Determinants of Trademarking: Evidence from Arizona and New Mexico Startups

Wang, Haoying and Dou, Shuming

New Mexico Tech

18 April 2018

Determinants of Trademarking: Evidence from Arizona and New Mexico Startups

Haoying Wang

(haoying.wang@nmt.edu)

Department of Business and Technology Management

New Mexico Tech

801 Leroy Pl, Socorro, NM 87801, USA

Shuming Dou

Department of Petroleum Engineering

New Mexico Tech

Abstract:

Trademarks are considered an important indicator of entrepreneurial innovation, especially among nontechnology-based service firms and startups. Therefore, it is essential to understand the motivations and drivers behind trademark applications to get a grasp of firm innovation behavior. This study focuses on the trademark decisions of startup firms. The paper assembles a unique dataset of startup firms linking firm trademark application and registration information with firm characteristics. The goal is to empirically examine the determinants of startup trademark decisions. The key results show that firm size is important, and startups of 51-200 employees have the highest propensity of seeking trademarks. Startup location, firm age, and firm type also matter. Within our study area, for example, startups in the Phoenix metro area are significantly more likely to file trademark applications than those in the Albuquerque metro area. Technology-related startups find trademarks less attractive compared to other startups.

Keywords:

Startup, Trademarks, Intellectual Property, SMEs

JEL Classifications: L26, O34, O39

1. Introduction

The continuing economic growth in the US southwest has provided plenty of opportunities for entrepreneurship and innovation activities. The growth is powered by both rapid population expansion and increasing government investments in the region. The area that is of particular interest is the lower southwest including Arizona (AZ) and New Mexico (NM). With a semi-arid climate and frequently constrained water supply, the economic growth potential in traditional sectors such as the food and agriculture sector is often concerned. With the spread of new information technologies and the expanding population, however, new opportunities in technology development and innovative services have inspired and attracted entrepreneurs locally and across the country. For example, the Phoenix metropolitan area in Arizona has become one of the hottest technology startup hubs across the nation in the last decade (Shandrow, 2017). Driving by the expansion of two national labs and defense-related sectors, the Albuquerque metropolitan area in New Mexico has also experienced a startup boom in the last decade (Albuquerque Innovation Central, 2017). A question often raised regarding the startup booms is will the startups bring new innovation blood to the economic engine and could eventually lead to long-term economic prosperity. The answer to this question hinges on the relationship between startups and economic innovation, which is further subject to other factors and conditions. Shan et al. (1994), for instance, show that the innovation output of startup firms depend on the interactive cooperation of the network to which the startups belong. In this paper, instead of answering the question directly, we seek to understand under what conditions and what kind of startups reveal relatively stronger propensity of innovation provided that trademarks can be used as a valid precursor and signal of innovative motivation. In the literature, trademarks are found to be a robust indicator of innovation among SMEs and other types of firms (e.g.

Mendonça et al., 2004; Flikkema et al., 2014). Empirically, this paper examines the relationship between trademark activities and startup characteristics in Arizona and New Mexico. The goal is to identify significant determinants of trademark activities that may help explaining startup innovation behavior.

Trademarks can be valuable and important in many ways to startup survival and growth. First, trademarks play an important role in product and/or service differentiation for startups. Product differentiation is important to the survival of SMEs (Leitner and Güldenberg, 2010). Often time, the SMEs can only focus on developing and marketing one or a few products/services. Strategically differentiating their products or services becomes very important to revenue potential and business survival. Trademarking may just be the right tool that serves the goal. Second, startups frequently face limited resources in marketing and business development. Due to such a resource scarcity, trademarks become an inexpensive way for startups to build and leverage assets, penetrate new markets, and signal business vitality (Block et al., 2015). Third, trademarks are an indicator of a startup's market orientation and revenue potential (Sandner and Block, 2011; Block et al., 2014). A clear market orientation and a strong revenue potential can help attract venture capital and build customer loyalty. In general, from a resource-based view (RBV), it is safe to argue that trademarks can play a critical and strategic role in startup planning and business development. Lastly, intellectual properties like trademarks and patents may also give the startup a better value when being acquired by another company (Farre-Mensa et al., 2016).

Beyond the valuation of trademarks from the firm strategic management perspective, a better grasp on the determinants of startup trademark activities can certainly help local governments and entrepreneurial infrastructures such as startup incubators and accelerators to effectively guide and support startup firms. Based on survey data from 600 SMEs in innovative industries, Block et al. (2015) find that SMEs seek trademark registration for three main reasons: protection, marketing, and exchange. For nontechnology-based service firms and startups, another important motive of trademarking is to signal the firm's innovative capability. According to another survey study of firms in the creative and cultural industries by Castaldi (2018), 'to show that we came up with something new' and 'to strengthen the company image' are among the most important motives of firm trademark decisions. In addition, firm characteristics are often linked to firm innovativeness and spending in innovative activities (Lefebvre and Lefebvre, 1992; Barker III and Mueller, 2002). The influence of firm characteristics in trademark decisions has been understudied in the literature. One of the contributions of this study is to examine the influence of several important firm characteristics on the propensity of trademark filing and registration decisions among startups.

It is worth noting that, in some industries, trademarks are more important and cost-efficient than patents as both a signal of innovation and a business asset. First of all, trademarks are less expensive to obtain compared to patents in almost all regards. Without accounting for any legal expenses, the current application fee for a trademark in the US is only around $275[1]. And the legal process of registering a trademark is much less involved compared to patents. Second, patents are difficult to establish and justify in some industries simply because of the nature of the business. In the fast-growing service industries, for example, trademarks/service marks have become a popular choice for branding and intellectual property protection compared to using patents as in the realm of technologies and physical products (Schmoch and Gauch, 2009). Third,

[1] https://www.uspto.gov/trademarks-application-process/filing-online/trademark-application-fee-structure, accessed April 1, 2018.

as Barnes (2006) argues, a trademark is not ‘intellectual’ *per se*. It has a very loose relation in disseminating new knowledge and enriching public domain (Barnes, 2006). This actually fits the profiles of many service industries and other industries whose tangible ‘intellectual assets’ are difficult to format and protect. In this case, it may be more efficient and useful for firms to accumulate the ‘intellectual assets’ into a valuable equity. Among different alternatives, trademark and service mark are a straightforward choice.

