



Paper to be presented at the DRUID Academy Conference 2017 at University of Southern Denmark, Odense, Denmark on January 18-20, 2017

New venture initial funding in contexts of latent category labels

Jonathan Sitruk
SKEMA Business School
Entrepreneurship
jonathan.sitruk@skema.edu

Abstract

New venture initial funding in contexts of latent category labels Jonathan Sitruk - jonathan.sitruk@skema.edu
SKEMA Business School, 5th year PhD student Expected graduation date 2017 State of the art The liability of newness hinders financing because new ventures are less capable of manifesting the underlying quality of their credentials and offering. Categorization scholars have observed how offerings that span categories tend to be discounted (Zuckerman, 1999) and that acceptance of a group's norms increases chances of success (Kovács & Hannan, 2015). Research gap The underlying assumption of categorization theory is that clearly belonging to a category facilitates an investor's ability to judge the quality of a venture and increases a producer's ability to attribute sufficient resources to succeed within a given market category. However, previous research has focused on situations where categories are clearly visible through the use of labels (Kovács & Hannan, 2015).

We know little of what happens to project evaluation and funding when categories are not easily visible because labels are not clearly applied. How do producers communicate categorical belonging and typicality when clear labels are absent? Does the audience understand these signals? Theoretical arguments When category labels are absent, we posit at the core of our theorizing that audiences use several elements to judge producers. First, they identify latent categories that are present in the narratives of product descriptions allowing them to tentatively identify producers' categorical belonging. Second, they observe the way producers express themselves making subsequent judgment calls on the typicality of the offering with regards to the categorical rules. Method and Data We capture these project features through the narratives that project holders provide about their offering. Narratives allow new ventures to secure more funding by providing a clearer depiction of who they are as an organization (Wry et al., 2013; Martens et al., 2007; Lounsbury & Glynn, 2001). We apply machine learning semantic techniques to analyze the descriptions of crawled crowdfunding data. We focus on the music category and observe how the use of latent labels (i.e.: music genres) and overall semantic typicality of descriptions across music genres influences project success and funder composition. We analyze these data through probit and network analysis. Results and Contribution We find support for our contention concerning funding success: projects that are more typical and rely on latent labels are more successful on average. Analysis concerning funder composition are still preliminary but promising. Finally, the insights gained from this work pertaining to the success factors and strategies used by new ventures in lieu of clear categories may be valuable to both the entrepreneurship and categorization literatures, and provide notable managerial implications. Selected references Durand, R., L. Paoletta. 2013. Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. *J. Manag. Stud.*

50(6) 1100–1123. Fleming, L. 2001. Recombinant Uncertainty in Technological Search. *Manage. Sci.* 47(1) 117–132. Kovács, B., M. T. Hannan. 2015. The Space of Categories and the Consequences of Category Spanning. *Sociol. Sci.* 2 252–286. Kovács, B., R. Johnson. 2014. Contrasting alternative explanations for the consequences of category spanning: A study of restaurant reviews and menus in San Francisco. *Strateg. Organ.* 12(1) 7–37. Lounsbury, M., M. A. Glynn. 2001. Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strateg. Manag. J.* 22(6–7) 545–564. Martens, M. L., J. E. Jennings, P. D. Jennings. 2007. Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition. *Acad. Manag. J.* 50(5) 1107–1132. Mollick, E. 2014. The dynamics of crowdfunding: An exploratory study. *J. Bus. Ventur.* 29(1) 1–16. Wry, T., M. Lounsbury, M. A. Glynn. 2013. Legitimizing Nascent Collective Identities: Coordinating Cultural Entrepreneurship. *Organ. Sci.* 22(3) 226–245. Zuckerman, E. 1999. The Categorical Imperative: Securities Analysts and the Illegitimacy Discount. *Am. J. Sociol.* 104(5) 1398–1438. Zuckerman, E. W. 2015. Optimal Distinctiveness Revisited: An Integrative Framework for Understanding the Balance between Differentiation and Conformity in Individual and Organizational Identities. *Handb. Organ. Identity.*

New venture initial funding in contexts of latent category labels

Jonathan Sitruk
SKEMA Business School, Strategy,
Entrepreneurship & Economics Department
Tel: 04 93 95 45 63
e-mail: jonathan.sitruk@skema.edu

WORKING PAPER
December 2016

This article focuses on the determinants of initial new venture funding in contexts of latent category labels. More precisely, we explore contexts where competition and uncertainty are high forcing entrepreneurs to devise alternative ways to communicate about their project. We call on categorization theory to decipher what entrepreneurs do in such situations. We find that entrepreneurs will use coherent communication narratives as a means to signal their typicality level within this uncertain context. The study uses novel machine learning and semantic measures applied to over 2,800 crowdfunded projects from the Music category presented on Indiegogo from 2009 to early 2015.

1 Introduction

Past research has shown that entrepreneurs carrying projects that are representative of a category are more successful in accessing initial funding because their audience succeeds in interpreting their offer (Hannan et al. 2007). Entrepreneurs experience difficulties in securing this financing because their liability of newness and subsequent lack of legitimacy makes it more difficult for them to manifest the underlying quality of their credentials and offer (Aspelund et al. 2005) often leading them to fail (Brush et al. 2008). Thus, gaining access to these initial funds is an important milestone for any organization and identifying strategies aimed at improving resource procurement have been a driving motivation of the entrepreneurship and strategy research streams (Stinchcombe 1965). For instance, scholars have explored how the viability and clarity of an offering are crucial for new venture success (Barry and Elmes 1997, Cornelissen and Clarke 2010, Lounsbury and Glynn 2001, O'Connor 2002). Investors face bounded rationality in understanding how clear a product offering is and they must rely on limited cues to infer new venture quality. Originating from the field of social psychology, the categorization literature addresses this situation by positing that people have limited cognitive abilities leading them to use sense-making tools (Brewer 1991) to alleviate the informational difference they have regarding project holders. Categories allow audiences to have a clearer depiction of the expectations that producers of a given industry should follow. Although rich in its insights, this literature has tended to focus mainly on situations where category labels are clearly visible (Kovács and Hannan 2015, Reschke 2015).

Kovács and Hannan (2015) explain that in standard contexts a market intermediary (e.g.: a website manager) is the one who generally attributes labels to producers or products. As such, most research in categorization theory has observed contexts where category labels that are either self-appointed (e.g.: restaurant study by Kovács and Johnson (2014)) or appointed by a third party (e.g.: wine study by Negro et al. (2015)). Regardless of the appointer, these studies take place in contexts where category membership is clearly and easily identifiable by the audience. The importance of these clear labels lies on the main assumption that individuals understand the underlying meaning of a category label as it serves as an anchor allowing audience members to contextualize a producer or a product

in a given schema that has its rules and expectations (Hannan et al. 2007, Zuckerman 1999). We lack a deeper understanding of how audiences derive typicality of a producer and evaluate them in contexts where categories are not easily visible or clearly applied. Phrased differently, we aim to understand how producers can use categorization tenets to their advantage in such contexts to secure higher chances of financing. Addressing these issues is important for firms evolving in highly uncertain environments such as in nascent and/or innovative industries.

We posit at the center of our theorizing that audiences use several elements to judge if producers are worthy of investment. Audiences first observe latent category labels or cues present in the phrasing of product description allowing audiences to tentatively identify the categorical belonging of producers or products (Zuckerman 1999). However, given the latent and poor quality signal of these labels, audiences may not address them in a traditional fashion. Indeed, they need to be more thorough in their evaluation and detect whether the way producers express themselves is in accordance with a categorical schema leading them to better understand a producer's ability to deliver on their product offering. From a producer's side, we observe how producers strategically use coherent communication narratives allowing audiences to comprehend their level of typicality, which in turn leads to higher levels of funding.

