Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

How consumption, communication and network position propel the transition from product user to entrepreneur: An empirical analysis of a technology platform for music apps

MortenM H.J. Fenger
Aarhus University/BSS
Business Administration
mhjfenger@badm.au.dk

Lars Frederiksen
Aarhus
Business Administration
l.frederiksen@asb.dk

Hans Jørn Juhl
Aarhus University/BSS
Business Administration
hjj@asb.dk

Joachim Scholderer
Aarhus University/BSS
Business Administration
sch@badm.au.dk

Abstract

Despite increasing interest in technological platforms as innovation ecosystems, existing research offers little predictive insight into how firms can identify users who are likely to become successful third-party providers. In our analysis, we combine (a) the complete sales history of all applications related to a technological platform and (b) the complete history of communications in all user and developer communities related to the platform. We employ logistic, extended Cox, and Poisson regression models, combined with social network analysis. We show that the prior network positions of users, their prior communication histories, and their prior user experiences from consumption can jointly predict the transition from user to third-party developer (i.e. indicating entrepreneurial intention), the launch of a first platform application (i.e. entrepreneurial action) and the sales of the application (i.e. entrepreneurial success). Our study contributes to the literature in strategic management on the dynamics of innovation on technological platforms, explicitly linking the production and consumption sides of the two-sided market. It also adds to the entrepreneurship literature by showing how the entrepreneurial process manifests itself in the context of technological platforms.

Jelcodes:O32,O31

How consumption, communication and network position propel users to become entrepreneurs: An empirical analysis of a technology platform for music apps

February, 2015

Total word count: 10,762

Abstract:

Despite increasing interest in technological platforms as innovation ecosystems, existing research offers little predictive insight into how firms can identify users who are likely to become successful third-party providers. In our analysis, we combine (a) the complete sales history of all applications related to a technological platform and (b) the complete history of communications in all user and developer communities related to the platform. We employ logistic, extended Cox, and Poisson regression models, combined with social network analysis. We show that the prior network positions of users, their prior communication histories, and their prior user experiences from consumption can jointly predict the transition from user to third-party developer (i.e. indicating entrepreneurial intention), the launch of a first platform application (i.e. entrepreneurial action) and the sales of the application (i.e. entrepreneurial success). Our study contributes to the literature in strategic management on the dynamics of innovation on technological platforms, explicitly linking the production and consumption sides of the two-sided market. It also adds to the entrepreneurship literature by showing how the entrepreneurial process manifests itself in the context of technological platforms.

INTRODUCTION

New and exciting ways of organizing innovation on technological platforms offer opportunities to pose different empirical and conceptual questions about how key processes work in such innovation ecosystems (Adner, 2012; Yoo, Boland, Lyytinen & Majchrzak, 2012). Today we witness how technological platforms dominate in an increasing number of businesses (Gawer & Cusumano, 2002; Adner & Kapoor, 2010). A platform is built upon an “ecosystem” owned by a company that provides the technological architecture that allows different types of users and complementary business partners to connect and benefit from the platform’s base functionality (Suarez, 2012). Unlike traditional value chains, where businesses interact with non-overlapping group of upstream suppliers and downstream buyers, platforms function as two-sided markets where third-party providers offer competing goods and services on one side of the market, and products compete to find end-users on the other side (Parker & Van Alstyne, 2005; Gawer, 2014). The two sides of the market are intrinsically connected and thus influence each other’s potential success and failure (Rochet & Tirole, 2003). For platform owners, this organization poses particular challenges in terms of the strategic management of innovation (Eisenmann, Park & van Alstyne, 2006).

Examples of technological platforms are many. Traditional ones include industrial platforms where modular products support economies of scale and scope by systematically re-using components within a product family, based on well-designed interfaces within a common sub-system or design architecture (Wheelright & Clack, 1992). Today, people are perhaps more familiar with the rising number of applications (apps) populating IT-based company-specific platforms. The number of different apps running on, for example, the Apple iOS and Google Android platforms is staggering, and over time specialization in type and quality is increasing. Consider also that the "app economy" has grown unbelievably fast[1]. Yet, despite the rapidly increasing economic significance of apps and platforms we have scant insights into how the underlying dynamics of innovation are working in such systems (Thomas, Autio & Gann, 2014), especially concerning the identification and involvement of third party providers (Piezunka, 2011).

Technology platforms allow for significant network effects and hence for reaping economies of scope of innovation (Katz & Shapiro, 1986). Third-party providers are crucial for platform-owning firms as they expand the available features and functionalities of the initial product portfolio (Boudreau, 2012) and generate opportunities for additional network effects for the proprietary platform owner (Mollick, 2012). And third party app developers can make it into big business themselves[2]. We define third-party providers as individuals or firms who, on behalf of the platform owner, develop applications targeted at end-users of the platform. Third-party providers are often affiliated with a platform owner via arm's-length contracts (Boudreau & Lakhani, 2009). The third-party provider is rarely compensated directly for the development

[1] The consulting house Asymco estimates that, in 2014, the rents generated by Apple iOS apps alone surpassed the rents generated in the US from all film production in Hollywood, resulting in revenues of over $10BN for app developers (http://www.asymco.com/2015/01/22/bigger-than-hollywood/). It is further estimated that the global revenues from apps were $86.3Bn in 2013, supporting the creation of almost one million jobs in the EU countries (Gigaom Research, 2014).

[2] Consider, for example, how the app "Angry Birds" expanded almost overnight into a major diversified business.

work; instead, the provider can enter a marketplace for applications that typically offers greater reach than would otherwise be available to the provider[3] (Ghazawneh & Henfridsson, 2013).

Relevant previous research has focused on entry strategies, platform quality, and consumer expectations of platform functions (Cennamo & Santalo, 2013; Zhu & Iansiti, 2012), control strategies for platform-owning companies (Boudreau, 2010), and competition between third-party developers (Boudreau & Jeppesen, 2014; Eisenmann, Parker & Van Alstyne, 2006). Key aspects of innovation and entrepreneurship on platforms are addressed less in this relatively infant research domain (Gawer, 2014). In order to generate successful platform strategies it is, however, imperative for platform owners to understand the entrepreneurial process on platforms much better: to identify and recruit third-party developers is the lifeblood of sustaining a platform ecosystem. This is especially challenging since third-party developers self-select into their developer role and so begin their entrepreneurial journey towards the launch of a product (e.g., an app) and the generation of sales. Where most research previous almost exclusively focuses on strategic management by the platform owner, add a more operational level by also focusing in the individual third-party developer.

In this paper we aim to provide a more detailed understanding of the production side of technology platforms. We explore whether structural conditions surrounding individual behavior are useful for predicting the entrepreneurial process of third-party providers of apps. Hereby we also seek to improve knowledge of the mechanisms supporting the traditional economic logic of entrepreneurship; that is, how individuals seek entrepreneurial rents from spotting and evaluating market opportunities, inventing and commercializing their inventions (Schumpeter, 1934). To

[3] In the following, we use the terms “third-party provider” and “third-party developer” interchangeably.

make such inferences we capture data prior to the individual decision to become an app developer and to the launching and successful selling of an app on an online platform.

Furthermore, research on the provision of innovation on platforms tends to focus on supply-side explanations only, for example non-pecuniary motivations of third-party providers (Mollick, 2012). However, we specifically link data on the behavior of individuals on *both* sides of the market into our explanation. Our paper demonstrates how (a) understanding the individual relationships and structural positions in an online community connected to a platform and (b) patterns of individual use of platform offerings both provide relevant input to a predictive model of the likelihood that a given user will make the transition to becoming an entrepreneur. Hence, we empirically investigate how nascent entrepreneurs, who are notoriously difficult to identify a priori, may become informed about market opportunities and enact innovation in the new and exciting context of the app economy, a phenomenon with dynamics of which we so far have only negligible empirical knowledge. Being able to predict users' entrepreneurial progress is of vital importance for platform owners. This relates not only to knowing early in the process which user-entrepreneurs are likely generate profits to the platform owner, but also to the platform owner being able to incentivize particular individuals who show some developer intention to continue their way towards user-entrepreneurship.

Primarily focused on how third-party providers progress through the entrepreneurial process, we add two pieces of explanation to the growing literature on platforms and users as innovators and entrepreneurs (von Hippel, 2005, Nambisan and Baron, 2012). First, we add the component of individual patterns of *consumption* of platform offerings; that is, buying behavior as a "real" demand-side driver catalyzing or accelerating the move to becoming a third-party app provider.

