Dissecting the Gender Gap in Entrepreneurial Ambitions and Success: Evidence from Crowdfunding

by Anna May

Submitted to Central European University Department of Economics

In partial fulfillment of the requirements for the degree of Master of Arts in Economic Policy in Global Markets

Supervisor: Gábor Békés

Budapest, Hungary 2018

Acknowledgments

I would like to thank Gábor Békés for all his support and enthusiasm about this project.

I am also grateful to Jonathan Lansey, who granted me access to his data, and Ancsa and Johannes who took me on an incredible journey during my CEU studies which I hope that will not end here.

Finally, I am lucky to have the unconditional love and support of Barna and my family.

Abstract

Despite the growing number of women-owned businesses, women still lack equal opportunities for raising external capital. An alternative form of traditional private equity sources, crowdfunding, is expected to democratize access to capital: the transparent nature of online crowdfunding marketplaces might mitigate the effect of discrimination and biased social structures against women and minorities. In my thesis, I seek to examine which parts of the gender gap in entrepreneurial participation and success can be closed by using this new form of financing, and what are the characteristics of crowdfunding platforms that can be adopted by venture capitalist and business angels to make their decisions more equitable. I analyze the gender differences in risk-taking and domain choice of entrepreneurs through examining more than 15 thousand product development projects initiated on the largest reward-based crowdfunding website, Kickstarter. I find no gender difference in risk-taking and success in terms of setting and reaching the desired target amount, but I find that men collect/receive more premium over their goal. Using topic modelling on the textual descriptions of the campaigns, I show that women and men engage in different fields, and women tend to choose less lucrative domains. Domain choice explains a substantial part of the gender gap in premium on Kickstarter. Based on these findings, as a first step towards narrowing the gender gap in access to capital, traditional sources of financing should adapt more transparent and equitable policies as seen on crowdfunding platforms. However, for long term change, regulators need to take into account deeper underlying social structures that affect men and women's domain choices and that cannot be addressed solely by changing entrepreneurship policies.

Table of Contents

A	cknov	wledgments	ii
A	bstra	ct	iii
\mathbf{Li}	st of	Tables	\mathbf{vi}
Li	st of	Figures	ix
1	Intr	roduction	1
2	Rel	ated work on women and entrepreneurship	6
	2.1	Gender differences funding opportunities	6
	2.2	Gender differences in entrepreneurial ambitions	7
		2.2.1 Risk-taking	7
		2.2.2 Self-selection to domains in entrepreneurship	8
	2.3	Gender-based discrimination	9
	2.4	Summary	9
3	Bac	kground	11
	3.1	About crowdfunding	11
	3.2	Gender differences in the crowdfunding literature	13
4	Dat	a and Features	15
	4.1	About Kickstarter	15
	4.2	Data collection	16
	4.3	Gender recognition	17
	4.4	Focusing on product development projects	19

5	Ger	ıder di	fferences in descriptive statistics	20
	5.1	Decon	aposing the difference in all dollars raised	21
		5.1.1	Difference in participation rates	21
		5.1.2	Differences in ambitions: goal set	21
		5.1.3	Differences in success: funding rates	23
	5.2	Gende	er differences in categories	24
6	Ger	nder di	fferences in topics	27
	6.1	Topic	models	27
		6.1.1	Topics of the Technology category	28
		6.1.2	Topics of the Design category	29
		6.1.3	Topics of the Crafts category	30
		6.1.4	Topics of the Games category	31
		6.1.5	Topics of the Fashion category	32
	6.2	Topics	s, interest, gender representation	32
	6.3	Projec	et novelty	35
	6.4	Contro	ols for project management	35
		6.4.1	Number of reward levels	36
		6.4.2	Number of team members	37
		6.4.3	Length of the project description and risk section	37
7	Res	ults		38
	7.1	Predic	ting success	38
	7.2	Predic	eting premium	41
8	Dis	cussior	1	44
	8.1	Dissec	ting the gender gap in dollars raised	44
		8.1.1	Gender gap in funding ratio	45
	8.2	Self-se	election and domain	45
	8.3	Succes	ss of women on online crowdfunding platforms	47
	8.4	Policy	implications	49
9	Cor	nclusio	n	50
A	App	pendix		53

Bibliography

List of Tables

4.1	Result of the gender recognition method Gender Guesser	18
5.1	Dissecting the gender gap in all dollars raised on Kickstarter	20
5.2	Descriptive statistics for the variables goal and pledged	22
5.3	Funding ratios for successful men and women (descriptive statistics)	23
5.4	Descriptive statistics per gender and category	26
6.1	Gender representation in topics of the Technology category	29
6.2	Gender representation in topics of the Design category $\ldots \ldots \ldots \ldots$	30
6.3	Gender representation in topics of the Crafts category	31
6.4	Gender representation in topics of the Games category	32
6.5	Gender representation in topics of the Fashion category	33
7.1	Predicting success of Kickstarter campaigns	40
7.2	Predicting premium on Kickstarter projects	42
A.1	Main descriptives for the initial unrestricted sample of 94 896 Kickstarter	
	projects	53
A.2	Main descriptives for the final sample of 15 059 product development Kick-	53
A.2	projects Main descriptives for the final sample of 15 059 product development Kick- starter projects	53 53
A.2 A.3	projects Main descriptives for the final sample of 15 059 product development Kick-starter projects Dissecting the gender gap in all dollars raised on Kickstarter (for the users	53 53
A.2 A.3	projects	53 53 54
A.2 A.3 A.4	projectsMain descriptives for the final sample of 15 059 product development Kick-starter projectsDissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the top 5%)Dissecting the gender gap in all dollars raised on Kickstarter (for the users	53 53 54
A.2 A.3 A.4	projectsMain descriptives for the final sample of 15 059 product development Kick-starter projectsDissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the top 5%)Dissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the bottom 95%)	53535454
A.2 A.3 A.4 A.5	projectsMain descriptives for the final sample of 15 059 product development Kick-starter projectsDissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the top 5%)Dissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the bottom 95%)Topic within the Technology category with their most frequent keywords, k	53 53 54 54
A.2 A.3 A.4 A.5	projectsMain descriptives for the final sample of 15 059 product development Kick-starter projectsDissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the top 5%)Dissecting the gender gap in all dollars raised on Kickstarter (for the userswith goals set in the bottom 95%)Topic within the Technology category with their most frequent keywords, k= 3	 53 53 54 54 54 54
A.2A.3A.4A.5A.6	projectsMain descriptives for the final sample of 15 059 product development Kick- starter projectsStarter projectsDissecting the gender gap in all dollars raised on Kickstarter (for the users with goals set in the top 5%)Dissecting the gender gap in all dollars raised on Kickstarter (for the users with goals set in the bottom 95%)Topic within the Technology category with their most frequent keywords, k $= 3 \dots$ Topic within the Design category with their most frequent keywords, k = 12	 53 53 54 54 54 54 54
 A.2 A.3 A.4 A.5 A.6 A.7 	projects	 53 53 54 54 54 54 54 55

A.9	Topic within the Fashion category with their most frequent keywords, $k = 4$	55
A.10	Sample project short descriptions from two topics of the Technology cate-	
	gory: community software and device	56
A.11	Sample project short descriptions from two topics of the Design category:	
	lifestyle and wallets/organizers	57
A.12	Sample project short descriptions from two topics of the Crafts category:	
	candles/gifts and woodworking $\ldots \ldots \ldots$	58
A.13	Predicting % funded for failed projects on Kickstarter	59

List of Figures

5.1	Density of goal set and dollars pledged per gender	22
5.2	Density of premium (% funded over goal) for successful projects \ldots .	22
5.3	Funding rates and women's participation per goal decile	25
5.4	Average dollars pledged and share of men and women per categories	26
6.1	Share of women in topic and funding rates	33
6.2	Share of women and difference between men's and women's funding rates	
	(in women's percentages) \ldots	34
6.3	Topic entropy of women's and men's projects	36

Chapter 1

Introduction

Studying gender disparity in entrepreneurship and particularly in access to external capital sources has been the interest of economist for a long time. Although 39% of the companies in the US have woman majority owner, less than 3% of the firms who received venture capital funding were founded by at least a women [American Express Open, 2017, Olsen, 2018]. These numbers show that the times when women did not have the ambitions to start ventures are over, but funding opportunities have not accommodated to these changes. Women entrepreneurs frequently mention the lack of access to capital as one of the most important barriers to venturing [Coleman and Robb, 2009].

Recently, a new form of fundraising emerged called crowdfunding. Instead of relying on a few sophisticated investors, crowdfunding platforms enable entrepreneurs to collect the necessary funds for their venture from a large number of individuals with a small contribution each [Mollick, 2014, Belleflamme et al., 2015]. Crowdfunding is found to be a more transparent and equitable source of financing that promises the democratization of access to capital by mitigating gender-based discrimination and bias against underrepresented groups [Horvat and Papamarkou, 2017].

Further, crowdfunding marketplaces offer the opportunity for social science researchers to observe the characteristics of entrepreneurial projects, the determinants of success and the community's judgment about them. Most crowdfunding platforms provide opportunities for project owners to publish a description of their campaign, which can be analyzed with text mining methods to process the information content in an automated way. Thus, examining such online platforms can hold interesting conclusions which can be used to choose the right policies to equalize the access to traditional forms of external financing such as angel investment and venture capital, too. Women and men entrepreneurs are found to be very similar in terms of education and their ambitions to make their idea happen [Cohoon et al., 2010], yet women seem to be less successful than men: ventures led by women are more likely to fail and are expected to have less revenue, sales and profit [Coleman and Robb, 2009].

Previous research suggests that there are common confounders behind women-led enterprises' lower performance and weaker access to capital. First, women are seen more risk-averse than men, especially in financial decisions [Powell and Ansic, 1997]. Consequently, women are less likely to raise external capital, and they also ask for less money when they do [Coleman and Robb, 2009].

Second, women and men are not only expected to choose gender-congruent occupations [Blau and Kahn, 2017], but they are also more likely to start ventures in fields with traditionally female and male domains in accordance with their gender. In the US, 50% of the businesses owned by women can be found in three industries: human-capital-intensive services such as hair salons, pet care businesses (23%), health care and social assistance such as child day care and home health services (15%) and professional, scientific and technical services (12% of all women-owned firms) [American Express Open, 2017]. Domains where the share of women is the highest tend to be less profitable than fields with high representation of men [Loscocco and Robinson, 1991].

In my thesis I chose to analyze gender differences in risk-taking and domain choice of entrepreneurs through examining more than 15 thousand Kickstarter product development projects. Kickstarter is the largest reward-based crowdfunding website established in 2009¹. Observing patterns of project-level features and their connection to success on an online crowdfunding platform helps to disentangle various external factors behind women's worse access to capital such as the interests of investors or discrimination and internal factors such as domain choice and goal set.

I analyze the numeric and textual features of my sample of Kickstarter projects within the categories (fields) Design, Crafts, Fashion, Games and Technology. Although founders launch Kickstarter projects in more artistic categories such as Music and Publishing, I chose to focus on campaigns that are strictly concerned with a development of a product. The evaluation of artistic ventures would go beyond the scope of this thesis.

Previous work has examined gender differences on various types of crowdfunding platforms. An interesting but understudied feature of crowdfunding platforms is that although

¹https://www.kickstarter.com/about?ref=global-footer

according to previous research, men are slightly more likely to fail², but if they do succeed, they overshoot their goal by significantly more than women [Marom et al., 2013]. To fill these gaps in the existing literature, I will examine the funding rates of failed and successful projects to find out why men raise more premium with successful projects.

RQ1: Why do men raise more premium over their goal than women do?

Although sizable differences have been found between women's representation in categories, to my knowledge, no research focuses on gender-based self-selection *within* these categories. Broad categories such as Design or Crafts can include a variety of products ranging from pet accessories to seismographs in the Design category and from jewelry to woodwork in the Crafts category. The typical goal amount and profitability of projects also varies within categories, so controlling only for these wide categories can leave certain gender differences hidden.

RQ2: Do women entrepreneurs' choices of topic explain their lower premium?

To uncover gender differences within the aforementioned broad categories, I apply a probabilistic text mining method called topic modelling on the free text descriptions of the projects to create more fine-grained fields [Blei et al., 2003]. We see that even within categories, women choose to undertake projects in stereotypically female domains such as building community platforms, while men are more likely to engage in masculine projects such as computer games or applications. Further, we can observe that domains with high representation of women have lower average funding rates: women start ventures in less profitable fields on the platform.

Using the fine-grained topic classification, I also compute a score describing campaign uniqueness called *topic entropy*, which measures the extent to which the particular project defies the topic classification. Horvát et al. uses this statistic to measure campaign novelty [Horvát et al., 2018]. My findings suggest that there are no significant gender differences in campaign uniqueness: men and women are expected to propose equally novel campaigns.