Previous studies have mostly focused on the significance and implications of deeming trademarks as an indicator of innovation and the linkage in between. For example, Mendonça et al. (2004) presented a comprehensive review and discussion on using trademarks as an indicator of firm innovation and technological progress. The key observation from their study is that there has been a positive correlation between the use of patents and the use of trademarks, which suggests using trademark applications and registrations as a complementary indicator of innovation. This paper expands the spectrum of the knowledge by further cracking into firm characteristics. Firm characteristics data, in general, is difficult to compile. This is particularly true in the case of startup firms. This study scrapes startup characteristics data across nine major websites including seven company business information websites and two websites that track company web domain history, which creates a unique dataset of startup characteristics. The empirical analysis relies on a discrete choice framework to model the decision of filing a trademark application. The key findings include: (1) the larger the startup size, the more likely it applies for a trademark, with startups of employment size 51-200 having the highest propensity; (2) AZ startups on average have a significantly higher propensity of applying for trademarks compared to NM startups; (3) the more mature a startup, the more likely it applies for a trademark, which suggests that startups often choose to strategically wait; (4) relative to other sectors, technology-related startups have a much lower propensity of trademarking, which may imply a preference of patents over trademarks among technology startups.

The remainder of the paper is organized as the follows: Second 2 discusses conceptual framework and empirical methodology. Section 3 describes data collection and the variables. Section 4 presents regression analysis results and their interpretation. Section 5 discusses the implications of the results and potential future research directions. Section 6 concludes.

2. Conceptual Framework and Empirical Method

A startup's decision of filing and registering trademarks is essentially an economic one. That is, the NPV of having a trademark should be large enough over the time horizon that the startup is strategically planning. If the planning horizon is short term and the goal is to maximize the startup's acquisition value, then the firm may want to build brand value and revenue potential early on. In this case, trademarking at the very beginning may be a rational choice. On the other hand, if the planning horizon is long term, then the startup may choose to strategically wait for the optimal time to start investing in brand-building and intellectual protection. A few related questions follow. Does firm type matter (e.g. nonprofit vs. privately held)? Does the industry or business category matter? What is the right time to apply for a trademark, conditional on the location of the startup? Is there an optimal firm size to reach before trademarking? To answer these questions, this paper uses a logistic discrete choice model (logit) based on random utility theory to explore the empirical data. Similar models have been used in the literature to study trademark and patent data (e.g. Frenz and Prevezer, 2012; Lasagni, 2012; De Vries et al., 2017).

Specifically, we derive the empirical model from a binary discrete choice framework of trademarking decision. A startup can choose to file a trademark application or not to maximize its utility (derived from the NPV of having a trademark or from doing the alternative). Suppose there are N startups in the sample, and the trademarking decision of the $i-th$ startup can be represented by the following latent model:

$$\phi_i^* = X_i\boldsymbol{\beta} + \varepsilon_i \tag{1}$$

where X_i is a matrix of variables which affect the expected outcome ϕ_i^* (i.e. net utility or achievement from having a trademark). ε_i is an idiosyncratic error associated with startup i, which is private knowledge and known to the firm management but unobservable to the researcher. β is the vector of parameters to be estimated. The researcher does not observe ϕ_i^*, and can only observe the discrete choice $\emptyset_i$ made by the startup, which is based on the following decision rule:

$$\begin{cases} \phi_i = 1 & if \ \phi_i^* > 0 \\ \phi_i = 0 & if \ \phi_i^* \leq 0 \end{cases} \tag{2}$$

The above model represents the startup's benefit-cost analysis of deciding to apply for a trademark if the net utility from doing so is positive. Since we cannot measure this net utility and what we can only observe is the actual decision the startup has made, the models in (1) and (2) cannot be estimated with usual linear regression methods. Instead, we can estimate a discrete choice model with a dependent variable representing the observable outcome $\emptyset_i$. One necessary assumption for this purpose is that ε_i is independently distributed with the logistic distribution (Maddala, 1983). Given the assumption, it follows that:

$$
\begin{aligned}
\Pr(\phi_i = 1 \mid X) &= \Pr(\phi_i^* > 0 \mid X) \\
&= \Pr(X_i\beta + \varepsilon_i > 0 \mid X) \qquad (3) \\
&= F(X_i\beta)
\end{aligned}
$$

where $F(\cdot)$ is the cumulative distribution function (CDF) for a logistic variable. Using this conditional probability, we can construct a log-likelihood function and estimate β – the parameters of interest. The log-likelihood estimation follows by maximizing the following function:

$$
\ln L = \sum_{i=1}^{N} \left\{ \phi_i \ln F(X_i\beta) + (1 - \phi_i) \ln\left[1 - F(X_i\beta)\right] \right\} \qquad (3)
$$

Note that the elements of the direct estimate $\hat{\beta}$ from the maximum likelihood estimation above cannot be interpreted simply as the marginal impact of the variable on the likelihood of applying for a trademark. For a particular coefficient estimate $\hat{\beta}_k$, it can be shown that the marginal effect of the estimate is $f(X\hat{\beta})\hat{\beta}_k$, where X is usually set at a vector of given values (e.g. mean or quantiles). $f(\cdot)$ is the logistic probability density function (PDF). The interpretation of marginal effects will be further elaborated in the results section.

