Business Accelerators: evidence from Start-Up Chile[1]

Juanita Gonzalez-Uribe[2] and Michael Leatherbee[3]

We estimate the effect of business accelerators on start-up performance. We focus on the case of Start-Up Chile (SUP), an accelerator sponsored by the Chilean government, which provides 40,000 USD (equity free) in seed capital, a work visa, and free office space to 100 participants every 4 months. We collect information on all start-up applicants since inception of the programme in 2010, and track their subsequent performance using web-based metrics such as presence in fund raising sites like AngelList and Crunchbase, and social-media activity in Facebook and Linkedin. We use the original ranking of applications by external judges, together with the discrete jump in the probability of selection around the 100-th (ranked) company, as an instrumental variable for selection into SUP. Our results do not allow us to rule out the possibility that participation in SUP has no impact on subsequent start-up performance. We contemplate several explanations for the findings, and argue that the most likely explanation is that there is heterogeneity in the effect of SUP, and that the average quality of start-ups near the 100-th company threshold is too low for them to benefit from the accelerator services. Using data from the "the Highway", an accelerator within the accelerator only available to top-performing participants, we find stronger evidence that business accelerators positively impact start-up performance.

[1] Su Wang provided excellent research assistance. Financial support from Abraaj Group is gratefully acknowledged.
[2] London School of Economics, J.Gonzalez-Uribe@lse.ac.uk
[3] Pontificia Universidad Católica de Chile, mile@stanford.edu

Over the past decade, a new method of incubating technology start-ups has increasingly gained participation in the entrepreneurial ecosystem: business accelerators. From only one programme in 2005, Y Combinator in Silicon Valley, there are now dozens worldwide, including Techstars which operates in several cities in the U.S., Seedcamp, originally London-based and currently pan-European, and several government-based intitiatives such a the pioneering Start-up Chile (SUP).

Despite their increasing popularity, formal analysis of these programmes is lacking due to data- and methodology-related challenges. Because participants are early-stage start-ups, they are often not legally incorporated, and are thus missing from standard business data sources. In addition, the probability that these early stage start-ups "pivot" is so large, that is challenging to even define, let alone adequately measure, post-application performance. Finally, because simple comparisons across participants and non-participants confound the effect of the programme with the higher quality of teams that succeed in the competitive application processes, researchers are prevented from more in-depth analysis, unless selection processes are random or rules-based.

In this paper, we overcome these data and measurement challenges and provide the first formal analysis of an accelerator programme by focusing on SUP, an accelerator promoted by the Chilean government since late 2010. In SUP, selected participants receive a grant for U$40,000 (equity free), a one year work visa (i.e., the programme is open to Chilean and non-Chilean teams), and a space to work and interact with other teams for six months in downtown Santiago de Chile. Approximately 600 businesses apply every four months for a total of 100 coveted spots in the program.

SUP has an advantage for researchers over other accelerator programmes in that it makes the selection decision through a well-documented, rules-based process. Each round,

applications are scored by external judges using three criteria: the quality of the founding team, the merits of the project, and the expected impact of the project on Chile's entrepreneurial environment. Based on the judges' scores, applications are ranked, and the first 100 start-ups are selected into the programme.

Our analysis exploits very detailed, manually collected data at the start-up level for teams that applied to SUP during the 2010-2013 period. SUP provided us access to confidential records of the companies that applied to the programme, the evaluation scores from the panel of judges, and the selection decisions made. Based on these records, we collected information on subsequent start-up performance using surveys and extensive web-searches on the businesses and the teams' leaders in fund raising sites such as AngeList, Techcrunch, social media sites like Facebook, Linkedin, and in web-page tracking sites like Google Insights and Alexa.

Our empirical analysis starts by comparing the characteristics as well as subsequent performance of participants and non-participants. Participants are overall very similar to non-participants. They differ mostly on geographical origin (those selected include a disproportionate amount of North-Americans) and industry (participants are less focused on social media, and more focused on e-commerce, energy and clean technology and finance). Participants are in general more successful as measured by web-presence in fund raising sites and social media, and web-traffic, and there are no significant differences across these groups in actual find-raising or employee growth.

We then narrow down our comparison to participants and non-participant start-ups that ranked close to the 100^{th} company threshold. In detail we compare characteristics and subsequent performance of start-ups that ranked just above this threshold (ranked 85-100), and thus barely made it into the programme, with the start-ups that ranked just below the

threshold (ranked 101-115), and thus were almost not rejected. The identification relies on start-ups ranking close to the threshold, having similar unobservable growth opportunities. Once we focus on the small window we find no significant difference across groups.

Finally, we estimate the impact of SUP using RD approaches that exploit the jump in the probability of selection into the programme around the 100-th start-up threshold. We use both, parametric and semi-parametric methods following Lee and Lemiex (2010) and Van der Klaauw (2008). The underlying identification assumption is that growth opportunities of start-ups are continuously related to the ranking. Similar to the results from our simple comparison across companies above and below the 100^{th} company threshold, we cannot rule out that SUP has no impact on start-up performance.

We contemplate several explanations for the findings, and argue that the most likely explanation is that there is heterogeneity in the effect of SUP and that the average quality of start-ups near the 100-th company threshold is too low, for these companies to benefit from the accelerator services.

We then exploit detailed data on "The Highway", an accelerator within SUP, which is only available for start-ups that were ranked by external judges among the top performers two months into the programme. Highway participants are allowed access to the best mentors and meet frequently with their technical supervisors in SUP. We compare participants in and out of the Highway that were ranked similarly by external judges, and find stronger evidence that accelerator services positively impact start-up performance.