In my sample of product development projects, project owners raised USD 143 million altogether, from which USD 22 million was raised by women and USD 121 million was

 $^{^{2}}$ A project fails on Kickstarter if contributions does not reach the goal set by the project owner.

raised by men. The difference in money pledged can be decomposed to three factors: the participation rate of women on the platform (explains 76% of the difference), the goal they set (explains 2% of the gap) and the ratio of average dollars pledged over the average goal (accounts for 22% of the gap). I call the ratio of dollars pledged over goal set *funding ratio*. Descriptive statistics show that men and women are almost equally likely to reach their target (succeed) on Kickstarter. However, men who succeeded are expected to raise much higher premium over their target: successful women's funding ratio is 2.35 on average, while the same measure is 4.58 for men, thus, the fundamental difference between men's and women's projects in terms of performance is in the premium they raise over their goal.

I build linear regression models to assess the importance of gender on collecting external capital on Kickstarter. Gender is an insignificant explanatory factor for project success as a binary variable (reaching its goal), but it is a significant factor influencing the premium project owners get above their goal set if the project succeeds, even after controlling for goal amount, project duration, project management controls and the fine-grained topics. However, the topics can explain a substantive share of gender differences in premium raised: just by including topic controls instead of the broader category controls, the gender premium gap shrinks from 58% to 45%, which also comes with a 5% points increase in the model's R-squared. Interestingly, campaign uniqueness has an insignificant coefficient in the model explaining success differences, but it is a significant negative predictor of premium raised over goal: investors seems to punish projects that are too experimental with their mixed domain choice.

I conclude that gender differences in terms of participation on the online crowdfunding platform Kickstarter are smaller than in traditional forms of fundraising such as in the venture capital industry. Women are also less risk-averse than in the case of traditional ventures, also, they are just as likely to succeed as men are. Still, the highly successful projects that overshoot their goal are mostly initiated by men.

Further, my findings suggests that domain choice is an important determinant of project success on Kickstarter, just as in the "offline" world, meaning that gender differences in interest are such a factor behind women's lower entrepreneurial success which is hard to change. Traditional forms of external capital sources can use the example of crowdfunding platforms to adopt more inclusive and gender-neutral structures where women are willing to take the risk of reaching out for external sources of financing.

My thesis is organized as follows: Chapter 2 summarizes the literature on women entrepreneurs, Chapter 3 describes how crowdfunding works and references the most important existing research on gender disparities in crowdfunding. Chapter 4 introduces my data source and the gender identification process, in Chapter 5, I decompose the gender difference in dollars raised on Kickstarter. In Chapter 6, I create a textual model to create fine-grained topic classification for the project descriptions and introduce my campaign uniqueness measure, in Chapter 7, I show the results of my linear models that predict success and premium over goal. In Chapter 8, I discuss my findings about gender differences on Kickstarter, connect them to disparities in traditional financing sources and add my policy recommendations, in Chapter 9 I summarize my findings and conclude.

Chapter 2

Related work on women and entrepreneurship

More than two-thirds of all companies in the US have a women majority owner as of 2017 after a steady increase in the recent decades [American Express Open, 2017]. These enterprises create more than 23 million jobs, 16% of all the workplaces in the US [Brush et al., 2018]. Survey data shows that successful women and men entrepreneurs are quite similar in most aspects, such as the level of education and their intentions to create wealth using the potential of their business idea [Cohoon et al., 2010]. However, women tend to have smaller businesses than men, and they also experience difficulties accessing external funding: they frequently list insufficient access to capital as the most important issue they face [Coleman and Robb, 2009].

2.1 Gender differences funding opportunities

Women-led ventures frequently face difficulties in accessing the necessary funding for their goals. In general, women start their businesses with a lower initial amount of capital, they are less likely to opt for external financing opportunities [Orser et al., 2006, Sara and Peter, 1998], and they are expected to raise less money when they do decide to seek additional funding [Coleman and Robb, 2009]. In the venture capital industry, women-led ventures rarely receive funding: 15% of all venture capital investment went to firms with at least a women on its executive board between 2011 and 2013, but only 2.7% of all companies receiving venture capital were lead by women [Brush et al., 2018]. More recently, a venture capital database called Pitchbook reported similar numbers: in 2017, 2.2% of all venture

capital investments went to companies founded by women alone [Olsen, 2018]. According to Brush et al., there are three major reasons behind women's low representation among venture capital or angel investment recipients: i) structural barriers which are rooted in the institutional framework of venture capital, ii) human capital needed to start and successfully manage ventures and iii) strategic choice of obtaining additional funding to fuel growth [Brush et al., 2003].

In the next paragraphs, I am going to review the literature of potential confounders that affect the success of women-led ventures, primarily through their access to sufficient funding opportunities. I start with women's ambitions and strategic decisions to seek external funding, then I continue with an important human capital issue, the self-selection to certain professions, lastly, I move on to the potential structural barriers and genderbased discrimination. There are additional confounders which can possibly affect women's access to capital such as managerial abilities, knowlegde and networking [Brush et al., 2003], but examining these characteristics is beyond the scope of this thesis.

2.2 Gender differences in entrepreneurial ambitions

2.2.1 Risk-taking

The willingness to take financial risk is an important determinant of seeking external funding. Men entrepreneurs were found to be less concerned about hazard in various other experiments, which can be explained by the following factors according to Brindley: i) sociopolitical factors such as attitudes towards power and influence, ii) financial factors such as attitudes towards monetary risk and iii) women's decisions about growth paths of the venture [Brindley, 2005]. Interestingly, these factors overlap with the determinants of women's lower representation among venture capital recipients by Brush et al. as cited above ([Brush et al., 2003]). Further, women were found to have a lower preference for risk in financial matters, regardless of domain and costs, thus, they are likely to opt for strategies that provide the maximum level of security [Powell and Ansic, 1997]. Previous research shows that female-led SME-s are less profitable than male-led firms, even after controlling for the age and size of the business and industry. Watson and Robinson explained this difference by the gender disparity in willingness to take risk: gender differences in performance diminished after controlling for the variability of profits Watson and Robinson, 2003. As a consequence, women frequently start smaller ventures with less capital requirements [Powell and Ansic, 1997] or ask for less money when seeking external

funding [Zarya, 2017], which also coincides with women being more likely to find in industries where more human capital and less financial funds are needed such as services rather than manufacturing (see the next section).

2.2.2 Self-selection to domains in entrepreneurship

Half of the gender wage gap between 1980 and 2010 in the US can be explained by women and men working in different occupations and industries [Blau and Kahn, 2017], as occupations where the share of men working is higher tend to offer higher wages on all education levels [Shaw et al., 2016, Washington Center of Equitable Growth, 2017]. The sectoral selfselection can be observed in the case of the women entrepreneurs as well. In Europe, the share of women entrepreneurs is the highest in the sectors human health and social work activities (60%), other service activities (59%) and education (55%), while it is the lowest in Water supply, sewerage and waste management (8%), Transportation and Storage (7%) and Construction (3%) sectors [European Commission, 2014].

In addition to the industry-level disparities, Loscocco and Robinson found evidence for second-tier gender self-selection: even within industries, women were more likely to own traditionally female-typed businesses, for instance, within the health and medical services industry, the firms owned by women concentrated in the nursing subcategory, while men were the most likely to open medical offices. In general, subcategories where men's participation was higher could offer more lucrative business opportunities, and also women entering these fields could produce revenue streams that were similar to men's Loscocco and Robinson, 1991. According to Mayer, the labor market segmentation of women is an *ex ante* factor influencing women's decisions where to position their new ventures. Further, the author found that the second-tier gender segmentation is also present in the high-tech industry: women are more likely to start traditionally female-type ventures such as consulting, while men are more likely to enter male-typed subsectors such as hightech manufacturing [Mayer, 2008]. As shown in the next section, venture capitalists are predominantly male, and they often have no interest in some fields where women frequently start ventures such as fashion [Zarya, 2017], so the domain choice affects women's funding opportunities through the gender composition of those who decide about investments.

2.3 Gender-based discrimination

As shown in the previous paragraphs, among other confounders, differences in risk-taking, ambitions and field choice may partially explain the gender differences of ventures obtaining and receiving funding, but there is also evidence for discrimination against women, especially in the decision process of venture capitalist and other external funders.

Orser et al. found that although there is strong evidence for little or no discrimination in the approval rates of debt financing, it can be a possible explanation for women's disparity in equity financing. Discrimination can prevent women entrepreneurs from accessing capital in two ways: the first is perceived discrimination which discourages them from seeking financing in the first place, the second type of discrimination manifests through higher turn-down rates for women [Orser et al., 2006]. The source of bias against women entrepreneurs can stem from the role incongruity between gender stereotypes and an idealized entrepreneur's characteristics, which then causes external evaluations to be biased against women [Lee and Huang, 2018]. There are examples of both experimental and data-driven evidence for gender discrimination in entrepreneurship in the literature ([Lee and Huang, 2018, Guillén et al.,]). Further, the venture capital industry itself has a women problem: at the top 100 venture capital forms, women's share was only 8% as of 2017 [Teare and Desmond, 2017]. Thus, a potential reason for women's low representation among venture capital recipients is that women entrepreneurs are not part of venture capitalist network [Zarya, 2017].

2.4 Summary

This section reviewed the most important differences in challenges of ventures led by women are men which also affect their access to external financing and their profitability. As a starting point, previous research showed that women are less likely to seek external funding and they also ask for less money. These practices can be partially explained by women's different attitude towards risk-taking. Also, women and men operate businesses in different industries, and even within industries, there are substantial differences in the type of ventures they start. The choice of field and domain is largely influenced by self-stereotyping and as a consequence, segregation. Engaging in fields with lower capital requirements and less lucrative business opportunities make ventures led by women perform worse than the ones led by men. Furthermore, ventures led by women still face discriminatory practices when they try to seek capital, which is partially due to the manly perception of entrepreneurship as a profession and the almost all-male composition of venture capitalists and business angels.

Consequently, if one wants to raise the profitability of women-led ventures through better access to financing, one has to solve the underlying issues: women's stance towards risk-taking, their self-selection to lower-paying industries and the discrimination against them from the side of investors. As is many cases, it would be hard to run controlled labor experiments on men and women entrepreneurs in an artificial environment where discriminatory factors in external funding decisions are nonexistent and where they were free to choose what venture to start. In the lack of an appropriate experimental setup or sufficient data collection possibilities, social scientist often turn to platforms where the subjects' interactions and decisions are well-documented and easily accessible: online crowdfunding marketplaces are such environments to analyze. In the next section, I outline the key characteristics of such platforms.

Chapter 3

Background

3.1 About crowdfunding

What is crowdfunding? Crowdfunding is a relatively new form of obtaining the necessary capital from fractured small-amount investors instead of a few sophisticated investors through internet platforms to make various projects happen. Mollick provides the following narrow definition about this form of microfinancing: "Crowdfunding refers to the efforts by entrepreneurial individuals and groups – cultural, social, and for-profit – to fund their ventures by drawing on relatively small contributions from a relatively large number of individuals using the internet, without standard financial intermediaries." [Mollick, 2014], or in Belleflame's interpretation, "an open call to provide financial resources" [Belleflamme et al., 2015]. ArtistShare, a platform that enables creative artists to seek funding for their artistic works is considered to be the first crowfunding marketplace on the internet¹. It was followed by Indiegogo in 2008², Kickstarter in 2009³ and GoFundMe in 2010⁴, just to mention the few biggest. The size of the industry was estimated to be \$16.2 billion in 2014, and it has been rapidly growing since then [Statista, 2018]. On Kickstarter alone, contributors pledged \$3.7 billion during its ten-years operation.

Enabled by the emergence of well-functioning online platforms, crowdfunding became a viable form of starting a project or a venture using external funds. The size, scope and

¹http://www.artistshare.com/about

²https://www.indiegogo.com/about/our-story

³https://www.kickstarter.com/about?ref=global-footer

⁴https://www.gofundme.com/about-us

nature of the crowdfunding projects can highly vary, and specialized platforms emerged for different kinds of projects. A possible classification based on what is given in exchange for the funds is the following: equity, reward-based, and donation-based crowdfunding platforms. In equity crowdfunding, project supporters become shareholders in the venture in exchange for their contribution. In the case of reward-based crowdfunding, contributors do not receive a stake of the company, but another form of benefit defined by the project owner. This reward can range from a simple thank you card to receiving the endproduct of the particular project. Project owners on donation-based platforms do not offer personal benefits in exchange for the support. Equity crowdfunding can be an alternative investment opportunity, and it is usually preferred by ventures with higher capital requirement. Reward-based platforms draw the attention of supporters who are interested in the endproduct or service offered, donation-based projects seek funding from backers who want to see the particular initiative materialize [Belleflamme et al., 2015]. A common feature of both schemes is the presence of community benefits: the supporters enjoy helping projects to come alive, they obtain novel products before it is sold for wider audiences, or they experience a closer involvement in the business development of the supported project [Belleflamme et al., 2014, Kuppuswamy and Bayus, 2017].