We capture the project features through the narratives that project holders provide about their offering. Indeed, narratives have been shown to be a strong communication tool and "help individuals understand and describe who they are" (Martens et al. 2007, p. 1110). Furthermore, narratives allow "individuals [to] "construct" themselves as "characters" whose attributes can be revealed and communicated[, shaping] [...] not only how individuals view themselves, but also how others view them" (Martens et al. 2007, p. 1110). Several scholars have studied how narratives allow new ventures to secure more funding by providing a clearer depiction of who they are as an organization (Lounsbury and Glynn 2001, Martens et al. 2007, Reinecke et al. 2012, Santos and Eisenhardt 2004, Wry et al. 2013) Other researchers have shown that narratives where selective

entrepreneurial traits are present are important determinants of venture funding (Allison et al. 2013, Cardon et al. 2009, 2012, Short et al. 2010).

Our analysis takes place in the context of crowdfunding. Crowdfunding is a fairly new phenomenon and scholars in the field of management and entrepreneurship have taken great interest in understanding it. One of the most prominent focuses taken is to understand determinants of success of crowdfunding projects (Mollick 2014) fitting our research interests. Researchers have taken different theoretical foundations to explain what are success factors of different crowdfunding projects. For instance, (Agrawal et al. 2013) and Mollick (2014) take it as an asymmetry of information problem and explain that signals such as use of videos, images, type of financing, length of projects, number of project holders help reduce the latter. Others suggest (Agrawal et al. 2011, 2013) that crowdfunding success is mostly due to the amount of friends and family investments (i.e.: love money) in early stages of projects that lead to a herding behavior by other funders. We take a different perspective in this paper by using different socio-psychological theories to explain how a project can increase its chances of success. In the third point on Indiegogo's tips to having a successful campaign they recommend that "the more you can demonstrate you've thought through as many aspects of your campaign as possible [...] the better your audience can visualize where their money will go, which builds trust". However, for the audience or potential investors to understand where their money will go, they must first clearly understand what it is that a project is offering. We propose to explore this idea by calling on socio-psychological theories.

We contribute to the categorization literature by exploring settings where category labels are latent and not provided by a third party leading producers to adopt different strategies. We note how coherent communication narratives play a role in making audience members see projects as typical. Our contribution to the entrepreneurial literature is to provide a new take on the determinants of success of initial venture funding in contexts where competition is high and clearly expressing one's identity is trying. We further contribute to the strategy literature by providing a categorization strategy that producers can use to increase their chances of being seen as typical within industries of high uncertainty.

In the next sections, we first provide a literature review and develop a set of hypotheses related to how a project's clarity and differentiation impacts success. Second, we provide background on the crowdfunding context. Third, we explain our data. Fourth and fifth we explain our methods and results. We finish we a conclusion.

2 Theoretical background:

2.1 Categorization, categories, labels, and schemata

Originating from the field of social psychology, scholars hold that people have limited cognitive abilities making them unable to understand everything that they try to evaluate without using some sense making tools (Brewer 1991). Indeed, we do not relearn how to distinguish the difference between an action movie, a smartphone, and a horse every time we encounter them in our everyday lives. One of these sense making tools is categorization (Brewer 1991) which essentially consists of fitting objects, people, businesses into boxes for which we have a clear understanding and clear prototype(s). This is more efficient and cognitively feasible given our limited capacities. We thus recognize an action movie because we may at the broadest level categorize it as a movie (as opposed to a phone, or an animal). Within the movie realm, we may acknowledge that the movie we are watching produces a higher level of stress and adrenaline than other types of movies (e.g.: comedy, drama). These categories allow us to clearly identify that this particular movie is an action movie. It is important to note how categories can occur in different scopes (i.e.: global level vs. more precise) but our understanding relies on this categorization. Cognitively speaking, when we see an action movie and try to make sense of it, we try to see where it fits, understand the “what is it” question (Martens, et al., 2007) by seeing how it compares to the prototypical ideal of the category that we are comparing it to.

This categorization does not happen in a void and consists of a socialized process agreed upon by the audience members of a field. Producers first collaborate to identify a shared understanding of a given category in which they are engaged by defining the labels and schemata they agree on (Hsu and Hannan 2011, Wry and Lounsbury 2013). In a second step and once a category’s identity is legitimized, producers – whose identity has been derived from the categorical labels it has been assigned – may attempt to improve it by showing how their offer provides greater value than other members of a category (Zuckerman 1999, 2015). In essence, “categories convey the cultural 'codes' that are associated with belonging to a particular category” (Vergne and Wry. 2014, p. 63) and when a producer meets the expectations, she has greater chances of being financed. However, projects that cannot be positioned within any category will trouble investors

making them unable to evaluate it with regards to any category schema (Brewer 1991).

Hannan et al. (2007) explain that schemata and labels are the constituents of a category and allow producers to gain an (organizational) identity. A schema is built from the taken-for-granted defaults (i.e.: expectations) of a category: its rules, norms, and canons, while a label is what audiences and producers use to signal a given category. One category may use several labels but only one schema. When a producer is assigned a given category label, her identity is tied to the anticipated compliance to the defaults of that category. Said differently, audience members will assume that the producer meets (to a minimal degree) the rules of a given category providing her an understandable and legitimate identity which “consists of a pair given by (1) a composite label and (2) all that is taken for granted on the basis of this (composite) label alone” (Hannan et al. 2007, p. 104).

The underlying assumption is that audience members do not need to investigate whether a producer is respecting their categorical engagements. This assumption relies on the notion that if a producer is attributed a label, audience members use this label and its underlying schema as a sense-making instrument and do not need to go through the efforts of categorizing the producer themselves. In other words, the label provides a signal from which all the taken-for-granted expectations are derived and audiences use this as a foundation to evaluate the typicality of a producer in a given category.

However, if the producer is caught violating the expectations (or defaults) of the category for which she has received the label(s), she will be sanctioned accordingly. Audiences will give less trust to the violating producer (Hannan et al. 2007, p. 106) leading to a drop in their grade of membership in the category. The grade of membership is the “degree to which the object shares values with members of other categories” (Hannan et al. 2007, p. 108) and with representativeness makes up typicality. A lower grade of membership either means that a producer 1) loses typicality regarding a focal category and by extension 2) spans between several categories causing her to have less appeal with

regards to competitors, and a subsequent heightened mortality rate (Hannan et al. 2007, p. 106).

2.2 Typicality and Spanning

Organizational typicality consists of two constructs of different importance that people use in “assigning objects to categories” (Hannan et al. 2007, p. 108). The first is certainty of membership or the “degree to which the object shares feature values with members of other categories” (Hannan et al. 2007, p. 107). This grade of membership has been measured in the past by observing different use of category labels and can be equated to a construct playing a role at the inter-category level of analysis. Audience members will thus seek to understand to what extent a producer is part of one or more categories. The second is representativeness or the “degree to which the entity shares feature values with other category members” (Hannan et al. 2007, p. 107-108). This construct plays a role at the intra-category level of analysis where audience members judge the similarity that a producer has with other members of that category and its prototype.

Theory holds that typical producers are more likely to receive funds because the label they are attributed – and its underlying schema defaults – is appropriate. They behave in accordance to the expectations of that category and do not violate any of its norms. On the other hand atypical projects may have a low grade of membership in a focal category leading them to be atypical of such a category.