In future versions of the paper we plan to include results from a detailed survey on applicants regarding their experience in SUP, and including information about the most useful aspects of the programme. We will also analyse subsequent performance at the founder level, by examining which types of jobs are subsequently held by founders. Finally, we will

also present suggestive evidence of the more general impact of SUP on the Chilean entrepreneurial ecosystem, by comparing registering rates of start-ups in Chile across industries targeted and not targeted by SUP.

Our paper has policy implications. Business accelerators are becoming more frequent as policy tools to sponsor entrepreneurship. In particular, the SUP model has already been adapted in several countries such as the U.S., Brazil and Peru, and is in the process of being adapted to several other countries (e.g., Canada, Denmark and Spain, among others). Understanding whether these programmes affect start-up performance and how, can help policy makers adapt more successfully this type of programs to the idiosyncrasies of their countries.

The rest of this paper is as follows. In Section 1 we describe the SUP programme and its selection process. In Section 2 we detail the empirical approach to assess the causal effect of SUP on subsequent performance of participants. We present and discuss results in Section3. In Section 4 we conclude.

1. Start-up Chile

SUP is a government-sponsored program launched in August 2010 to attract early-stage, high-potential entrepreneurs to bootstrap their ventures in Chile. The programme is run by the Ministry of Economy and is executed by the Chilean Economic Development Agency (CORFO), the leading organization for promoting innovation and entrepreneurship in the country. Its main long-term goal is to convert Chile into an innovation and entrepreneurial hub in Latin America not only by bringing in more entrepreneurs, but also by creating a much better-developed ecosystem of supporting institutions—including venture capital firms and angel investors.

SUP offers four main benefits to participants. First, SUP provides selected start-ups with $40,000 equity-free seed capital. The capital is staged: 50% is delivered at the beginning of the programme, and the remaining 50%, 3 months after. The second instalment is conditional on pre-determined performance milestones.[4] The staging of capital provides incentives to entrepreneurs to provide effort, and accountability of participants' expenditures.

Second, SUP sponsors a temporary one-year work visa for accepted participants in order to attract foreign entrepreneurs. The programme also helps participants settle in Chile through a "buddy system". The buddy-system pairs entrepreneurs with local members of the Santiago business community based on background interests and language. Local buddies advice participants on opening Chilean bank accounts, registering with the police, obtaining a local ID, and securing housing and mobile phones, in addition to checking in with participants once or twice a month throughout the entrepreneurs' stay in the country.

Third, SUP provides free, shared office space in downtown Santiago, equipped with WiFi, for all start-ups. Workshops on think-tanking and pitch-training based on peer to- peer teaching are held on-site. Start-ups also have access to SUP's network of mentors.

Starting in 2012, SUP expanded its programme to include more accelerator-type activities such as national and international pitch competitions. It created an accelerator within the accelerator known as the Highway, which provides additional resources to participants including access to the most renowned mentors and frequent monitoring by the SUP staff. Participants are carefully selected into the Highway after a pitch competition, in which external and internal judges rank participants. Roughly the first 20 participants of each generation have classified into the Highway since SUP's fourth generation.

[4] In the inception of the programme, capital disbursements were neither pre-expense nor staged. This system was implemented in the first semester of 2013.

The SUP program, in turn, requires accepted entrepreneurs to stay in Chile for the six-month duration of the program, and contribute to the building of an entrepreneurial culture in Chile. During their stay, entrepreneurs have to accumulate 4,000 in "Return Value Agenda" (RVA) points, a system to measure the social contribution of participants in the Chilean entrepreneurial ecosystem. Participants have the option to attend, organize or innovate in social-related activities. Attendance refers to participation in local events, such as meetings and conferences at which entrepreneurs make themselves available to share knowledge and to network with locals. Organization can include giving a talk at a school, presenting a pitch to a local investor, or mentoring a local entrepreneur or student. Innovation refers to initiatives that actively engage the Chilean business community, such as starting a new business with a Chilean partner or patenting a product in Chile.

1.1. Selection SUP

Selection into SUP is a two-part process that takes place every four months. First, entrepreneurs apply to the programme and their applications are ranked by external judges. SUP outsources this first part to Younoodle, a consulting start-up in California, which provides and objective evaluation of the merit of the start-ups outside the particular context of the Chilean economy.

Entrepreneurs fill in their applications through an open surve, and then Younoodle resorts to Silicon Valley experts (3-4 judges per application) who evaluate applications using three criteria: the quality of the founding team, the merits of the project, and the impact that it is likely to have on Chile's entrepreneurial environment. Using the experts' judging sheets, applicants are ranked. No ties are permitted; if companies tie in their judges' score they are randomly ranked.

The second part of the selection process is handled by CORFO, which makes the final decision based on Younnodle's ranking. A threshold is pre-specified each round (normally 100), and roughly, only companies that rank above the threshold are selected.[5] The threshold is decided on by the government before the application process begins and is a function of government's budget.

The start-ups cannot precisely manipulate their ranking. Because start-ups do not know the judges' scoring rules, and are unlikely to learn about these rules from past SUP participants, it is improbable that start-ups have room for manipulating their scores around the 100-th company cut-off. In addition, the judges are unlikely to manipulate the scores, as no judge evaluates all applications and only observe the very few he/she is asked to score.

As it is common in government-sponsored programmes, however, the selection committee at CORFO does not strictly follow the selection rule and thus not all participants who rank above the 100th company threshold end up participating in the programme. Indeed, of the top 100 ranked applicants, about 75% of them are selected into SUP. The remaining 25% are selected by a committee among applicants ranked between 101 and 300 based on qualitative attributes of the applications. The setting, thus, is akin to a fuzzy RD design (Van der Klauuw, 2008), where, although there is no 100% compliance of the selection rule, there is nonetheless a discrete jump in the probability of selection which can be exploited for identification. In Figure 1 we document the jump in selection around the 100th company threshold.