When do projects succeed? There are two typical definitions of success on crowdfunding platforms. If the platform adapts a so-called *All-or-Nothing* model, project owners only receive the funds collected in case they reach their pre-defined goal. If they fail to reach to goal within the campaign period, the supporters receive their money back from the platfrom provider. On platforms using a flexible funding policy (*Keep-It-All* model), project owners can always keep the collected amount of money regardless of its magnitude. Most of the well-known crowdfunding websites such as Kickstarter and Indiegogo use the former model, as making the platform users set a certain goal requires planning and commitment from their side: if the goal is unrealistic, the project will not succeed and no funds will be raised. Also, previous research confirmed that setting higher goals lowers the probability of reaching it. [Belleflamme et al., 2015, Mollick, 2014]

Crowdfunding and venture capital Crowdfunding can serve as an alternative for angel investors, banks or venture capital funds early-stage ventures, where the entry barriers are lower, but the success of the capital accumulation is equally unpredictable [Kuppuswamy and Bayus, 2017, Vulkan et al., 2016]. Instead of relying on expert opinions, the quality of the new ventures on the crowdfunding platform is assessed by a larger group of

(presumably) amateurs. Mollick found that the unsophisticated funders on a crowdfunding platform rely on similar signals of quality to venture capitalist when choosing which project to support, such as a history of successful projects by the entrepreneur, demonstratable third-party endorsements and preparedness (e. g. working prototypes). However, there are important differences between the choices of crowdfunders and venture capitalists: crowdfunding tends to be less biased in terms of gender and geographical location of the investors. In 2017, only 2.2% of all venture capital money went to start-ups with female founders [Olsen, 2018], and the supported ventures are highly concentrated to a few geographical areas. On the contrary, in the absence of face-to-face interactions, crowdfunders are less likely to favor geographical locations or discriminate against women and minorities. [Mollick, 2013]

3.2 Gender differences in the crowdfunding literature

As shown in the previous section, crowdfunding platforms have the potential to democratize the access to capital in the absence of the bias of venture capitalists and other traditional sources of equity. In recent years, numerous articles examined if crowdfunding platforms were truly able to mitigate real-life inequalities in participation and success of women and underrepresented groups in entrepreneurship and access to funding.

Gender, participation and domain selection Marom et al. showed that on Kickstarter, about one-third of all project owners and 44% of the investors are women, as opposed to the 2% of female start-up funders, 4% of female CEO-s among venture capital recipients and 9% women venture capitalists [Marom et al., 2013, Olsen, 2018]. Another reward-based crowdfunding platform called Indiegogo exhibited similar participation rates of women among project owners [Gorbatai and Nelson, 2015]. On a leading equity crowdfunding platform in the UK, women's participation rate is substantially lower: 14% of the entrepreneurs and 15% of the investors are female [Horvat and Papamarkou, 2017]. Although women's participation on these platforms is higher than in the venture capital industry, they are not distributed evenly across different project categories: on Kickstarter, women project owners frequently launch projects in the Dance, Fashion and Food categories, while men's share in the Games and Technology categories is around 85-90% [Marom et al., 2013]. On Indiegogo, 35% of the small business campaigns are led by a woman, while only 10% of the technology projects have female owner [Gorbatai and Nelson, 2015]. Women entrepreneurs prefer to launch campaigns in the Food&Drink, Health and Cosumer Products categories, and there are more men than women in the Finance, E-Commerce&Markets and IT&Telecom categories on the previously mentioned equity crowdfunding website [Horvat and Papamarkou, 2017].

Gender, ambitions and success On Kickstarter, Marom et al. found that women are expected to set lower goals, but they are more likely to reach them. In the same time, the funding rate (amount of money collected over the goal set) is more volatile for men: they are less likely to reach the goal, but when they reach it, they are expected to raise higher premium over the goal set [Marom et al., 2013]. Horvat and Papamarkou found that although women reach their goal more often than men, gender is not a significant predictor of campaign success. On the examined equity crowdfunding platform they also noted that the distributions of the asked and received funds are not different for men and women [Horvat and Papamarkou, 2017].

Other gender differences Gorbatai and Nelson observed that the language of Indiegogo campaigns is a confounder of women's success on the platform, because the usage of positive and inclusive language is a rewarded feature of campaign texts, and women use these text features more extensively. Also, women are expected to use [Gorbatai and Nelson, 2015]. Also, women are less likely to launch a second campaign after an initial project that failed or succeeded by sizable margins, and the fact that women do not try again after a failure affects the number of female-led projects adversely [Kuppuswamy and Mollick, 2016]. Lin and Pursiainen found that on Kickstarter, men tend to set higher goals than women which marks their overconfidence, and they are also more likely to fail. However, in a successive campaigns, men's goals converge to women's, which means that men's relative overconfidence diminishes for serial entrepreneurs. [Lin and Pursiainen, 2017]

Chapter 4

Data and Features

In this section I describe my dataset used for analysing gender differences in entrepreneurial ambitions and success in fundraising on an online crowdfunding platform, Kickstarter. In the first section, I describe how Kickstarter works, next, I present my data source and collection method, the gender recognition process, and finally, I show how I restricted my data to the final sample.

4.1 About Kickstarter

Kickstarter is one of the most popular crowdfunding platforms. In the ten years since its foundation, 15 million people helped to make 144 153 projects happen by supporting them with \$3.7 billion USD altogether ¹. The platform uses a reward-based funding structure with an All-Or-Nothing policy. The goal of the platform is to make creative projects alive with the help of a social-network like community of backers ². According to the founder of the website, Perry Chen, the platform is somewhere in the "middle ground between patronage and commerce", which means that supporters do not and should not expect perfect end-products in exchange for their contribution. However, a share of Kickstarter projects continue as highly successful real-life ventures such as a smart watch called Pebble ³, which

¹https://www.kickstarter.com/help/stats

²Frequently used expressions: Kickstarter users and creators developed their own dictionary for the most frequent expressions connected to projects on the platform. The dollars pledged means the amount of dollars collected. Project owners are also called founders, and their supporters are frequently denoted as funders or backers.

 $^{^{3}} https://www.kickstarter.com/projects/getpebble/pebble-time-awesome-smartwatch-no-compromises?ref=discovery$

had an initial goal of \$500 000, but managed to collect more than \$20 million from almost 80 000 supporters, or a tabletop game called Exploding Kittens, which is sold now on Amazon⁴. Some projects are made by professional teams and are easy to imagine in commercial production, other ones are rather charity-typed projects with no real product in return for supports. To filter out less product-like projects, I will focus only on certain categories in my research (see Chapter 4).

Kickstarter lets its users to define a project goal and different reward levels, and to categorize their projects in one of the 15 main categories. Most projects belong to the Music, Film & Video and other artistic categories, but there are also less art-centered categories such as Technology or Design. ⁵ On the project page, founders publish textual information, video and pictures about the project, connect it to social media accounts and external webpages, they can share updates, respond to comments and questions to keep in touch with the funding audience. Importantly, most often the only information backers have about the project they support is displayed on the project site, which provides a excellent opportunity for social science researchers to extract the factors that influence funders when they make their funding decision.

Although I do not argue that every aspect of Kickstarter campaigns mirrors an offline entrepreneurial process and the way entrepreneurs seek funding, the way project owners decide about starting a new project dedicating time and money to it, taking risks by using supporters' funds to make their own idea happen does resemble early-stage entrepreneurial ventures in a substaintial part of Kickstarter projects.

4.2 Data collection

The source of the data is a custom collection of 134 844 projects webpages from Kickstarter, which I received from Jonathan Lansey who collected it for his own research purposes. I parsed the webpages to extract numeric and textual information from the html code. The sample of projects is non-comprehensive, but it is sizable compared to previous studies on the same website [Marom et al., 2013, Mollick, 2014]. All projects in the sample were launched between April 2009 and August 2015. First, I restricted my sample to the projects denominated in USD to keep them comparable, and I also removed projects which were

⁴https://www.kickstarter.com/projects/elanlee/exploding-kittens

⁵https://www.kickstarter.com/help/stats

marked as cancelled or suspended (I only kept failed and successful projects), which left me with 94 896 projects (I report descriptive statistics of these projects in Appendix A.1 to make it comparable to my final sample described in Section 4.4).

4.3 Gender recognition

As a starting point of any analysis of gender differences, it is essential to have a solid method for identifying the gender of the observed users. Kickstarter as a platform does not require users to disclose their genders on their profile page, however, there are widely used methods in the computational social science literature to infer gender from online profiles [Karimi et al., 2016]. For instance, Jadidi et al. researched systemic disparities in the carrier paths of men and women computer scientists using automatized gender recognition methods [Jadidi et al., 2017], Vasilescu et al. uses a similar technique to observe and compare the activity and success of men and women programmers on an platform for developers [Vasilescu et al., 2012]. On online platforms where the username usually matches the real name of the user, as in the case of crowdfunding websites, researchers frequently apply dictionary-based gender inference methods, which assign a gender to each username based on the likelihood of the surname belonging to a man or a woman in the US Social Security database [Marom et al., 2013, Gorbatai and Nelson, 2015, Horvat and Papamarkou, 2017].

Similarly to the cited articles about gender differences in crowdfunding, I also applied a dictionary-based tool called *Gender Guesser* [Michael and Gecko, 2007]. The software uses a custom dictionary of 40 000 first names and the corresponding gender, which covers most of the European, American, Chinese, Japanese, Indian etc. surnames validated by native speakers of each language. The dictionary also includes a frequency for each name and gender on a scale between 1 (rare) and 13 (extremely common). Using this dictionary, a Python program named *Gender Guesser* assigns one of the following labels to each of the first names supplied to it based on the probability of that name belonging to a male or a female: *unknown* (name not found), *androgynous*, *male*, *female*, *mostly_male*, or *mostly_female* [Michael, 2016]. (Androgynous means that the name is given to a male or female with the same probability, unknown means that the name was not found in the dictionary).

Out of the 94 898 project owners, Gender Guesser could successfully identify 42 289 men and 16 308 women, which means that the share of women among the safely identified users

gender	number of occurrences
male	42289
unknown	27367
female	16308
$mostly_male$	4204
$mostly_female$	3855
androgynous	875

Table 4.1: Result of the gender recognition method Gender Guesser

The program could safely identify 42289 male and 16308 female users.

is 28% (Table 4.1). I discarded the projects with *unknown*, *mostly_male*, *mostly_female* or *androgynous* project owners from any further analysis.

Accuracy of the automatized gender inference To validate the gender recognition method, I drew a random sample of 100 projects from those who were identified as male or female, and manually checked the gender of the project owners based on their Kickstarter profile which includes a picture and a short biography, and also visited the listed social media profiles if the gender was ambiguous after reading throught the informations present on Kickstarter. The GenderGuesser correctly identified 24 women and 72 men in the sample, while I was not able to identify the gender of 4 users based on their profile (they were all classified as male by GenderGuesser). Based on this random sample, GenderGuesser identifies women with 100% and men with 94.7% accuracy.

"Unknown" users I also check the users who were categorized as "unknown". I followed the same procedure described in the previous paragraph to check the gender of 100 randomly selected users. In 24% of the cases, the user labelled as "unknown" was a women, in 46%, the use was a men, and in 30% of the cases, the gender of the user remained unknown even after reviewing all the available information on Kickstarter and on the connected social media profiles. In some of these cases, the project owners created an account with the name of the project team or company, in other cases, they used abbreviations and gender-neutral names as a username. The share of women among the unknown users is not substantially different from their share among the identified ones, which means that to the best of my knowledge, I could not detect any bias in the algorithm or either gender's username choice. Thus, I accept the results of my gender recognition method.

4.4 Focusing on product development projects

After discarding users with unknown or ambigous gender, my sample consists of 58 597 projects with a gender label. Further, I take a subset of my projects to analyze which belong to the categories *Technology*, *Design* and *Crafts*. I chose to restrict my analysis to these projects because these are not artistic projects like the ones in the categories *Art*, *Publishing* or *Music*, but rather yield a product-like output which can be bought by the backers. I argue that finding gender differences and evaluating the results of an artistic process such as composing, novel writing or poetry goes beyond the limitations of this thesis and has little to with differences in entrepreneurship. On the contrary, the categories *Technology*, *Design* and *Crafts* mostly contain end-products which could be produced in the framework of a venture as well, thus, gender differences in these categories are expected to be easier to compare with offline entrepreneurial activities. In the followings I will call this subgroup of users the *product development subset*. In addition, I also dropped all the projects with goals less or equal to \$10, because these projects that they can hardly be seen as entrepreneurial endeavours. Eventually, I ended up with a sample of 15 057 projects spreading to the five afordementioned categories.

As shown in Appendix A.2 and A.1, projects in the product development subsample have higher averages goal but similar average dollar pledged. Product development projects are less volatile in terms of funding ratios (goal over pledged), also, the descriptives of these projects show less skewness than the overall sample of 94 896 Kickstarter projects. Interestingly, the probability of succeeding (reaching the goal set) is lower in the product development subgroup. To sum up, the most important features of the product development projects range over narrower intervals and they are less skewed than all the projects on Kickstarter.

Chapter 5

Gender differences in descriptive statistics

The investment structure of Kickstarter resembles those of the venture capital industry the most among offline investment opportunities, which is know for women being underrepresented among the investment recipients (see Section 3.1 and 2). In the following paragraphs, I introduce and compare the representation and success of women in fundraising on Kickstarter to offline statistics of women entrepreneurs.