We consider this original since most of the literature focuses exclusively on the mere *use* of products, typically based on self-reported information on intensity and time of usage. Second, a significant aspect of our study is that we only use *behavioral* data and thus do not rely on data from focus groups, surveys or interviews that may be biased in several ways as well as expensive to obtain. Relevant behavioral data, for example on communication (e.g., in user and developer fora) and consumption (e.g., download and registration of software development kits, purchases of apps) by users, are typically stored on the systems of the platform owner and therefore easily available to the company.

We employ a unique data set obtained from a platform-owning company in the music software industry. While the company has for more than a decade invited users to innovate their products, supported by a strategy of "selective revealing" (Alexy et al., 2013) and a vibrant online user community, the company allowed users in July 2012 to develop and sell apps to complement their software products. The company chose an innovation model where users interested in developing apps could be granted developer rights. These rights include access to a restricted online community where developers can share information and create ideas, and access to a software development kit and component library to support user's programming activities and reduce development cost (von Hippel & Katz, 2002). We regard the act of obtaining developer rights as a clear indication that the end-user intends to develop an app. This, in turn, we regard as an indicator of an underlying ambition to launch and sell apps and thus becoming an entrepreneur. From the online community linked to the platform, we extracted all communications that occurred between users over a period of more than 10 years. We then linked this data to individual-level transaction data for all users. We included two types of transactions: (a) all online app purchases by customers over an 18-month period, beginning with the opening of the platform for third-party apps in July 2012 and (b) all purchases by customers of proprietary software products offered by the platform owner over a period of 10 years.

Given the explorative nature of our study and the limited amount of knowledge on the phenomenon, we follow Helfat's (2007) "stylized facts" approach and refrain from generating hypotheses that would be inappropriately specific[4] at this stage. The paper progresses as follows: First we offer a conceptual framework which guides our investigation through generating three research questions. Second, we explain the qualities of our data as well as the methods we apply. Next, we provide research results followed by a discussion section which links our findings to the current literatures on technological platforms and entrepreneurship, outlines limitations of our study and offers conclusion.

CONCEPTUAL FRAMEWORK

How should firms develop their technological platform? Theory suggests that platforms can be developed as two-sided markets and that this can be done simultaneously from both sides (Parker & Van Alstyne, 2005). Still, firms face strategic challenges with regard to deciding either to promote the demand-side of the platform and its content through marketing campaigns or discounts, or to focus on involving third-party developers and thereby develop the supply-side of the platform (Gawer & Cusumano, 2012; Boudreau, 2012). If a firm follows the latter strategy, two important questions emerge: (a) how to create a suitable base of developers who can differentiate the platform products and (b) how to make sure that developers launch apps that

[4] "To put it bluntly, the current state of affairs where researchers feel they have to come up with hypotheses in order to justify empirical work is counterproductive. It would make a lot more sense to simply identify a study as an investigation of a potential empirical regularity and then explain the motivation behind the investigation" (Helfat, 2007:188).

meet market demand. The imperative question then is whether firms, early on, can use a simple set of leading indicators to predict which users may become developers and which of these continue their entrepreneurial journey into market entry and making sales? Hence, identifying successful third-party providers is important for the development of the platform innovation ecosystem.

A line of research that is useful here focuses on customers or product users engaged in innovation, supported by platform-related online communities (Jeppesen & Frederiksen, 2006). Under certain conditions, some users may progress from sharing their advice and possible inventions with other users (but without remuneration) to becoming entrepreneurs and commercializing their inventions. Our research aims to shed light on who these users are and why they create and sell their inventions. Explanations of user innovation typically revolve around ability profiles (i.e., lead users; Shah & Tripsas, 2007, von Hippel, 2005) and various external stimuli such as monetary rewards, firm or peer recognition (Shah, 2006; Jeppesen & Frederiksen, 2006). Only recently were relational arguments introduced as an explanation for user innovations and their success (Dahlander & Frederiksen, 2012). This relational view suggests that both (a) communication patterns such as how much and to how many unique others a user is connected and (n) structural position in the community network have predictive value for potential engagement in innovation. User innovation studies are important but have remained largely descriptive since most are based on qualitative or cross-sectional data.

Firms owning technological platforms must strategically manage the “input” side of their two-sided markets. We are interested in predicting the transition from being a user consuming products on one side of the market into being a developer and thus launching and selling products on the other side. While Thornton (1999) argues that the entrepreneurship literature has little to offer regarding demand-side effects on entrepreneurial entry, White (1981) claims from an institutionalist perspective that the emergence of markets and thus opportunities for consumption *and* selling (e.g., platforms as two-side markets) is an initial condition for entrepreneurship to unfold. Still, in the setting of technological platforms and third party providers, individual users move through an entrepreneurial process of initially showing entrepreneurial intention by registering as a developer (Krueger, Reilly & Carsrud, 2000). Therefore, in our setting, entrepreneurial intent is not a purely cognitive process of opportunity recognition and evaluation and thus about understanding the demand-side uncertainties and the trade-offs against opportunity costs (Eckhardt & Shane, 2003). Entrepreneurial intent also includes user activities such as obtaining authorization from the platform owner to access relevant proprietary tools and information to enable the development of apps. The next step is that the developer actually creates a functional product (i.e., she or he designs the functions and features of an app and writes the code), obtains approval by the platform owner's quality control unit and then launches the product for sale on the platform market. This market entry represents entrepreneurial action for the developer (McMullen & Shepherd, 2006). The last leg in the process is to generate sales, which represent entrepreneurial success (Klein, 2008). However, what makes our story different from the traditional entrepreneurship story is that in the context of a technological platform it is not only the user making the transitioning to entrepreneur who profits from commercializing his invention (Shah & Tripsas, 2007) but also the platform-owning company. Hence, platform owners have an intrinsic interest in understanding the transition from product user to app entrepreneur.

Behavioral explanations

Consumption

Typically consumption is understood as shaped by our needs and thus revealing our preferences. Yet, at times, consumption provides more than pure utility, for example opportunities to engage and to learn (Holbrook & Hirschman, 1982). The experiences derived from using a product may help users develop their creativity and identify their own skills (Bem, 1972). Bandura (1977) explains how the discovery of one's own expertise may lead to increased self-efficacy. And particularly in the early stages of the entrepreneurial process, self-efficacy has been shown to be a key motivational factor, increasing the likelihood that entrepreneurial intentions will be formed and enacted (Ardichvili, Cardoso & Ray, 2003; McMullen & Shepherd, 2006). If user experience can be linked to entrepreneurial intent and action, one might also speculate as to whether there might be particular patterns of use and consumption that are more likely to lead to entrepreneurial intent and action than others. Such patterns might involve the number of different platform products owned by a user, the time a user has owned these products and gathered experience with them, or a combination of these. Also in later stages of the entrepreneurial process, when a user has made the transition to becoming a developer, platform apps created by others may continue to serve as a source of learning. Comparing one's own code to code written by others, for example, may offer opportunities to increase one's analytical understanding of efficient code structures in general. Interacting with other developers may lead to additional knowledge transfer, for example, about technology trajectories or market trends. Although the process may not unfold in the manner of a coherent action plan (Autio et al., 2013), these arguments suggest that platform users with a high number of purchases of and expenditure on platforms apps, and a long history of use of the platform, may be more likely to become app entrepreneurs, and more likely to create successful apps, than users for whom this is not the case.

Communication patterns and network position

Entrepreneurship research has for two decades emphasised the importance of an individual's social relationships and network position as a catalyst for the process into entrepreneurship (Birley, 1986; Hoang & Antoncic, 2003; Nanda & Sørensen, 2010). Also, recent research robustly show that users' relationships and structural position in networks in online communities organized around particular platform products influence innovative and entrepreneurial behavior (von Hippel, 2007; Dahlander & Frederiksen, 2012; Autio, Dahlander & Frederiksen, 2013). Several specific mechanisms may be at work: First, users who are connected to a larger number of unique individuals may simply have access to more information than others, which may enable them to make better-informed choices about becoming a developer and differentiating their product. Second, users who span otherwise non-connected individuals may obtain opportunities through information arbitrage about, for example, future exchange opportunities. Such information access also means that they may be able to better assess how ideas for potential apps are evaluated by other users (i.e., potential consumers). Additionally, they are also be better positioned to acquire novel information serving as input to the opportunity recognition process. Third, there are status hierarchies in online communities (Dahlander & Frederiksen, 2012). Some individuals enjoy particularly prestigious social positions and may therefore be preferred by other users as a source of advice on problem solving and new ideas for future product use and development (Morrison, Roberts & Midley, 2000). Through their status position, particular individuals may be able to influence technological trajectories and thereby shape the future of demand (Autio et al., 2013). Each of these mechanisms may make it easier for individuals with such communication patterns and network positions to understand potential markets, recognize opportunities, and reduce uncertainty in general. Jointly, they offer these individuals specific incentives to venture into becoming developers and subsequently to launch and sell their products. These arguments are well-aligned with Podolny's argument that networks behave like markets where relations between individuals take the form of both pipes for information flows as well as prisms for assigning status (Podolny, 2001). Hence, we propose that certain users in the online community, due to their communication patterns and structural position in the communication network, have a higher probability to make the transition to becoming third-party providers than other users.