3. Data

3.1 Data Collection

The data sample used for regression analysis is constructed in three parts and from multiple sources. All the data are publicly available. The first part of the data is the list of startup firms in AZ and NM, which is downloaded directly from the Angel's List (https://angel.co/). A quick overview of the Angel's List startup database reveals that almost all of the startup firms have some kind of unique graphical or textual design as their logo. It suggests a lot of them may have been taking the advantage of the common law trademark rights[2]. Among the original list of startups, we cleaned out all the 'fake' startups. This small portion of the startups is usually sole proprietorships who want to promote themselves through the List and its web traffic. For example, a doctor's office or clinic which operates a traditional business may list itself as a startup to increase its publicity. Startups without a registered business address or a valid physical address are also excluded, which we verified through Google search and public company databases such as Crunchbase (www.crunchbase.com) and Manta (www.manta.com). The second part of the data is the trademark filing and registration information matched to the startups in the list, which we compiled together by directly searching the US Patent and Trademark Office (USPTO) Trademark Electronic Search System (TESS). Sometimes, it is possible that a firm's name or slogan has been registered as a word mark by another firm. To eliminate such cases, we search for the trademark information by checking both firm name and registered owner address (at least to the level of city and state). The trademark information of a particular firm if available contains at least filing/application data, which is uniquely identified by a trademark serial number and the filing date. For some startups, both the filing information and the registration information are available. In these cases, the registration information is uniquely identified by a trademark registration number and the registration date. In rare cases, registered trademarks may be dead or abandoned. For the latter, an abandoned date is usually reported.

[2] The common law trademark rights are acquired through the use of a trademark even if a firm or entity does not register the mark with the USPTO. According to the USPTO, 'common law rights arise from actual use of a mark for particular goods or services and may allow the common law user to successfully challenge another party's use in court."

The third part of the data is the most important to this study – the firm characteristics. We collected this data by scraping major company business information websites and startup websites that focus on the US market. Specifically, we searched and extracted firm information from seven major websites. They include Crunchbase, Manta, Angel's List, LinkedIn (www.linkedin.com), Bloomberg (www.bloomberg.com), Bizapedia (www.bizapedia.com), and Buzzfile (www.buzzfile.com). From these websites, we were able to collect and validate the following information on each startup: (1) business name; (2) business location - city and state; (3) firm founded time (year); (4) industry or business category; (5) firm size in terms of employment; (6) the number of followers on the Angel's List. From this information, we then construct all the variables needed for the regression analysis. The details of the variables will be discussed in the next section. One thing to note is regarding the firm founded time. For a small number of startups, their founded time is difficult to find through searching these websites. Instead, we use two web domain tracking websites to determine the time of the earliest activity (the Whois) on the startup's published websites. By cross-checking on these two tracking sites, DomainTools (www.domaintools.com) and Web Archive (https://web.archive.org), we can get a very good approximate founded year for each of these startups.

3.2 Variables

Based on the firm information collected, we derive the following explanatory variables which are hypothesised to be important in explaining the startups' trademarking decisions.

Firm type – the startups are grouped into three types: nonprofit (1), private company (2, including sole proprietorship, partnership, and privately held firms), and public company (3). Most of the startups fall into the private company category. Note that the distinctions among sole proprietorship, partnership, and privately held may not be drawn precisely on some of the websites. Therefore, it makes a more reliable variable by grouping them together.

Firm size - the variable is coded based on the reported employment size range. Specifically, firm size = 1 if with an employment size of 1-10; firm size = 2 if with an employment size of 11-50; firm size = 3 if with an employment size of 51-200; firm size = 4 if with an employment size of 201-500; firm size = 5 if with an employment size above 500. Most of the startups are SMEs as expected. In some cases, for example, the service industries, firm size may be large (4 or 5). But it is only a small portion of the sample.

Firm age - which is directly computed as the difference between current year (2018) and the firm founded year. Given that the dependent variable is defined as the trademark status of the firm (filed or not) as of 2018, firm age is expected to be a relevant explanatory variable. A related variable here is the year at which the startup filed a trademark application if ever. The variable becomes relevant when examining the determinants of successful trademark registrations among all trademark applications filed. This will be further explored in the results section.

The number of followers on the Angel's List - this variable simply indicates the popularity of a startup or the attention it has received on the website. Since the Angel's List also has a lot of registered investors, this variable may help to reveal the relationship between investor influence and trademark decisions. The number of followers is also re-coded to eliminate the impacts of outliers. Specifically, number of followers = 1 if with 0-5 followers; number of followers = 2 if with 6-15 followers; number of followers = 3 if with 16-50 followers; number of followers = 4 if with 51-100 followers; number of followers = 5 if with above 100 followers.

State indicator - this indicator variable is designed to capture the average difference in the propensities of applying for a trademark between AZ startups and NM startups.

Metropolitan area indicator - similarly, this indicator variable is designed to capture the average difference in the propensities of applying for a trademark between metro startups and non-metro startups. In this study, metro startups are defined as those located in the Phoenix metropolitan area and the Albuquerque metropolitan area. The rest of the startups are defined as non-metro ones.

Technology-related firm indicator, *Internet-related firm indicator* - it is expected that startups in different industries may have a different propensity of applying for a trademark or a different preference between trademark and patent. These two indicators are designed to capture such differences. The startups are categorized into technology-related, Internet-related, or others based on the business category (tags) they choose on the Angel's List. Examples of technology-related startups are medical devices, manufacturing, biotech, and nanotechnology. Examples of Internet-related startups are mobile health, crowdsourcing, big data, and mobile commerce. Others include mostly service businesses and other not well-categorized businesses.

3.3 Descriptive Statistics

Table 1 summarizes the descriptive statistics of the binary dependent variable and all the explanatory variables defined above. There are 632 startups in total included in the sample. Among which and as of April 1, 2018, 177 have filed a trademark application. Among the 177 filings, 120 trademarks are officially registered with the USPTO as of April 1, 2018. A few additional notable patterns can be directly observed from the summary statistics. First, most of the startups are small in terms of employment size as expected. Second, the average startup is about six years old. Third, on average the startups apply for trademarks between year 2 and 3. Lastly, within the sample, close to 20% are technology-related startups; and a little less than 30% are Internet-related startups.

