). Text messaging, personality, and the social context. *Journal of Research in Personality*, 45(1), 92-99.
- Huberman, B. A., & Wu, F. (2008). The economics of attention: maximizing user value in information-rich environments. *Advances in Complex Systems*, 11(04), 487-496.
- Jensen, C., Potts, C., & Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behaviour. *International Journal Human-Computer Studies*, 63(1-2) 203-227.
- Jiang, Y., Chen, Z., & Wyer, R. S. (2014). Impact of money on emotional expression. *Journal of Experimental Social Psychology*, 55, 228-233.
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11(2), 143-157.
- Kampouris, P. & Evans, A. (2013). TL;DR (Too Long; Didn't Read)!: A study on the effects of length and fluency on the helpfulness of online reviews. (Master's Thesis, Tilburg University, Netherlands).
- Kasule, D. (2011). Textbook readability and ESL learners. *Reading & Writing*, 2(1), 63-76.
- Kickstarter in Canada!* (2013, September 9). [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-in-canada-0>
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel* (No. RBR-8-75). Naval Technical Training Command Milington TN Research Branch.
- Kirchler, E., & Maciejovsky, B. (2002). Simultaneous over- and under confidence: Evidence from experimental asset markets. *Journal of Risk and Uncertainty*, 25(1), 65-85.
- Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing—An ERP study. *Biological Psychology*, 80(1), 75-83.
- Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly*, 10(1), 62-102.
- Koellinger, P., Minniti, M., & Schade, C. (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, 28(4), 502-527.
- Krasnova, H., Kolesnikova, E., & Guenther, O. (2009). "It won't happen to me!": Self-disclosure in online social networks. *Amcis 2009 Proceedings*, 343.
- Kuppuswamy, V., & Bayus, B. L. (2015). A Review of Crowdfunding Research and Findings. In preparation for Golder, P. and Mitra, D. (eds.) Handbook of New Product Development Research, Forthcoming. Available at SSRN.
- Lea, S. E. G. & Webley, P. (2006). Money as a tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences*, 29(2) 161-209.



- Lea, S. E. G. (2008). Evolutionary Psychology and Economic Psychology, in A. Lewis (ed.), *The Cambridge Handbook of Psychology and Economic Behaviour*. Cambridge: Cambridge University Press: 512-526.
- Lee, B. K., & Lee, W. N. (2004). The effect of information overload on consumer choice quality in an on-line environment. *Psychology & Marketing*, 21(3), 159-183.
- Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings carryover effects of emotions on economic decisions. *Psychological Science*, 15(5), 337-341.
- Letchford, A., Preis, T., & Moat, H. S. (2016). The advantage of simple paper abstracts. *Journal of Informetrics*, 10(1), 1-8.
- Liu, Z. (2005). Reading behaviour in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation*, 61(6) 700-712.
- Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, 140-151.
- Lockwood, F. (2005). *Estimating student workload, readability and implications for student learning and progression*. Proceedings of the Open and Distance Learning Association of Australia. Retrieved from [http://wikieducator.org/images/5/5a/Odlaa2005Lockwoodestimating\\_student\\_workload.pdf](http://wikieducator.org/images/5/5a/Odlaa2005Lockwoodestimating_student_workload.pdf)
- Lund, E., & Gram, I. T. (1998). Response rate according to title and length of questionnaire. *Scandinavian Journal of Public Health*, 26(2), 154-160.
- Lurie, N. H. (2004). Decision making in information-rich environments: The role of information structure. *Journal of Consumer Research*, 30(4), 473-486.
- Luth, H. A. (2010). *Behavioural economics in consumer policy: The economic analysis of standard terms in consumer contracts revisited* (dissertation). Available from urn:NBN:nl:ui:15-1765/19572.
- Malcolm, M., Rosenboim, M., & Shavit, T. (2015). Does attachment to money affect reciprocity behaviour? *Applied Economics Letters*, 22(4), 251-254.
- Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of consumer research*, 8(4), 419-430.
- Marom, D., Robb, A., & Sade, O. (2016). Gender Dynamics in Crowdfunding (Kickstarter): Evidence on Entrepreneurs, Investors, Deals and Taste-Based Discrimination. *Social Science Research Network, Working Paper*. Retrieved from <http://dx.doi.org/10.2139/ssrn.2442954>
- Martin, L. N., & Delgado, M. R. (2011). The influence of emotion regulation on decision-making under risk. *Journal of Cognitive Neuroscience*, 23(9), 2569-2581.
- Masson, M. E. J., & Waldron, M. A. (1994). Comprehension of legal contracts by non-experts: Effectiveness of plain language redrafting. *Applied Cognitive Psychology*, 8(1), 67-85. doi:10.1002/acp.2350080107