The arguments above motivate three overall research questions that will guide our empirical analysis. First we ask: to which degree can transition from product user to third-party developer (i.e., entrepreneurial intention) be predicted? That is, do community-active users who are about to register as a developer, already differ from other users in terms of their community activity, network position, and consumption history before they display such intention (RQ1)? Second, we investigate to what extent developers who will launch an app (i.e., entrepreneurial action) can be identified beforehand, based on leading indicators. That is, do community-active developers who are about to launch their first app already differ from other developers in terms of their community activity, network position, and consumption history before launch activity can be observed (RQ2)? Finally, we investigate how well sales of apps launched by third-party providers can be predicted. That is, do developers whose apps sell successfully already differ in terms of their community activity, network position, and consumption history from other developers who launch an app before the sales begin (RQ3)?

DATA AND METHODS

Research setting

Our research setting is a firm-owned music software platform. The firm offers a range of complementary products for producing, processing and recording music. Related to the platform is a vibrant online community that allows users to interact, discuss future technological features and functions, and collaboratively solve problems related to the products offered by the platform-owning firm. Users can buy the proprietary company software (i.e., the base products of the platform) and various types of extension apps in a secure online store operated by the company. For users who wish to become app developers, the firm offers a software development kit and code libraries (similar to software platforms such as Apple iOS or Google Android) and access to a gated online forum for developers. Only registered developers are allowed to offer their apps on the online marketplace (see Figure 5 in Appendix) where they compete with other user-developed and company-developed apps, grouped into instrument categories such as guitar or drums, and genres such as metal or house. Price and user ratings are displayed. All apps have a file format that is reserved for products of the platform-owning firm and is incompatible with competing platforms.

The company opened for user-contributed apps in July 2012. The number of users registered as app developers reached 1,657 by mid-December 2013. Between early 2012 and mid-December 2013, a total of 59,747 app sales were recorded (average price: USD 46, min: USD 3, max: USD 211). The total app sales for the first eleven months of 2013 amounted to USD 2.8mn with 11,284 users active as app consumers (average expenditure per consumer: USD 248). The distribution of app sales in the first eleven months of 2013 is shown in Figure 1a. The distribution of app launches per between January 2012 and February 2014 is shown in Figure 1b.

Besides operating the partly user-driven app market, the company also hosts an online community. In total 1,085,613 posts were created by 27,297 unique individuals since the discussion forum first was launched in February 2002. 9,466 of these individuals were identified as active in the period December 2011 to December 2013. Figure 2 shows the development of posting activity up to December 2013. For the year December 2012 to December 2013 a total of 225,908 posts were made in the community. This is more than 20% of all posts ever made. Each post is nested in a thread within the community. When posting in the community, individuals have the option to either start new threads or respond to existing threads. The distribution of individuals per thread is right-skewed with 22% of new threads never answered, 23% only including 2 individuals, 77% with max 5 individuals, 10% +10 individuals, 1% +30 individuals, and 0.1% +70 individuals. It is important to notice that 68% of the active community participants are registered users of at least one proprietary software product offered by the platform owner.

Data sources

The data material for our analysis was obtained in the form of two snapshots of the platform owner's database. The database contains sales transaction data, together with a complete user forum extraction, lists of users registered as app developers and users registered as members of app producing companies, and multiple other related datasets regarding individuals and products. The database snapshots were extracted in the end of December 2012 and December 2013, respectively. The different data sources were received as raw data tables, and it was possible to link the tables based on user IDs and product IDs, enabling detailed data extractions of sales and and communication metrics on the individual user level. The data tables used in this paper were the following: *App sales list*, *Product extensions sales list*, *Extraction of all posts from user* *community forum*, *List of product registrations by user ID* (purchased products), *Demographic data for userIDs* (only geography variables are available), *List of users registered in a user-companies* (not dated), *List of free product extension downloads by user IDs*, *Registered app developer user IDs* (not dated), *Date where the user companies were registered*, *App product details* (Title, description and release dates), *List of user-company names with company IDs*, *List of user company IDs' launched apps*, and *Price list for apps*.

A total of 1,411 individual users were registered as developers in December 2012 and 1,657 in December 2013, an increase by 246 new developers over one year. Only 278 (20%) of the 1,411 developers registered in December 2012 were also community members, compared to 378 (23%) in December 2013. Although these numbers suggest that entrepreneurial intentions are to a considerable extent formed "outside" the user community, they also indicate that 4.4% of the community members were developers in December 2012 (278 out of 6,292)[5] whereas this was only the case for 0.3% of the users from outside the community (1,133 out of 448,860). This suggests that it may be much more relevant and possible for the platform owner to identify future developers from within the community. Indeed, the probability that an active community member becomes a developer is more than 50 times higher than the probability that a user who is not a community member becomes a developer (for related research, see Dobrev & Barnet, 2005).

Measures

Dependent variables

Entrepreneurial intention was measured a binary indicator that took the value 1 when a user (a) signed up as a developer on the platform-owning firm's website and received the software

[5] While this seems a very low ratio of entrepreneurship relative to the overall community population we emphasis that entrepreneurship is a rare event. Also our ratio is not different from other settings, for example, Nanda & Sørensen (2010) find a similarly low entry rate for entrepreneurship in Denmark and Dobrev & Barnett (2005) show a similar low rate of entrepreneurial entry in their sample of MBA alumni of an elite U.S. business school.

development kit required to program apps for the company's proprietary software platform or (b) registered a new/existing app-producing user company in the hosting company's database, and 0 otherwise.

Entrepreneurial action. The date of a user's first app launch is not available from the datasets and thus an alternative approach for identifying entrepreneurial action was taken. The first database snapshot from the platform owner was received in mid-December 2012, and the second snapshot in mid-December 2013. Neither the registration date as a developer, nor date a user registers as part of an app-developing company, were available due to the SQL architecture implemented by the company. However, combining these two data snapshots allowed identification of the individuals who transitioned from a user to become a developer during the period of time. By December 2012 a total of 455,152 unique user IDs were were recorded in the company's database, and 6,292 of these had showed activity in the community defined as having either started new discussion threads or responded to existing posts by other users. In Dec2013 these figures had increased to 554,767 and 10,532 users, respectively.

Entrepreneurial *action* is defined as a user launching or being involved in launching his first app. Thus this represents the entry step into the market but not its magnitude. Showing action is conditioned on first becoming a developer given the setup by the platform owning company, where only registered app developers have access to the required software development kit, and only registered user companies have access to market their apps on the webshop likewise operated by the platform owner.

Entrepreneurial *success* can be measured in multiple ways, for example, by developers' total app sales value, total unit sales, average user rating of launched app(s), number of user ratings, or a combination hereof. The most tangible measure of success is perhaps value generated.

However, since there were no sales registered (yet) for several apps by December 2013, we used the total unit sales registered for all apps launched by the respective developer as the measure.

Explanatory variables

Social network metrics were calculated based on interactions in the user community. Due to the architecture of the online community, each post is either a new thread or a response to a an existing post. This enabled us to construct a social network from the posts, with users represented by nodes and posts represented by directed edges[6]. The following network metrics were calculated for each of the 10,532 community-active users for the periods December 2011 to December 2012 and December 2012 to Dec2013, respectively: Number of posts by user in the analysis period, Number of threads started by user in the analysis period, Community tenure of user since his or her first post (not limited by the analysis period), Number of other users active in the same threads as the user, Number of unique other developers that user had received posts from, Nodal degree centrality (Wasserman & Faust, 1994) of the user in the community, and Network prestige (indegree; Alexander, 1963). Prestige has previously been shown to influence decision-making for entrepreneurial activities (Van Praag, 1999). Burt's measure of constraint was also calculated, indicating the user's opportunity for information arbitrage. This measure was introduced by Burt (2004) as a measure of structural hole spanning in social networks by actors. This measure represents an index between 0 and 1 for information arbitrage, where 0is completely unconstrained and 1 is completely constrained. Ability to span structural holes influences an individual's social capital (Walker, Kogut, & Shan, 1997) and social capital has been shown to have a positive impact on entrepreneurship entry and success (Westlund &

[6] The poster was interpreted as the *sender*, and the poster of the post to which the sender was responding was interpreted as the *receiver*. All directed edges were summarized into a weighted network for the given analysis period. A similar approach was used by Dahlander and Frederiksen (2012).