Table 1: Summary statistics of variables

Variable	N	Mean	Std. Dev.	Min	Max
Filed a trademark (yes = 1, no = 0)	632	0.28	0.45	0	1
Startup state (AZ = 0, NM = 1)	632	0.39	0.49	0	1
Startup location (metro = 1, non-metro = 0)	632	0.80	0.40	0	1
Number of followers in the Angel's List	632	1.81	0.97	1	5
Startup age (year)	632	6.03	4.50	0	39
Startup type	632	2.01	0.26	1	3
Startup employment size	632	1.36	0.69	1	5
Trademark filing time (calendar year)	177	2013.14	3.43	1999	2018
Technology-related startup (yes = 1)	632	0.19	0.39	0	1
Internet-related startup (yes = 1)	632	0.27	0.44	0	1

4. Analysis and Results

Table 2 and Table 3 show the main estimation results with the proposed logit model framework. Given the nature of the estimates in discrete choice models as mentioned before, we present the marginal effects of the estimates in this section for the sake of easy interpretation. The corresponding original coefficient estimates can be found in the Appendix. In Table 2, the analysis focuses on the following determinants of trademark decision: potential state-level fixed effects, being located in a metropolitan area, the number of followers on the Angel's List; startup age in years, startup firm type, and startup employment size. To ensure the robustness of the empirical results, different model specifications are estimated. The specifications are designed based on the way of including different fixed effects and how the employment size variable (in cardinal numbers) is specified. Specification (1) explores the state fixed effects (the difference between AZ and NM on average). Specification (2) focuses on the differences between metro and non-metro areas. Specification (3) examines all the combinations of the state fixed effects and the metropolitan effects, where the combination of New Mexico - Non-metro is omitted to avoid collinearity. Specification (4) focuses on the fixed effect of each employment size level, where the employment size 1 (1-10 employees) group is omitted, similarly to avoid collinearity. Specification (5) tries to see if there are any interaction effects between startup size and startup age.

Table 2: Logit regression of trademark decision (trademark = 1): marginal effects

Variable	Model (1)	(2)	(3)	(4)	(5)
Startup state (AZ = 0, NM = 1)	-0.1510*** (0.0390)				
Startup location (metro = 1, non-metro = 0)		0.0431 (0.0461)			
Startup - Arizona, Metro			0.1009* (0.0532)	0.1074** (0.0528)	0.0999* (0.0532)
Startup - Arizona, Non-metro			0.0005 (0.0908)	0.0003 (0.0903)	0.0021 (0.0908)
Startup - New Mexico, Metro			-0.0980 (0.0630)	-0.0950 (0.0626)	-0.0967 (0.0630)
Number of followers	-0.0171 (0.0187)	0.0051 (0.0180)	-0.0174 (0.0186)	-0.0234 (0.0185)	-0.0190 (0.0187)
Startup age (year)	0.0073* (0.0038)	0.0080** (0.0038)	0.0068* (0.0038)	0.0071* (0.0038)	0.0077** (0.0039)
Startup type 2 - private company	-0.0316 (0.1034)	-0.0067 (0.1044)	-0.0293 (0.1031)	-0.0224 (0.1020)	-0.0244 (0.1033)
Startup type 3 - public company	-0.1000 (0.1364)	-0.0386 (0.1376)	-0.0914 (0.1363)	-0.0267 (0.1343)	-0.0822 (0.1363)
Startup employment size (continuous)	0.0797*** (0.0248)	0.0849*** (0.0251)	0.0784*** (0.0249)		0.0867*** (0.0265)
Startup employment size (=2)				0.1454*** (0.0380)	
Startup employment size (=3)				0.2769*** (0.0798)	
Startup employment size (=4)				0.0017 (0.1377)	
Startup employment size (=5)				-0.1015 (0.2271)	
Interaction: startup age * startup size	No	No	No	No	Yes
Number of observations			632		
Pseudo R^2	0.0499	0.0315	0.0554	0.0726	0.0563

Note: throughout the article, asterisks (*, **, ***) indicate statistical significance at 10%, 5%, and 1% level, respectively, unless otherwise noted. Standard errors in the table are computed with the Delta method.

Across all five specifications, we can make the following three observations based on the statistical significance of the results. First, startups in AZ on average have a higher propensity of filing a trademark application. The difference between AZ startups and NM startups is about 15%, which is an extremely significant result. Second, the age of startups does matter. As a startup firm gets older (more mature and established), its likelihood of filing a trademark application increases. The rate of increase is close to 0.8% per year and it is statistically significant at 10% level. This is an intuitive result given that most of the startups are too busy with other priorities (e.g. attract funds and product/service development) at the beginning stages. Lastly, the firm size as measured by employment size is important. The relationship between firm size and the likelihood of filing a trademark application, however, is not strictly linear. Although on average the likelihood of filing a trademark application increases by roughly 8% as startup employment size jumps by one level, the groups with higher propensities are the small to medium size ones, namely the group 2 and 3 (with 11-200 employees). Note that, we have considered the potential interaction effects between startup size and startup age. The coefficient estimate for the interaction term is insignificant. If we look at the marginal effects of startup age and startup size, the change of the former is very small. The marginal effect of startup size, however, has roughly a half percent increase compared to the other three estimates (specifications 1-3). Unfortunately, given a standard error estimate above 2% across specifications the half percent increase cannot be considered significant. Therefore, we can conclude that there are no significant interaction effects between startup size and startup age.