- Mayseless, O., & Kruglanski, A. W. (1987). What makes you so sure? effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes*, 39(2), 162-183.
- McCloskey, D. N., & Ziliak, S. T. (1996). The standard error of regressions. *Journal of Economic Literature*, 34(1), 97-114.
- McGregor, M. (2013, October 15). Kickstarter in Australia and New Zealand [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-in-australia-and-new-zealand>
- McGregor, M. (2014a, April 8). Kickstarter in the Netherlands [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-in-the-netherlands>
- McGregor, M. (2014b, June 4). Introducing the Creator Handbook [Web log post]. Retrieved from <https://www.kickstarter.com/blog/introducing-the-creator-handbook>
- McGregor, M. (2014c, September 16). Kickstarter in Scandinavia and Ireland! [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-in-scandinavia-and-ireland>
- McInnes, N., & Haglund, B. J. (2011). Readability of online health information: implications for health literacy. *Informatics for health and social care*, 36(4), 173-189.
- McLaughlin, G. H. (1969). SMOG grading – a new readability formula. *Journal of Reading*, 12(8), 639-646.
- Mendes-Da-Silva, W., Rossoni, L., Conte, B. S., Gattaz, C. C., & Francisco, E. R. (2016). The impacts of fundraising periods and geographic distance on financing music production via crowdfunding in Brazil. *Journal of Cultural Economics*, 40(1), 75-99.
- Mitra, T., & Gilbert, E. (2014, February). The language that gets people to give: Phrases that predict success on Kickstarter. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 49-61). ACM.
- Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *The American Economic Review*, 96(1), 222-235.
- Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, 29(1), 1-16.
- Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502-517.
- Moore, D. A., & Kim, T. G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology*, 85(6), 1121-1135.
- Myers, D. G. (1998). *Social psychology* (5<sup>th</sup> ed.). McGraw-Hill, New York.
- Neight-Allen. (2013, May). django-kickstarter-scraper. Github. [Web log post] Retrieved from <https://github.com/neight-allen/django-kickstarter-scraper#db-dump>.
- Osherenko, A., & André, E. (2007, September). Lexical affect sensing: Are affect dictionaries necessary to analyze affect?. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 230-241). Springer Berlin Heidelberg.

- Park, Y. J., & Santos-Pinto, L. (2010). Overconfidence in tournaments: Evidence from the field. *Theory and Decision*, 69(1), 143-166
- Peng, C. C. (2015). Textbook Readability and Student Performance in Online Introductory Corporate Finance Classes. *Journal of Educators Online*, 12(2), 35-49.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A. L., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: LIWC.net.
- Porter, M.F. (1980). An algorithm for suffix stripping, *Program*, 14, 130-137.
- Powers, R. D., Sumner, W. A., & Kears, B. E. (1958). A recalculation of four adult readability formulas. *Journal of Educational Psychology*, 49(2), 99.
- Profatilov, D. A., Bykova, O. N., & Olkhovskaya, M. O. (2015). Crowdfunding: online charity or a modern tool for innovative projects implementation? *Asian Social Science*, 11(3), 146-151.
- Putrevu, S., Tan, J., & Lord, K. R. (2004). Consumer responses to complex advertisements: The moderating role of need for cognition, knowledge, and gender. *Journal of Current Issues & Research in Advertising*, 26(1), 9-24.
- Reutner, L., & Wänke, M. (2013). For my own benefit or for the benefit of others: Reminders of money moderate the effects of self-related versus other-related persuasive arguments. *Social Psychological and Personality Science*, 4(2), 220-223.
- Risoldi Cochrane, Z., Gregory, P., & Wilson, A. (2012). Readability of consumer health information on the Internet: A comparison of US government-funded and commercially funded websites. *Journal of health communication*, 17(9), 1003-1010.
- Robinson, A. T., & Marino, L. D. (2015; 2013). Overconfidence and risk perceptions: Do they really matter for venture creation decisions? *International Entrepreneurship and Management Journal*, 11(1), 149-168.
- Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: Is shorter better? A review and meta-analysis. *Value in Health*, 14(8), 1101-1108.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755-1758.
- Schacht, A., & Sommer, W. (2009). Emotions in word and face processing: early and late cortical responses. *Brain and Cognition*, 69(3), 538-550.
- Shafqat, W., Lee, S., Malik, S., & Kim, H. (2016). The language of deceivers: Linguistic features of crowdfunding scams. *Proceedings of the 25th International Conference Companion on World Wide Web, Montréal, Québec, Canada*.
- Shenk, D. (2003). Information Overload, Concept of, in D. H. Johnston (ed.), *Encyclopaedia of International Media and Communications*. New York: Elsevier: 395-405.
- Simon, H. (1971). Designing organizations for an information-rich world, in M. Greenberger (ed.), *Computers, communications, and the public interest*. Baltimore, MD: The Johns Hopkins Press: 37-72.
- Smith, E. A., & Senter, R. J. (1967). *Automated readability index*. Cincinnati University OH.