Denoting acceptance by SUP as A, and the external judges' score by S, the selection process just described can be characterized as follows:

[5] The threshold has been 100 in every generation, except the second generation where the threshold was set at 150.

$$\boldsymbol{A_i = f\left(S_i, I(S_i > \overline{S})\right) + \varepsilon_i} \qquad (1)$$

where the selection status, A_i, corresponds to a binary variable indicating whether the start-up was selected or not by SUP, and the unobserved component ε_i captures all other relevant characteristics of the start-up (and possibly the judges and CORFO officers) which influenced the SUP selection decision.

The evaluation problem of SUP can be described by the selection equation (1) together with the following outcome equation:

$$O_{ir} = \alpha A_i + X_{ir}{}' + \mu_r + \epsilon_{ir} \qquad (2)$$

where the outcome measure, O_i, corresponds to any of the following: web-based presence, we-traffic, and social media preference and X_i is a vector of start-up characteristics. The term μ_r is an i.i.d. random generation component. The remaining error component ϵ_{ir} is specific to start-ups. The first error component is introduced to parameterize possible within-generation correlation in outcomes. The SUP participation coefficient, α, is the coefficient of interest.

It is clear that OLS estimation of the outcome equation (2) would likely lead to a biased estimate for α as the error in the selection equation, ε_i, is likely correlated with the error in the outcome equation, ϵ_{ir}. We use a fuzzy RD approach to address this selection bias. Instrumental variables (IV) estimates of equation (2) use the discontinuity in the relation between selection by SUP and the S index (captured by f), to identify the causal effect of selection by SUP. The identifying assumption is that any other relationship between the S index and start-up outcome is controlled by smooth functions of the S index in the vector of covariates. In practice this includes linear, polynomial and piecewise linear functions of S_i.

2. Data

We were given access to applicant records for seven generations of SUP. In total we have information on 2,394 applicants, 531 and 1,653 participants and non-participants, respectively. Table I includes summary statistics.[6]

We collected information on subsequent for the applicant start-ups using extensive web-searches. Table A1. in the Appendix has a list of the outcome variables and their sources and Panel C in Table I includes summary statistics of these variables.

Due to data issues, in some of the tables below we present results restricting the sample to Generations 1 through 6. In total, 210 firms participate in the Program in Generation 7, representing 9% observations in our full sample.

3. Results

We begin by comparing the characteristics as well as subsequent performance across participants and non-participants, in Panel A of Tables II and III-IV, respectively. Table II shows that at the application, participants are very similar to non-participants. They differ mostly on geographical origin (those selected include a disproportionate amount of North-Americans) and industry (participants are less focused on social media, and more focused on e-commerce, energy and clean technology and finance). Tables III and IV show that participants are in general more successful as measured by web-presence in fund raising sites and social media, and web-traffic, and there are no significant differences across these groups in actual find-raising or employee growth.

[6] Results from the Global Entrepreneurship Monitor (GEM) report provide a basis for comparison between the entrepreneurs that apply to SUP, and the average Chilean entrepreneur. According to the latest GEM (2012), the average Chilean entrepreneur is 37.5 years old, is twice as likely to be male than female, has studies beyond those that are compulsory, and has a business that serves the consumer sector. The survey on micro entrepreneurship (EME) also provides a basis of comparison for the composition of Chilean SUP entrepreneurs. According to the EME of 2012 the average Chilean micro entrepreneur is male (69%), has between 45 and 59 years of age (39%), is responsible for a home (74%), has basic to mid-level education (67%) and its business belongs to the sectors: retail, restaurant and hotel (34%), agriculture and fishing (24%) and manufacturing (13%).

We then narrow down our comparison to participants and non-participant start-ups that ranked close to the 100th company threshold. We summarize results in Panel B of Tables II-IV. In detail we compare characteristics and subsequent performance of start-ups that ranked just above this threshold (ranked 85-100), and thus barely made it into the programme, with the start-ups that ranked just below the threshold (ranked 101-115), and thus were almost not rejected. The identification relies on start-ups ranking close to the threshold, having similar unobservable growth opportunities. Table II shows that participants above and below the cut-off point are indistinguishable at the time of application. Tables III and IV show that once we focus on the smaller window, participants and non-participants are also indistinguishable in terms of subsequent performance.

Table V presents results from the RD approach using as dependent variables an indicator of whether the start-up has an AngeList profile and the number of followers in AngeList. Figure 1 illustrates the jump in the probability of selection around the 100th company cut-off. The jump is not very drastic, but is statistically significant as shown in Table V, where most of the first stages are significant and the F-test is higher than 10. Results from the McCrary (2008) density test are presented in Figure 2. Table V shows that using the RD approach, we cannot rule out that SUP has no effect on the subsequent performance of start-ups.

In unreported regressions, preliminary analysis comparing subsequent performance across participants and non-participants from the "the Highway, provides stronger evidence that accelerator services positively impact start-up performance.

In future versions of the paper we plan to include results from a detailed survey on participants regarding their experience in SUP including information about the most useful aspects of the programme. We will also analyse subsequent performance at the founder level, by examining which types of jobs are subsequently held by founders. Finally, we will also

present suggestive evidence of the impact of SUP on the Chilean entrepreneurial ecosystem, by comparing registering rates of start-ups in Chile across industries targeted and not targeted by SUP.

4. Conclusions

In this paper we provide new evidence about an essential question in entrepreneurial finance, do early-stage financiers add any value to companies. We focus on the experience of SUP, an example of an accelerator, the newest method of incubating technology start-ups. We quantify the causal impact of SUP by simultaneously exploiting novel, rich micro-data and addressing concerns about unobserved heterogeneity. We find that we cannot rule out that accelerators have an impact on start-up performance. We contemplate several explanations for the findings, and argue that the most likely explanation is that there is heterogeneity in the effect of SUP and that the average quality of start-ups near the 100-th company threshold is too low, for them to benefit from the accelerator services. Using data from the "the Highway", an accelerator within the accelerator only available for start-ups that were ranked by external judges among the top performers two months into the programme, we find stronger evidence that accelerator services positively impact start-up performance.