Although both grade of membership and representativeness/similarity are important in judging organizational typicality, research has shown that audiences tend to only consider representativeness once the task is inductive; i.e.: making predictions about unobserved features (Hannan et al. 2007). This is because representativeness influences confidence in inductions so that if a product is representative of a category, an audience will feel more comfortable making predictions about the product’s unobserved features. This is especially important when producers span multiple categories.

Category spanning has traditionally been viewed as having a negative effect on organizational success. Category spanning is the idea of belonging to more than one category at the same level of analysis. Scholars have contended that spanning creates confusion for investors who, as the amount of category belonging increases, find new ventures increasingly ambiguous (Zuckerman 1999). In line with categorization literature, as a venture increases the amount of categories it is attributed to, the prototypes against which an investor has to evaluate the venture increase and it becomes increasingly problematic to do so. Spanning also carries a negative bias from a producer's perspective. Indeed, a producer spanning different categories is, by definition, less specialized in each category that they belong to (Kovács and Johnson 2014). Furthermore, several studies have shown that specialization has a positive impact on success (Hsu 2007). The schemata that are called upon by using multiple category labels increase the difficulty that audiences have in understanding what elements of a producer to focus on. Indeed, multiple schemata decreases the grade of membership in any one category leading to producers lacking representativeness. Producers become overall less typical, leading them to be penalized by audiences' lowered appeal and confidence about their inductions regarding producer unobserved features resulting in a lack of identity and difficulty to access funds.

Accepting the underlying tenets of categorization theory means to accept the importance that category labels play in the audience's identification of what schemata (i.e.: rules) to apply to a given producer. Categories are created when labels are seen as legitimate in large part because they emanate from an influential body that "possesses authority over some realm" (Hannan et al. 2007, p. 112). (Kovács and Hannan 2015) explain that this authority generally comes from market intermediaries (like website managers) who decide to attribute a certain label – and its underlying implied expectations – to a producers' offering. Therefore, categorization works at its best when labels defining a category are clearly present to indicate to investors what category a product belongs to and what rules it is constrained to follow if it does not want to be penalized (Kovács and Hannan 2015, Reschke 2015). However, not all contexts have such a clear third party labeling approach. Little research in categorization theory analyzes contexts where labels are not clear or inexistent (Reschke 2015) leading us to have

a blurred understanding of how audiences evaluate and fund producers when categories are not easily visible (i.e.: where labels are not clearly applied).

We posit as the core of our theorizing that in contexts void of clear category labels, audiences use several elements to judge if producers are worthy of investment. One the one hand, they call on narrative cues present in the phrasing of product description which allow them to evaluate the product offering. On the other, they observe and evaluate the way producers express themselves to make decisions on their ability to deliver on their product offering. More precisely, we believe that typicality is still an important element of producer evaluation and ultimate success. However, it is evaluated indirectly by using coherent narratives and the way producers portray themselves. The following sections provide our hypotheses.

3 Hypotheses

Typicality is an important construct of categorization theory. Audiences need to see producers as representative of a category (i.e.: typical) in order for the latter to gain that category's identity, legitimacy, and resulting access to funds (Hannan et al. 2007, p. 112). Labels and schema allow audiences to determine a producer's representativeness in standard contexts. However, these labels no longer play the same signaling role (more on this below) in contexts void of clear category labels leaving audiences at a loss. Our contention is that the way audiences (i.e.: investors) evaluate projects differs when labels representing categories are not clearly attributed by a third party (Kovács and Hannan 2015) leading to harder project evaluation (Reschke 2015). Before explaining how this evaluation differs, we start by positioning our argumentation within this tenet of categorization theory. Regardless of how audiences determine a producer's categorical belonging and typicality, if audiences interpret producers as being representative of a category, they will have a higher chance of receiving funds. For this reason, we propose the following baseline hypothesis:

H1: In an environment void of clear category labels, producers who are interpreted as typical tend to be more successful

With hypothesis 1 in mind, we suggest that in contexts void of clear categories attributed by a third party, the rules of the game differ compared to situations where category labels are explicitly presented. This is especially true given how audiences principally use labels as a category signal to limit cognitive investment. We suggest that in contexts of latent category labels, their indicative nature loses its power and they no longer are sufficient signals to encompass the expectations of a category due to their costless nature. Indeed, in order for signals to communicate as intended, they need to be too costly for lower quality agents to reproduce in order to be indicative of product quality (Milgrom and Roberts 1992). In the case signals are easy to reproduce, they become less trustworthy leading investors to have to perform thorough due diligence in order to identify quality projects (Puranam et al. 2006, Zider 1998).

Thus, although producers in an environment void of categories can easily use labels because of their low cost, these are not highly indicative of the categorical membership or the quality of the offering. Instead of relying entirely on category labels and situations where typicality (i.e.: representativeness) is easily evaluated, producers must devise other indirect strategies to communicate their categorical belonging and level of typicality. This situation is further exacerbated by the means by which the signal is delivered. In standard contexts, an investor will immediately identify the label of an offering and use their understanding of the underlying schema expectations associated with such a label to assume the behavior that a representative producer should adopt and evaluate a project with respect to the prototype of the category given by the label. If a producer is in line with these expectations – that is if she is typical of this category and her representativeness is appropriate – she has higher chances of procuring resources. If she strays too much from it, she will be seen as a violator and be penalized accordingly.

Investors do not get this luxury in situations void of categories. The label is concealed within a project description and audiences cannot immediately rely on established schema. Only when they analyze a producer's offering more thoroughly - by reading through a project description – may they potentially become aware of latent labels. Simultaneously, they use these identified labels as anchors and may derive a project's level of typicality thanks to the words and

tonality used in their narratives. Narratives are strong communication tools that “help individuals understand and describe who they are” (Martens et al. 2007, p. 1110). They allow “individuals [to] “construct” themselves as “characters” whose attributes can be revealed and communicated[, shaping] [...] not only how individuals view themselves, but also how others view them” (Martens et al. 2007, p. 1110). Scholars have studied how narratives allow new ventures to secure more funding by providing a clearer depiction of who they are as an organization (Lounsbury and Glynn 2001, Martens et al. 2007, Reinecke et al. 2012, Santos and Eisenhardt 2004, Wry et al. 2013). The assumption is that a description that is clearly anchored in a category will use the adequate terminology and writing patterns. Audience members should find it easier to identify whether a project’s narrative corresponds to a given category if this description is representative of that category. Consequently, this label limited context reduces labels’ conventional strength and needs to be combined with producer narratives in order to indicate categorical membership, underlying expectations, (Durand and Jourdan 2012, Durand and Paoletta 2013, Zuckerman 1999, 2015) and quality of the offering (Negro et al. 2015).

We need to note that some research has recently recognized that producers of the same category must show individuation through differentiation by being relatively different from others within their category (Brewer 1991, Durand and Jourdan 2012, Kovács and Hannan 2015, Überbacher 2014). However, in contexts where labels no longer play the same signaling role, expressing one’s differentiation will only exacerbate the already difficult evaluation process audiences have to undergo. This is why more representative narratives should increase a producer’s chances of success as it offers a communication anchor that audiences will find easier to interpret than labels alone.