Differences in all dollars raised In my sample of 15 059 Kickstarter projects, project owners raised USD 143 million altogether. 16% of this amount was raised by women project owners (USD 22 million), meaning that projects led by women raised proportionally more funds than female CEO-s in the venture capital industry (2.2%, see Subsection 2.1). However, on Kickstarter, men still raised 536% of the amount raised by women.

	women	men	% difference	% of difference in sum of money pledged
sum of money pledged (000 USD)	22 470	120 474	436	100
population (count)	3 292	$11 \ 767$	257	76
average goal (000 USD)	49.6	51.8	4	2
average funding ratio (%)	14	20	40	22

Table 5.1: Dissecting the gender gap in all dollars raised on Kickstarter

Average funding ratio means the average money pledged over average goal.

5.1 Decomposing the difference in all dollars raised

The 436% difference in money pledged (in women's percentages) can be decomposed to three factors: the participation rate of women on the platform, the average goal they set and the ratio of average dollars pledged over the average goal (funding ratio). Table 5.1 shows this decomposition¹.

5.1.1 Difference in participation rates

76% of the total gender difference in dollars raised is merely due to the difference in the number of women who seek financing on Kickstarter: my sample contains only 3 292 women (22%) who could be safely identified as opposed to the 11 767 men (78%). The share of women in the Kickstarter sample is higher than among the venture capital fund recipients but lower than the share of all companies with a women majority owner. To relate this to numbers from the offline world, according to Pitchbook, in 2017, 4.4% of all venture capital transactions targeted companies with a female primary leader in the US [Olsen, 2018]. Hence, my findings suggest that proportionally more women participate on Kickstarter than in the venture capital industry.

5.1.2 Differences in ambitions: goal set

For the observed product development projects, the differences in average goal set are quite small, and this difference in averages accounts for only 3% of the gap in all dollars raised. However, a closer look at the distribution of the project goals show that women and men slightly differ in the size of target they set.

As described in Section 4.1, Kickstarter operates with an All-Or-Nothing policy, meaning that the founder only gets the amount pledged if (s)he reaches the goal set. Thus, setting higher goals involves risk-taking: the project owner might lose all the money collected if the target is too high. The related literature provides evidence for women being

¹Tables A.3 and A.4 in Appendix show that for the users with goals from the bottom 95% of the distribution, the breakdown of the differences is similar to that of all the users, but the share of difference explained is distributed more evenly between the difference in average goal set and average funding ratio (14 and 15%, respectively). For the top 5% of users in terms of goal set, men outnumber women almost 7 times, and this difference accounts for the majority of the difference in sum of money pledged. In this upper subset of users, the few women present tend to set higher average goals and reach them with higher probability. I conclude that although statistics of the top 5% of users diminish the gap in average goal set which is present for the bottom 95% of users, in general, they do not distort the descriptives.

000 USD	women's goal	men's goal	women's pledged	men's pledged
# of observations	3 292	11 767	3 292	11 767
min	0.015	0.011	0	0
median	5	8	0.62	0.74
mean	49.6	51.8	7	10
sd	1 751	1 200	24	36
max	100 000	100 000	447	1 092
skewness	57	66	10	13

Table 5.2: Descriptive statistics for the variables goal and pledged

Goal means the target of the fundraising in dollars, pledged is the amount of money collected until the end of the campaign period.



Figure 5.1: Density of goal set and dollars pledged per gender

Figure 5.2: Density of premium (% funded over goal) for successful projects

more risk-averse in financial decisions than men (see Subsection 2.2.1), which seems to be the case on Kickstarter, too. Figure 5.1 shows the distribution of goals and dollars pledged by gender. The density plot of men's goals is similar to women's, but it is slightly shifted to the right. The average and the median men sets higher goals than the average and median women. Table 5.2 shows detailed descriptive statistics. The two-sample Kolmogorov-Smirnov test proves that the distribution of men's and women's goals are different (D = 0.099, p-value < 2.2×10^{16}). Not only is men's average goal higher, but they set higher target values than women along the whole distribution, meaning that men are slightly more ambitious with their project goals than women.

5.1.3 Differences in success: funding rates

A considerable part (22%) of the gender gap in all dollars pledged is due to the difference in the funding rates of men and women: although they set the same goals on average, women collect a smaller share of that goal. Women pledge USD 6 826, while men pledge USD 10 238, 50% more on average (Table 5.2). Similarly to the goal statistics, the distribution of men's dollars pledged resembles that of women, but it is shifted to the right (Figure 5.1). According to the two-sample Kolmogorov-Smirnov test, the distributions of men's dollar pledged are different (D = 0.44, p-value $< 2.2 \times 10^{16}$).

Probability of success Interestingly, in my sample, there is no considerable difference between the probability of succeeding (reaching the goal) for women and men (they succeed with a probability of 37.8 and 37.6%, respectively). This finding contradicts the statistics about the success of women-led ventures in the offline world (see Chapter 2). Since there are no large differences in the success of women- and men-led ventures, the source of the funding gap is to be found in the entrepreneurs' funding rates above their goals.

Table 5.3: Funding ratios for successful men and women (descriptive statistics)

	% funding ratio	n	mean	sd	median	min	max	skew	se
•	women	1246	2.35	5.08	1.19	1	85.21	9.88	0.14
	men	4426	4.58	27.79	1.57	1	1340.90	38.03	0.42

Funding ratio means the dollars pledged over the goal set. The probability of succeeding is essentially the same for women and men: 37.8% and 37.6%, respectively.

Funding rates Table 5.3 shows that a project that reached its goal is expected to gain 2.35 times its funding goal if the project owner is a women, and 4.58, almost twice as much of its funding goal if the project owner is a men. Men's funding rate distribution is more skewed and has higher variance then women's, but the median funding rate is also higher for men than women. The two-sample Kolmogorov-Smirnov test shows that the distribution of men's and women's funding rates are different (p-value $< 4.9 \times 10^{13}$). These statistics mean that for the same goals, men are expected to raise more money from their supporters.

Funding rates and participation per decile bins A substantial part of the gender difference in dollars pledged is due to the gap in the funding rates of men and women.

However, due to the skewness in the distribution of goal set and dollars pledged, some extremely successful men might drive up the averages in favor of men. To rule out this possibility, I divided my entries to below and above-median projects based on the goal they set and observed the funding rates of women and men in those groups. Figure 5.3 shows the funding rates (distribution of the pledged per goal ratio) for below- and above-median projects. The size of the gap in funding rates varies depending on the size of the goal set. For below-median projects, men have substantially higher funding rates. In the case of above-median projects women have slightly higher average funding rates, but their share among all project owners is lower in this group. For high goals, women tend to be more successful, but they are also underrepresented. For lower goals, men have systematically higher funding rates.

Premium over goal Figure 5.2 shows the distribution of goal premium (percentage funded over goal) for successful projects: men's density curve is shifted to the right of women's. Although ambitions and success are more equal for men and women on Kick-starter than in the case of offline enterprises, men still seem to be rewarded more: they are significantly more likely to overshoot their goal. To shed light on the differences in projects that can show this, I am going to analyze category-level differences across project profitability.

5.2 Gender differences in categories

As described in Section 2.2.2, women entrepreneurs are more likely to start in service or retail type of ventures, while men are expected to be highly represented in fields such as technology and engineering. These differences are not only constrained on the industry level: even within industries, women choose subdomains which are fundamentally different from men's, and less profitable at the same time. In the next sections, I am going to compare the distribution of women to men in Kickstarter categories.

Kickstarter lets the founders assign the projects they launch into 15 categories. The participation rates of women varies across categories. Among all categories, the share of women is highest in Dance, Fashion and Food, while it is lowest in Games, Comics and Technology [Marom et al., 2013]. As discussed previously 4.4, I chose to focus on product type categories, namely Crafts, Fashion, Design, Technology and Games. In the categories Crafts and Fashion, women and men are about equally well-represented, however, in the



Figure 5.3: Funding rates and women's participation per goal decile

categories Design, Technology and Games, women take up no more than 20% of all the project owners. Figure 5.4 shows that in categories with high share of women, projects are expected to pledge less than in the Design or Games categories.

The probability of succeeding (the number of projects that reach their goal) is almost always the same for men and women in each category, however, there are considerable differences in the size of the pledged over goal ratio between men and women (see Table 5.4). In each category, men are expected to have higher pledge to goal ratios than women. The gap is the smallest in the Fashion category (successful women and men users pledge 1.68 and 2.37 times their goal, respectively), and the largest in the Design category (pledge to goal ratios of 1.13 and 2.69 for women and men).



Figure 5.4: Average dollars pledged and share of men and women per categories

category	gender	%women in category	% succeeded	mean goal 000 USD	mean pledged 000 USD	mean funding rate for the successful
Crafts	woman	49	0.30	4.92	1.33	2.70
Crafts	man	49	0.27	27.95	1.42	3.88
Design	woman	19	0.43	20.59	10.66	2.47
Design	man	19	0.44	23.17	14.96	5.94
Fashion	woman	41	0.38	100.69	5.48	1.68
Fashion	man	41	0.31	14.23	7.81	2.37
Games	woman	10	0.55	45.03	14.17	3.55
Games	man	10	0.51	57.42	10.68	4.50
Technology	woman	11	0.28	40.30	7.99	2.02
Technology	man	11	0.24	96.25	8.76	4.07

Table 5.4: Descriptive statistics per gender and category

Chapter 6

Gender differences in topics

We saw that there are considerable differences in gender representation across categories. However, it is important to note that Kickstarter's categories such as Design or Technology are not telling about the nature of the product itself: for instance, the Design category can contain projects that are aimed to create a dog leash, a musical instrument, or a stargazer telescope as well. I argue that this broad categorization absorbs differences in the topic of the projects across genders. Although projects are *categorized*, its *topic* remains unknown.

6.1 Topic models

To overcome this difficulty, I turn to an unsupervised learning method called topic modelling. This text mining method labels individual documents that are parts of a larger collection called corpus with a set of "substantively meaningful coding categories called topics" [Mohr and Bogdanov, 2013]. Topic models are used to process and synthetize large amounts of textual data by clustering documents of the collection based on the words they contain. The method treats documents bag-of-words, meaning that it captures frequent co-occurrences of words and assumes that they represent some unobserved topic regardless of deeper complexities of the language. This approach compares the frequency of tokens on the document level to their overall frequency in the corpus. The algorithm takes the number of final clusters k - topics - as an input parameter, and matches keywords to the topics from the entire corpus based on their relative frequency in the individual documents. Finally, it associates each document with a k-dimensional topic vector that contains the percentage topic representation in the text for each of k topics. I apply the most widely used probabilistic topic model called *Latent Dirichlet allocation* algorithm [Blei et al.,
2003].

I applied topic modelling for the text descriptions of the projects per category, so in each topic model, only the descriptions from a single category such as Design or Technology were present. Individual descriptions were the documents, while the corpus consisted of all descriptions in the given category. Although global topic models for all the projects in the sample could have been equally interesting (where all descriptions would have constituted the corpus, regardless of category), my goal was to refine the initial Kickstarter categorization, so I kept my models within the categories. The next sections describe the results of these models per category. The number of topics I found in each of the categories differ: I altered the number of topics k until the resulting k topics were homogeneous both in the keywords used and in the descriptions assigned to it.

LDA does not assign a single topic to each project description, but a k-dimensional topic vector with percentage contributions of each topic to the particular description. Qualitative topic labels can be assigned based on the keywords of the topic: for instance, the keywords *app*, *information*, *data*, *web*, *mobile* within the Technology category are labelled as *application*. To create a single label to each description, I assigned the label of the topic to the project when it was represented in more than equal to 50% of the description. In cases where there was not any topic with at least 50% representation in the project description, I labelled the project as *mixed*.

After labelling the project descriptions this way, I read 20 randomly chosen projects descriptions in each label bin, and reviewed the label name and consistency. In cases where the projects assigned to the label did not constitute a homogeneous bin, I re-ran the topic model in the category to get more meaningful labels. Appendix A.5 to A.9 contains the most frequent words connected to each topic, and Appendix A.10 to A.12 shows the short descriptions of some sample projects assigned to six topics from three different categories. The final topic labels exhibit consistency among the project descriptions within the topic and they denote distinct topics within the category. Next, I explored the gender participation rates within the newly defined topics.