References

Hahn, J., Todd, P., Van Der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica* 69, 201–209

Kerr, William, Joshua Lerner and Antoinette Schoar, 2014, The Consequences of Entrepreneurial Finance: Evidence from Angel Financings, *Review of Financial Studies,* 27(1) pp. 20-55.

Lelarge, Claire, David Sraer and David Thesmar, 2013, Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Program in NBER volume on "International Differences in Entrepreneurship" edited by Joshua Lerner and Antoinette Schoar, University of Chicago Press.

Lee, David and Thomas Lemieux, 2010. Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, vol. 48(2), pages 281-355, June.

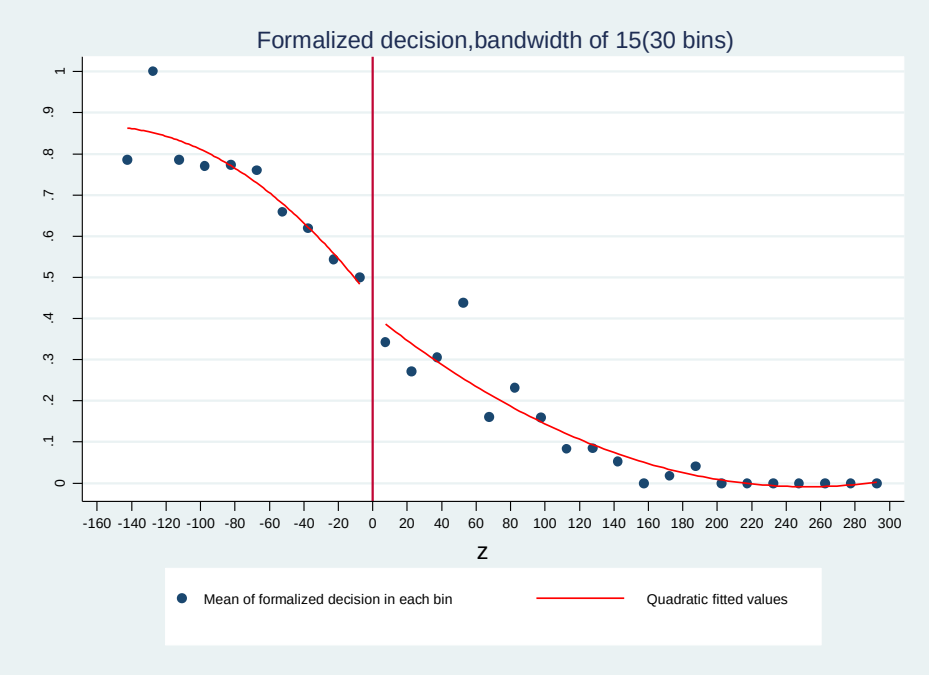
Van Der Klaauw, W., 2002. Estimating the effect of financial aid offers on college enrollment: a regression-discontinuity approach. *International Economic Review* 43, 1249–1287.

Van Der Klaauw, W., 2008. Regression-Discontinuity Analysis: A Survey of Recent Developments in Economics, Labour: *Review of Labour Economics and Industrial Relations*, Vol. 22 (2), 2008, p.219-245.

Figure 1

Panel A and B show the mean of Formalized against the midpoint of each bin with bandwidth of 15 and bin<=20. Panel A includes all generations (N=1910) and Panel B excludes Generation 7 (N=1751). The blue points in both graphs are the mean of dummy variable Formalized evaluated at the midpoint of each bin on the both sides of the cutoff point. The red line in both graphs draws the quadratic fitted formalized values for midpoint of each bin at the left side and the right side of the cutoff point separately. We observe the jump of Formalized decision and both plots show no other unexpected jumps.