To summarize, labels used in these contexts void of clear category labels are not sufficient in serving their initial intended purpose of indicating appropriate taken-for-granted expectations of a category. Rather, they must be combined with an audiences’ more thorough analysis of a project leading us to suggest that audiences derive producer typicality by observing the use of latent labels present in coherent communication narratives. We hypothesized that category labels are not sufficient

in environments void of clear category labels and coherent communication narratives are what permit audiences to evaluate the quality and typicality of a given producer. Said differently, producer typicality only resides in the vision that audiences of a category have of it. We endogenize the construct and hypothesize that producers use coherent communication narratives in contexts of no clear category labels to allow audiences to evaluate their typicality. Therefore:

***H2:** In an environment void of clear category labels, coherent communication narratives have a positive impact on audiences' judgment of producer typicality*

Consistent with categorization theory, we have hypothesized that funders who are interpreted as typical are more likely to be successful because they meet the expectations of a given category (hypothesis one). In contexts where category labels are not clear, evaluation of typicality no longer resides in the sole acceptance of category labels. Rather, latent category labels are weak signals that audience members use as an anchor to undergo more thorough evaluation of a producers' representativeness within a category (hypothesis three). Given the importance that coherent communication narratives have in such contexts for audience's ability to evaluate producer representatives, we propose that audiences' judgment of producer typicality moderates the impact of coherent communication narratives on success.

***H3:** In an environment void of clear category labels, coherent communication narratives' impact on success is moderated by audiences' judgment of producer typicality*

4 The crowdfunding context

4.1 Definition of crowdfunding

Lambert and Schwienbacher (2010, p.6) define crowdfunding as “an open call, essentially through the Internet, for the provision of financial resources either in the form of donation or in exchange for some form of reward and/or voting rights in order to support initiatives for specific purposes”. Hemer (2011) complements Lambert and Schwienbacher (2010) definition in a quasi-identical fashion.

Although it has grown in popularity over the past 5-10 years, as witnessed by the increased media coverage it has received, crowdfunding is actually an old phenomenon. Classic examples of pre-Internet crowdfunding are numerous; such as Pulitzer’s crowdfunding campaign (CFC) to get New Yorkers to donate money to fund the Statue of Liberty’s pedestal base or even Mozart’s pre-selling of tickets to finance concerts (Hemer 2011, Niederer 2014).

Crowdfunding has been revitalized, has seen substantial growth, and has become increasingly popular in the advent of Web 2.0, the Internet, and “viral networking and marketing” which has allowed mobilizing a “large number of users in specific Web communities within a relatively short period of time” (Hemer, 2011, p.8). In addition, Schwienbacher & Larralde (2010) explain that the three main characteristics of Web 2.0 provide the technological ability and a “participatory” Internet that moves beyond the previous “passive” one in which users were merely just observers. These main characteristics are: 1) collaboration permitting combination of knowledge and resources, 2) openness permitting easy multiple contributions, and 3) ease of access increasing participation (Lee et al., 2008).

4.2 Types of Crowdfunding

Several types of crowdfunding exist and have emerged over time. Ranked in order of simplicity both regarding the level of legal complexity, and asymmetry of information, these are: donation-, reward-, lending-, and equity-based crowdfunding (Ahlers et al. 2015, Hemer 2011, Schweinbacher and Larralde 2010). Our study focuses on reward crowdfunding

Other studies have given a wider typology of crowdfunding types (adding for example equity-like crowdfunding) while the crowdfunding industry report shows the recent emergence of royalty crowdfunding as well as many mixed models (Crowdsourcing.org 2012, Kevin Berg Kartaszewicz-Grell 2013). In this study, we will only focus on the former four named crowdfunding models. In this type of crowdfunding, the creator provides a reward or gift to the funder in exchange for their contribution. This can be a T-shirt, or a cap, or even an early version of the product in preordering schemes (e.g.: Kickstarter or Indiegogo – the platform from which we have acquired data). Reward crowdfunding has increasingly been seen as a pre-sales crowdfunding platform.

4.3 Prior research in crowdfunding

Recent research attempting to explain the determinants of success in crowdfunding has concentrated on the idea of asymmetry of information that exists between founders and backers and has used signaling theory as a basis. Consistent with this theory, apparent quality of a project seems to be an important signal that funders observe to lower their asymmetry of information with regards proposed projects. Interestingly, what is considered to be a signal of quality by funders varies depending on the type of CFP. It would appear that funders seem to use many signals emitted by creators and projects to infer the quality of a project in order to try and circumvent the aforementioned asymmetry of information with varying degrees of success.

For instance, in peer-to-peer lending, group leaders serve as signals for other investors to identify quality investments because the latter trust that the former perform the appropriate screening of the investments they are pursuing (Hildebrand et al., 2013). On the other hand, in equity CFPs, Ahlers and Cumming (2012) attempt to shed light on this exact issue; “given different start-ups with similar observable characteristics, what leads small investors to invest in certain start-ups and not in others?” (Ahlers et al. 2015, p. 2). The authors claim that although asymmetry of information in crowdfunding is quite important, “small investors have seemingly been able to infer the quality of start-ups by interpreting the information provided on the platform” (Ahlers & Cumming, 2012, p. 18). In staying close to signals of quality traditionally used in entrepreneurial finance, the

authors conclude that the main signals funders use to determine the quality of a project are the number of board members, the entrepreneur's education, the network quality, the envisaged exit type (IPO or trade sale), the provision of financial forecast, and the business age. We focus hereafter on signals used in reward crowdfunding platform.

Mollick (2014) associates preparedness of a project holder as an overall signal of quality and shows its positive impact on project success. He measures this preparedness by looking at three different criteria: whether projects had a video, the speed of updates (i.e.: activity creators had on the platform), and spelling errors. Second, the organizational form of projects may be an indicator of quality. Indeed, observing that projects that are tagged as not-for-profit on CFPS are more successful, they suggest that these project holders may be "more prone to commit to high quality products or services" (Schweinbacher and Larralde 2010, p. 10) because they put "less emphasis on profit making, they may therefore focus more on quality, which may be an important requirement for attracting donations" (Schweinbacher and Larralde 2010, p. 10). Third, social capital has been identified as a determinant of project success on several different types of CFPs (Agrawal, et al., 2011; Colombo, Franzoni, & Rossi-Lamastra, 2014; Freedman & Jin, 2014; Mollick, 2014; Zvilichovsky et al., 2013).

Although many scholars have already discussed the importance of signaling quality for project success, we propose to observe how categorization and psychological traits influence crowdfunding success. Next section explores our data collection procedure.

5 Data Collection:

The data collected are from the crowdfunding platform (CFP) Indiegogo. The platform was launched in 2008 and is now the second largest reward CFP in the U.S., after Kickstarter. Projects are divided into 22 categories such as music, dance, small business, film, technology, and animals and accepts projects from all around the world, written in different languages, and asking for one of 5 currencies (USD, EU, GBP, AUD, CAD). The platform is available in 4 languages and accepts projects from verified non-for-profits, individuals, groups, and companies to name a few. Projects normally last up to 60 or 120 days depending on the type of financing that has been chosen: All-or-Nothing (AON) vs. Keep-it-All (KIA). Project holders create their project and explain what perks (gifts or products), if any, they are willing to give for the different funding levels they propose.

Each project is made up of 4 pages: *Story*, *Updates*, *Comments*, and *Funders* and sometimes a fifth one named *Gallery*. The *Story* page is similar to the homepage of any website and is where most information about a project is found (Annex 1.1). A project holder will provide information such as the title of their campaign; the category; the city & country; any external links that they want to provide such as social networks (Facebook, Twitter, YouTube, and LinkedIn), personal websites or blogs, or even community specific websites (like IMDB for movies); the project's amount goal & percentage completion; time left; and perks that they propose; and project leaders' Indiegogo profile links.

The *Updates* page allows project holders to discuss relevant updates to their project during or even after a campaign (Annex 1.2). The *Comments* page provides one with access to the amount of comments, the comment text, the comment author, and their link (Annex 1.3). This is the same for the *Funders* page but with the amount that they have decided to fund (Annex 1.4). It is important to note that the amount pledged as well as the person name can be anonymous, and does not necessarily have a link as explained above). Both *Comments* and *Funders* pages are sequentially ordered in a reverse chronological manner (i.e.: last amount funded appearing first).