6.1.1 Topics of the Technology category

With the help of LDA I identified three distinct topics within the Technology category. (Table A.5 in Appendix shows the most frequent keywords per topic.) The first topic contains words associated with devices, the second with software with a focus on communities, the third with applications. I label these categories accordingly. In the Technology category, there were 88 projects that received the "mixed" label, only 2% of all the projects in the category (See Table 6.1). The "application" topic contains the most projects within the category, with relatively low success probability (15%). The second largest topic, the "device", exhibits the highest success and funding rates (I only report the funding rates for the successful projects), with 34% of the projects succeeding and with an average funding rate of 651%. The "community software" category, where the representation of women is the highest (apart of the "mixed" category), has the lowest expected funding rate for successful projects (308%).

label	count	% women	mean goal in topic [*]	prob. of success in topic	success diff.**	mean funding rate in topic ^{***}	funding rate diff.**
mixed	17	19	46 827	35	-18	354	-35
community software	153	18	183 315	27	-17	307	166
application	194	12	52 363	16	-26	326	128
device	64	5	77 679	35	-12	455	42

Table 6.1: Gender representation in topics of the Technology category

Mean funding rate in topic means the average of individual goal per pledged ratios within the topic, funding rate difference means the difference between men's and women's average funding rate in women's percentages. (*000 USD, ** women = 100%, everything except count is expressed in % terms, *** for successful projects)

6.1.2 Topics of the Design category

The Design category contains a wide range of Kickstarter projects from building community spaces to manufacturing leather goods. Thus, the LDA model could identify 12 distinct topics. The list of topics can be found in Appendix, Table A.6. Within the Design category, the size of the topic groups is relatively evenly distributed. The representation of women is the highest in the topics pets (31%), prints (29%) and lifestyle (29%), while the least women can be found in the topics accessories for devices (8%), wallets/organizers (8%), handcrafted products and photography/devices (10%). The most successful topics in terms of reaching the goal are handcrafted products, wallets/organizers, accessories for devices and playing cards, while the most overfunded topics are handcrafted products, prints and pets.

label	count	% women	mean goal in topic*	prob. of success in topic	success diff.**	mean funding rate in topic***	funding rate diff.**
pets	81	31	22 629	27	11	355	32
prints	117	29	13 851	47	-16	1030	694
lifestyle	95	29	22 951	30	-11	262	124
building/ community space	50	25	37 140	41	-16	222	102
playing cards	52	23	22 094	51	13	375	47
lighting/furniture	33	18	22 238	42	48	377	-61
mixed	90	16	20 063	43	-7	330	120
biking	38	16	30 255	37	-32	303	36
liquid containers	33	12	20 587	47	13	509	-35
handcrafted products	18	10	12 314	65	-4	1592	888
photography/devices	25	10	45 151	38	-14	269	41
wallets/organizers	26	8	14 828	59	-4	522	134
accessories for devices	21	8	21 505	52	-10	363	3

Table 6.2: Gender representation in topics of the Design category

Mean funding rate in topic means the average of individual goal per pledged ratios within the topic, funding rate difference means the difference between men's and women's average funding rate in women's percentages. (*000 USD, ** women = 100%, everything except count is expressed in % terms, *** for successful projects)

6.1.3 Topics of the Crafts category

LDA identified six topics within the Crafts category. Five topic are clearly separable: cosmetics, woodwork, community space, candles/gifts and jewelry, and there is a topic with a wide range of products in it: the similarity between these projects is that in the description, the project owner states that (s)he wants to level up his or her hobby. Among the observed categories, the share of women is the highest in Crafts, and women are wellrepresented in most of the topics within the category as well. In the topics cosmetics, candles/gifts, miscellaneous hobby and mixed, women make up the majority of project owners. The probability of succeeding is the highest in the woodwork and miscellaneous hobby topics, while jewelry, woodwork and miscellaneous hobby are the most overfunded topics.

label	count	% women	mean goal in topic*	prob. of success in topic	success diff.**	mean funding rate in topic***	funding rate diff.**
cosmetics	121	62	19 181	21	-12	344	-35
candles/gifts	202	61	7 328	21	14	235	75
mixed	30	56	4 846	33	-20	156	49
miscellaneous hobby	39	53	7 193	46	-12	670	1
jewelry	91	46	6 893	32	-7	294	61
community space	215	46	9 445	27	-46	184	5
woodworking	28	18	76 941	48	-13	561	30

Table 6.3: Gender representation in topics of the Crafts category

Mean funding rate in topic means the average of individual goal per pledged ratios within the topic, funding rate difference means the difference between men's and women's average funding rate in women's percentages. (*000 USD, ** women = 100%, everything except count is expressed in % terms, *** for successful projects)

6.1.4 Topics of the Games category

The Games category is the one where the fewest women are present among all observed categories, the share of women is only 10%. The topic modelling phase could differentiate between three distinct topics: computer games, board games, roleplay/other. The third one is a bigger and less pure topic containing numerous projects that did not fit in any of the former two topics. The probability of succeeding is similar in the computer games, board games and mixed topics (66-67%), but it is substantially lower in the roleplay/other topic. The funding rate of successful projects is by far the highest in the computer games and board games subtopics of the games category from all observed projects with 2986%and 2681%, respectively. The reason behind this phenomenon can be that in this category, funders often back products because they are enthusiastic about the endproduct - the game - itself, and even though the project already reached its goal, they continue buying the rewards to buy the game. Also, funders can think that the probability of a particular game reaching the phase when it can be enter large-scale production is low, so they engage in buying the game through Kickstarter, where it is more or less assured that they actually receive it. The strikingly high funding rates raise the question why women appear in such a low numbers in this seemingly profitable category.

label	count	% women	mean goal in topic [*]	prob. of success in topic	success diff.**	mean funding rate in topic***	funding rate diff.**
computer games	99	10	9552	69	3	572	33
roleplay	162	10	106 006	32	-22	275	-15
boardgames	60	8	$13 \ 465$	69	-8	458	59
mixed	7	6	10 936	68	62	309	68

Table 6.4: Gender representation in topics of the Games category

Mean funding rate in topic means the average of individual goal per pledged ratios within the topic, funding rate difference means the difference between men's and women's average funding rate in women's percentages. (*000 USD, ** women = 100%, everything except count is expressed in % terms, *** for successful projects)

6.1.5 Topics of the Fashion category

The representation of women in the Fashion category is the second highest among the observed categories, (41%). The LDA algorithm identified three topics connected to different pieces of clothing: leather goods, jewelry, women's apparel, the fourth category contains words about brand building or website design (presumably for existing products). The latter is the topic with the highest number of projects. Women are well-represented in all categories, but the highest number of women can be found in the women's apparel category. The probability of succeeding is similar in the mixed, leather goods, jewelry and women's apparel topics, and substantially lower in the branding topic. The average funding rates are substantially higher in the leather goods and jewelry topics than in the remainder (295 and 266%, respectively).

6.2 Topics, interest, gender representation

The previous paragraphs showed that there are substantial gender differences in the choice of topics even within categories. To highlight a few of these: within the Design category, I observed the highest share of women in the pets and lifestyle categories, while men preferred topics such as wallets/organizers and accessories for devices such as mobile phones or tablets. In the Crafts category, women were most likely to be found in the cosmetics and candles/gift topic, while men had the highest share in the woodwork topic. In the Technology category, women were most likely to build software for community building and education, and men engaged more in developing new devices or applications. In Appendix

label	count	% women	mean goal in topic*	prob. of success in topic	success diff.**	mean funding rate in topic***	funding rate diff.**
women's apparel	362	51	152 992	39	-8	169	65
leather goods	177	43	21 208	45	2	296	121
mixed	51	43	12 786	36	3	163	49
jewelry	187	38	13 319	40	-2	266	-17
branding	354	34	11 902	24	-43	136	6

Table 6.5: Gender representation in topics of the Fashion category

Mean funding rate in topic means the average of individual goal per pledged ratios within the topic, funding rate difference means the difference between men's and women's average funding rate in women's percentages. (*000 USD, ** women = 100%, everything except count is expressed in % terms, *** for successful projects)



Figure 6.1: Share of women in topic and funding rates

A.10 to A.12, I show the short description of 8 sample projects chosen randomly from each of the aforementioned topics.

These findings are in-line with the offline statistics showed in Section 2.2. On Kickstarter, women are likely to start ventures in stereotypically female domains than in masculine ones. Also, *manly* projects such as computer games, handcrafted products or applications are the most profitable topics among all observed. Figure 6.1 shows that in topics in which women's share is higher, the mean funding rate of the topic is expected to be lower.



Figure 6.2: Share of women and difference between men's and women's funding rates (in women's percentages)

Thus, as in the case of offline enterprises, women are more likely to start businesses in fields which are less profitable. Also, the figure proves the viability of topic distinction: in large categories, such as Design, funding rates and women's share are considerably different across topics. These differences would remain hidden if I only controlled for categories in my modelling phase (see Section 7).

Research on offline ventures showed that women who enter *masculine* fields tend to start enterprises that are similarly lucrative to men's. However, figure 6.2 shows that this is not the case on Kickstarter: there is no straightforward pattern in the relationship between the share of women in the topic and the difference in funding rates (expressed in women's percentages).

Evidence from topic modelling shows that fields chosen by women are systematically less profitable. I will include the controls for topics to my final models to control for domain choice which is expected to explain a large part of the gap in funding rates across projects of men and women entrepreneurs.

6.3 Project novelty

I used a simple method to assign a single topic to each project: I chose the topic with the highest percentage share in the project description given by the LDA algorithm if the representation of the particular topic in the whole project description exceeded 50%, otherwise, it was labelled as *mixed*. To control for projects where the description mentions more than one topic with a high frequency, I compute a measure called *topic entropy* as seen at [Horvát et al., 2018]. Low topic entropy means that the description complies well with the topic categorization created by the LDA algorithm, high topic entropy implies that the project represents a mixture of multiple LDA topics. Given the k-dimensional topic vector of the project where the j-th entry in the vector represents the share of the j-th topic in the description, the topic entropy is calculated as follows:

$$topic \ entropy_i = \frac{-\sum_{j=1}^{k} T_i(j) \log(T_i(j))}{\log(k)}$$
(6.1)

The score is normalized by the log number of topics so it ranges between 0 and 1. Horvat et al. propose that since topic entropy shows to what extent the project defies categorization based on the corpus of all descriptions, it is a measure of novelty (it represents a unique or rarely used combination of topics) [Horvát et al., 2018]. Although there are no large differences in the averages of women's and men's topic entropy scores (27.5 and 28.5%, respectively), Figure 6.3 shows that the distribution of these entropy scores is different across genders: men are more likely to have entropy scores close to 0 (descriptions which perfectly fit in one of the topics defined by the topic model), women frequently have entropy scores close to the average, and men are more likely to have above-average entropy scores as well. The violin plot suggests that there are more men on the extremes of the topic entropy distribution, meaning that men are better represented among exceptionally pure and exceptionally impure projects as well.

I include the topic entropy to my models in a binary form: I control the project having below- or above-median topic entropies. This feature measures if investors reward unique combinations of the topics found by the topic model.

6.4 Controls for project management

In addition to gender differences in topic choice, women and men can also differ in how they manage their projects. Seemengly good project management can be a feature which is



Figure 6.3: Topic entropy of women's and men's projects

rewarded by the potential investors. An et al. found that recurring backers on Kickstarter usually favor projects where the founder has proven management skills: s(he) interacts with the backers frequently, posts updates and defines fine-grained reward levels [An et al., 2014].

6.4.1 Number of reward levels

On Kickstarter, the project owners can define reward levels: backers receive different products (*rewards*) for different prices. The rewards can range from listing the name of the backer on a website, sending a thank you postcard to actually receiving the end product of the project. Frequently, project owners allow the backers to purchase different versions of the final product for different prices. Thus, providing a wide range of reward options shows professionalism.

In my sample, the number of reward options ranges between 1 and 146, with a mean of 8 options. I created four bins: 1, 1-5, 5-10 and more than 10 reward levels. There is little difference between men and women in setting the number of reward levels. Having only one reward level is uncommon: when there is only one reward option available, it is mostly something symbolic: a t-shirt or a thank you message, which means that there is no real product development involved in the project.

6.4.2 Number of team members

A large share of Kickstarter projects is not an individual endeavor: the project is a joint work of a team. However, Kickstarter only lists one single project owner, so the number of people truly involved remains unknown. As I have the descriptions of the projects available, I create a control variable for a team being involved: I mark the projects that listed the words *team, we, CEO, CTO, manager, co-founder(s), founders* as team projects. In my sample, 61.8% of women fall into this category as opposed to 70% of men, which means that men project owners are somewhat more likely to list team members. There is no hard evidence that these women who own the projects on Kickstarter are necessarily team leaders, but they do take financial responsibility for the funds collected on Kickstarter. Projects made by teams are more likely to succeed: individual projects succeed in 28% of the cases, while team projects succeed in 42% of the cases.

6.4.3 Length of the project description and risk section

The length of the project description measured in word count shows how much effort the project owner put in describing the project, also, the relative length of the risk section is an indicator of the project owner's awareness of the caveats of the implementation. The average project description's length is 161 words excluding stopwords, the average risk description length is only 35 words without stopwords. The average description length is 144 words for women project owners and 1.66 words for men. Although women have shorter descriptions, they have slightly higher relative risk section lengths: 36% as opposed to 37% for men.

Chapter 7

Results

Based on the descriptive statistics I can conclude that there are no large gender gaps in the average success rate of product-related projects, however, there are big differences in the extent of overshooting, i.e. the amount of money raised in percentages of the goal. Men are expected to *overshoot* their goal with a multiplier of 4.58 (358%), while for women, the same number is only 2.35 (135%) (see Table 5.3). In the followings I will show that success differences can be explained by other confounders, but gender is an important factor in the level of overshooting.

Also, previous research mostly involved category controls in the models, however, now I will use the newly defined topic controls that allows to compare more similar projects in my models.