Another important determinant to explore is the industry or business category that a startup associates itself with. Table 3 reports results of exploring technology-related startups and Internet-related startups. From all three specifications, we can see that the technology-related startups have a significantly lower propensity of filing a trademark application compared to the rest of the startups (mostly service businesses and traditional businesses). The magnitude is almost 10% and it is highly significant. For the Internet-related startups, their propensity of filing a trademark application is no difference from the rest of the startups. This is an interesting and important result. It essentially suggests that the technology-related startups may have a strong preference for patents over trademarks (Graham et al., 2009). Though it is intuitively true, patent data for all startups are necessary to verify the conjecture, which points to an interesting aspect of future research.

Table 3: Logit regression exploring the differences among startup industries: marginal effects

Variable	Model		
	(1)	(2)	(3)
Startup - Arizona, Metro	0.1022** (0.0524)	0.1058** (0.0528)	0.1022** (0.0525)
Startup - Arizona, Non-metro	-0.0115 (0.0901)	-0.0013 (0.0903)	-0.0115 (0.0901)
Startup - New Mexico, Metro	-0.1024* (0.0623)	-0.0950 (0.0625)	-0.1024* (0.0623)
Number of followers	-0.0234 (0.0184)	-0.0238 (0.0185)	-0.0234 (0.0184)
Startup age (year)	0.0060 (0.0037)	0.0071* (0.0038)	0.0060 (0.0038)
Startup type 2 - private company	-0.0294 (0.1009)	-0.0277 (0.1023)	-0.0293 (0.1013)
Startup type 3 - public company	-0.0204 (0.1335)	-0.0298 (0.1343)	-0.0204 (0.1336)
Startup employment size (=2)	0.1634*** (0.0387)	0.1471*** (0.0381)	0.1634*** (0.0387)
Startup employment size (=3)	0.2958*** (0.0803)	0.2799*** (0.0800)	0.2958*** (0.0804)
Startup employment size (=4)	0.0277 (0.1385)	0.0060 (0.1379)	0.0276 (0.1385)
Startup employment size (=5)	-0.0419 (0.2265)	-0.0959 (0.2273)	-0.0419 (0.2265)
Technology-related startup (yes = 1)	-0.0986** (0.0472)		-0.0987** (0.0489)
Internet-related startup (yes = 1)		0.0227 (0.0387)	-0.0002 (0.0400)
Number of observations		632	
Pseudo R^2	0.0787	0.0731	0.0787

It is worth noting that the decision of filing a trademark application and the chance of getting a trademark registered (succeeding a trademark registration with the USPTO) is supposed to be different. The question is how it is different. This sub-section devotes to answering the question. Table 4 presents the estimation of logit models with trademark registration (Yes=1, No=0) as the dependent variable. Note that the logit models here are conditional on the decision of filing a trademark application by a startup. The data sample has therefore been reduced from 632 to 177. Figure 1 shows the distribution of these 177 registrations over time. In the new discrete choice model framework, the chance of getting a trademark registered is explained by a similar set of variables except for one new variable - the trademark filing time (measured in the calendar year). Intuitively, the later (the closer to 2018) a trademark application is filed, the less likely it gets registered because the trademark review and approval process takes several steps and it can be time-consuming when there is a queue at USPTO. Therefore, it is hypothesised that this variable has a negative sign.

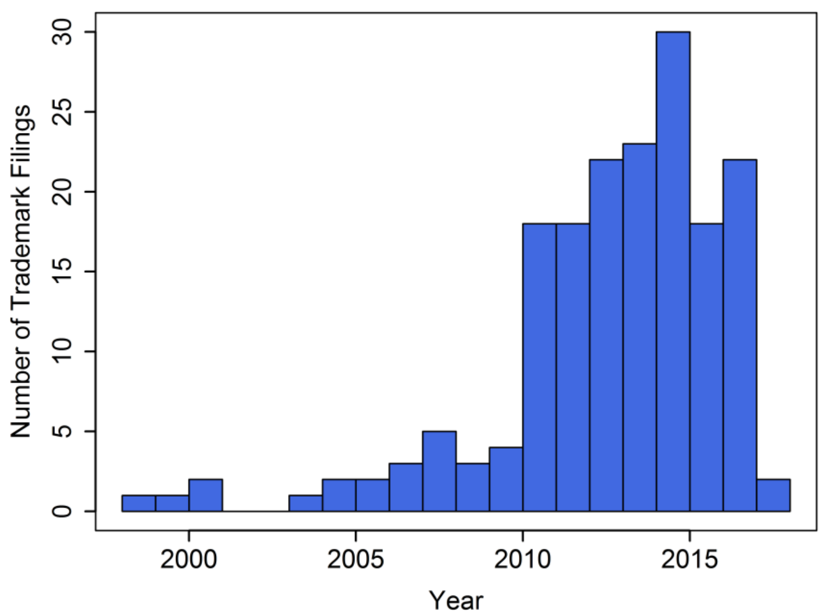

Figure 1: Distribution of trademark application years among startups

Table 4: Logit regression of trademark registration among filings: marginal effects

Variable	Model (1)	(2)	(3)
Startup state (AZ = 0, NM = 1)	-0.0475 (0.0790)		
Startup location (metro = 1, non-metro = 0)		0.0331 (0.0897)	
Startup - Arizona, Metro			0.0780 (0.1034)
Startup - Arizona, Non-metro			0.1620 (0.1934)
Startup - New Mexico, Metro			0.0689 (0.1304)
Number of followers	0.0701* (0.0393)	0.0725* (0.0390)	0.0689* (0.0391)
Trademark filing time (year)	-0.0281** (0.0116)	-0.0285** (0.0116)	-0.0278** (0.0115)
Startup type 2 - private company	-0.0893 (0.2381)	-0.0817 (0.2386)	-0.0998 (0.2383)
Startup type 3 - public company	-0.3227 (0.2769)	-0.3042 (0.2752)	-0.3368 (0.2772)
Startup employment size (continuous)	0.0111 (0.0461)	0.0127 (0.0460)	0.0143 (0.0461)
Technology-related startup (yes = 1)	-0.0268 (0.1018)	-0.0214 (0.1015)	-0.0161 (0.1028)
Internet-related startup (yes = 1)	-0.0697 (0.0782)	-0.0628 (0.0777)	-0.0660 (0.0786)
Number of observations		177	
Pseudo R^2	0.0605	0.0595	0.0628