5.1 Data Collection Timeframe

The data were collected by means of a web crawler from December 2014 to February 2015 from the Indiegogo platform. We collected both ongoing and finished projects and have all the projects starting in 2009 and until the beginning of 2015. We have access to all the information on the story, funder, and comment pages.

5.2 Empirical Setting

We gather information on more than 300,000 finished and active Indiegogo projects but test our hypotheses by focusing on the 14,000 projects from the *Music* category. We clean the data further by removing any projects that were created but not edited. Producers may create a project on Indiegogo but not edit any of the words present on the template provided by the platform. We consider that these projects were created with no real intent of the producer to go through with her idea explaining why we drop these observations. We also remove observations that have less than 150 words in them. Indeed, past research papers in management that have analyzed narratives have looked at diverse documents such as annual reports (ADD), patent abstracts (ADD), or entrepreneurial pitches (ADD). The shortest ones are patent abstracts that are at most 150 words for US patents. Taking this into consideration, we believe that expressing one's project goal and categorical belonging would be quite hard beneath a threshold and decide to drop projects that use less than 150 words in their description leaving us with around 8,000 projects. To test our hypotheses linked to funders being perceived as typical by audience members (hypotheses 1 and 3), we focus our attention on projects for which we have at least one funder. This means that we remove any projects that did not receive any funding. Furthermore, our methodology calls on identifying the level of typicality of a project deriving it from the level of typicality of their funders; that is to say what kind of funders does a producer attract. We claim that the more typical this pool of financiers, the more typical a project is. This finally leaves us with over 2,800 observations to test our hypotheses¹.

In order to test our hypotheses, we focus on over 2,800 projects from the *Music* category. The reason we do this is twofold. First, research has been done on

¹ We describe in greater detail in our methods section how we create this variable.

categorization in music (Montauti and Wezel, 2014) giving us more confidence to use this tested context. Second, and in part because of our first point, using the context of music allows us to easily produce independently generated music subcategories (i.e.: genres) that we can then use to assign projects to. This allows limiting endogeneity problems when creating the dictionary. We develop on the way we build our musical genre dictionary in the following sections.

To create our constructs, we analyze the descriptions of music projects through a vector space modeling (VSM) approaches called Doc2Vec. The next section explains what these processes are and what data is extracted. In a last step, the data is used to run our econometric tests in Stata.

5.3 Vector Space Modeling

5.3.1 Introduction

Using the data collected from our crawler, we use a vector space modeling (VSM) approach called document to vector (Doc2Vec) on the description of over 2,800 Music projects. We focus on music projects in order to test our hypotheses concerning category belonging and evolution.

VSM is seen as robust in Information Retrieval Theory (Manning et al., 2009) and is the backbone of modern-day search engines (Turney and Pantel 2010). Computer science research labs in machine learning and technology startups in the field of real-time information retrieval and sentiment analysis from social media websites currently use the Doc2Vec approach. To implement these VSM, we use what is commonly referred to as machine or deep learning where the aim is to teach the computer to recognize the cooccurrence of words by going over a corpus, derive their likelihood of occurrence and, in the case of Doc2Vec, determine their meaning. The following sections will explain the process in more detail and the data that were derived from it.

5.3.2 Process

Before explaining the mechanisms of Doc2Vec, it is important to understand its origins. Vector space modeling represents each word used within a compilation of texts in a high-dimensional space. The natural language processing environment

calls this compilation of texts a “corpus”. In our study, this corpus consists of all the music project descriptions from Indiegogo. In traditional VSM like Term Frequency-Inverse Document Frequency (TF-IDF) or Latent Dirichlet Allocation (LDA), the dimensions of a word that has been vectorized correspond to all of the words that are in the corpus.

This means that the machine will observe the cooccurrence of each word through the use of the “bag-of-words” natural language processing technique (Turney and Pantel 2010) and convert words to their corresponding vector space dimension. More precisely, VSM would add all the words that are in our corpus of 2,800 projects to a dictionary and give the cooccurrence matrix of all the words of the matrix with one another (cooccurrence value between -1 and 1). The method computes the maximum average log probabilities of each pair of co-occurring words. For example, the word “jazz” would be vectorized in a dimension equal to the amount of words that exist in the created dictionary. Although this has been used quite a lot in the past (Turney and Pantel 2010), the traditional approach requires large computer processing power and does not necessarily render the most appropriate results. Indeed, vectorizing a word with respect to all the words that occur in a corpus may produce cooccurrences that are not logically significant. For instance, the word “jazz” may occur frequently with the word “computer” but this cooccurrence does not teach us much. Doc2Vec attempts to solve both of these issues and offers added benefits.

In Doc2Vec, the “bag-of-words” approach is also used but it is incorporated within the framework. Contrary to the traditional VSM approaches that run cooccurrence measures of all words with each other, Doc2Vec only compares cooccurrence of words with the words that precede (T-1, 2, 3...) and follow (T+1, 2, 3...) them in all their appearances across the corpus. Thus, the computation is no longer based on all words by rather it becomes based on a window of words (or training context) preceding and trailing the central word on which one observes the maximum average log probabilities of each pair of co-occurring words.

$$\frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j} | w_t)$$

Where T is the amount of training words, w is the word, and c is the size of the

training context which we here choose to be a function of the centered word w_t , i.e. the window previously discussed. For more information on the method refer to (Mikolov et al. 2000).

The advantage is to observe the word in a context and see, for example, all the words that directly precede or follow the word “jazz”. In Doc2Vec, this context is a user modifiable window. The user then also decides of the dimensionality that they want to give each word, though most studies agree to use above 100 dimensions to have consistent results but no more than 300 because cooccurrences are only marginally improved after this threshold (Rehurek 2011). For these reasons, we chose a dimensionality of 300 and a window of 10 (Rehurek 2011, Pennington et al.). Thus the machine looks for all the cooccurrences of word T with words $T-1/10$ and $T+1/10$ until reaching at most 300 words.

Doc2Vec greatly increases computational speed, allows for a higher context-specificity, and provide other interesting benefits. Using Doc2Vec also allows us to 1) capture proximity of words with words like any other traditional VSM, but also allows to 2) vectorize and capture proximity of whole sentences (i.e.: crowdfunding project descriptions) with other sentences, and most interestingly for us 3) compare whole sentences with specific words. For example, we are able to see the proximity of the word “jazz” with all music projects presented on Idiegogo. These improvements over traditional VSM approaches explain why computer science research labs and startups in the field of natural language processing have adopted Doc2Vec and by extension why we chose to use it.

6 Methods

6.1 Dependent variable

To be consistent with prior research in crowdfunding, we use a dummy variable as the dependent variable to measure success of a venture on the crowdfunding platform. We define *success* of a crowdfunded project as 1 if it has achieved at least 100% of its funding goal, 0 otherwise. This measure has been used in several other studies working on crowdfunding platforms (Belleflamme et al. 2015, Mollick 2014). We do realize that Indiegogo is particular because of three different funding decisions (Flexible Funding, Fixed Funding, and InDemand) may influence this variable and we include it as a control variable (discussed later).

6.2 Independent variables

Coherent communication narratives (Coherent Narrative). To measure how the overall degree of typicality impacts project success, we use a Herfindhal index that measures the average concentration that a project description has across all latent categorical labels. We do so by first measuring the average semantic proximity that a project has to a given music category through our machine learning approach. We then measure the weighted normalized average proximity that the project has across all the labels it uses in its description resulting in our coherent narrative measure.