7.1 Predicting success

First I predict success (reaching the goal set) by only controlling for the project owner's gender (7.1), then for gender and goal (7.2), the project's duration, the number of reward opportunities listed, its category, the year when the project was launched. I also control for team projects and the length of both the project description and the risk section $(7.3)^1$. In regression (7.4), I use topic controls instead of category controls. I estimate simple linear regressions to predict my binary outcome variable. My topic choice allows me to compare the models that explain success and the ones that explain premium, although the first

¹I will refer to the the variables duration, # of rewards, start year, lists team, log description words and log risk words as controls in the equations

group of models (7.1-7.5) have a binary LHS variable, while the second group of models (7.6-7.10) have a continuous dependent variable.

$$success_i = \beta_0 + \beta_1 gender_i + \epsilon_i \tag{7.1}$$

$$success_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \epsilon_i \tag{7.2}$$

$$success_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \beta_3 category_i + controls_i + \epsilon_i \tag{7.3}$$

$$success_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \beta_3 topic_i + controls_i + \epsilon_i \tag{7.4}$$

$$success_i = \beta_0 + \beta_1 gender_i + \beta_2 log goal_i + \beta_3 topic_i + \beta_4 topic entropy_i + controls_i + \epsilon_i$$
 (7.5)

Table 7.1 shows that in all settings, the gender coefficient is either close to zero or insignificant, which means that women or men are not trusted more only by their gender. Regression (3) shows that comparing two projects with the same duration, number of rewards, descriptions lengths within the same category, the project owned by a woman and a man is equally likely to succeed. Regression (4) shows that exchanging category controls for topic controls resulted in 2% points increase in the R-squared of the regression, meaning that the text-based topic controls reduce the uncertainty about project outcome more than category controls. In regression (5), I add the above-median topic entropy (campaign uniqueness) control the model in regression (4), but its coefficient is insignificant, hence it does not seem to play a role in campaign.

These results show that there are no fundamental gender differences in success on Kickstarter, which is contradictory to the findings of Marom et al. in all categories on Kickstarter [Marom et al., 2013], but are in line with the results of Horvát and Papamarkou in their equity crowdfunding study [Horvat and Papamarkou, 2017]. I emphasize that as opposed to a large fraction of Kickstarter projects, the projects within my subset exhibit features that can make them more similar to equity crowdfunding projects than to other Kickstarter campaigns in the sense that these products are easy to imagine on the market or as the product of a start-up seeking venture capital.

		De	ependent var	iable:	
			successful		
	(1)	(2)	(3)	(4)	(5)
gender	-0.002	0.02**	-0.01	-0.02^{***}	-0.02^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
log goal		-0.06^{***}	-0.09^{***}	-0.08^{***}	-0.08^{***}
		(0.002)	(0.002)	(0.002)	(0.002)
lists we			0.03^{***}	0.03^{***}	0.03^{***}
			(0.01)	(0.01)	(0.01)
log description length			0.12***	0.11***	0.11***
			(0.005)	(0.005)	(0.005)
log risk section length			-0.02^{***}	-0.02^{***}	-0.02^{***}
			(0.004)	(0.003)	(0.003)
topic entropy above median			× ,	× ,	-0.01
1 10					(0.01)
category control			YES		~ /
topic control				YES	YES
start year control			YES	YES	YES
reward control			YES	YES	YES
Constant	0.38^{***}	0.93***	281.17***	274.08***	274.13***
	(0.01)	(0.02)	(8.08)	(8.01)	(8.01)
Observations	15,029	15,029	15,029	15,029	15,029
\mathbb{R}^2	0.0000	0.05	0.36	0.38	0.38

Table 7.1: Predicting success of Kickstarter campaigns

Note:

*p<0.1; **p<0.05; ***p<0.01

7.2 Predicting premium

After predicting success I turn to the feature that showed a wider gender gap: the difference in premium. Premium means the amount of money pledged above the goal set by the project owner. Among the successful product-type projects, women are expected to raise 2.35 times their goal set, while the average man raises 4.58 times his initial goal. This section seeks to disentangle this premium gap across genders. For this analysis, I restrict my sample to the 5663 successful projects ². Again, I use a simple linear regression to estimate premium for projects.

Although the *success* gender gap can be explained by controls such as the topic of the project, the year when it was initiated and the number of rewards, Table 7.2 shows that this is not the case for the *premium* gap. After introducing the same controls as in Table 7.1 in regression (3), the gender coefficient stays significant and favors men. That is, comparing two projects with the same duration, starting year, number of rewards and within the same category, the one with a male project owner is expected to have 57% higher premium than the one with a female project owner (Table 7.2, regression (3)). However, when I introduce topic controls instead of category controls (regression (4)), the size of the gender funding gap decreases from 58% to 45% of the goal. Also, the explanatory power of my model grows: there is a 5% increase in R-squared compared to regression (3).

In regression (5), I add the above-median topic entropy control, and the coefficient is significant and negative in both. The sign of this coefficient means that campaigns with above-median topic entropy (high campaign text uniqueness) are expected to raise less premium compared to a similar project with below-median topic entropy. Note that the coefficient of topic entropy was insignificant in the success regressions, and also in this regression, there is no substantial increase in the R-squared of the model, suggesting that campaign novelty (at least measured this way) does not play an important role in determining the premium project owners get.

$$premium_i = \beta_0 + \beta_1 gender_i + \epsilon_i \tag{7.6}$$

$$premium_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \epsilon_i \tag{7.7}$$

 $^{^2\}mathrm{in}$ Table A.13 in Appendix, I also show the results of the same regression for projects that did not succeed

$$premium_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \beta_3 category_i + controls_i + \epsilon_i$$
(7.8)

$$premium_i = \beta_0 + \beta_1 gender_i + \beta_2 log \ goal_i + \beta_3 topic_i + controls_i + \epsilon_i$$
(7.9)

 $premium_{i} = \beta_{0} + \beta_{1}gender_{i} + \beta_{2}loggoal_{i} + \beta_{3}topic_{i} + \beta_{4}topicentropy_{i} + controls_{i} + \epsilon_{i}$ (7.10)

		De	pendent vari	able:	
			log_premiun	1	
	(1)	(2)	(3)	(4)	(5)
gender	0.86***	0.89***	0.58***	0.45***	0.44***
log goal	(0.06)	(0.06) -0.26^{***}	(0.06) -0.44^{***}	(0.06) -0.45^{***}	(0.06) -0.45^{***}
00		(0.02)	(0.02)	(0.02)	(0.02)
lists we			-0.005	0.07	0.07
1 1			(0.06)	(0.06)	(0.06)
log description length			0.53^{***}	0.48^{***}	0.48^{***}
log risk section length			(0.04) 0.03	(0.04)	(0.04) 0.01
log lisk section length			(0.02)	(0.01)	(0.01)
topic entropy above median			(0.02)	(0.0_)	-0.09^{*}
			VDO		(0.05)
category control			YES	VEC	VEC
topic control			VEC	I ES VES	YES
reward control			I ES VES	I ES VES	I ES VES
Constant	-1 44***	0 66***	306 41***	290 72***	291 86***
	(0.05)	(0.14)	(52.32)	(50.75)	(50.74)
Observations	5,663	5,663	5,663	5,663	5,663
R ²	0.03	0.08	0.17	0.23	0.23

Table 7.2: Predicting premium on Kickstarter projects

Note:

p < 0.1; p < 0.05; p < 0.01

Results show that introducing topic controls could significantly decrease the size of the unexplained gender gap in premium raised over goal. This finding reinforces my conclusions from the descriptive statistics and feature creation, namely that there is observable gender self-selection across categories and even within categories on Kickstarter, and this selfselection largely explains the disparities in success on the platform.

Chapter 8

Discussion

In this section I put my findings about women entrepreneurs on the crowdfunding website Kickstarter in the broader context of women's overall entrepreneurial ambitions and success, and compare my results to previous research about gender disparities in online communities and offline ventures.

8.1 Dissecting the gender gap in dollars raised

In my sample of 15 059 Kickstarter projects, project owners raised USD 143 million altogether. 16% of this amount was raised by women project owners (USD 22 million), the remaining 82% was raised by men entrepreneurs. These numbers indicate that projects led by women raised proportionally more funds than female CEO-s in the venture capital industry (2.2%, see Subsection 2.1).

The gender gap in dollars pledged can be decomposed to three factors: 76% of the gap in all the dollars raised is due to the difference in participation rates of women and men on the platform, differences in the average goal they set accounts for only 2% of the gap, and the remaining 22% is due to the difference in men's and women's funding rates.

Section 3.2 summarizes the participation rates of women on other crowdfunding platforms. Statistics on different online marketplaces show consistently higher gender representation than the venture capital industry. In the next sections I will focus on the performance of those entrepreneurs who decided to enter the platform, although the motivations to enter could also make an interesting research topic.

8.1.1 Gender gap in funding ratio

Men and women are equally likely to succeed, meaning that the source of the difference in funding rates is the premium entrepreneurs raise over their goal. This finding is contrary to previous observations on other samples: Horvát and Papamarkou reports a 5% success gap in favor of women on an equity crowdfunding platform [Horvat and Papamarkou, 2017], Marom et al. also found an 8% gender gap in the probability of succeeding of a sample of all projects on Kickstarter [Marom et al., 2013]. Research on crowdfunding platforms consistently show that women are at least as likely to succeed as men, contrary to the statistics of offline enterprises, where women-led firms are more likely to fail (see [Watson and Robinson, 2003, Coleman and Robb, 2009] and Subsection 2.3). This phenomenon suggests that factors causing women-led enterprises to underperform men's in terms and external capital raised such as different ambitions, attitude to risk and discrimination are not that salient on online crowdfunding platforms.

Among projects that succeed, women are expected to raise 2.35 of their funding goal, while the average men raises 4.58 of his funding goal. Marom et al. also find considerable differences in women's and men's premium on Kickstarter, but the authors argue that after controlling for categories and goal set, the difference between men and women in success as a binary variable is persistent, while in overshooting, the difference disappears [Marom et al., 2013]. These results are contradictory to my findings on the product development subset.

My findings suggests that the main difference in women's and men's performance lies in size of the premium they raise over their goal. To explain this difference, I focused on the content of the textual project descriptions, the entrepreneur's domain choice and other features of the campaigns.

8.2 Self-selection and domain

Chapter 6 discussed that women and men start ventures in different domains. Women's participation was the highest in the topics cosmetics, candles/gifts, women's apparel and jewelry, while it was the lowest in topics devices, wallets/organizers and boardgames. Figure 6.1 shows that topics where women's representation is higher tend to have lower mean funding ratios, meaning that topic preferred by women are less profitable. Also, the results of the explanatory models in Chapter 7 indicate that a substantial part of the gender gap in premium over goal set can be explained by domain choice: the 57% premium gap

shrinks to 44% after introducing topic controls instead of the broader category controls. Some domains are more likely to be overfunded, while in some fields, projects are expected to gain little extra support.

On Kickstarter, many project owners start a campaign to make their dream project come true. Therefore, it is a reasonable assumption that women and men undertake ventures on Kickstarter driven by their genuine interest in the domain, without carefully considering economic prospects. This assumption leads to the conclusion that women are more interested in domains such as fashion and cosmetics, and men are more interested in inventing new devices. Also, the domains in which women are interested are likely to be less profitable. As described in Section 2.2, women entrepreneurs are expected to choose less capital-intensive and lucrative businesses not only on online crowdfunding platforms but also when they decide to start real-life ventures. Consequently, women's domain choice also explains a part of the difference in performance of women and men-led ventures [Loscocco and Robinson, 1991, European Commission, 2017].

One could conclude that women self-select themselves to less lucrative industries and fields. However, it is important to note that these interests are likely to be affected by the education women and men receive and the professional experiences they have. Interestingly, the single largest factor behind the gender wage gap is also women's self-selection to lower paying professions: Half of the gender wage gap between 1980 and 2010 in the US can be explained by women and men working in different occupations and industries [Blau and Kahn, 2017], as occupations where the share of men working is higher tend to offer higher wages on all education levels [Shaw et al., 2016, Washington Center of Equitable Growth, 2017].

The career choices by man and women are largely determined by cultural factors, such as women's stance towards STEM occupations. Although meta-analysis shows the gender differences in mathematical competences are small and declined over time, women are still less likely to choose STEM professions (Science, Technology, Engineering, and Math) [Hyde et al., 1990]. Multiple studies show that men with the same demonstrated abilities (such as test scores) are likely to perceive higher self-competencies in mathematics, while they do not assess their verbal competences higher than women. Also, the higher a student's selfassessment in mathematics, the likely that they choose quantitative professions. [Correll, 2001, Wang et al., 2013]

However, self-selection across professions might be the result of choice beyond constraining social structures. In countries with higher levels of gender inequality, favorable economic conditions and high-quality education, the STEM graduation gap tends to be higher, despite of having no considerable differences in mathematical abilities across genders [Stoet and Geary, 2018]. In the absence of economic pressure to choose STEM professions, women often opt for less mathematics-intensive occupations.

The policy implications of these findings are that there might be inherent gender differences in interest, but previous research emphasizes the effect of existing social structures on the formation of such disparities. To increase the number of women entrepreneurs in domains connected to STEM professions, equitable opportunities have to be given them throughout their education years so they can choose their professions freely, which should also affect the number of women entrepreneurs in STEM fields.