Looking at Table 4 and comparing to the logit models of trademark filing decision (Table 2 and Table 3): (1) the state-level heterogeneous effects disappear; (2) startup employment size is no longer an important factor; (3) industry and business category do not matter anymore. These major shifts in results could be due to the large reduction in the sample size, but it is more likely due to the change of the model structure (how the dependent variable is defined). Even with a much smaller sample size, an interesting and important result emerges. The number of followers on the Angel's List (the same website also has a lot of registered investors) has a significant positive impact on the chance of succeeding the trademark registration. A one-level increase in the number of followers (as defined in the *variables* section) can lead to roughly a 7% jump in the chance of succeeding the trademark registration. As elaborated before, this variable indicates the popularity of a particular startup on the Angel's List website. To a large extent, a larger value implies receiving more attention and interests from potential investors (e.g. venture capital) compared to startups with fewer followers. This is consistent with the findings in the literature that investor influence and advisory are one of the key drivers in seeking patent and innovation (Engel and Keilbach, 2007; De Vries et al., 2017). In addition, the results in Table 4 support the trivial hypothesis that the later a trademark application is filed, the less likely it gets registered, which is straightforward.

5. Discussion

As the economy grows, especially following the rapid expansion of innovative service industries, the trademark applications have soared in the US and around the world (Block et al., 2015). According to the World Intellectual Property Organization (WIPO), from 2000 to 2016 (roughly corresponding to our study period), the total number of trademark applications in the US has increased from 292,464 to 393,210 (an increase of 34%)[3]. Though the startups account only for a small portion of the increase, they have the potential to leverage a much larger economic value through innovation. At the same time, failure is quite common among startups. According to the US Bureau of Labor Statistics (BLS), over half of the businesses have died in the first 5-6 years[4]. The survival rate for small businesses which includes most of the startups is even lower. Trademark is a potential tool that can help reduce startup failures and increase small business survival rates.

[3] https://www3.wipo.int/ipstats/index.htm?tab=trademark, accessed April 1, 2018.

[4] https://www.bls.gov/bdm/entrepreneurship/bdm_chart3.htm, accessed April 1, 2018.

Related to the findings in this study, several insights can be drawn. Hopefully, they can shed some light on startup planning and strategic management and be found useful among startup founders, investors, as well as startup incubator and accelerator managers. First, a geographically-based corporate environment that supports startups is important for startup innovation. Such an environment implies strong network effects and rich learning opportunities (Baum et al., 2000). It helps startups avoid common mistakes in the early stages. As a whole or a startup cluster, it can also increase the chance of matching investment funds for each startup. Phoenix metropolitan area is a good example of such an environment. Second, we have observed a significant positive relationship between startup age and trademark application. It can be interpreted as that startups are more likely to wait until being well-established or reaching certain scale before claiming a registered trademark. The US Small Business Administration (SBA) finds that as small businesses get older (hence better established) the survival rate flattens out (SBA, 2012). This suggests an interesting but likely sophisticated relationship among startup age, business survival, and trademark decisions, which points to a potential direction for future research. Third, being consistent with the findings in this paper, trademarking is an important strategy for SMEs in practice. Large startups and very small businesses care much less. In other words, there may be an optimal range in terms of firm size for startups to become seriously considering trademarks. To some extent, this echoes the previous elaboration on startup age and trademark applications. One potential explanation is that for individuals and very small businesses the potential litigation risk over intellectual properties can be costly (Lanjouw and Schankerman, 2004). The insignificant effects found associated with larger firm sizes may be simply due to a small number of observations falling in those categories in this study. Given that the focus of this study is startups who are mostly SMEs, any further attempts to interpret the results on the tails (very small or very large firms) become a stretch. Lastly, it is well-known that many service firms do not patent and the patent system has been deployed mostly to deal with more tangible innovations (Miles et al., 2000). This study confirms this early observation. In addition, we find that Internet-related and digital industries have more interests and needs for trademarks compared to technology-related industries. The software industry, for example, makes a good case here. One explanation is that most of the software firms especially the smaller ones do not consider patenting as a primary strategy for attaining competitive advantage (Samuelson, 2010).

Previous research has shown that trademarks are a good indicator of innovation and corporate image (e.g. Mendonça, 2004; Flikkema et al., 2014), together with the findings in this study, important implications for local government economic development agencies and small business administrative agencies can be drawn. First of all, creating and facilitating a network of SMEs and startups are important. As mentioned earlier, this helps to generate learning and spillover effects among firms - a necessary and fundamental component for cultivating large industrial clusters. Second, innovation and therefore economic growth can emerge in very different ways, which necessitates the co-existing of industries when promoting startups (Meuer et al., 2015). Focusing all of the support on one industry (e.g. biotech or e-commerce) may create a comparative advantage in the short run but could seriously undermine the engine of economic growth in the long run.

6. Conclusions

This paper has demonstrated the feasibility of scraping public information on the Internet to provide insights into important entrepreneurship and innovation policy issues. Some of the questions associated can be extremely difficult to answer without otherwise collecting information from each startup directly. The advantage of our approach is that we can potentially include the entire population of the startups into the analysis at a low cost. Though we could not be as comprehensive as a survey study of startups which has been done in the literature in terms of the information collected, the scraped data allows us to show a complete picture of the startup landscape in the study region. Our empirical analysis, focusing on the startups in AZ and NM, reveals several important results regarding the innovation behavior of startups. Specifically, (1) the propensities of innovation as indicated by trademark activities vary geographically and it is likely driven by the spatial heterogeneities in the corporate environment for SMEs. (2) Startups tend to use trademarks as both an intellectual property safeguard and an innovation strategy conditional on factors such as firm age and employment size. (3) Technology-related startups find trademarks less attractive in comparison to other startups.