- Sobkowicz, P., Thelwall, M., Buckley, K., Paltoglou, G., & Sobkowicz, A. (2013). Lognormal distributions of user post lengths in Internet discussions—a consequence of the Weber-Fechner law?. *EPJ Data Science*, 2(1), 1.
- Soll, J. D., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 299-314.
- Song, H., & Schwarz, N. (2008). If it's hard to read, it's hard to do processing fluency affects effort prediction and motivation. *Psychological Science*, 19(10), 986-988.
- Spache, G. (1953). A new readability formula for primary-grade reading materials. *Elementary School Journal*, 53, 410-413.
- Spache, G. D. (1968). *Good reading for poor readers*. (Rvd. 9th Ed.) Champaign, Illinois: Garrard.
- Stark, D. P., & Choplin, J. M. (2009). License to deceive: Enforcing contractual myths despite consumer psychological realities, a. *New York University Journal of Law & Business*, 5 (1): 617-744.
- Strack, F., and Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437-446.
- Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *Journal of Marketing*, 71(3), 171-193.
- Strickler, Y. (2011). *Shortening the Maximum Project Length*. [Web log post] <https://www.kickstarter.com/blog/shortening-the-maximum-project-length>
- Strickler, Y. (2014a, June 4). Introducing launch now and simplified rules [Web log post]. Retrieved from <https://www.kickstarter.com/blog/introducing-launch-now-and-simplified-rules-0>
- Strickler, Y. (2014b, September 20). An update to our terms of use [Web log post]. Retrieved from <https://www.kickstarter.com/blog/an-update-to-our-terms-of-use>
- Strickler, Y. (2015, April 28). Hallo Deutschland! [Web log post]. Retrieved from <https://www.kickstarter.com/blog/hallo-deutschland>
- Strickler, Y., Chen, P., & Adler, C. (2012, November 1). Kickstarter in the UK [Web log post]. Retrieved from <https://www.kickstarter.com/blog/kickstarter-in-the-uk-0>
- Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47(2), 143-148.
- Swift, J. (1841). *The Works of Jonathan Swift*. London: William Clowes and Sons.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Titmuss, Richard M. (1970). *The Gift Relationship*. London: Allen and Unwin.

- Tong, L., Zheng, Y., & Zhao, P. (2013). Is money really the root of all evil? the impact of priming money on consumer choice. *Marketing Letters*, 24(2), 119-129.
- Torgler, B., & Piatti, M. (2016). *A Century of American Economic Review: Insights on Critical Factors in Journal Publishing*. Basingstoke: Palgrave Macmillan.
- Townsend, D. M., Busenitz, L. W., & Arthurs, J. D. (2010). To start or not to start: Outcome and ability expectations in the decision to start a new venture. *Journal of Business Venturing*, 25(2), 192-202.
- Trachtman, R. (1999). The Money Taboo: Its Effects in Everyday Life and in the Practice of Psychotherapy. *Clinical Social Work Journal*, 27(3), 275–288.
- Trusiak, H. C. (2016). *What's happening in the crowd? : analysis of crowdfunding contributor behaviors using the theory of planned behaviour* (Masters dissertation). Retrieved from <https://sdsu-dspace.calstate.edu/handle/10211.3/169215>
- Van Boom, W. H., Desmet, P., & Van Dam, M. (2016). “If it’s easy to read, it’s easy to claim”—the effect of the readability of insurance contracts on consumer expectations and conflict behaviour. *Journal of Consumer Policy*, 39(2), 187-197.
- Venugopal, V., & Bagadia, S. (2015). Understanding the dynamics of crowdfunding: Kickstarter Edits. CS224N Final Project.
- Voelker, T. A., & McGlashan, R. (2013). What is crowdfunding? Bringing the power of kickstarter to your entrepreneurship research and teaching activities. *Small Business Institute Journal*, 9(2), 11-22.
- Vohs, K. D, Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science*, 314(5802), 1154-1156.
- Weinberg, B. A. (2009). A model of overconfidence. *Pacific Economic Review*, 14(4), 502-515.
- Weinberger, C. J., Evans, J. A., & Allesina, S. (2015). Ten simple (empirical) rules for writing science. *PLOS Computational Biology*, 11(4), e1004205.
- Welcoming Austria, Belgium, Italy, Luxembourg, and Switzerland to the Kickstarter Community!* (2015, June 16). [Web log post]. Retrieved from <https://www.kickstarter.com/blog/welcoming-austria-belgium-italy-luxembourg-and-switzerland-to-th>
- Wheeler, L. R., & Smith, E. H. (1954). A practical readability formula for the classroom teacher in the primary grades. *Elementary English*, 31(7), 397-399.
- Williams, P. (2005). Consumer understanding and use of health claims for foods. *Nutrition reviews*, 63(7), 256-264.
- Wilson, M. (2009). Readability and patient education materials used for low-income populations. *Clinical Nurse Specialist*, 23(1), 33-40.
- Wilson, T. D., Houston, C., E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 124(4), 387-402.

- Wright, P. (1981). "The instructions clearly state..." can't people read?. *Applied ergonomics*, 12(3), 131-141.
- Wu, P. F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing*, 30(11), 971-984.
- Xiao, E., & Houser, D. (2009). Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange. *Journal of Economic Psychology*, 30(3), 393-404.
- Xu, A., Yang, X., Rao, H., Fu, W. T., Huang, S. W., & Bailey, B. P. (2014). Show me the money!: An analysis of project updates during crowdfunding campaigns. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Canada*.
- Zhou, M., Du, Q., Fan, W., Qiao, Z., Wang, G., & Zhang, X. (2015). Money talks: A predictive model on crowdfunding success using project description. *Twenty-first Americas Conference on Information Systems, Puerto Rico, 2015*.