Panel A



Panel B

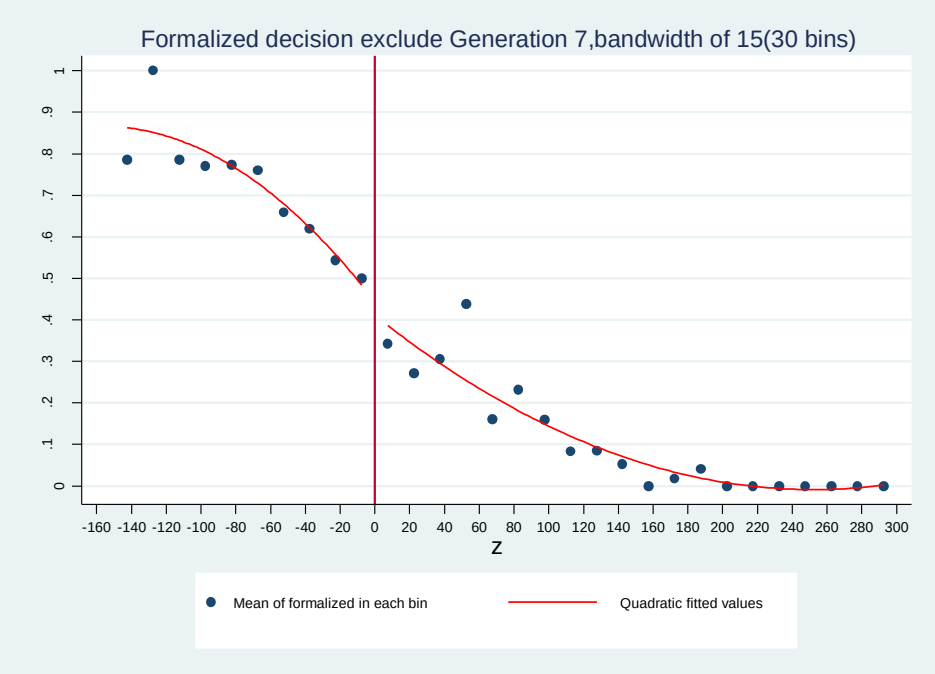
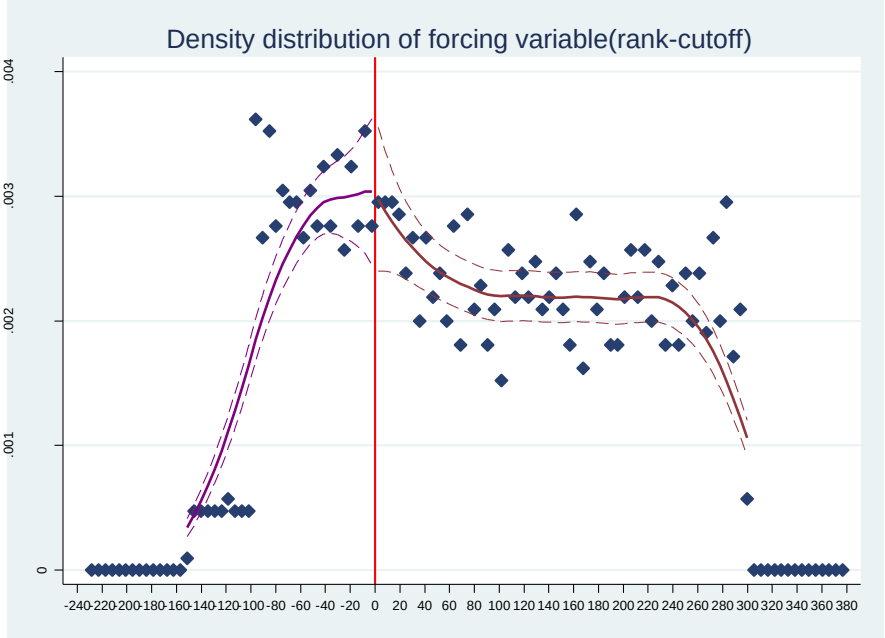


Figure 2

Figure 3 shows the Density of the forcing variable, the rank-cutoff for both sets of samples computed over all generations in Panel A (N=1910) and excluding Generation 7 in Panel B (N=1751) with bandwidth of 15 and bin<=20. The diamond blue point again is the density of each bin computed at the midpoint of each bin. The purple line on the left side and the Maroon line on the right side of the cutoff are soothed lines estimated by polynomial model of McCrary tests using Stata procedure -DCdensity-.The dashed lines are the 95 percent confidence bands. Both plots show no evidence of discontinuity at the cutoff.

Panel A



Panel B

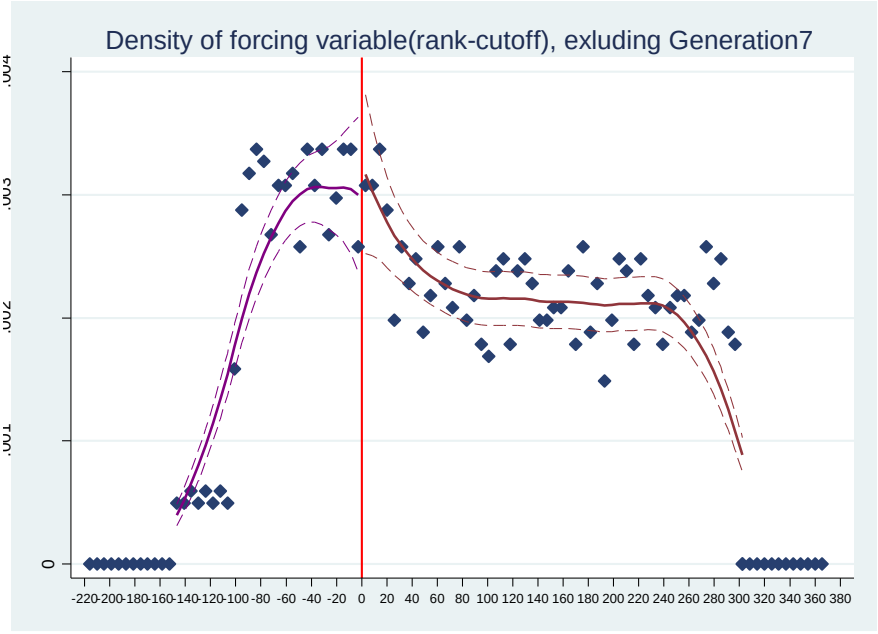


Figure 3

Panel A and B show the mean of Angel followers against the midpoint of each bin with bandwidth of 15 and bin<=20. Panel A includes all generations (N=1910) and Panel B excludes Generation 7 (N=1751). The blue points in both graphs are the mean of Angel followers evaluated at the midpoint of each bin on the both sides of the cutoff point. The red line in both graphs draws the quadratic fitted formalized values for midpoint of each bin estimated at the left side and the right side of the cutoff point separately. We observe the jump of Angel followers and both plots show no other unexpected jumps. There is not much difference between including and excluding Generation 7, but the quadratic line fits the points better in Panel B. The left side of the cutoff line also suggests the second or higher order functional form of regression in both plots.