8.3 Success of women on online crowdfunding platforms

As shown in Chapter 5, women are better represented on Kickstarter than in the venture capital industry. Also, as opposed to women entrepreneurs seeking external financing, they are fundamentally different from men in the amount of money they ask for, and they are also almost equally likely to get it. There can be multiple reasons behind this phenomena, such as the difference in riskiness of starting a Kickstarter campaign and seeking external funding from offline sources, or the absence of discriminatory practices on online crowdfunding platforms.

Risk on Kickstarter Women are found to be less prone to take financial risk (see Subsection 2.2.1), and as a result, woman-led ventures seek external financing much more rarely than men do. However, on Kickstarter, the share of women is higher than among venture capital recipients, and they also ask for about the same amount of money as men do. One possible explanation can be that on Kickstarter, woman-led projects are not as exposed to financial risk as companies seeking venture capital financing are. On Kickstarter, they do not have to take financial responsibility for their projects¹, but they are trusted by the community of backers that they will do their best to deliver. Consequently, this ease from the threat of indebtedness can explain women's behavior on the website.

¹https://www.kickstarter.com/trust?ref=whoisresponsiblefaq

Success on Kickstarter After controlling for goal size, project topic, duration, number of rewards and other project-level characteristics, women are still found to be just as likely to reach their funding goal as men are (see Section 7.1). This finding suggests that genderbased discrimination is not that of a threat on crowdfunding platforms as in the case of traditional ventures. On Kickstarter, the gender of the founder is not hidden, as it can be usually determined from the owner's name and avatar picture, but it is not emphasized either. Most probably, project backer's majority does not explicitly differentiates between projects led by men or women. Whereas in venture capital financing, fund managers decide about the investment based on a lot of personal characteristics of the founder team.

One must add that even on Kickstarter, there are sizable differences in the crowd's support: men are expected to overshoot their campaign goal more than women are (see Section 7.2), which is expected be the result of unobserved differences between women's and men's projects.

8.4 Policy implications

Crowdfunding platforms exhibit characteristics which create more equitable and inclusive environments for both women and men entrepreneurs. Policymakers should find a way to adopt elements of crowdfunding's beneficial environment to other offline external financing institutions such as the venture capital industry. In the following paragraphs I summarize the key recommendations of this thesis.

- Educate women about financial risk-taking. Studies show that women are less likely to take financial risks than men. Whatever the reason for this, the result is that women use less external capital affects their ventures' performance. By comparison, men and women asked for the same amount of funding on Kickstarter, a lower-risk environment. As one cannot eliminate the risks of starting an enterprise, women should be educated about the potential payoffs of taking more risk: women-led enterprises should be supported with subsidized consulting services about opportunities to grow their businesses more quickly.
- Narrow the STEM education gap. Women's underrepresentation in STEM professions is a significant factor explaining the lower success possibilities of women-led ventures overall. Women in STEM should face equitable treatment throughout their education and beyond. Programs promoting women choosing mathematics-intensive professions should be started and supported.
- Make funding decisions more inclusive. To help women-led enterprises succeed, women should be represented both as decision makers and as recipients in the allocation of financial resources. In a comprehensive policy framework, private equity ventures with women in management or investing in women's enterprises should receive tax benefits. Further, regular reports should be published about the performance of these fund managers and ventures to lead by example.

According to Henry et al., "women's entrepreneurship policies focus primarily on individuallevel challenges ("do-it-yourself solutions") rather than institutional (cultural and normative) level interventions". Evidence from a more equal platform, Kickstarter, suggests that women do not lack the ambition or capability to invent and raise funds for good ideas. It is rather the old institutional structures of external financing are put obstacles in the way of women-led businesses. Thus, policies are needed that promote normative changes in how women are perceived as entrepreneurs.

Chapter 9

Conclusion

In my thesis I analyzed the performance and success of more than 15 thousand entrepreneurial ventures concerned with product development on Kickstarter to examine women and men entrepreneurs' ambitions, willingness to take risk and domain choice.

I found that as opposed to traditional sources of financing, women and men who enter crowdfunding platforms set themselves similar goals. Since Kickstarter is an all-or-nothing game, meaning that those projects that does not reach their goals lose the money that was given to them, setting higher goals indicates higher willingness to take risks. Therefore I conclude that women and men have similar risk preferences on Kickstarter.

Further, in my sample, men and women entrepreneurs are equally likely to succeed (reach their goals), but men are expected to overshoot their goal more than women: for successful projects, women are expected to raise 2.35 times their funding goal, while the average man raises 4.58 times his target. This difference means that the most successful projects are still owned by men entrepreneurs, and the community of Kickstarter funders is expected to get enthusiastic about men's projects than women's.

I built probabilistic topic modes on the textual descriptions of the projects to get a clear picture about the domain of the campaigns. I found that there are substantial differences in men's and women's participation in these custom topics, also, the topics differ in profitability. Women were more likely to enter domains which were less profitable than men. I calculated a measure of campaign uniqueness based on the extent to which it defied the topic categorization called topic entropy as seen at [Horvát et al., 2018], which exhibited no substantial gender differences.

Next, by estimating linear models explaining the premium over goal per project, I found that even in top of Kickstarter's built-in category controls, topic controls explain a

substantial part of the premium gap between men or women (the 58% premium gap shrinks to 45% after exchanging broader category controls for topic controls, with a 5% points increase in R-squared). The coefficient on the topic entropy control was significant and negative, meaning that entrepreneurs are punished by investors for being too experimental in their domain choice.

Limitations The scope of my thesis was to disentangle various external factors behind women's worse access to capital such as the interests of investors or discrimination and internal factors such as domain choice and goal set. Therefore, I could not examine every aspects of the differences between women and men entrepreneurs. Brush et al. suggests that women have worse access to networks of external financing, and they also underperform men in using their networking opportunities. Also, differences in men's and women's social responsibilities such as parenting affect their choices to undertake ventures [Brush et al., 2003, European Commission, 2017]. Regarding my choice of data source, Kickstarter does not provide a comprehensive picture of all women and men entrepreneurs, and it is not and it is not expected to be a perfect alternative of other forms of private equity, so I cannot perfectly generalize the patterns of gender differences in crowdfunding to other sources of financing. Also, even within Kickstarter, my sample of 15 thousand projects is not comprehensive, and the last projects in my sample are dated in 2015, so it might be worth to revisit my results on a sample of more recent projects as well.

Final remarks I conclude that on Kickstarter, more women participate than in traditional forms of private equity, and they are also more likely to be successful in raising the desired funds. They are also willing to take the risk of setting just as high goals as men do. Still, successful men are expected to overshoot their funding goal more than women do, which can be partially explained by men choosing more lucrative fields to start project in.

Crowdfunding platforms exhibits features which makes them a more equitable place for women. Traditional sources of capital are advised to adopt practices which mitigate the barriers of women entrepreneurs who wants to enter the private equity market such as low level of discrimination. However, the example of Kickstarter projects show that differences in domain choices of men and women - let it be the result of differing interests or structural barriers discouraging women from certain fields - still explain a substantial part of the gender gap in funding. Thus, solving the gender gap in entrepreneurial success cannot be separated from solving the issue of women's low representation in occupations which are perceived as masculine such as STEM professions.

Appendix A

Appendix

Table A.1:	Main	descriptives	for	the	initial	unrestricted	sample	of	94	896	Kickstarter
projects											

	mean	sd	median	min	max	skew
goal (000 USD)	34	858	5	0	100 000	93.0
pledged (000 USD)	9	52	1.5	0	10 267	109.6
% funded	300	17 725	100	0	$4\ 153\ 501$	172.0
probability of success	0.50	0.50	1	0	1	-0.01

Table A.2: Main descriptives for the final sample of 15 059 product development Kickstarter projects

	mean	sd	median	min	max	skew
goal (000 USD)	51	1 340	7.500	0.011	100 000	64.8
pledged (000 USD)	9	34	0.7	0	1 092	13.5
% funded	159	1 527	14	0	13 490	67.9
probability of success	0.38	0.49	0	0	1	0.5

Table A.3: Dissecting the gender	gap in all dollars	raised on Kickstarter	(for the users with
goals set in the top 5%)			

	women	men	% difference	% of difference in sum of money pledged
sum of money pledged (000 USD)	2 944	8 850	201	100
population (count)	85	661	678	185
average goal (000 USD)	1 527	697	-54	-71
average funding ratio (%)	2	2	0	-15

Table A.4: Dissecting the gender gap in all dollars raised on Kickstarter (for the users with goals set in the bottom 95%)

	women	men	% difference	% of difference in sum of money pledged
sum of money pledged (000 USD)	19 525	111 624	472	100
population (count)	3 207	11 106	246	71
average goal (000 USD)	10	13	28	14
average funding ratio (%)	58	75	29	15

Table A.5: Topic within the Technology category with their most frequent keywords, $\mathbf{k}=3$

topic	keywords	label
1	power, system, light, board, device	device
2	student, team, school, community, software	community software
3	app, information, data, web, mobile	application

Table A.6: Topic within the Design category with their most frequent keywords, k = 12

topic	keywords	label
1	card, deck, poster, printed, add	playing cards
2	bottle, coffee, beer, water, glass	liquid containers
3	community, space, golf, building, home	building/community space
4	desk, light, lamp, stand, base	lighting/furniture
5	bike, bicycle, ride, bikes, plastic	biking
6	iphone, phone, ipad, light, device	accessoires for devices
7	wallet, bag, pocket, tool, leather	wallets/organizers
8	pet, baby, dog, home, cat	pets
9	designs, print, art, paper, book	prints
10	pen, wood, guitar, black, clock	handcrafted products
11	camera, system, lens, watch, business	photography/devices
12	kids, children, toy, life, family	lifestyle

topic	keywords	label
1	soap, natural, skin, oil, essential	cosmetics
2	level, bowl, diameter, picture, plush	miscellaneous hobby
3	wood, pen, black, piece, tool	woodworking
4	yarn, local, shop, build, community	community space
5	candles, cards, day, wax, family	candles/gifts
6	jewelry, colors, rings, silver, leather	jewelry

Table A.7: Topic within the Crafts category with their most frequent keywords, $\mathbf{k}=6$

Table A.8: Topic within the Games category with their most frequent keywords, $\mathbf{k}=3$

topic	keywords	label
1	experience, team, development, gaming, character	computer games
2	dice, board, rules, friends, hand	boardgames
3	book, pdf, character, adventure, additional	roleplay

Table A.9: Topic within the Fashion category with their most frequent keywords, k = 4

topic	keywords	label
1	leather, shoe, bag, wallet, pocket	leather goods
2	color, jewelry, size, black, silver	jewelry
3	brand, company, art, money, website	branding
4	fabric, dress, style, women, fit	women's apparel

Table A.10: Sample project short descriptions from two topics of the Technology category: community software and device

category	label	short description
Technology	community software	Ever had trouble catching a cab? Frustrated because cabdrivers just don't see you? Your solution is here!
Technology	community software	Our kids need better Science Kits like the great ones from the 50's and 60's you can help do that and launch a makerspace too.
Technology	community software	Free educational mobile app designed as a tool for students or cinephiles alike! Make the Paris of films your classroom and playground!
Technology	community software	Dedicated to heroes who answered the call to serve their country and community.
Technology	community software	Songwriters need a place to post and sell their songs - musicians need a place to find and buy music - let's make the connection.
Technology	community software	The goal is to create a mobile planetarium program that will train my students to present Astonomy programs to elementary schools.
Technology	community software	The LilyBean Animator will bring stuffed animals to life to en- tertain, educate, and energize children like nothing you've seen before.
Technology	community software	The QuivExchange is an online social trading site allowing ath- letes to buy, sell, trade or exchange extreme sports equipment.
Technology	device	Keep your new Apple Watch free from everyday nicks and scratches with the ability to still use the crown wheel and push button.
Technology	device	If you are like me, you hate filling the water reservoir on your coffee maker. Why not automate that with Touch-N-Fill TM ?
Technology	device	This thermostat lets you reduce your carbon footprint and di- rectly control how much you spend heating and cooling your home.
Technology	device	W-sitting solution! Provides immediate alarm when child is W-sitting. The FixSit (patent pending).
Technology	device	Using a centrifuge, 2 LEDs connected to 2 button batteries and a spectrograph, one can observe redshift if time is bent.
Technology	device	A complete hydraulic hose kit designed to give a permanent solution that you create yourself - quickly and easily.
Technology	device	Affordable Hydroponic and Aquaponic Control from your Smart- Phone, Tablet, or PC. Starting at \$349
TechnologydeviceWith this device attached to an aircraft, you will be able t cate the approximate location of a crashed aircraft at sea c land		

category	label	short description
Design	lifestyle	This shelter is a new idea in the way you enjoy the great outdoors, such as watching sporting events, camping, on the beach, anywhere
Design	lifestyle	Made for summer camping & festivals! Keeps you cooler and drier in the morning. Easily attaches directly to your tent, like a rain fly.
Design	lifestyle	Mason Surf Designs is a classic-inspired traditional surfboard brand for the outdoor, coastal adventure person.
Design	lifestyle	Geleeo: A unique gel pad designed to keep kids cool when out on their strollers
Design	lifestyle	Craftsmanship & design meet to bring you a quality bottle opener that is both personal and unique.
Design	lifestyle	The world's coolest, most exciting action sport toy for 2015. Jump & land snowboard tricks like PRO! Designed in USA. Simply genius!
Design	lifestyle	A tool that enables blind people to tell apart their medications, bot- tles of wine, canned foods, and other indistinguishable items.
Design	lifestyle	Experience the pleasure & sophistication of a portable charcoal grill Nearly smokeless fire creates a fun & healthy atmosphere
Design	wallets/organizers	Creating a useful messenger bag out of used bicycle innertubes.
Design	wallets/organizers	Easily slide your cards & cash in and out. Store your earbuds on your phone stand, and put your keys in a matching 3 or 6 key orga- nizer
Design	wallets/organizers	A modern iteration of the classic 1980s Duffle Bag. Available in a two-tone blue and grey color combination.
Design	wallets/organizers	Love the handiness of drawstring bags but hate all the annoying problems that come with them!? The Solution: The Grind Bag TM
Design	wallets/organizers	Meet TAB: A slim and stylish carry-on that fits right into the air- plane seatback pocket and organizes all travel gear.
Design	wallets/organizers	A keychain/wallet combo. Let's face it! Slim wallets are great but we don't just carry our wallets WHAT ABOUT OUR KEYS!!!
Design	wallets/organizers	A minimalist silicone bracelet with a function. Gokey securely con- ceals your key while you are on the go.
Design	wallets/organizers	These Clever Laptop and iPad sleeves are ready for a Big production run, right here in the USA!