Audience perception of producer typicality (Typicality). To measure how audiences perceive producer typicality, we first observe the distribution of project funders' investments on the crowdfunding platform by creating a Herfindahl index for each funder. We essentially observe how concentrated each funder is in a specific musical genre. Then and by project, we sum these distributions and divide the result by the amount of funders in that project. This measure is used both as an independent variable (hypotheses 1, 2, and 4) and a dependent one (hypothesis 2)

6.3 Control variables

We control for various variables that have been used in previous studies as standard controls (Belleflamme et al. 2015, Mollick 2014) as well as add some related to our research². These are:

² Given our dependent variable "coherent narratives", we do not include all of the controls used in

Project Financing (type of financing). We control for the type of financing that the project decides to use with the dummy variable *type of financing*. Firstly, we drop the financing decision InDemand because this concerns projects that have already been successful but are still asking for funds after the end of the campaign. This category was recently introduced (January 2015) for projects that were successful and do not serve our purpose here and could potentially bias our results. We thus drop 39 observations. We then control for the two remaining financing types: Flexible Funding and Fixed Funding. In the former, the project holders will get any amount of money collected regardless of whether they reach the monetary goal that they have set. In the latter, project holders will only receive the funding if they reach or exceed the monetary goal set at the beginning of their campaign.

Project Duration (Project Duration). The length of time (days) a project ran for. Indiegogo recommends that project holders to have projects running at least 30 days but no longer than 60 days.

Amount of activity (amount of activity). The amount of times (integer) the project holder updated their project during the campaign. More active projects are more likely to succeed.

Amount of comments (Amount of comments). The amount of comments (integer) funders gave on the project's dedicated thumbnail. More comments may show that the project is receiving attention.

Pictures posted (pictures posted). A dummy variable that reports whether a project provides photos in the dedicated photo thumbnail.

Use of videos (use of videos). A dummy variable that reports whether a project provides a video. Having a video is highly encouraged by Indiegogo and has been showed to have a strong positive impact on project success.

crowdfunding research because the ones relating to project description (description length and amount of latent genre labels) are bad controls or “variables that are themselves outcome variables in the notional experiment at hand” (Angrist and Pischke 2008, p. 47).

Goal amount (goal amount). The log dollar goal set by the founder (integer).

Perks exist (perks exist). A dummy variable reporting whether a project provides rewards to funders. Not providing any rewards may deter funders from financing a project.

Sponsored (sponsored). A dummy variable reporting whether a project has a company or association as a supporter. The partnership is clearly mentioned on the project's page with a link to the partner's profile. On the sponsor's page, one can see the projects that this partner has helped funded or sponsored. It is fair to suggest that this may influence funders' willingness to finance.

6.4 Model

Due to an overrepresentation of zeros in our dependent variable *success* we use a tobit regression to test our hypotheses. We also use a probit regression as a robustness test given that most of the research done in crowdfunding uses this approach.

7 Results

Table 1 presents the correlation results of our dependent variables (*success*) independent variables (*typicality and coherent narratives*), and control variables (*type of financing, project duration, activity level, number of comments, posted photos, posted video, goal, rewards, sponsors, and whether the project is a verified non-profit*).

[Insert Table 1]

We observe that most variables are positively correlated to our dependent variable except for *goal amount* and *duration* that are negatively correlated. This is consistent with prior research (Belleflamme et al. 2015). We do however notice that our constructs are not correlated with success. Correlations are low with most of our controls; number of *comments* and *type of financing* are the most correlated to success with 0.326 and 0.246 respectively. Overall, the correlations that we observe are consistent with what we would expect from the literature.

[Insert Table 2]

Table 2 present the results of our first hypothesis in which we expect that typicality of a project (indirectly observed by audiences' judgment of producer typicality) should have a positive impact on resource procurement. We first run the model with our controls as a baseline. We then test our first hypothesis by adding typicality and as expected, it significantly and positively predicts access to funds. Level of typicality of a project has a positive impact on its ability to procure funds due to an audience's capability to understand how a producer's offering is representative of a given category. At the core of our paper, our aim is to show how audiences derive producer typicality in contexts of unclear category labeling by evaluating how strategically crafted narratives influence their judgment. We find that coherent narratives have a positive and significant impact on an audience's perception of producer typicality as seen in Table 3. This provides support to our second hypothesis.

In the last model of table 2, we observe that coherent narratives positively influence typicality that in turn influences resource procurement leading us to support our claim that typicality moderates the impact of coherent narratives on success, providing support to our third hypothesis.

[Insert Table 3]

8 Discussion

New ventures have a hard time securing financing. Entrepreneurship scholars have explored this liability of newness and have found that new ventures often have difficulty showing the quality of their venture (Stinchcombe 1965). Crowdfunding projects are also victim of the same situation and one could argue that it is actually accentuated by the fact that entrepreneurs and financiers do not directly communicate to each other and are potentially scattered all across the globe. Past research in crowdfunding has attempted to explain the determinants of success of projects presented on these platforms (Agrawal et al. 2011, Colombo et al. 2014, Freedman and Jin 2014, Mollick 2014, Zvilichovsky et al. 2013). We decide to take a novel approach, slightly deviating from signaling by observing how categorization, as expressed through project narratives, influence crowdfunding success.

The paper is not immune to certain limitations. The music context may be unique and the results may not be replicable in others. We call on others to see whether coherent narratives can be used as strategic tools in other contexts.

This paper contributes both to theory and practice. First, we contribute to the entrepreneurship literature by showing that in high uncertainty environments, producers can use categorization and narratives as a strategy to gain access to more funds. Second, we contribute to the categorization literature by providing a better understanding to what occurs in situations where categories are not clearly appointed by third party actors. Third, our results contribute to the growing field of crowdfunding by providing alternative constructs determining crowdfunding success. Fourth, we advance the field of management studies by using novel textual analysis that goes beyond the traditional word count and spelling mistake.

Furthermore, our vector space modeling approach allows to contextualize words as opposed to taking any co-occurrence as a sign of being related. Adopting these methods from the computer science research field that are used in data mining and computer deep learning makes us confident as to the validity and value of the approach that we take. We believe that much research could benefit from such alternative methods.

Finally, we provide interesting insights to entrepreneurs who want to finance their projects. Indeed, they should first pay particular attention to clearly explain their *raison d'être* by being very precise in using the vocabulary of the category they belong to. They should also show how they are different by using latent labels of categories their project connects to. Finally, if they are very typical, they should show some intuitive thinking in order to increase the differentiation of their project allowing them to signal originality.