Bolton, 2001). Therefore we link structural hole spanning activities in the online community to the entrepreneurship process.

In some cases a given network metric may be missing or undefined for a given user. A user may, for example, be active in the community by starting a lot of threads, but never receive any responses and thus not be part of the network. In such cases, the missing values for the community and network metrics were set to zero in the datasets, with the exception of Burt's measure of constraint, which was left as missing since it is not defined for isolates in networks (Burt, 2004).

Besides the community and network features extracted from the community, various measures of individuals' consumption history were extracted from the data. These variables include: Number of unique proprietary base products of the platform owner registered by the user, Days since first unique proprietary base product from the platform owner was registered by the user, Number of times the user had registered such products (some products are registered multiple times for the same user), Number of apps purchased by the user, Total value of app purchases by the user, Number of free apps downloaded by the user. Where no consumption was detected for a user by a given consumption metric, we imputed a value of zero.

Control variables

Various additional measures were calculated. One of these was the duration between the first and the second app purchase, a measure often used to identify key customers for novel products (Cardozo, Smith, & Viswanathan 1988). This variable was missing for approximately two thirds of the community-active users, and for those where it could be calculated, it was insignificant when added to the logistic regression model for RQ1.

We also calculated control variables related to community post content. We applied supervised and unsupervised text mining techniques to the posts users made in 2012. This way we derived clusters of similar writing styles and topics discussed by user, inspired by previous research on transition into entrepreneurship by Dobrev and Barnett (2005). However, none of the cluster membership variables constructed in this way has a significant relationship with our dependent variables. We further identified keywords such as “SDK” (short for *software development kit*, which in this context is software used to develop apps), and “platform” with high discriminative power between regular users and developers in the community. The same applied for the highly discriminating keywords.

In modeling a developers’ app sales we include the average price of the apps as a predictor. However, app prices in a professional community may be correlated with app quality, and thus we make no attempt to interpret the effect of app prices on sales as causal.

Estimation strategy

Three different modeling approaches are chosen to address each of our research questions separately[7]. The main purpose is to understand what drives entrepreneurial intention, action, and success, at each stage of the entrepreneurial process. Further, the desired sample for RQ3 would not be nested in the previous RQs but further include developers who are community inactive, as this allows for us to quantify the importance of community membership for app sales. We restrict our focus to those individuals identified as developers by December 2013, whom we can also identify in the community the previous year as communicating users. We focus on the 59 users we identified as communicating users in December 2012 and as developers in December 2013.

[7] It was considered to model the three RQs in one nested Heckman selection model, but this idea was abandoned because directly conditioning interpretations on nested models would not necessarily reflect the managerial reality in which our models would be implemented and interpreted.

We compare these 59 users to the 5,955 users who were also active in the community between December 2011 and December 2012.

To predict entrepreneurial intention (RQ1) - operationalized as the transition from user to registered developer in the period December 2012 to December 2013, we formulated a logistic regression model. To predict entrepreneurial action (RQ2), we formulated an extended Cox regression model, using the same input variables. For each quarter between January 2012 and October 2013, community activities in the previous six months were evaluated, and community activity measures and network metrics were calculated[8].

A simplified approach was taken in the prediction of app sales (RQ3)[9]. The observations were the 103 developers who had apps on the market by December 2013 and could already in December 2012 be identified as users from their community activity or purchase history. Their app sales in the next year were the dependent variable, and their consumption and communication metrics (as evaluated by December 2012) were the explanatory variables. Before estimating the full model, a reduced model was estimated taking as input only a dummy for community membership[10] together with the consumption measures and the developer's mean app price. Missing network and communication metrics for these developers were imputed by zeros, except for Burt's measure of constraint for which no meaningful imputation values exists. This variable was thus omitted in the further analysis.

[8] The rationale for this was that we assume users to be more active in the community in the period leading up to the launch of their app, and six months was deemed a reasonable time period for developing an app. This was verified in various interviews conducted with employees from the platform owning company.

[9] Whereas in RQ2 it is both possible and meaningful to dynamically analyze app launches, in RQ3 it is less meaningful to do so. This is because it is unclear when in the process to evaluate product- and user characteristics. In order to obtain network metrics, time must be treated as discrete periods, similarly to RQ2. We expect each incident of an app sale to be partly predictable by the developer's network activities up to the sales incident. However, given the discrete time periods and a potentially strong correlation between an apps launch time and its sales per time period, sales- and network activities may peak within the same time period, hampering estimation of any time dependence.

[10] The rationale for including a dummy for community membership was that, among the developers who had launched apps, there were also developers who had *not* been active in the online community.

The data were processed and analyzed in SAS 9.4, SAS Text Miner 13.2, and R 3.0.3, using the packages 'network' (version 1.9.0), 'sna' (version 2.3-1), 'igraph' (0.6.6), and 'Zelig' (4.0-11). In the model outputs we only report significant variables, in order to provide a clear overview of the significant predictors of the different stages of entrepreneurship. Logarithmic variable transformations were performed for all explanatory variables except Burt's measure of constraint which was logit transformed. Univariate[11] and bivariate descriptive statistics[12] are reported in Table 1.

RESULTS

In an initial step, we assessed if user-developers actually differed from regular users. Figure 3 shows a profile plot of differences in terms of standardized variables between those who were registered as developers by December 2013 (1,657) and those who were not (554,767). A MANOVA revealed that the multivariate difference between the groups ($p < .001$). While no causality can be inferred from this result, it suggests that developers may be a different type of users, and that future developers might share similarities with current developers.

RQ1 - Predicting entrepreneurial intention

Table 2 reports the results of the logistic regression predicting new developers in the year December 2012- December 2013 from the set of regular users in the community and their

[11] Table 1 also reports summary statistics for each explanatory variable in its level form. The skewness is partly remediated by the applied transformations. In order to check for robustness, the logistic regression model in RQ1 was also estimated using rank transformations of the covariates, where similar patterns of parameter estimates and significance were obtained.

[12] Many of the correlation coefficients are high, which is expected due to similarities in many of these measures, for example, the network measures; degree centrality, prestige, the summary of number of posts, and to a certain degree the number of inputs received from app developers in the community are all highly correlated, which is no surprise, since the first measures the sum of all ties, the second measure only incoming ties, the third the number of outgoing ties, and the latter a subset of incoming ties. This suggests that only one of these may be significant in a given model, as including multiple of these measures will likely result in high levels of multicollinearity. However, Wasserman & Faust (1994) recommends using both variables in social network analysis, because they conceptually attempt to measure different structural properties of the network. The same applies for the variable set consisting of the number of product registrations and number of different products owned, and the products sets number of apps purchased and app purchase value.

communication-, network- and consumption metrics by December 2012. Out of 3,911 users in the community in December 2012 with more than one post made, 50 made the transition into becoming developers. These transitions indeed represent relatively rare events, and the below results were confirmed for robustness using estimation methods intended for rare events as mentioned in the methods section. Here we report the results obtained in the naïvely reduced model, where insignificant predictors were removed one at the time, until every remaining variable was significant. The number of apps purchased was positively related to the probability of users transitioning into developer status ($p<.001$). Also the number of software products purchased by the users from the platform owner had a positive effect ($p<.05$). Prestige was found to be positively related to entrepreneurial intention ($p<.001$), and likewise was the number of unique contacts that a user had in the community identified as a positive predictor of entrepreneurial intention ($p<.05$). The other considered variables were not significant. The Max-rescaled R-squared measure of the model suggests that it predict 12.5% of the variation in who among the community members will venture into entrepreneurship within the next year of previously unseen data. As this measure is not informative regarding the actual predictive power of the model, Figure 4 adds to this overview by summarizing the ROC curve for the model. The ROC curve shows how the model successfully identifies a subset of the new developers who are particularly salient for the model (Hanley & McNeil, 1982).

RQ2 - Predicting entrepreneurial action

Out of 1,657 developers registered by December 2013, 1,534 had not launched their first app by Dec2012. 123 developers launched their first app between December 2012 and December 2013, and these events were modeled using an extended Cox regression model. In Table 3 only significant predictors are summarized. Notice first that prestige again had a positive parameter estimate ($p<0.001$), which suggests that those with many incoming ties are more likely to launch apps. Degree centrality has a negative impact on the likelihood of launching apps ($p<0.001$). This suggests that when keeping incoming ties fixed, then additional ties (by definition: outgoing) has a *negative* effect on the probability of launching first app ($p<0.001$). We also find that the parameter estimate for product tenure is negative. This suggests that among developers additional product tenure is negatively correlated to launching an app ($p=0.012$). We find that starting more threads is likewise negatively correlated with the launch of developers' first apps ($p=0.032$).