# Appendices

## Appendix A: Tables

**Table A1.1** Descriptive statistics per project

	N	Mean	SD	Min.	Max.
Backers		88.57	843.28	0	91585
Raised (\$)		6423.31	68773.82	0	10300000
Word count		494.90	465.61	4	32135
Funded	37962	515.81	458.69	4	26305
Not Funded	43930	476.83	470.77	4	32135
Text character count		2807.17	2663.27	19	191344
Goal (\$)		15276.92	221083.8	.01	21500000
Duration (days)		37.42	16.03	1	91.95
Before rule change	21541	46.02	21.45	1	91.95
After rule change a	60351	34.35	12.19	1	60.04
Funded					
No	43930	.54			
Yes	37962	.46			
Update dummy					
No	75414	.92			
Yes	6478	.08			
U.S. dummy					
No	5729	.07			
Yes	76613	.93			

<sup>a</sup> In June 2011, Kickstarter adjusted the maximum project duration from 90 days to 60 days (Strickler, 2011).

Notes: N is 85,743 unless otherwise stated. Geographic location is excluded as non-informative.

**Table A1.2** Search terms to identify ‘edits’

General edit	Reminder	Progress
Update*	days to go	ve reached
UPDATE*	hours to go	we reached
	weeks to go	goal reached!
	days left	GOAL REACHED
	hours left	goal met!
	weeks left	GOAL MET
	days remain	reached our goal
	hours remain	we did it!
	weeks remain	we made it!
		WE DID IT
		WE MADE IT
		funding reached
		we are funded!
		WE ARE FUNDED
		e are FUNDED
		has been reached

Note: Case insensitive unless word/phrase contains at least one upper case letter.

**Table A2.1** Readability Indices

Abbreviation	Name	Reference	Elements	Google Scholar citation
ARI	Automated Readability Index	Smith & Senter (1967)	W, S <sub>t</sub> , C	91
ARI <sub>simple</sub>	Automated Readability Index Simple	Smith & Senter (1967)	W, S <sub>t</sub> , C	91
COLEMAN	Coleman	Coleman (1965)	W, W <sup>1Sy</sup>	23
CL	Coleman-Liau	Coleman & Liau (1975)	W, S <sub>t</sub> , C	284
CL <sub>GRADE</sub>	Coleman-Liau	Coleman & Liau (1975)	W, S <sub>t</sub> , C	284
CL <sub>SHORT</sub>	Coleman-Liau	Coleman & Liau (1975)	W, S <sub>t</sub> , C	284
DC	Dale-Chall	Dale & Chall (1948)	W, S <sub>t</sub> , W <sub>-WL</sub>	1416
DB <sub>1</sub>	Danielson-Bryan 1	Danielson & Bryan (1963)	C, S <sub>t</sub>	30
DB <sub>2</sub>	Danielson-Bryan 2	Danielson & Bryan (1963)	C, S <sub>t</sub>	30
ELF	Easy Listening Formula	Fang (1966)	W <sub>2Sy</sub> , S <sub>t</sub>	32
<b>FRES</b>	<b>Flesch Reading-Ease</b>	<b>Flesch (1948)</b>	<b>W, S<sub>t</sub>, C, S<sub>y</sub></b>	<b>2907</b>
FK	Flesch-Kincaid Grade Level	Kincaid et al. (1975)	W, S <sub>t</sub> , C, S <sub>y</sub>	1216
FRES <sub>PSK</sub>	Flesch Powers-Sumner-Kearl	Powers et al. (1958)	W, S <sub>t</sub> , C, S <sub>y</sub>	115
FOG	Frequency of Gobbledygook	Gunning (1952)	W, S <sub>t</sub> , W <sub>3Sy</sub>	1281
FOG <sub>PSK</sub>	Frequency of Gobbledygook (Powers-Sumner-Kearl)	Powers et al. (1958)	W, S <sub>t</sub> , W <sub>3Sy</sub>	115
FORCAST	FORCAST	Caylor et al. (1973)	W, W <sup>1Sy</sup>	80
FORCAST <sub>RGL</sub>	FORCAST Readability Grade Level	Klare (1974)	W, W <sup>1Sy</sup>	659
nWS	Neue Wiener Sachtextformeln	Bamberger & Vanecek (1984)	W, W <sub>3Sy</sub>	72
RIX	Anderson's Readability Index	Anderson (1983)	W <sup>7C</sup> , S <sub>t</sub>	92
SMOG	Simple Measure of Gobbledygook	McLaughlin (1969)	W <sub>3Sy</sub> , S <sub>t</sub>	1484
SMOG <sub>simple</sub>	Simple Measure of Gobbledygook Simplified	McLaughlin (1969)	W <sub>3Sy</sub> , S <sub>t</sub>	1484
SP	Spache Revised Formula	Spache (1974)	W, S <sub>t</sub> , W <sub>-WL</sub>	367
SP <sub>OLD</sub>	Spache	Spache (1953)	W, S <sub>t</sub> , W <sub>-WL</sub>	367
WS	Wheeler-Smith	Wheeler & Smith (1954)	W, W <sub>2Sy</sub> , S <sub>t</sub>	27