Panel A



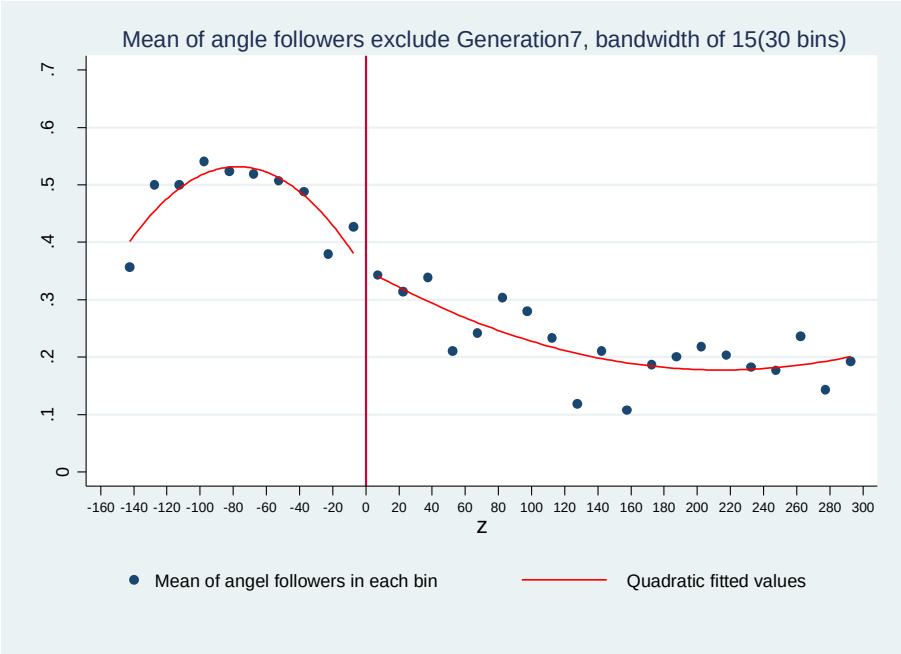


Table I - Composition sample and Summary Statistics

Table I represents the summary statistics of our sample. The full sample includes 2394 observations. Panel A shows the basic characteristics by generation. Column (2) is the mean of leader's age of startups. Column (3)-(5) are percentage of male, startups with no money raised and startups incorporated before SUP in all observations respectively. The column (6) is the percentage of each generation. Panel B shows the distribution of observations across different categories for the continent, industry and the field of leaders, including the frequency and the percentage. Panel C summarize the current outcomes of the startups, including the mean, standard error, median and the non-missing observation. Extensive margin includes the Number of startups listed in AngelList, Facebook, LinkedIn and Crunchbase. Intensive margin includes the Number of followers in AngelList, LinkedIn , Number of Facebook likes, Google Search Interest, Alexa global rank , US rank and Num of Alexa Site Link. Capital raised includes the AngelList founding, Crunchbase funding and the company size in LinkedIn.

Panel A : Basic characteristics by generation

Panel B : Continent, Industry and Field of Leader

B.1 Continent

	Africa	Asia	Europe	N. America	Oceania	S. America
Freq	18	189	448	639	36	1,059
Percent	0.75	7.91	18.75	26.75	1.51	44.33

B.2 Industry

	Consulting	E-commerce	Education	Energy & Clean Tech	Finance	Healthcare & Biotech	IT & Software
Freq	3	424	161	76	58	81	384
Pct	0.13	18.35	6.97	3.29	2.51	3.5	16.62

Generation	Age (mean)	Gender Male(pct)	Money Raised No(pct)	Incorporated Yes(pct)	Obs (pct)
G1	0	93.33	85.83	0.00	5.01
G2	0	86.44	0.00	0.00	15.71
G3	0	82.15	72.36	0.00	13.58
G4	82	42.26	78.22	46.73	14.04
G5	30.37	35.73	75.24	44.05	22.10
G6	30.10	84.34	76.71	39.16	20.80
G7	30.12	86.67	96.19	29.52	8.77
Obs.	1220	2394	2005	2394	2394

	Media	Mobile & Wireless	Natural Resources	Social Enterprise	Social Media	Tourism	Other
Freq	94	211	35	116	282	107	279
Pct	4.07	9.13	1.51	5.02	12.2	4.63	12.07

B.3 Field of leader

	Arts and Humanities	Business	Engineering	Law	Natural Science	Social Science
Freq	359	676	751	42	149	151
Pct	16.88	31.77	35.29	1.97	7	7.1

Panel C: Current situation

C.1 Extensive margin

	Num of startups AngelList	Num of startups Facebook	Num of startups LinkedIn	Num of startups Crunchbase
Freq	673	1,296	836	485
Pct/mean	28.11	54.14	34.92	20.26
Nonmissing Obs	2,394	2,394	2394	2,394

C.2 Intensive margin

	Num of followers AngelList	Num of followers LinkedIn	Num of Facebook Likes	Google Search t	Alexa global rank	Alexa US Rank	Num of Alexa Site Link
Mean	31.29	73.58	6455.58	17.01	2232470.00	4434.49	26.23
sd	75.88	290.20	52547.92	29.49	6256374.00	33967.31	235.41
Median	11	14	275	0	0	0	0
Nonmissing Obs	669	822	1203	2260	2390	2390	2390

C.3 Capital Raised

	AngelList Funding	Crunchbase Funding	LinkedIn Company Size
mean	708971.60	324208.30	42.41
sd	1886691.00	1421060.00	407.42
median	200000	40000	10
Nonmissing Obs	155	240	757

Table II- Comparison of covariates

Table II shows the comparison of covariates. Panel A compares participants with nonparticipants of SUP, participants are startups formalized in SUP and nonparticipants are not formalized startups in SUP. Panel B compares the group above and below the cutoff point of discontinuity, including the mean and percentage of basic characteristics, the continent, industry and the field of leader's distribution comparisons. Observations with rank in just one bin to the left and to the right side of cutoff point with bin bandwidth of 15 are included in the comparison in Panel B. In each Panel, we summarize the same characteristics for each group similar to Table I. In addition, we add the row "p-value Kolmogorov-Smirnov test" of null the distributions of two groups are different at the end of each category variable. For comparison between participants and nonparticipants we can reject the null whereas the comparison between groups above and below the cutoff point we cannot reject the null.