The key implications of the findings can be summarized into two aspects. For early-stage entrepreneurs, startup planning is not just about the technology or the business idea itself. Protecting the technology or the idea and growing values around it are also critical to the survival of a startup in the long run. Trademark is an important tool and asset to reach the goal. For entrepreneurship infrastructures (e.g. government agencies, business incubators and accelerators, and startup loan programs), the focus should not stay on each of the individual startups. Creating and growing a network of startups is the big picture. It is usually difficult for an individual startup to promote a startup alliance or an innovation cluster. The entrepreneurship infrastructure has a critical role to play here.

This study has several limitations that leave hints to future research. First, this study basically uses a cross-sectional approach to a more-or-less dynamic problem. Though for this particular study and its research question the cross-sectional approach serves the goal, for more specific and more structural research questions such as what is the optimal time to file a trademark application a dynamic modeling approach becomes necessary. Another limitation is related to the data sample. The startup information used in this study was compiled from multiple sources. The information among these sources may not be perfectly consistent with other each. Addressing the related measurement errors issue without knowing the source of the errors goes beyond the scope of this paper. However, a new study dedicated to the issue can become very valuable. It will definitely help expanding the bandwidth of entrepreneurship and innovation research in the era of information technology.

Reference

Albuquerque Innovation Central. (2017) Albuquerque is emerging as one of the hottest new 'Startup Hub' cities. http://innovationcentralabq.com/albuquerque-emerging-one-hottest-new-startup-hub-cities-challenging-silicon-valley/, accessed April 1, 2018.

Baum, J.A., Calabrese, T., and Silverman, B.S. (2000) Don't go it alone: Alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, 21, pp. 267-294.

Barker III, V.L. and Mueller, G.C. (2002) CEO characteristics and firm R&D spending. *Management Science*, 48(6), pp. 782-801.

Barnes, D.W. (2006) A new economics of trademarks. *Northwestern Journal of Technology and Intellectual Property*, 5 (1), pp. 22.

Block, J.H., De Vries, G., Schumann, J.H., and Sandner, P. (2014) Trademarks and venture capital valuation. *Journal of Business Venturing*, 29(4), pp. 525-542.

Block, J.H., Fisch, C.O., Hahn, A., and Sandner, P.G. (2015) Why do SMEs file trademarks? Insights from firms in innovative industries. *Research Policy*, 44(10), pp. 1915-1930.

Castaldi, C. (2018) To trademark or not to trademark: the case of the creative and cultural industries. *Research Policy*, 47(3), pp. 606-616.

De Vries, G., Pennings, E., Block, J.H., and Fisch, C. (2017) Trademark or patent? The effects of market concentration, customer type and venture capital financing on start-ups' initial IP applications. *Industry and Innovation*, 24(4), pp. 325-345.

Engel, D. and Keilbach, M. (2007) Firm-level implications of early stage venture capital investment - an empirical investigation. *Journal of Empirical Finance*, 14(2), pp. 150-167.

Farre-Mensa, J., Hegde, D., and Ljungqvist, A. (2016) The bright side of patents. Working Paper No. w21959. Cambridge, MA: National Bureau of Economic Research.

Flikkema, M., De Man, A.P., and Castaldi, C. (2014) Are trademark counts a valid indicator of innovation? Results of an in-depth study of new benelux trademarks filed by SMEs. *Industry and Innovation*, 21(4), pp. 310-331.

Frenz, M. and Prevezer, M. (2012) What can CIS data tell us about technological regimes and persistence of innovation? *Industry and Innovation*, 19(4), pp. 285-306.

Graham, S.J., Merges, R.P., Samuelson, P., and Sichelman, T. (2009) High technology entrepreneurs and the patent system: results of the 2008 Berkeley Patent Survey. *Berkeley Technology Law Journal*, 24(4), pp. 1255-1327.

Lanjouw, J.O. and Schankerman, M. (2004) Protecting intellectual property rights: are small firms handicapped? *Journal of Law and Economics*, 47(1), pp. 45-74.

Lasagni, A. (2012) How can external relationships enhance innovation in SMEs? New evidence for Europe. *Journal of Small Business Management*, 50(2), pp. 310-339.

Lefebvre, E. and Lefebvre, L.A. (1992) Firm innovativeness and CEO characteristics in small manufacturing firms. *Journal of Engineering and Technology Management*, 9(3-4), pp. 243-277.

Leitner, K.H. and Güldenberg, S. (2010) Generic strategies and firm performance in SMEs: a longitudinal study of Austrian SMEs. *Small Business Economics*, 35(2), pp. 169-189.

Maddala, G.S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge MA: Cambridge University Press.

Mendonça, S., Pereira, T.S., and Godinho, M.M. (2004) Trademarks as an indicator of innovation and industrial change. *Research Policy*, 33(9), pp. 1385-1404.

Meuer, J., Rupietta, C., and Backes-Gellner, U. (2015) Layers of co-existing innovation systems. *Research Policy*, 44(4), pp. 888-910.

Miles, I., Andersen, B., Boden, M., and Howells, J. (2000) Service production and intellectual property, *International Journal of Technology Management*, 20(1/2), pp. 95-115.

Samuelson, P. (2010) Why do software startups patent (or not)? *Communications of the ACM*, 53(11), pp. 30-32.

Sandner, P. G. and J. H. Block. (2011) The market value of R&D, patents and trademarks. *Research Policy*, 40(7), 969-985.

SBA. (2012) Do economic or industry factors affect business survival? *SBA Small Business Facts*, June 2012. https://www.sba.gov/sites/default/files/Business-Survival.pdf, accessed April 1, 2018.