*Note:* W = number of words, S<sub>t</sub> = number of sentences, C = number of characters, S<sub>y</sub> = number of syllables, W<sub>#Sy</sub> = words with at least # syllables, W<sup>1Sy</sup> = words with exactly 1 syllable, W<sup>#C</sup> = words with at least # characters, W<sub>-WL</sub> = number of words not in a certain word list. Google Scholar citation count as of 4<sup>th</sup> August, 2016.



**Table A2.2** Number of projects beyond turning point

Readability measure	ln(Raised)		ln(Backers)	
	# projects	percentage	# projects	percentage
ARI	2922	4.19%	3422	4.91%
ARI <sub>simple</sub>	3009	4.31%	3562	5.11%
COLEMAN	1115	1.60%	1768	2.53%
CL	122	0.17%	189	0.27%
DC	687	0.98%	832	1.19%
DB <sub>1</sub>	1486	2.13%	1643	2.36%
DB <sub>2</sub>	850	1.22%	1068	1.53%
ELF	1581	2.27%	1752	2.51%
<b>FRES</b>	<b>508</b>	<b>0.73%</b>	<b>406</b>	<b>0.58%</b>
FRES <sub>PSK</sub>	1154	1.65%	1073	1.54%
FOG	3884	5.57%	5204	7.46%
FOG <sub>PSK</sub>	7171	10.28%	12609	18.08%
FORCAST	1115	1.60%	1768	2.53%
FORCAST <sub>RGL</sub>	1115	1.60%	1768	2.53%
Fucks	3508	5.03%	5070	7.27%
nWS	1346	1.93%	2094	3.00%
RIX	1513	2.17%	1823	2.61%
SMOG	1585	2.27%	2212	3.17%
SMOG <sub>simple</sub>	1585	2.27%	2212	3.17%
SP	1915	2.75%	2048	2.94%
SP <sub>OLD</sub>	2080	2.98%	2213	3.17%
WS	1581	2.27%	1752	2.51%
<b>Average</b>	<b>1901</b>	<b>2.72%</b>	<b>2568</b>	<b>3.68%</b>



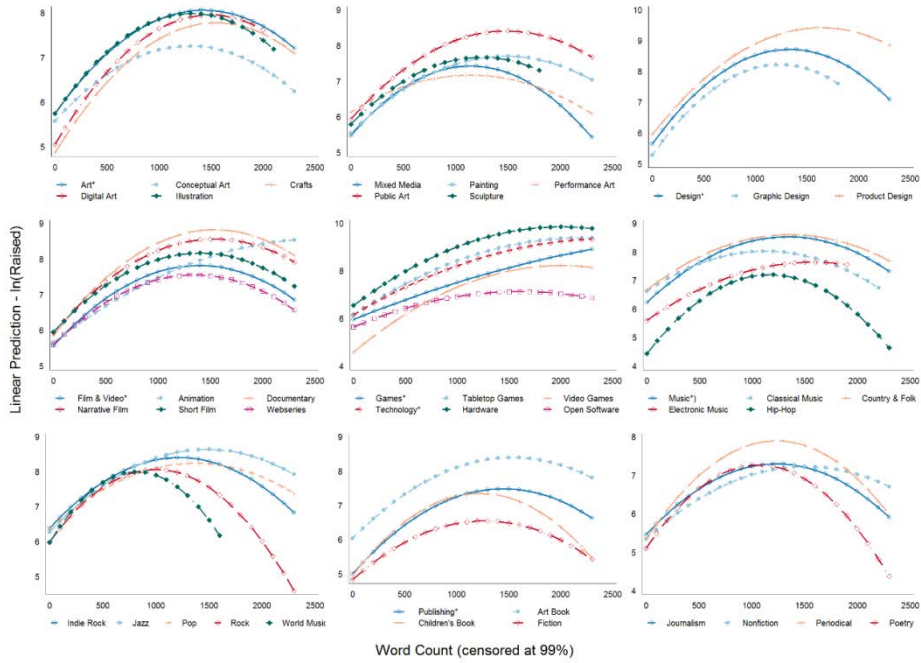
**Table A3.1** Confidence Descriptive statistics