Panel A. Participants and Nonparticipants

Basic Characteristics				
	Nonparticipants		Participants	
	Freq	Pct/(Mean)	Freq	Pct/(Mean)
Age	1,003	30.13	216	30.60
Gender(Male)	1,157	63.26	480	84.96
Money Raised(No)	1,259	80.34	314	71.69
Incorporated(Yes)	142	27.61	505	25.13
Continent				
	Nonparticipants		Participants	
	Freq	Pct	Freq	Pct
Africa	16	0.88	2	0.35
Asia	143	7.84	46	8.14
Europe	338	18.53	110	19.47
N. America	456	25	183	32.39
Oceania	24	1.32	12	2.12
S. America	847	46.44	212	37.52
Total	1824	100	565	100
p-value Kolmogorov-Smirnov test				0.002
Industry				
	Nonparticipants		Participants	
	Freq	Pct	Freq	Pct
Consulting	3	0.17	0	0.00
E-commerce	306	17.42	118	21.34
Education	119	6.77	42	7.59
Energy & Clean Tech	51	2.90	25	4.52
Finance	39	2.22	19	3.44
Healthcare & Biotech	54	3.07	27	4.88
IT & Software	301	17.13	83	15.01
Media	76	4.33	18	3.25
Mobile & Wireless	159	9.05	52	9.40
Natural Resources	26	1.48	9	1.63
Social Enterprise	82	4.67	34	6.15
Social Media	245	13.94	37	6.69
Tourism	84	4.78	23	4.16
Other	212	12.07	66	11.93
Total	1757	100	553	100
p-value Kolmogorov-Smirnov test				0.007
Field of Leader				
	Nonparticipants		Participants	
	Freq	Pct	Freq	Pct
Arts and Humanities	304	18.18	70	13.11

Business	517	30.92	182	34.08
Engineering	600	35.89	183	35.49
Law	34	2.03	8	1.90
Natural Science	109	6.52	44	6.94
Social Science	108	6.46	47	7.03
Total	1672	100	534	100
p-value Kolmogorov-Smirnov test				0.229

Panel B: Groups above and below cutoff point

Basic Characteristics				
	Below		Above	
	Freq	Pct/Mean	Freq	Pct/Mean
Age	35	30	25	29.92
Gender (Male)	75	78.95	63	78.75
Money Raised (No)	66	83.54	52	78.79
Incorporated (Yes)	18	18.95	18	22.50
Continent				
	Below		Above	
	Freq	Pct	Freq	Pct
Africa	0	0.00	0	0.00
Asia	6	6.38	2	2.50
Europe	12	12.77	14	17.50
N. America	28	29.79	24	30.00
Oceania	1	1.06	2	2.50
S. America	47	50.00	38	47.50
Total	94	100	80	100
p-value Kolmogorov-Smirnov test				1.000
Industry				
	Below		Above	
	Freq	Pct	Freq	Pct
Consulting	0	0	0	0
E-commerce	21	22.58	18	23.38
Education	4	4.30	8	10.39
Energy & Clean Tech	6	6.45	2	2.60
Finance	2	2.15	2	2.60
Healthcare & Biotech	2	2.15	2	2.60
IT & Software	17	18.28	8	10.39
Media	5	5.38	4	5.19
Mobile & Wireless	2	2.15	10	12.99
Natural Resources	1	1.08	1	1.30
Social Enterprise	4	4.30	4	5.19

Social Media	10	10.75	4	5.19
Tourism	4	4.30	1	1.30
Other	15	16.13	13	16.88
Total	93	100	77	100
p-value Kolmogorov-Smirnov test				0.883
Field of Leader				
	Below		Above	
	Freq	Pct	Freq	Pct
Arts and Humanities	15	16.30	15	20.55
Business	35	38.04	22	30.14
Engineering	30	32.61	22	30.14
Law	1	1.09	3	4.11
Natural Science	5	5.43	6	8.22
Social Science	6	6.52	5	6.85
Total	92	100	73	100
p-value Kolmogorov-Smirnov test				0.997

Table III- Univariate Tests of Differences of information in AngelList in Sample

Table III presents the mean comparison of AngelList information between Participants and the non-Participants start-ups, reporting mean and standard errors of observed variables in two groups. The variables under observation are dummy variable Listed in AngelList and Number of followers in AngelList. The dummy variable Listed in Angel List equals 1 if the startup is listed in AngelList and 0 otherwise. Participant is an indicator that equals 1 if the start-up is formalized after first round competition in SUP. Panel A compares mean between participant and nonparticipant, Panel B further restrict observations included in the local regressions with bandwidth of 15, i.e. observations with rank in just one bin to the left and to the right side of cutoff point with bin bandwidth of 15. Observations in column (2)-(3) and column (6)-(7) in both panels exclude the observations from Generation 7. Standard errors are reported in brackets. t-statistic of tests under null hypothesis that the mean in different groups are equal is reported in the last row. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Participants and Nonparticipants

	Listed in AngelList				Number of Followers in AngelList			
	Subsample		Full Sample		Subsample		Full Sample	
Participant	1	0	1	0	1	0	1	0
Mean	0.58	0.174	0.6	0.183	39.739	22.51	41.326	21.099
	(0.021)	(0.009)	(0.021)	(0.009)	(4.580)	(4.553)	(4.266)	(3.953)
Observations	531	1653	565	1,829	306	286	337	332
Mean Diff.	-0.406		-0.417		-17.228		-20.227	
t-statistic	-19.832***		-20.981***		-2.664**		-3.476***	