RQ3 - Predicting entrepreneurial success

The dependent variable measuring entrepreneurial success was app launchers' aggregated app unit sales in the year December 2012 to December 2013, predicted based on characteristics by December 2012 and measured on the same set of predictors as introduced in the previous models. The dummy variable for community membership was first introduced and showed positive and significant ($p<0.001$). In order to obtain a more informative results for which communication- and network measures that had a predictive power for successful app sales, the dummy variable for community membership was replaced by the more detailed network- and communication metrics previously introduced. These were number of posts by the developer, number of threads started, and number of contacts in the community. Regarding the network- and communication measures, the number of community contacts held by the app producer had a positive effect on the user's total app sales for the period ($p=0.033$). Prestige likewise had a positive effect ($p<0.001$). Keeping other variables constant, an increase in number of threads started was found to have a negative effect on sales ($p<0.001$).

In terms of consumption metrics, we found that the number of products owned were positively related to app sales ($p<0.001$). We found that the number of apps purchased by the developers was very right skewed with many developers having no registered app purchases. We thus included this variable as a dummy for app purchases or not, and found that this variable had a positive effect on sales ($p<0.001$). The control variable "mean price of a developers apps" controlling for latent app characteristics, had a positive effect on expected sales ($p<0.001$).

A simple and interpretable pseudo R-square for Poisson regression was calculated by log-transforming the observed and predicted counts, then correlating them, and finally squaring this correlation to obtain the pseudo R-squared. This measure yielded 0.3649, suggesting that 36.5% of the future sales success can be predicted by these user characteristics at a previous point in time.

Summarizing our results across the three models, we find evidence that high levels of community prestige for a user predict entrepreneurship on the platform well. We also find that prior consumption (i.e. number of proprietary software products registered, and app purchases) adds significantly to predicting entry into the different stages of user entrepreneurship. Two additional interesting general findings are: a) that agenda setting, as represented by starting new threads, has a negative effect on entrepreneurship, while holding the other variables constant and b) Despite visible in the mean comparison in Figure 3, in none of our models are Burt's constraint was significant and thus users who span structural holes and so may enjoy benefits of information arbitrage are not more likely to enter into any stage of entrepreneurship than other users.

DISCUSSION AND CONCLUSION

For platform owners it is vital to be able to develop the supply-side of their innovation ecosystem. Our empirical analysis offers robust identification of those users who are likely to move from platform users into third party developers in the next time period as well as for predicting who, among users communicating in the platform community, are likely to progress into launching an app and be successful in selling apps.

Interestingly, our results from predicting first app launches indicates that developers launching their first apps are frequently being talked to by other users - and in particular so in threads started by others. This suggests that developers showing entrepreneurial action are developers successful in elaborating on agendas (threads) started by other users to an extent where their posts create sub-threads, as indicated by the positive prestige parameter estimate. Thus, such users can be opinion leaders by agenda moderating even without starting new threads. In the agenda moderating process the input they receive may provide developers with increased insight into promising new ideas for app development, estimate demand uncertainties, and/or these incoming ties may provide them with a feeling of prestige and being able to provide useful information to other users (Wasko & Faraj, 2000). This response from a potential market of users may in turn drive their confidence in being successful in their entrepreneurial endeavor. Still, the finding that technology probing and so agenda setting by starting new threads is negative significant in most of our models is surprising as well as in opposition to findings by Autio et al. (2013). Also, we find in Model 2 that product tenure has a negative effect on launching an app. This is surprising but indicates that the population of app developers who progress into market entry of their app are not long time product users and thus are not as such drawing on long-term experience with the software product of the platform as a basis for their app creation.

Finally, because our data is captured at the time of initiation of the platform it allows us to complement Mollick's (2012) study and thus offer empirical analysis about the formation stage of a marketplace for third party user generated content.

Theory contribution

Our research elaborates on three conceptual strands. First, we add to prior studies that demonstrate that user communities constitute an important determinant in explaining user entrepreneurship (Autio et al. 2013). We connect this literature to the current discussion about platform strategy and organization of innovation via platforms (Gawer, 2014; Boudreau, 2012). Second, our findings introduce a new type of explanation that should be included for obtaining a better understanding of the transition process into user entrepreneurship on platforms, namely, the consumption history of each user. By highlighting how consumption patterns play a role in predicting entrepreneurship relative to use experience, we believe, we are opening an original agenda for future research. This agenda is interested in explanations from the 'real' demand-side and thus to include insights on innovation arriving from the field of marketing (Hauser, Tellis & Griffin, 2006).

Second, the attention to user purchase and product registration offers a different approach into studies of user innovation than focusing mainly on individuals who self-report their product use, product related abilities or how certain stimuli affect their motivation for invention (von Hippel, 2005; Jeppesen & Frederiksen, 2006; Shah, 2006). Whereas research on user innovation yields important insights into the organization of innovation too often the focus has ignored overall behavioral measures. We believe our study serves as an inspiration for an alternative approach[13].

Still, related to the user innovation literature, one interpretation of our findings is that specific customers in consumer products with latent abilities and preference patterns benefit more strongly from adopting new apps because of their leading-edge needs and qualities (von Hippel, 1986), and they therefore consume more apps as they are better equipped and maybe appreciate

[13] We acknowledge that it is challenging to obtain data that allows for studies of how users over time consume, communicate as well as produce innovation but still encourage researchers to explore this new trajectory.

potential and features of a new app (Schreier, Oberhauser & Prügl, 2007). This interpretation is aligned with studies showing that lead users therefore have a higher probability for becoming user entrepreneurs (Autio et al. 2013).

Third, we contribute to studies in entrepreneurship by offering two contributions. First, research tends to have theorized about opportunities themselves (Shane & Venkataraman, 2000) rather than examined the conditions that prompt individuals to perceive opportunities, evaluate them, and act upon them (Choi & Shepherd, 2004). This has lead the conceptual lens to focus on the entrepreneurial process (McMullen & Shepherd, 2006). Yet, we explore empirically how the entrepreneurial process unfolds and, in particular, we demonstrate that only some parts of our behavioral explanations are useful for predicting progress along the entrepreneurial process (i.e. intention, action and success). Thus, while we caution to think that our explanations are equally useful for predicting the entrepreneurial process at each stage we still maintain that some explanations (i.e. community prestige and number of app purchase) are applicable for making prediction across all entrepreneurial stages. This also supports the view that entrepreneurship is perhaps not to understand as one homogenous strive towards commercialization of an invention but can usefully be thought of as a series of different stages where different factors determine the individual's choice for progressing into the next stage and thus for pursuing opportunities (McMullen & Shepherd, 2006;). Second, research shows that social relationships and network position influence individuals' decision to become an entrepreneur (Birley, 1986; Nanda & Sørensen, 2010). Yet, how such mechanisms work in an online setting is less evident. Research in entrepreneurship has only to a limited degree connected to the new type of innovation system that a technological platform constitutes (Nambisan & Baron, 2012). This is a system where both individual third party providers as entrepreneurs benefit from their efforts but also the platform owner benefits. This implies that entrepreneurship among users who become developers and potentially successful in selling their apps have also strategic importance on an extended organizational level, namely, for the platform owning firm.

Managerial implications

For managers interested in promoting innovation on their platform our research encourages to pay particular attention to the extent to which a user consumes and communicates in the online community connected to the platform. This, in order to be informed early on, about this user's possible progression into entrepreneurship and thus the development of the supply-side of the platform.

Also, our model predicts a number of false positives. While these users are wrongly predicted to become developers, at first glance makes prediction results less actionable for companies, we think that this result is important. This is because it suggests that a larger group of users, similar to future developers can be probed for entrepreneurial involvement. Thus, this indicates that these individuals may be more open to respond to strategic action from the platform owner such as recognition or pecuniary incentives to facilitate a transition into future entrepreneurship.

Limitations

As in most studies ours does come without limitations. Yet, these limitations serve to open up for formulating new research questions. For example, it is clear that our data has only to a limited degree allowed to control for individual level explanations such as individual ability and motivational drive. We are currently gathering survey data among developers on the platform to be able to assess to what degree our prediction model based on simple behavioral data is superior relative to survey based individual perspectives of motivational stimuli as well as individual level and types of experience and education.