Schmoch, U. and Gauch, S. (2009) Service marks as indicators for innovation in knowledge-based services. *Research Evaluation*, 18(4), pp. 323-335.

Shan, W., Walker, G., and Kogut, B. (1994) Interfirm cooperation and startup innovation in the biotechnology industry. *Strategic Management Journal*, 15(5), pp. 387-394.

Shandrow, K.L. (2017) Phoenix rising: how a tech startup oasis emerged in the Arizona desert. https://www.freeenterprise.com/how-phoenix-became-a-tech-startup-oasis-in-the-desert/, accessed April 1, 2018.

Appendix: Supplementary Results

Table A1: Logit regression of trademark decision: coefficient estimates

Variable	Model				
	(1)	(2)	(3)	(4)	(5)
Startup state (AZ = 0, NM = 1)	-0.7974*** (0.2127)				
Startup location (metro = 1, non-metro = 0)		0.2225 (0.2384)			
Startup - Arizona, Metro			0.5358* (0.2851)	0.5833** (0.2898)	0.5313* (0.2854)
Startup - Arizona, Non-metro			0.0025 (0.4822)	0.0018 (0.4906)	0.0114 (0.4824)
Startup - New Mexico, Metro			-0.5203 (0.3363)	-0.5160 (0.3414)	-0.5142 (0.3366)
Number of followers	-0.0903 (0.0988)	0.0265 (0.0930)	-0.0921 (0.0990)	-0.1271 (0.1008)	-0.1011 (0.0996)
Startup age (year)	0.0383* (0.0203)	0.0412** (0.0201)	0.0359* (0.0203)	0.0386* (0.0207)	0.0668* (0.0413)
Startup type 2 - private company	-0.1670 (0.5458)	-0.0344 (0.5392)	-0.1558 (0.5477)	-0.1218 (0.5542)	-0.1299 (0.5493)
Startup type 3 - public company	-0.5281 (0.7210)	-0.1993 (0.7110)	-0.4857 (0.7248)	-0.1451 (0.7298)	-0.4371 (0.7254)
Startup employment size (continuous)	0.4208*** (0.1347)	0.4389*** (0.1337)	0.4161*** (0.1357)		0.5768*** (0.2322)
Startup employment size (=2)				0.7898*** (0.2145)	
Startup employment size (=3)				1.5045*** (0.4484)	
Startup employment size (=4)				0.0093 (0.7481)	
Startup employment size (=5)				-0.5516 (1.2347)	
Startup age * Startup employment size					-0.0179 (0.0209)
Number of observations	632				
Pseudo R^2	0.0499	0.0315	0.0554	0.0726	0.0563

Table A2: Logit regression exploring the differences among startup industries: coefficient estimates

Variable	**Model**		
	(1)	**(2)**	**(3)**
Startup - Arizona, Metro	0.5596** (0.2898)	0.5751** (0.2901)	0.5597** (0.2901)
Startup - Arizona, Non-metro	-0.0630 (0.4934)	-0.0073 (0.4911)	-0.0630 (0.4934)
Startup - New Mexico, Metro	-0.5608* (0.3425)	-0.5163 (0.3414)	-0.5608* (0.3426)
Number of followers	-0.1281 (0.1012)	-0.1295 (0.1008)	-0.1281 (0.1013)
Startup age (year)	0.0329 (0.0209)	0.0387* (0.0207)	0.0329 (0.0209)
Startup type 2 - private company	-0.1609 (0.5526)	-0.1504 (0.5561)	-0.1606 (0.5546)
Startup type 3 - public company	-0.1118 (0.7307)	-0.1618 (0.7703)	-0.1115 (0.7316)
Startup employment size (=2)	0.8944*** (0.2215)	0.7996*** (0.2153)	0.8944*** (0.2215)
Startup employment size (=3)	1.6193*** (0.4564)	1.5218*** (0.4496)	1.6192*** (0.4567)
Startup employment size (=4)	0.1515 (0.7580)	0.0329 (0.7498)	0.1513 (0.7584)
Startup employment size (=5)	-0.2294 (1.2399)	-0.5213 (1.2361)	-0.2294 (1.2400)
Technology-related startup (yes = 1)	-0.5400** (0.2610)		-0.5404** (0.2704)
Internet-related startup (yes = 1)		0.1233 (0.2106)	-0.0013 (0.2190)
Number of observations	632		
Pseudo R^2	0.0787	0.0731	0.0787

Table A3: Logit regression of trademark registration among filings: coefficient estimates

Variable	**Model**		
	(1)	**(2)**	**(3)**
Startup state (AZ = 0, NM = 1)	-0.2346 (0.3917)		
Startup location (metro = 1, non-metro = 0)		0.1630 (0.4457)	
Startup - Arizona, Metro			0.3867 (0.5159)
Startup - Arizona, Non-metro			0.8030 (0.9658)
Startup - New Mexico, Metro			0.3413 (0.6483)
Number of followers	0.3461* (0.1998)	0.3573* (0.1984)	0.3415* (0.1993)
Trademark filing time (year)	-0.1390** (0.0602)	-0.1407** (0.0604)	-0.1375** (0.0599)
Startup type 2 - private company	-0.4410 (1.1770)	-0.4029 (1.1775)	-0.4946 (1.1827)
Startup type 3 - public company	-1.5932 (1.3849)	-1.4995 (1.3723)	-1.6689 (1.3934)
Startup employment size (continuous)	0.0549 (0.2277)	0.0627 (0.2267)	0.0708 (0.2289)
Technology-related startup (yes = 1)	-0.1324 (0.5032)	-0.1054 (0.5006)	-0.0797 (0.5097)
Internet-related startup (yes = 1)	-0.3441 (0.3896)	-0.3095 (0.3860)	-0.3270 (0.3925)
Number of observations	177		
Pseudo R^2	0.0605	0.0595	0.0628