Table A.11: Sample project short descriptions from two topics of the Design category: lifestyle and wallets/organizers

Table A.12: Sample project short descriptions from two topics of the Crafts category: candles/gifts and woodworking

category	label	short description
Crafts	candles/gifts	The Jacob's Ladder Tree ornament is made with 3D printing to bring the Jacob's Ladder toy (aka tumbling blocks) into your Holidays.
Crafts	candles/gifts	Love making candles for a hobby, now i'm trying to turn it into a small business, with your help.
Crafts	candles/gifts	Custom hand-made Mala's: Japa, Hand, Mantra, Zen, and Tibetan styles. Malas help users contemplate the meaning
Crafts	candles/gifts	Trendy handmade dorm decoration packages! The package above would include the wall sequins, canvases, photo strip and a pillow case.
Crafts	candles/gifts	Give your family a Christmas they will never forget! -Back us and get a Xmas tree! -Plant your own tree on our farm!
Crafts	candles/gifts	A candle with an inspiring mantra on the outside & a surprising mantra inside meant to be passed on. *Downloadable meditation included.
Crafts	candles/gifts	Bring the song of the breeze and beauty of colors to brighten your day with Autumn's Handmade Windchimes.
Crafts	candles/gifts	Our start-up of a line of 100% Natural Soy candles and wax melts/tarts. Made in America by American soy beans!
Crafts	woodworking	My father and I have a wood shop here in Richmond VA. We hand make cutting boards, bandsaw boxes, pens, bowls, and much more.
Crafts	woodworking	HeartWood is a line of wooden house and kitchen ware made primarily from material salvaged in and around New York City.
Crafts	woodworking	Fine men's and women's accessories and gifts created from reclaimed trop- ical hardwood, hand crafted in Seattle, WA.
Crafts	woodworking	A-Frame Dog Houses ~Expedition Boats ~ Camping Tools ~ Oak Back- Pack Frames ~ Outdoor Gear Sewing Repair Service ~ Ram Air Kites.
Crafts	woodworking	Wooden Toy Airplanes handcrafted in North Carolina from a variety of native woods.
Crafts	woodworking	I am a woodworking enthusiast. I enjoy recycling pallet wood and turning them into wonderful pieces.
Crafts	woodworking	I want to create beautiful furniture from reclaimed barnwood, beams, flooring and boards.
Crafts	woodworking	The Mill Springs Cabin needs a boardwalk from the existing sidewalk, around one of the exhibits, and ending in a deck.

	Dependent variable:					
		% funded				
	(1)	(2)	(3)	(4)	(5)	
gender	-0.38^{***}	-0.15^{**}	-0.14^{**}	-0.16^{**}	-0.16^{**}	
	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	
log goal		-0.40^{***}	-0.64^{***}	-0.65^{***}	-0.65^{***}	
		(0.02)	(0.02)	(0.02)	(0.02)	
lists we			0.13^{**}	0.17^{***}	0.17^{***}	
			(0.05)	(0.05)	(0.05)	
log description length			0.73^{***}	0.69^{***}	0.69^{***}	
			(0.04)	(0.04)	(0.04)	
log risk section length			0.26***	0.25***	0.25***	
			(0.04)	(0.04)	(0.04)	
topic entropy above median			()	· · · ·	-0.03	
1 10					(0.05)	
category control			YES		()	
topic control				YES	YES	
start year control			YES	YES	YES	
reward control			YES	YES	YES	
Constant	-4 40***	-0.94***	457 51***	455 71***	455 83***	
	(0.06)	(0.16)	(85.00)	(84.21)	(84.21)	
Observations	9,366	9,366	9,366	9,366	9,366	
\mathbb{R}^2	0.003	0.06	0.33	0.34	0.34	

Table A.13: Predicting % funded for failed projects on Kickstarter

Note:

*p<0.1; **p<0.05; ***p<0.01

Bibliography

- [American Express Open, 2017] American Express Open (2017). The 2017 state of womenowned business report.
- [An et al., 2014] An, J., Quercia, D., and Crowcroft, J. (2014). Recommending investors for crowdfunding projects. In *Proceedings of the 23rd international conference on World wide web*, pages 261–270. ACM.
- [Belleflamme et al., 2014] Belleflamme, P., Lambert, T., and Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of business venturing*, 29(5):585–609.
- [Belleflamme et al., 2015] Belleflamme, P., Omrani, N., and Peitz, M. (2015). The economics of crowdfunding platforms. *Information Economics and Policy*, 33:11–28.
- [Blau and Kahn, 2017] Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.
- [Blei et al., 2003] Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.
- [Brindley, 2005] Brindley, C. (2005). Barriers to women achieving their entrepreneurial potential: Women and risk. International Journal of Entrepreneurial Behavior & Research, 11(2):144–161.
- [Brush et al., 2018] Brush, C., Greene, P., Balachandra, L., and Davis, A. (2018). The gender gap in venture capital-progress, problems, and perspectives. *Venture Capital*, 20(2):115–136.
- [Brush et al., 2003] Brush, C. G., Carter, N. M., Gatewood, E., Greene, P. G., and Hart, M. (2003). Venture capital access: Is gender an issue. The emergence of entrepreneurship policy: Governance, start-ups, and growth in the US knowledge economy, pages 141–154.

- [Cohoon et al., 2010] Cohoon, J. M., Wadhwa, V., and Mitchell, L. (2010). Are successful women entrepreneurs different from men?
- [Coleman and Robb, 2009] Coleman, S. and Robb, A. (2009). A comparison of new firm financing by gender: evidence from the kauffman firm survey data. *Small Business Economics*, 33(4):397.
- [Correll, 2001] Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American journal of Sociology*, 106(6):1691–1730.
- [European Commission, 2014] European Commission (2014). Statistical data on women entrepreneurs in europe. Technical report, European Commission.
- [European Commission, 2017] European Commission (2017). Policy brief on women's entrepreneurship. Technical report, European Commission.
- [Gorbatai and Nelson, 2015] Gorbatai, A. and Nelson, L. (2015). The narrative advantage: Gender and the language of crowdfunding. *Haas School of Business UC Berkeley*. *Research Papers*.
- [Guillén et al.,] Guillén, L., Mayo, M., and Karelaia, N. Appearing self-confident and getting credit for it: Why it may be easier for men than women to gain influence at work. *Human Resource Management*.
- [Horvat and Papamarkou, 2017] Horvat, E.-A. and Papamarkou, T. (2017). Gender differences in equity crowdfunding. In *HCOMP 2017, Quebec City, QC, Canada*.
- [Horvát et al., 2018] Horvát, E.-Á., Wachs, J., Hannak, A., and Wang, R. (2018). The role of novelty in securing investors for equity crowdfunding campaigns. In *The 6th AAAI Conference on Human Computation and Crowdsourcing (HCOMP). AAAI*.
- [Hyde et al., 1990] Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., and Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of women quarterly*, 14(3):299–324.
- [Jadidi et al., 2017] Jadidi, M., Karimi, F., Lietz, H., and Wagner, C. (2017). Gender disparities in science? dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, page 1750011.

- [Karimi et al., 2016] Karimi, F., Wagner, C., Lemmerich, F., Jadidi, M., and Strohmaier, M. (2016). Inferring gender from names on the web: A comparative evaluation of gender detection methods. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 53–54. International World Wide Web Conferences Steering Committee.
- [Kuppuswamy and Bayus, 2017] Kuppuswamy, V. and Bayus, B. L. (2017). Does my contribution to your crowdfunding project matter? *Journal of Business Venturing*, 32(1):72–89.
- [Kuppuswamy and Mollick, 2016] Kuppuswamy, V. and Mollick, E. (2016). Second thoughts about second acts: Gender differences in serial founding rates.
- [Lee and Huang, 2018] Lee, M. and Huang, L. (2018). Gender bias, social impact framing, and evaluation of entrepreneurial ventures. *Organization Science*, 29(1):1–16.
- [Lin and Pursiainen, 2017] Lin, T.-C. and Pursiainen, V. (2017). Gender differences in crowdfunding-entrepreneur overconfidence and campaign outcomes. Technical report, Working paper.
- [Loscocco and Robinson, 1991] Loscocco, K. A. and Robinson, J. (1991). Barriers to women's small-business success in the united states. *Gender & Society*, 5(4):511–532.
- [Marom et al., 2013] Marom, D., Robb, A., and Sade, O. (2013). Gender dynamics in crowdfunding: Evidence on entrepreneurs, investors, and deals from kickstarter. In *Academic Symposium on Crowdfunding. UC Berkeley.*
- [Mayer, 2008] Mayer, H. (2008). Segmentation and segregation patterns of women-owned high-tech firms in four metropolitan regions in the united states. *Regional Studies*, 42(10):1357–1383.
- [Michael, 2016] Michael, J. (2016). Gender guesser.
- [Michael and Gecko, 2007] Michael, J. and Gecko, Z. (2007). Gender verification by forename.
- [Mohr and Bogdanov, 2013] Mohr, J. W. and Bogdanov, P. (2013). Introduction—topic models: What they are and why they matter.

- [Mollick, 2014] Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. Journal of business venturing, 29(1):1–16.
- [Mollick, 2013] Mollick, E. R. (2013). Swept away by the crowd? crowdfunding, venture capital, and the selection of entrepreneurs.
- [Olsen, 2018] Olsen, D. (2018). Do female founders get better results? here's what happened on my quest to find out.
- [Orser et al., 2006] Orser, B. J., Riding, A. L., and Manley, K. (2006). Women entrepreneurs and financial capital. *Entrepreneurship Theory and Practice*, 30(5):643–665.
- [Powell and Ansic, 1997] Powell, M. and Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of economic* psychology, 18(6):605–628.
- [Sara and Peter, 1998] Sara, C. and Peter, R. (1998). The financing of male-and femaleowned businesses. *Entrepreneurship & Regional Development*, 10(3):225-242.
- [Shaw et al., 2016] Shaw, E., Hegewisch, A., Williams-Baron, E., and Gault, B. (2016). Undervalued and underpaid in america: women in low-wage, female-dominated jobs. Technical Report IWPRD508, Institute for Women's Policy Research.
- [Statista, 2018] Statista (2018). Money raised by crowdfunding globally 2012-2015 statistic.
- [Stoet and Geary, 2018] Stoet, G. and Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological science*, page 0956797617741719.
- [Teare and Desmond, 2017] Teare, G. and Desmond, N. (2017). Announcing the 2017 update to the crunchbase women in venture report.
- [Vasilescu et al., 2012] Vasilescu, B., Capiluppi, A., and Serebrenik, A. (2012). Gender, representation and online participation: A quantitative study of stackoverflow. In Social Informatics (SocialInformatics), 2012 International Conference on, pages 332–338. IEEE.
- [Vulkan et al., 2016] Vulkan, N., Åstebro, T., and Sierra, M. F. (2016). Equity crowdfunding: A new phenomena. *Journal of Business Venturing Insights*, 5:37–49.
- [Wang et al., 2013] Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological science*, 24(5):770–775.
- [Washington Center of Equitable Growth, 2017] Washington Center of Equitable Growth (2017). Occupational segregation in the united states. Technical report, Washington Center of Equitable Growth.
- [Watson and Robinson, 2003] Watson, J. and Robinson, S. (2003). Adjusting for risk in comparing the performances of male-and female-controlled smes. *Journal of business venturing*, 18(6):773–788.
- [Zarya, 2017] Zarya, V. (2017). Venture capital's funding gender gap is actually getting worse.