While we do not have the benefit of random-assignment of users and developers or a natural experiment in our study, we undertake several additional tests to outline the sources of such endogeneity and to control for such spurious correlations.

Another limitation relates to generalizability of results since we focus solely on one platform in one specific industry domain. Like related studies (Autio et al. 2013) we acknowledge this limitation and how relative difficult it is to gather data for these large-scale empirical investigations in similar settings. Yet, we strongly support that more comparative analysis of the process of user entrepreneurship in various platform setting are conducted.

Limitations specifically regarding the estimation of the model seeking to answer RQ1: In predicting new developers, we argue to constrain our focus to community members by the valid argument that the likelihood of transitioning into entrepreneurship is almost 60 times higher for community members than other users. However, by that decision, we simultaneously disregard most new developers, by this choice. By this limited focus, we are able to use network- and communication metrics in the prediction process, but we fail to be able to explain what make users from outside the community venture into entrepreneurship. However, we do show, that developers from within the community are significantly more likely to launch their first app than developers from outside the community, indicating that it is indeed most important to understand the transitioning into entrepreneurship of users already active in the community.

Limitations specifically regarding the estimation of the model seeking to answer RQ2: In modeling app launches, there is an independence problem between some events, as some user companies consist of multiple developers. Thus, multiple developers may experience their first app launch in the same one app launch. The extent of this independence problem is hard to quantify. We may expect users within the same user company to be correlated on multiple explanatory variables, but the contrary may also be the case, if these users divide the work between them in such a way that e.g. one user is active in marketing within the community, whereas the other user is active in app purchases and development.

Limitations specifically regarding the estimation of the model seeking to answer RQ3: In the modeling of app sales characteristics of the apps may be included as predictors, as different types of apps may experience different demand and thus perform differently in terms of sales. This has not (yet) been implemented in our current analysis.

Conclusion

Our study breaks new ground for the understanding of the entrepreneurial process in a platform setting and thereby on how innovation comes about on the supply-side in such ecosystem. Therefore, the study informs strategic management of platform owners. By using a large-scale empirical study we emphasis, in particular, how behavioral explanations, which are low cost and easy accessible for the platform owner, provide robust prediction for who will progress through the different stages of entrepreneurship. Also, we contribute by offering an original type of explanation for entry and success in the entrepreneurial process by emphasizing how the consumption histories of users impact their subsequent decision for transitioning into entrepreneurship. This explanation complements other structural explanations for entrepreneurship on platforms that we offer, namely, user prestige in the online community connected to the platform. As the number of platforms rapidly increases and platforms gain significance in the economy our empirical study enriches our understanding of how the dynamics of innovation and entrepreneurship work in such ecosystems.

REFERENCES

Adner, R. (2012). *The wide lens: A new strategy for innovation*. Penguin UK.

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems: how the structure of technological interdependence affects firm performance in new technology generations. *Strategic Management Journal*, 31(3), 306-333.

Alexy, O., George, G., & Salter, A. J. (2013). Cui bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review*, 38(2), 270-291.

Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity identification and development. *Journal of Business venturing*, 18(1), 105-123.

Autio, E., Dahlander, L., & Frederiksen, L. (2013). Information exposure, opportunity evaluation and entrepreneurial action: An empirical investigation of an online user community. *Academy of Management Journal*, 56(5), 1348-1371.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.

Bem, D. J. (1972). Self-perception theory. *Advances in experimental social psychology*, 6, 1-62.

Birley, S. (1986). The role of networks in the entrepreneurial process. *Journal of Business Venturing*, 1(1), 107-117.

Boudreau, K. J. (2012). Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science*, 23(5), 1409-1427.

Boudreau, K. J., & Jeppesen, L. B. (2014). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal*.

Boudreau, K. J. & Lakhani, K. R. 2009. How to Manage Outside Innovation. *MIT Sloan Management Review*, 50(4): 68-76

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349-399.

Butts, C.T. (2013). sna: Tools for Social Network Analysis. R package version 2.3-1. http://CRAN.R-project.org/package=sna

Butts, C.T., Handcock, M.S., & Hunter, D.R. (2014). network: Classes for Relational Data. R package version 1.9.0. Irvine, CA. http://statnet.org

Cardozo, R. N., Smith, D. K., & Viswanathan, M. (1988). Identifying key customers for novel industrial products. *Journal of Product Innovation Management*, *5*(2), 102-113.

Cennamo, C. & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal*, 34(11): 1331-1350.

Choi, Y. R., & Shepherd, D. A. (2004). Entrepreneurs' decisions to exploit opportunities. *Journal of Management*, 30(3), 377-395.

Csardi, G,, Nepusz, T.: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.sf.net

Dahlander, L., & Frederiksen, L. (2012). The core and cosmopolitans: A relational view of innovation in user communities. *Organization Science*, 23(4), 988-1007.

Dobrev, S. D., & Barnett, W. P. (2005). Organizational roles and transition to entrepreneurship. *Academy of Management Journal*, 48(3), 433-449.

Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard Business Review*, 84(10), 92.

Eckhardt, J. T., & Shane, S. A. (2003). Opportunities and entrepreneurship. *Journal of Management*, 29(3), 333-349.

Friedman, J., Hastie, T., Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models vi Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. URL http://www.jstatsoft.org/v33/i01/.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, 43(7), 1239-1249.

Gawer, A., & Cusumano, M. A. (2002). *Platform leadership: How Intel, Microsoft, and Cisco drive industry innovation.* Boston: Harvard Business School Press.

Gawer, A., Cusumano, M. A., & Strategy, D. S. (2012). How companies become platform leaders. *MIT/Sloan Management Review*, 49.

Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third□party development: the boundary resources model. *Information Systems Journal*, 23(2), 173-192.

Gigaom Research (2014). Sizing the EU app economy, https://ec.europa.eu/digital-agenda/en/news/sizing-eu-app-economy

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.

Hauser, J., Tellis, G. J., & Griffin, A. (2006). Research on innovation: A review and agenda for marketing science. *Marketing Science*, *25*(6), 687-717.

Helfat, C. E. (2007). Stylized facts, empirical research and theory development in management. *Strategic Organization*, 5(2), 185-192.

Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, 132-140.

Hoang, H., & Antoncic, B. (2003). Network-based research in entrepreneurship: A critical review. *Journal of Business Venturing*, 18(2), 165-187.

Jeppesen, L. B., & Frederiksen, L. (2006). Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science*, 17(1), 45-63.

Katz, M. L., & Shapiro, C. (1986). Technology adoption in the presence of network externalities. *The Journal of Political Economy*, 822-841.

King, Gary, and Langche Zeng. 2001. Logistic Regression in Rare Events Data, *Political Analysis* 9: 137–163.

Klein, P. G. (2008). Opportunity discovery, entrepreneurial action, and economic organization. *Strategic Entrepreneurship Journal*, 2(3), 175-190.

Krueger, N. F., Reilly, M. D., & Carsrud, A. L. (2000). Competing models of entrepreneurial intentions. *Journal of Business Venturing*, 15(5), 411-432.

Lockhart, R., Taylor, J., Tibshirani, R. J., & Tibshirani, R. (2014). A significance test for the lasso. *Annals of statistics*, 42(2), 413.

McMullen, J. S., & Shepherd, D. A. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review*, 31(1), 132-152.

Mollick, E.R. (2012). Filthy Lucre: What Motivates the Commercialization of Innovations? (working paper, SSRN).

Morrison, P. D., Roberts, J. H., & Midgley, D. F. (2000). Opinion leadership amongst leading edge users. *Australasian Marketing Journal*, 8(1), 5-14.

Nambisan, S., & Baron, R. A. (2013). Entrepreneurship in Innovation Ecosystems: Entrepreneurs' Self□Regulatory Processes and Their Implications for New Venture Success. *Entrepreneurship Theory and Practice*, 37(5), 1071-1097.

Nanda, R., & Sørensen, J. B. (2010). Workplace peers and entrepreneurship. *Management Science*, 56(7), 1116-1126.

Owen, M., Imai, K., King, G., & Lau, O. (2013). Zelig: Everyone's Statistical Software. R package version 4.2-1. http://CRAN.R-project.org/package=Zelig

Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management Science*, 51(10), 1494-1504.

Piezunka, H. (2011). Technological platforms. *Journal für Betriebswirtschaft*, 61(2-3), 179-226.

Podolny, J. M. (2001). Networks as the Pipes and Prisms of the Market1. *American Journal of Sociology*, 107(1), 33-60.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990-1029.