	N	Mean	SD	Min.	Max.
Confidence					
(Duration ratio)		.52	.25	0	.98
Confidence <sup>2</sup>					
(Duration ratio)		.33	.21	0	.98
Confidence					
(Certainty ratio)	68163	.40	.22	0	1
Confidence <sup>2</sup>					
(Certainty ratio)	68163	.21	.21	0	1
Duration					
		38.41	16.61	1	91.95
No influence	5533	52.76	23.13	1	89.99
First Blog Post	16002	43.69	20.32	1	91.95
Rule Change	48064	34.99	12.49	1	60.04
Backers		77.11	742.61	0	87142
Goal (\$)		13938.2	198062.6	.01	21500000
Raised (\$)		5706.92	65384.29	0	10300000
Funded					
No	37308	.54			
Yes	32291	.46			
Word count		470.10	437.21	4	32135
Official category					
Comics	1784	.03			
Dance	1013	.01			
Design	2468	.04			
Fashion	2007	.03			
Film & video	18972	.27			
Food	2386	.03			
Games	3102	.04			
Music	16510	.24			
Photography	2268	.03			
Publishing	7713	.11			
Technology	1320	.02			
Theatre	3602	.05			
U.S. dummy					
No	3562	.5			
Yes	66037	.95			

*Notes:* *N* is 69,599 unless otherwise stated. Art is the reference group, with 6,454 observations. Latitude and longitude are excluded because they are not informative.

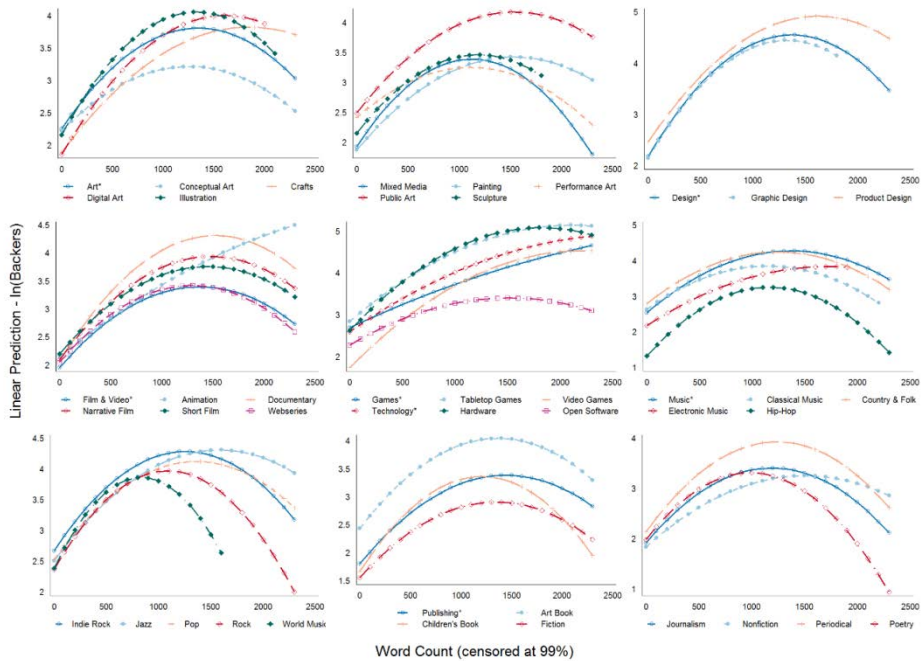
## Appendix B: Figures

**Figure B1.1** Marginal Effects of word count on  $\ln(\text{Raised})$  using sub-categories.



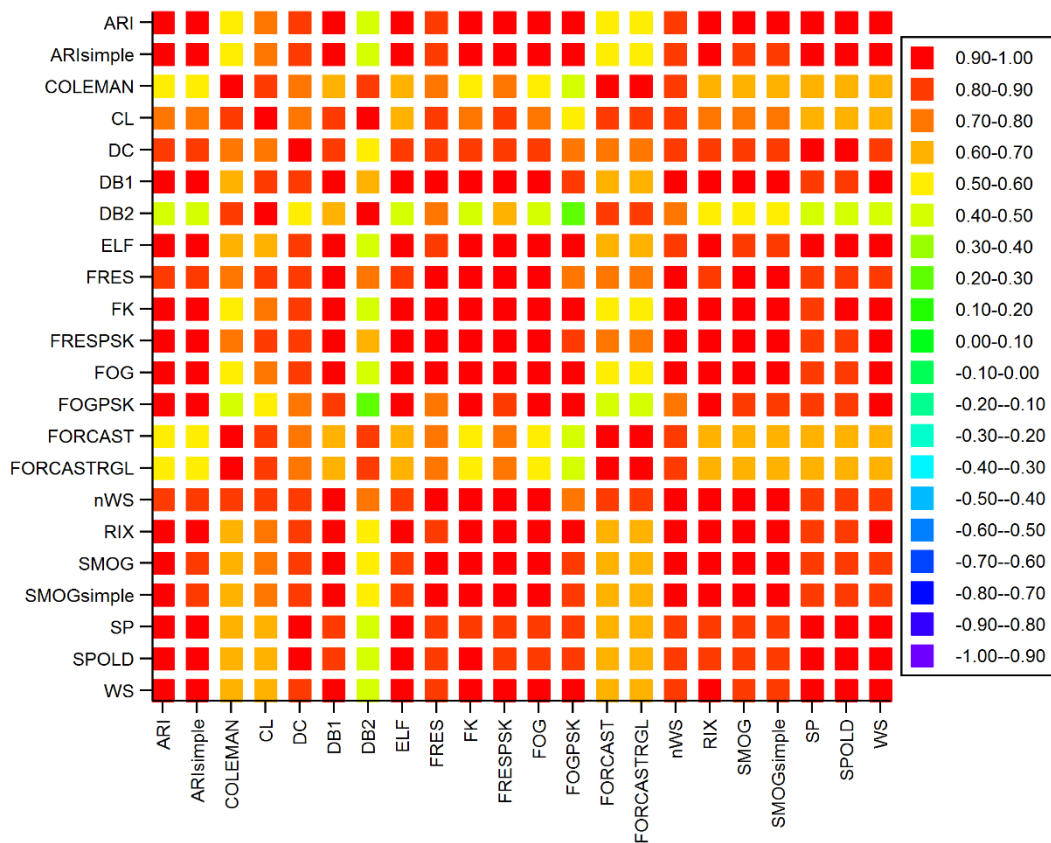
Note: \* are sub-categories that have not been specified.