9 Tables and Appendixes

Descriptive Statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
success	2,880	0.261	0.440	0	1
typicality	2,880	0.457	0.468	0	1
narrative	2,880	0.239	0.105	0	1
perks exist (1/0)	2,880	0.974	0.160	0	1
use of video (1/0)	2,880	0.793	0.405	0	1
type of financing (1/0)	2,880	0.0219	0.146	0	1
sponsored (1/0)	2,880	0.0212	0.144	0	1
goal amount (log)	2,880	8.339	1.113	6.215	13.82
pictures posted (1/0)	2,880	0.661	0.474	0	1
amount of activity (log)	2,880	0.798	0.961	0	4.454
amount of comments (log)	2,880	2.168	1.046	0	6.565
project duration (days)	2,880	42.48	15.59	14	87

Table 1 – Correlations

	success	typicality	narrative	perks exist (1/0)	use of video (1/0)	type of finan	sponsored (1/0)	goal amount (log	pictures poste	amount of activi	amount of comments (project duration (
success	1											
typicality	-0.0228	1										
narrative	-0.00313	0.0334	1									
perks exist (1/0)	0.0437*	0.0473*	0.0231	1								
use of video (1/0)	0.184***	0.0412*	-0.0357	0.0799***	1							
type of financing (1/0)	0.0509**	0.0242	-0.0236	0.108***	0.142***	1						
sponsored (1/0)	0.246***	-0.0123	0.0149	0.0246	0.0512**	-0.0172	1					
goal amount (log)	0.00577	0.0338	0.0640***	0.0242	0.0242	0.00384	-0.0220	1				
pictures posted (1/0)	-0.221***	0.126***	-0.0169	0.0591**	0.182***	0.192***	-0.0309	0.0825***	1			
amount of activity (log)	0.00408	0.0643***	-0.00833	0.0788***	0.281***	0.178***	-0.0332	0.0290	0.152***	1		
amount of comments (0.326***	0.0492**	0.0180	0.0696***	0.412***	0.160***	0.0696***	0.0431*	0.263***	0.188***	1	
project duration (days)	-0.117***	0.0455*	-0.0125	0.00534	0.0552**	-0.00451	-0.0436*	0.000587	0.173***	0.0241	0.00334	1

* p<0.05 ** p<0.01 *** p<0.001

Table 2 – Tobit regression with success as an outcome variable

VARIABLES	(1) model	(2) sigma	(3) model	(4) sigma	(5) model
typicality			0.06** (0.02)		0.06** (0.02)
narrative					-0.03 (0.03)
perks exist (1/0)	0.06 (0.05)		0.05 (0.05)		0.05 (0.05)
amount of activity (log)	0.08*** (0.02)		0.08*** (0.02)		0.08*** (0.02)
use of video (1/0)	0.12*** (0.03)		0.12*** (0.03)		0.12*** (0.03)
type of financing (1/0)	0.20*** (0.01)		0.20*** (0.01)		0.20*** (0.01)
sponsored (1/0)	0.02 (0.02)		0.02 (0.02)		0.02 (0.02)
goal amount (log)	-0.87*** (0.03)		-0.88*** (0.03)		-0.88*** (0.03)
pictures posted (1/0)	-0.09*** (0.03)		-0.10*** (0.03)		-0.10*** (0.03)
amount of comments (log)	0.89*** (0.03)		0.90*** (0.03)		0.90*** (0.03)
project duration (days)	-0.07** (0.04)		-0.08** (0.04)		-0.08** (0.04)
Constant	-1.29*** (0.08)	0.94*** (0.02)	-1.29*** (0.08)	0.93*** (0.02)	-1.29*** (0.08)
Observations	2,880	2,880	2,880	2,880	2,880

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3 – Tobit Regression with perception of typicality as the outcome variable

VARIABLES	(1) model	(2) sigma	(3) model	(4) sigma
narrative			0.04* (0.02)	
perks exist (1/0)	0.07* (0.04)		0.07* (0.04)	
amount of activity (log)	0.01 (0.02)		0.01 (0.02)	
use of video (1/0)	-0.00 (0.02)		-0.00 (0.02)	
type of financing (1/0)	-0.01 (0.02)		-0.01 (0.02)	
sponsored (1/0)	0.02 (0.02)		0.02 (0.02)	
goal amount (log)	0.15*** (0.02)		0.15*** (0.02)	
pictures posted (1/0)	0.04* (0.02)		0.04* (0.02)	
amount of comments (log)	-0.02 (0.02)		-0.02 (0.02)	
project duration (days)	0.03 (0.03)		0.03 (0.03)	
Constant	0.05 (0.05)	0.95*** (0.01)	0.06 (0.05)	0.95*** (0.01)
Observations	2,880	2,880	2,880	2,880

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

10 References

- Agrawal, A., C. Catalini, A. Goldfarb. 2011. *The Geography of Crowdfunding*. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0> [Accessed September 8, 2014].
- Agrawal, A., C. Catalini, A. Goldfarb. 2013. Some simple economics of crowdfunding. *NBER Work. Pap. No. 19133* **14**(April) 1–47. Available at: <http://www.nber.org/papers/w19133> [Accessed September 8, 2014].
- Ahlers, G. K. C., D. Cumming, C. Günther, D. Schweizer. 2015. Signaling in equity crowdfunding. *Entrep. Theory Pract.* **39**(4) 955–980. Available at: <http://www.ssrn.com/abstract=2161587> [Accessed September 8, 2014].
- Allison, T. H., A. F. McKenny, J. C. Short. 2013. The effect of entrepreneurial rhetoric on microlending investment: An examination of the warm-glow effect. *J. Bus. Ventur.* **28**(6) 690–707.
- Angrist, J. D., J. Pischke. 2008. Mostly Harmless Econometrics: An Empiricist's Companion. *Massachusetts Institute Technol. London Sch. Econ.* (March) 290.
- Barry, D., M. Elmes. 1997. Strategy retold: Toward a narrative view of strategic discourse. *Acad. Manag. Rev.* **22**(2) 429–452.
- Belleflamme, P., N. Omrani, M. Peitz. 2015. The economics of crowdfunding platforms. *Inf. Econ. Policy* **33** 11–28.
- Brewer, M. B. 1991. The social self: On being the same and different at the same time. *Personal. Soc. Psychol. Bull.* **17**(5) 475. Available at: <http://proquest.umi.com/pqdweb/?did=3796333&Fmt=7&clientId=46781&RQT=309&VName=PQD>.
- Brush, C. G., T. S. Manolova, L. F. Edelman. 2008. Separated by a common language? Entrepreneurship research across the atlantic. *Entrep. Theory Pract.* **32**(2) 249–266.
- Cardon, M. S., M. Der Foo, D. Shepherd, J. Wiklund. 2012. Exploring the Heart: Entrepreneurial emotion is a hot topic. *Entrep. Theory Pract.* **36**(1) 1–10.
- Cardon, M. S., J. Wincent, J. Singh, M. Drnovsek. 2009. The nature and experience of entrepreneurial passion. *Acad. Manag. Rev.* **34**(3) 511–532.
- Colombo, M. G., C. Franzoni, C. Rossi-Lamastra. 2014. Internal social capital and the attraction of early contributions in crowdfunding. *Entrep. Theory Pract.* 75–101.
- Cornelissen, J. P., J. S. Clarke. 2010. Imagining And Rationalizing Opportunities: Inductive Reasoning And The Creation And Justification Of New Ventures. *Acad. Manag. Rev.* **35**(4) 539–557. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=53502700&site=bsi-live>.
- Crowdsourcing.org. 2012. Crowdfunding industry report - market trends, composition and crowdfunding platforms. *Res. Rep.* (May) 1–30. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Crowdfunding+industry+report#2> [Accessed September 8, 2014].
- Durand, R., J. Jourdan. 2012. Jules or Jim: Alternative conformity to minority logics. *Acad. Manag. J.* **55**(6) 1295–1315.
- Durand, R., L. Paoletta. 2013. Category stretching: Reorienting research on categories in strategy, entrepreneurship, and organization theory. *J. Manag. Stud.* **50**(6) 1100–1123.
- Freedman, S., G. Z. Jin. 2014. The Information value of online social networks: lessons from peer-to-peer lending. *NBER Work. Pap.* Available <http://www.nber.org/papers/w19820> 1–53. Available at: <http://www.nber.org/papers/w19820>.
- Hannan, M. T., L. Pólos, G. R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes, and Ecologies*.
- Hemer, J. 2011. A snapshot on crowdfunding. *Enconstor* 39. Available at: <http://www.econstor.eu/handle/10419/52302> [Accessed September 13, 2014].
- Hildebrand, T., M. Puri, J. Rocholl. 2013. Adverse incentives in crowdfunding. Available *SSRN 1615483* (April) 43. Available at: <http://papers.ssrn.com/abstract=1615483> \n[http://www.esmt.org/sites/default/files/digital-measures/Adverse incentives in crowdfunding-1.pdf](http://www.esmt.org/sites/default/files/digital-measures/Adverse%20incentives%20in%20crowdfunding-1.pdf) [Accessed September 13, 2014].
- Hsu, D. H. 2007. Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Res. Policy* **36**(5) 722–741. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0048733307000716> [Accessed September 2, 2014].
- Hsu, G., M. T. Hannan. 2011. infKL and Organizational. *Organization* **16**(5) 474–490.
- Kevin Berg Kartaszewicz-Grell. 2013. 2013: Crowdfunding Industry Report. 96. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:No+Title#0> [Accessed September 8, 2014].