SAS Institute. (2014). SAS software. Copyright, SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

Sei, T. (2014). Infinitely imbalanced binomial regression and deformed exponential families. *Journal of Statistical Planning and Inference*, *149*, 116-124.

Schreier, M., Oberhauser, S. & Prügl, R.W. (2007) Lead users and the adoption and diffusion of new products: Insights from two extreme sports communities. *Marketing Letters*, 18 (1-2). pp. 15-30.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle* (Vol. 55). Transaction publishers.

Shah, S. K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, 52(7), 1000-1014.

Shah, S. K., & Tripsas, M. (2007). The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal*, 1(1□2), 123-140.

Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review*, 25(1), 217-226.

Suarez, F. F. (2012). Dethroning an established platform. *MIT Sloan Management Review*, 53(4).

Thomas, L. D., Autio, E., & Gann, D. M. (2014). Architectural leverage: putting platforms in context. The Academy of Management Perspectives, 28(2), 198-219.

Thornton, P. H. (1999). The sociology of entrepreneurship. *Annual Review of Sociology*, 19-46.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267-288.

von Hippel, E. (1986). Lead users: a source of novel product concepts. *Management Science*, 32(7), 791-805.

von Hippel, E. (2005). *Democratizing innovation*. Cambridge, Massachusetts.

von Hippel, E. (2007). Horizontal innovation networks—by and for users. *Industrial and Corporate Change*, 16(2), 293-315.

von Hippel, E., & Katz, R. (2002). Shifting innovation to users via toolkits. *Management Science*, 48(7), 821-833.

Walker, G., Kogut, B., & Shan, W. (1997). Social capital, structural holes and the formation of an industry network. *Organization science*, *8*(2), 109-125.

Wasko, M. M., & Faraj, S. (2000). "It is what one does": why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, 9(2), 155-173.

Wasserman, S. & Faust, K. (1994). *Social network analysis: Methods and applications*(Vol. 8). Cambridge university press.

Westlund, H., & Bolton, R. (2003). Local social capital and entrepreneurship.*Small business economics*, *21*(2), 77-113.

Wheelwright, S. C., & Clark, K. B. (1992). Creating project plans to focus product development. *Harvard Business Review*.

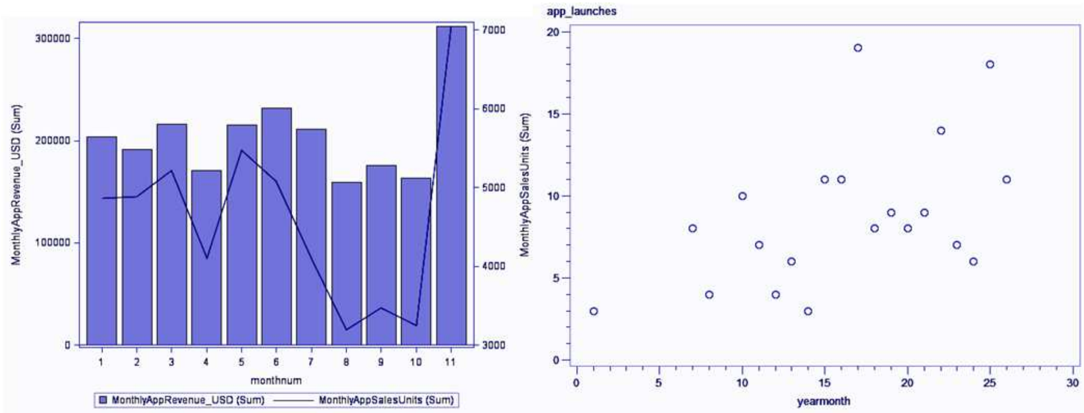
White, H. C. (1981). Where do markets come from? *Advances in Strategic Management*, 17(2), 323-350.

Yoo, Y., Boland Jr, R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, 23(5), 1398-1408.

Zhu, F., & Iansiti, M. (2012). Entry into platform□based markets. *Strategic Management Journal*, 33(1), 88-106.

APPENDIX

Figures



(a) (b)

Figure 1: Monthly app sales 2013 Jan-Nov (a) and monthly app launches Jan2012 to Feb2014 (b).

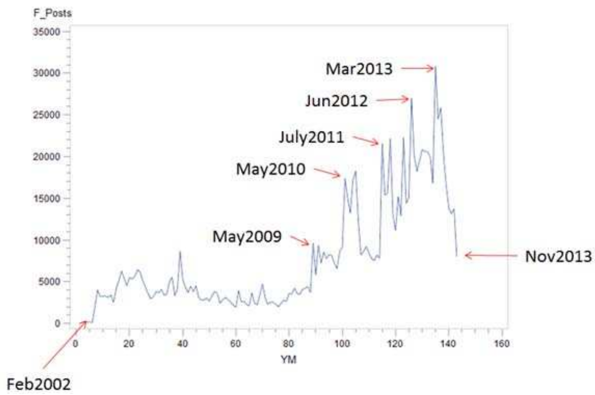


Figure 2: Monthly number of posts in forum, Feb 2002 - Nov 2013.

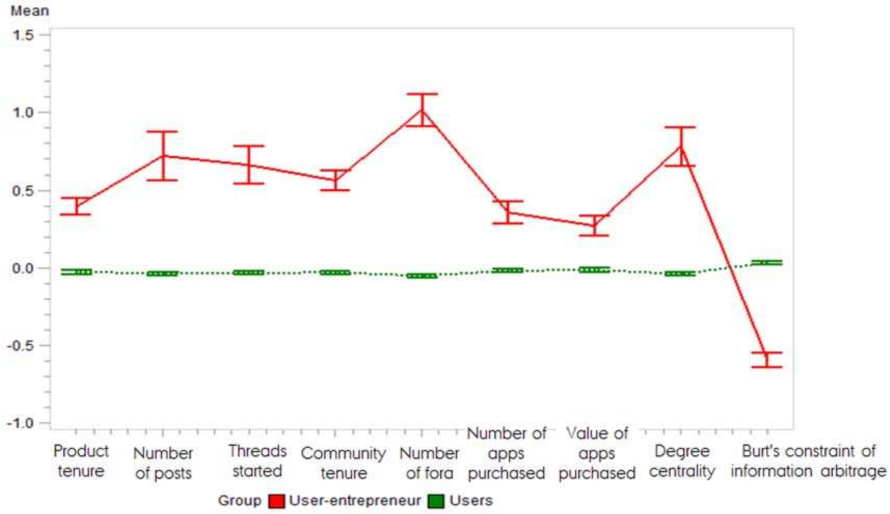


Figure 3: Profile plot of developers and users.

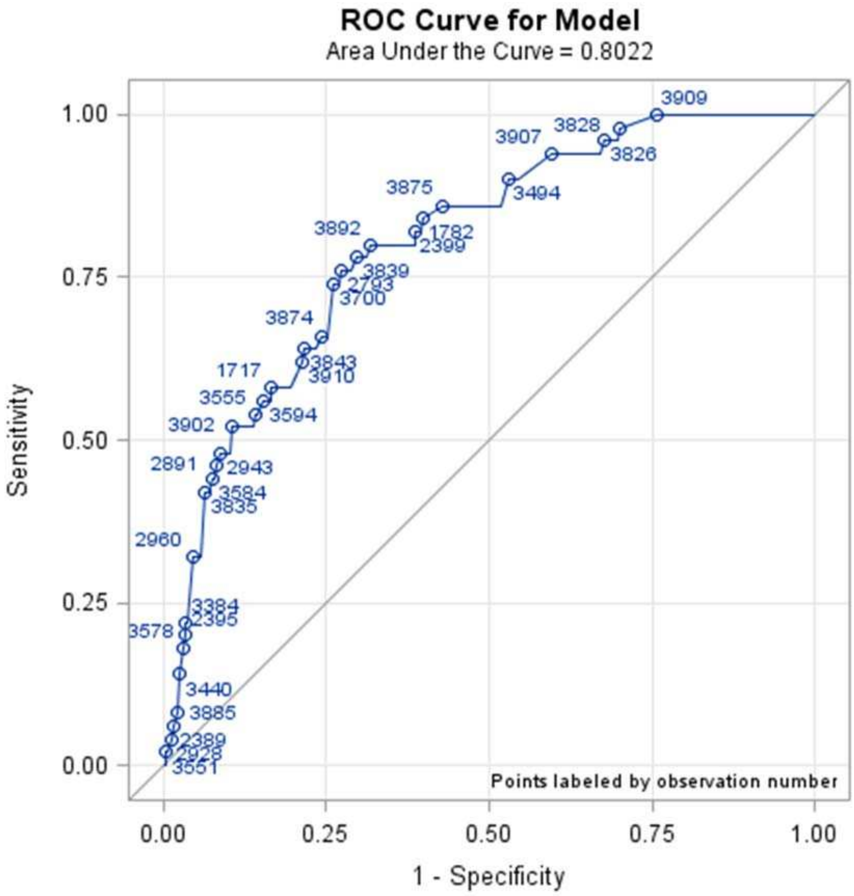


Figure 4: ROC curve for the prediction of new developers.