**Figure B1.2** Marginal Effects of word count on  $\ln(\text{Backers})$  using sub-categories.



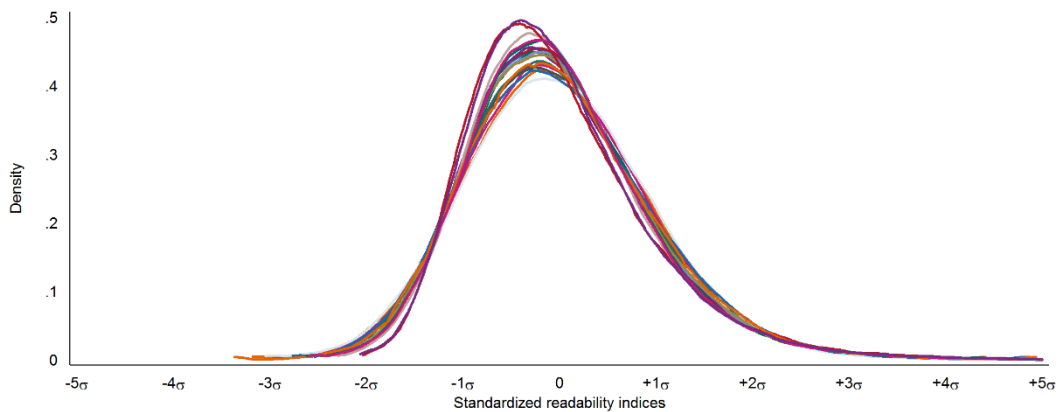
Note: \* are sub-categories that have not been specified.

**Figure B2.1** Correlation matrix between Readability indices



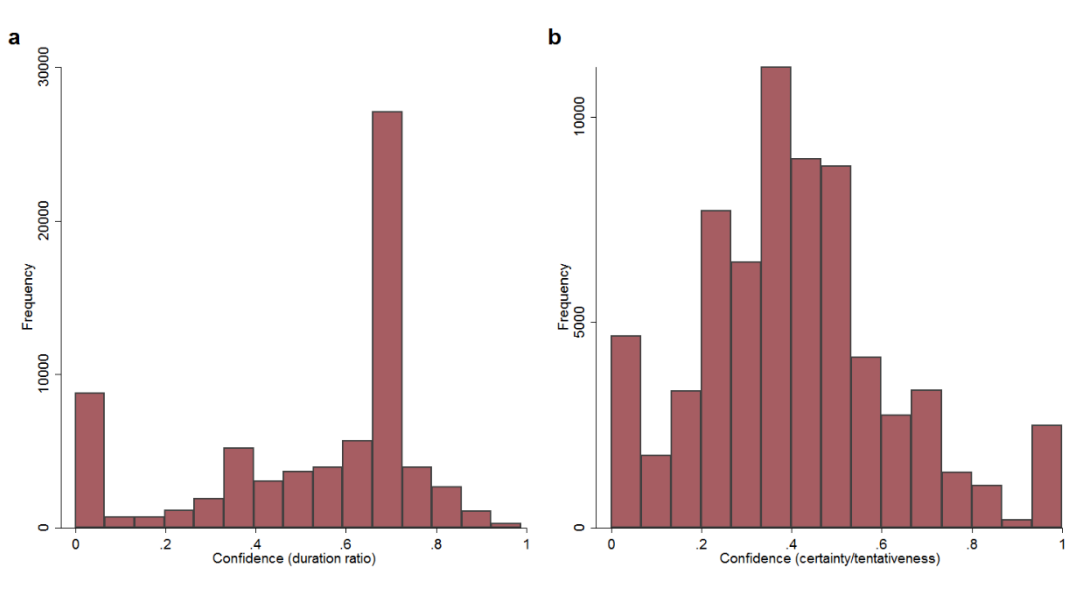
*Notes:* The following readability scores are reversed such that higher values equal to lower reading ease: *Coleman*, *Coleman-Liau*, *Dale-Bhall*, *Danielson-Bryan<sub>2</sub>*, and *Flesch*.