- Kovács, B., M. T. Hannan. 2015. The Space of Categories and the Consequences of Category Spanning. *Sociol. Sci.* **2** 252–286.
- Kovács, B., R. Johnson. 2014. Contrasting alternative explanations for the consequences of category spanning: A study of restaurant reviews and menus in San Francisco. *Strateg. Organ.* **12**(1) 7–37. Available at: <http://soq.sagepub.com/cgi/content/abstract/12/1/7>.
- Lambert, T., A. Schwienbacher. 2010. An Empirical Analysis of Crowdfunding. *SSRN Electron. J.* 1–23. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:An+Empirical+Analysis+of+Crowdfunding#0> [Accessed September 13, 2014].
- Lee, S. H., D. DeWester, S. R. Park. 2008. Web 2.0 and opportunities for small businesses. *Serv. Bus.* **2**(4) 335–345.
- Lounsbury, M., M. A. Glynn. 2001. Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strateg. Manag. J.* **22**(6–7) 545–564. Available at: <http://doi.wiley.com/10.1002/smj.188>.
- Martens, M. L., J. E. Jennings, P. D. Jennings. 2007. Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition. *Acad. Manag. J.* **50**(5) 1107–1132.
- Mikolov, T., K. Chen, G. Corrado, J. Dean. 2000. 10.1162/Jmlr.2003.3.4-5.951. *CrossRef List. Deleted DOIs* **1** 1–9. Available at: http://www.crossref.org/deleted_DOI.html.
- Milgrom, P., J. Roberts. 1992. Moral hazard and performance incentives. *Econ. Organ. Manag.* 166–203.
- Mollick, E. 2014. The dynamics of crowdfunding: An exploratory study. *J. Bus. Ventur.* **29**(1) 1–16. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S088390261300058X> [Accessed July 9, 2014].
- MONTAUTI, M., F. C. WEZEL. 2014. Blurred Maps and Sirens' Calls: Category Recombination and Entrepreneurial Inertia in the Market for Electronic Music, 1978-2011. *Acad. Manag. Annu. Meet. Proc.* 56–60. Available at: [10.5465/AMBPP.2014.163\http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=97634007&lang=zh-tw&site=ehost-live](http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=97634007&lang=zh-tw&site=ehost-live).
- Negro, G., M. T. Hannan, M. Fassiotto. 2015. Category Signaling and Reputation. *Organ. Sci. Publ.* **26**(2) 584–600.
- Niederer, P. 2014. *How Small Businesses Can ! Embrace Crowdfunding ! for Success !*
- O'Connor, E. 2002. Storied Business: Typology, Intertextuality, and Traffic in Entrepreneurial Narrative. *J. Bus. Commun.* **39**(1) 36–54.
- Palka, D. L., M. C. Rossman, A. S. VanAtten, C. D. Orphanides. 2008. *Effect of pingers on harbour porpoise (Phocoena phocoena) bycatch in the US Northeast gillnet fishery.*
- Puranam, P., B. C. Powell, H. Singh. 2006. Due diligence failure as a signal detection problem. *Strateg. Organ.* **4**(4) 319–348.
- Rehurek, R. 2011. Subspace Tracking for Latent Semantic Analysis. *Proc. Eur. Conf. Inf. Retr.* 289–300. Available at: http://link.springer.com/chapter/10.1007/978-3-642-20161-5_29
- Reinecke, J., S. Manning, O. von Hagen. 2012. The Emergence of a Standards Market: Multiplicity of Sustainability Standards in the Global Coffee Industry. *Organ. Stud.* **33**(5–6) 791–814. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84873859393&partnerID=40&md5=181423928d4cf39f72cc4f74172be6d9>.
- Reschke, B. 2015. Electronic Thesis and Dissertations by.
- Santos, F., K. Eisenhardt. 2004. Constructing markets and organizing boundaries: entrepreneurial action in nascent fields. *Acad. Manag. Proc.* J1–J6. Available at: <http://search.epnet.com/login.aspx?direct=true&db=bth&an=13863681>.
- Schweinbacher, A., B. Larralde. 2010. *Crowdfunding of Small Entrepreneurial Ventures.* Available at: <http://ssrn.com/abstract=1699183>.
- Short, J. C., J. C. Broberg, C. C. Coglisier, K. H. Brigham. 2010. Construct validation using computer-aided text analysis (CATA): An illustration using entrepreneurial orientation. *Organ. Res. Methods* **13**(2) 320–347.
- Stinchcombe, A. L. 1965. Social structure and organizations. *Econ. Meets Sociol. Strateg. Manag.* 229–259. Available at: (null):<Go to ISI>://000178666500009.
- Turney, P. D., P. Pantel. 2010. From frequency to meaning: Vector space models of semantics. *J. Artif. Intell. Res.* **37** 141–188.
- Überbacher, F. 2014. Legitimation of New Ventures: A Review and Research Programme. *J. Manag. Stud.* **51**(4) 667–698.
- Vergne, J. P., T. Wry. 2014. Categorizing categorization research: Review, integration, and future directions. *J. Manag. Stud.* **51**(1) 56–94. Available at:

- <http://doi.wiley.com/10.1111/joms.12044>.
- Wry, T., M. Lounsbury. 2013. Contextualizing the categorical imperative: Category linkages, technology focus, and resource acquisition in nanotechnology entrepreneurship. *J. Bus. Ventur.* **28**(1) 117–133. Available at: <http://dx.doi.org/10.1016/j.jbusvent.2012.03.001>.
- Wry, T., M. Lounsbury, M. A. Glynn. 2013. Legitimizing Nascent Collective Identities: Coordinating Cultural Entrepreneurship. *Organ. Sci.* **22**(3) 226–245. Available at: <Go to ISI>://WOS:000314086500013\n<http://pubsonline.informs.org/doi/abs/10.1287/orsc.1110.0730>\n<http://pubsonline.informs.org/doi/abs/10.1287/orsc.1100.0613>\n<http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db>.
- Zider, B. 1998. How venture capital works. *Harv. Bus. Rev.* **76**(6) 131–139.
- Zuckerman, E. 1999. The Categorical Imperative: Securities Analysts and the Illegitimacy Discount. *Am. J. Sociol.* **104**(5) 1398–1438.
- Zuckerman, E. W. 2015. Optimal Distinctiveness Revisited: An Integrative Framework for Understanding the Balance between Differentiation and Conformity in Individual and Organizational Identities. *Handb. Organ. Identity*.
- Zvilichovsky, D., Y. Inbar, O. Barzilay. 2013. *Playing Both Sides of the Market: Success and Reciprocity on Crowdfunding Platforms*. Available at: <http://ssrn.com/abstract=2304101>.