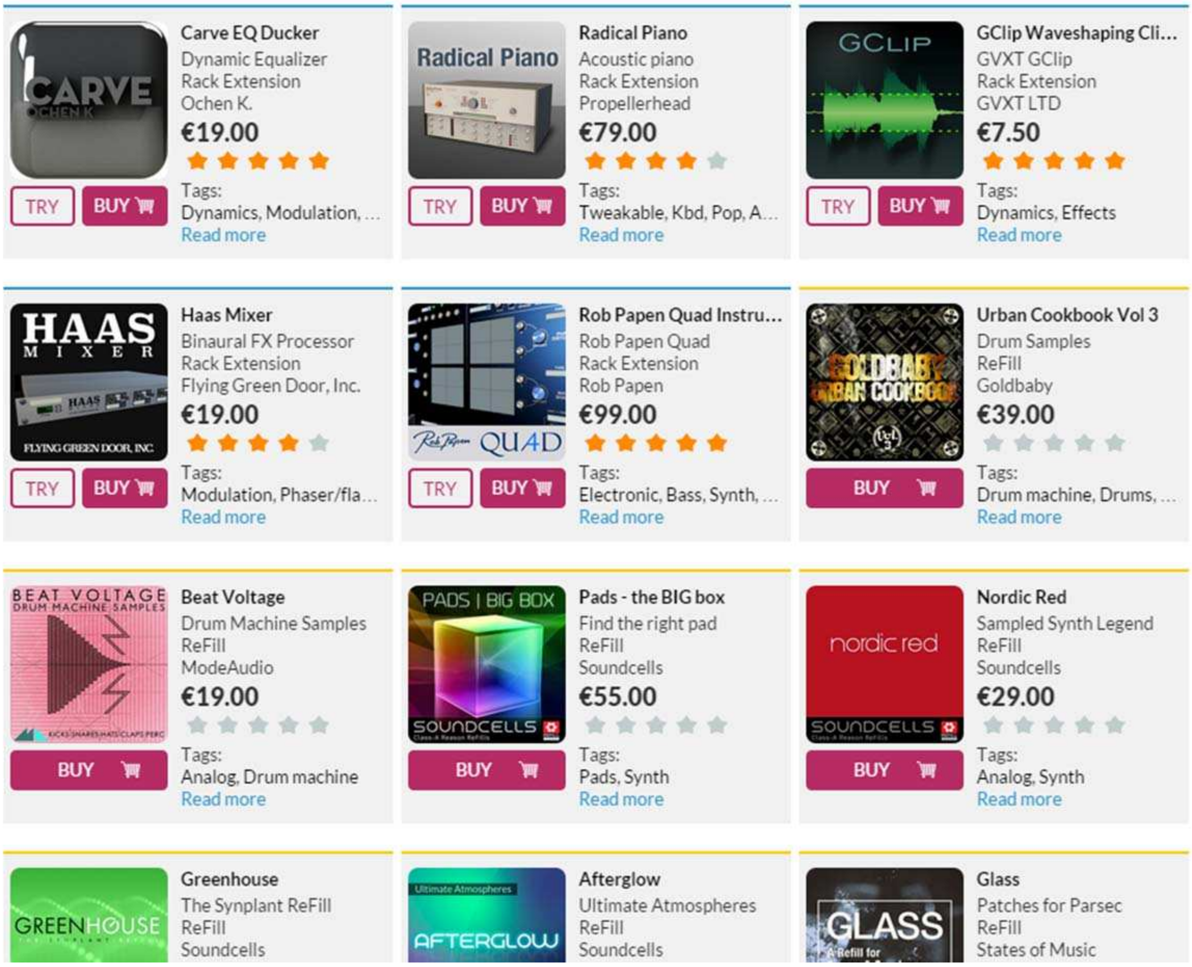


Figure 5: The platform's marketplace for user and company developed apps.

Table 1: Descriptive statistics and correlation table

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	N	Mean	Stddev	min	max	Skew	Kurtosis
1. Community posts														5,532	37.83	180.25	1	4,476	11.38	175.96
2. Threads started	0.666													5,532	3.26	9.09	0	165	8.00	86.03
3. Community tenure	0.203	0.239												5,532	601.00	106.84	384	750	-0.27	-1.07
4. Community contacts	0.137	0.105	0.092											5,532	2.59	7.38	0	82	4.96	29.30
5. Input from developers	0.989	0.669	0.208	0.147										4,763	37.48	165.25	1	3,359	10.49	147.35
6. Degree centrality	0.935	0.700	0.267	0.154	0.929									5,221	34.19	107.94	0	1,736	6.78	63.04
7. Prestige	0.925	0.729	0.273	0.156	0.928	0.996								5,221	17.03	50.62	0	706	6.28	52.16
8. Burts constraint	-0.228	-0.303	-0.300	-0.133	-0.227	-0.328	-0.343							5,148	0.42	0.35	0.0146	1.000	0.64	-1.04
9. Product registrations	0.156	0.113	0.030	0.056	0.173	0.184	0.189	-0.136						5,532	2.02	1.23	1	14	2.44	10.26
10. Products owned	0.140	0.111	0.032	0.057	0.152	0.169	0.174	-0.139	0.938					5,532	1.91	1.01	1	11	1.81	6.22
11. Product tenure	0.053	0.046	0.295	0.023	0.059	0.059	0.062	-0.036	0.406	0.421				5,532	587.33	78.15	384	750	0.22	-0.14
12. Number of apps purchased	0.177	0.171	0.036	0.048	0.171	0.250	0.251	-0.215	0.083	0.105	0.010			5,532	2.26	5.44	0	57	4.02	21.35
13. Value of apps purchased (USD)	0.154	0.152	0.039	0.038	0.149	0.221	0.222	-0.197	0.076	0.099	0.008	0.948		5,532	100.40	244.61	0	2,473	4.06	21.50
14. Downloads of free apps	0.001	0.035	-0.034	0.001	0.005	0.013	0.012	-0.009	0.084	0.106	0.085	0.078	0.052	5,532	5.23	13.65	0	236	5.44	41.67

* At N_{min}=4,763 correlations above |0.029| are significant at the 0.05 level, |0.037| at the 0.01 level, and |0.048| at the 0.001 level

Table2. Logistic regression on future Entrepreneurial intention

	Entrepreneurial intention		
Predictor variables	Full model	Naïvely reduced model	Regularized LASSO model†
Intercept	-7.4951***	-7.1508***	-6.1553
Community posts	-0.3483		
Threads started	0.1980		
Community tenure	0.0012		
Community contacts	0.3356	0.3896*	
Input from developers	0.2223		0.0334
Degree centrality	0.0118		
Prestige	0.5170	0.8167**	0.2705
Burts constraint	-0.1532		
Product registrations	-8.3604		
Products owned	8.7715	0.4924*	0.4932
Product tenure	0.0013		
Number of apps purchased	0.4658***	0.4683***	0.1741
Value of apps purchased			0.0263
Downloads of free apps	-0.1101		
Pseudo-R^2 (Max scaled)	0.1345	0.1248	
AIC	497.441	484.371	
-2 Log L	467.441	472.371	

*p<0.5; **p<0.01; ***p<0.001. Note: Significance of the naïvely reduced model has not been Bonferroni penalized

†: Only significant predictors reported. Estimation based on Tibshirani (1996), Friedman, Hastie, & Tibshirani (2010), and Lockhart, Taylor, Tibshirani, & Tibshirani (2014)

0

Table3. Poisson regression app unit sales

Predictor variables	Parameter estimates
Intercept	5.9312***
Threads started (range 1 to 4)	-0.0795***
Community contacts	0.0058*
Prestige	0.0537***
Products owned	0.0298***
Dummy for apps purchased	0.3456***
Developers' average selling price for his apps	0.0303***
Pseudo-R^2	0.3649

*p<0.5; **p<0.01; ***p<0.001.

Table3. Extended Cox regression on first app launches (time to event)

First app launch	
Predictor variables	Parameter estimates
Threads started	-0.0454*
Degree centrality	-0.0590***
Prestige	0.1292***
Product tenure	-0.0016*

*p<0.5; ***p<0.001.