**Figure. B2.2** Distribution of Readability indices (standardized z-score)

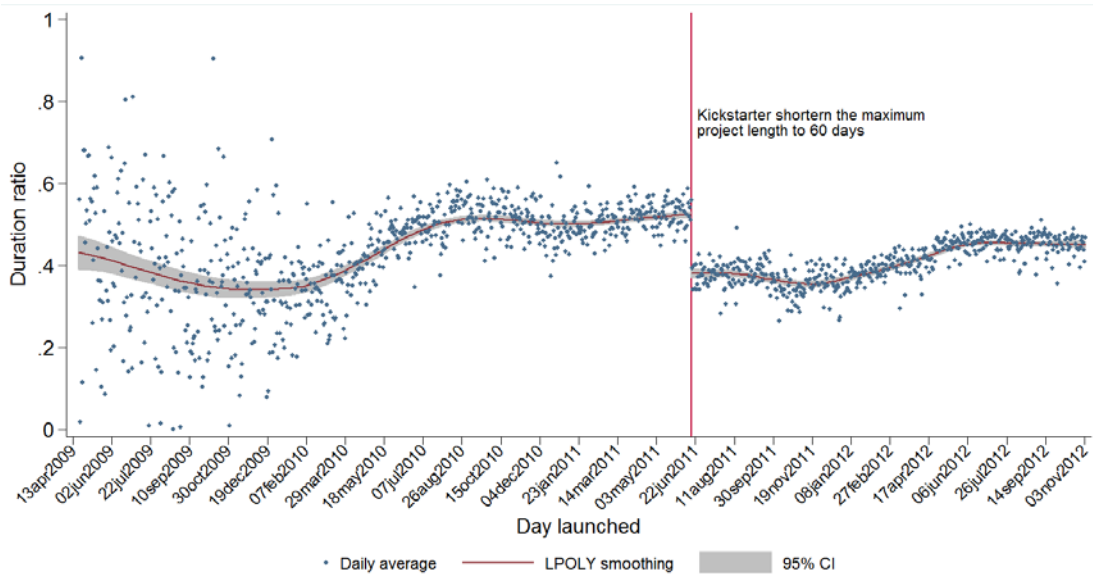


*Notes:* Higher values means lower reading ease. Raw score is standardised with mean equals to 0 and standard deviation equals to 1 (z-score)

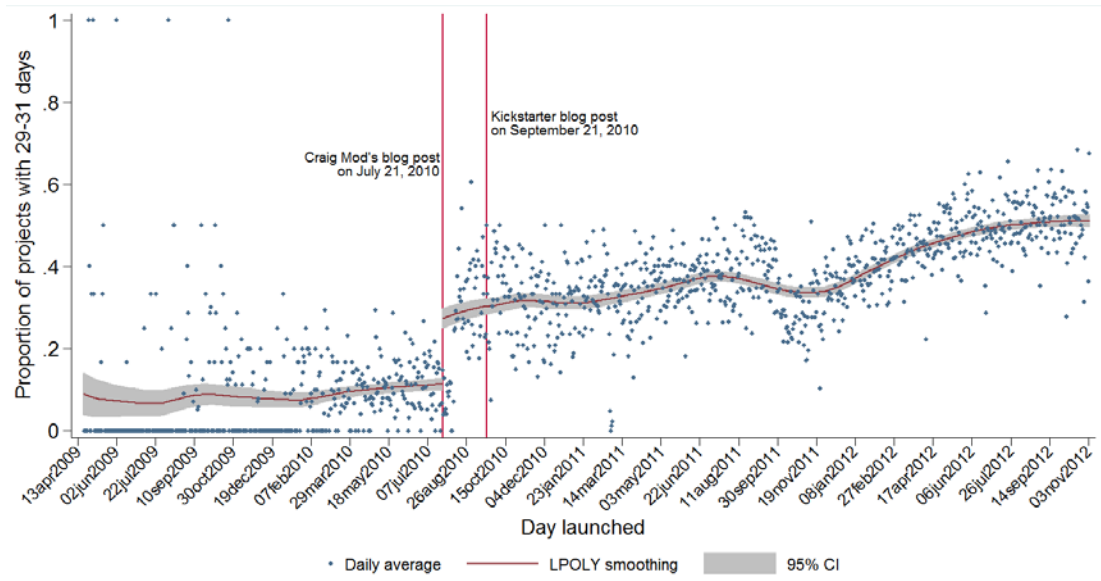
**Figure B3.1** Confidence distribution: Duration Ratio (a) and Certainty Ratio (b)



**Figure B3.2** Duration Ratio in different periods



**Figure B3.3** Proportion of projects with different Durations



## Appendix C

### **Sample of Money Terms and Porter Stemmed Terms**

The Money dictionary includes words such as ‘atm, atms’, ‘audit, audited, auditing’, ‘fund, funded, funds’ (Pennebaker et al., 2007), using Porter stemming these monetary terms are now represented by ‘atm’ , ‘audit’, and ‘fund’.