Panel B: Groups above and below cutoff point

	Listed in AngelList				Number of Followers in AngelList			
	Subsample		Full Sample		Subsample		Full Sample	
Participant	1	0	1	0	1	0	1	0
Mean	0.42	0.359	0.453	0.375	48.324	15.107	49.419	20.067
	(0.053)	(0.055)	(0.051)	(0.054)	(16.221)	(3.730)	(14.235)	(6.294)
Observations	88	78	95	80	37	28	43	30
Mean Diff.	-0.061		-0.078		-33.217		-29.352	
t-statistic	-0.807		-1.035		-1.752*		-1.643*	

Table IV-Univariate tests of outcome variables

Table AII shows the comparison of other outcome variables. Panel A compares participants with nonparticipants of SUP, participants are startups formalized in SUP and nonparticipants are not formalized startups in SUP. Panel B compares the group above and below the cutoff point of discontinuity. Observations with rank in just one bin to the left and to the right side of cutoff point with bin bandwidth of 15 are included in the comparison in Panel B. In each Panel, mean value of each group are listed in column (2) and (4). Observations of each group are listed in column (3) and (5). The mean difference and the P-value of the t-test are included in column (6).

Panel A. Participants and Nonparticipants

Panel B. groups above and beyond cutoff point

	Above cutoff	Obs	Below cutoff	Obs	Diff.	P-value
Num of startups Crunchbase	0.295	88	0.41	78	0.115	0.1230
Num of startups LinkedIn	0.42	88	0.462	78	0.041	0.5972
Num of startups Facebook	0.614	88	0.603	78	-0.011	0.8849
Num of Followers in LinkedIn	78.941	34	151.543	35	72.602	0.5364
Num of startups Crunchbase	0.463	531	0.131	1653	-0.333	0.0000
Num of Facebook Likes	6158.62	50	1660.889	45	-4497.731	0.2471
Num of startups LinkedIn	0.544	531	0.286	1653	-0.259	0.0000
Google Search	16.591	82	15.219	73	-1.372	0.7685
Num of startups Facebook	0.721	531	0.483	1653	-0.238	0.0000
Alexa global rank	1703655	88	2792993	78	1089338	0.1516
Alexa US Rank	88.905	284	68.974	463	-19.931	0.3796
Num of Alexa Site Link	37.761	88	16.282	78	-21.479	0.1436
Num of Facebook Likes	10282.9	366	5103.574	733	-5179.322	0.1407
AngelList Funding	517500	6	1030000	7	512500	0.5274
Crunchbase Funding	315000	17	386176.5	17	71176.47	0.8317
Google Search	17.08	488	16.852	1577	-0.227	0.8816
LinkedIn Company Size	37.606	33	22.758	33	-14.848	0.3825
Alexa global rank	3125514	531	1948604	1649	-1176910	0.0002
Alexa US Rank	9434.079	531	3389.29	1649	-6044.789	0.0006
Num of Alexa Site Link	41.637	531	23.84	1649	-17.797	0.1477
AngelList Funding	561097.7	88	806912.3	57	245814.6	0.3956
Crunchbase Funding	265472.2	180	319545.5	55	54073.23	0.7719
LinkedIn Company Size	22.55	269	59.017	420	36.466	0.2744

Table V-Regression results

Estimates stem from IV regressions. Observations are at the start-up-generation level. The dependent variables are dummy variable and the number of followers on the AngelList as specified on top of each panel. Dummy variable Listed in Angel List equals 1 if listed and 0 otherwise. Participants is a dummy that equals 1 if the start-up is Participants after first round competition in SUP. Participants is instrumented by a dummy variable Z that equals 1 if the start-up ranked above the generation cut-off, and is zero otherwise. Column (2), (6) and column (3), (7) report the polynomial regressions results with 1st and 2nd order of degree respectively. Column (4), (8) and column (5), (9) report the local regressions with bandwidth of 15 and bandwidth of 30 respectively. The observations of column (4)-(5), column (8)-(9) further restrict observations included in the local regressions with bandwidth of 15, i.e. observations with ranks just one bin to the left and to the right side of cutoff point with bin bandwidth of 15. Panel A and Panel B exclude observations from Generation 7. Panel B and Panel D regressions include the covariates variables: generation, industry, gender, age and capital raised before SUP. Standard errors are heteroskedasticity robust. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Regression results excluding Generation 7 and excluding covariates

Dependent Variable	Listed in AngelList				Number of Followers in AngelList			
	Polynomial		Local Regression		Polynomial		Local Regression	
	1st order	2nd order	bw=15	bw=30	1st order	2nd order	bw=15	bw=30
Participants	0.530***	0.516	0.431	0.363	6.788	91.248	222.019	98.073
	(0.152)	(0.623)	(0.481)	(0.220)	(37.913)	(178.425)	(193.194)	(66.646)
Constant	0.164***	0.138	0.210	0.211**	22.811	-31.420	-119.690	-33.541
	(0.047)	(0.234)	(0.206)	(0.097)	(19.723)	(116.530)	(132.720)	(42.695)
First stage								

Z	0.271***	0.102***	0.143***	0.218***	0.278***	0.086***	0.150	0.202***
	(0.029)	(0.039)	(0.075)	(0.053)	(0.055)	(0.073)	(0.116)	(0.087)
F-statistic	196.78	379.73	3.63	16.72	70.81	122.61	1.67	5.45
Observations	2,184	2,184	166	315	592	592	65	129
R^2	0.140	0.144	0.193	0.148	0.035			
Covariates?	No	No	No	No	No	No	No	No

Panel B: Regression results excluding Generation 7 and including covariates

Dependent Variable	Listed in AngelList				Number of Followers in AngelList			
	Polynomial		Local Regression		Polynomial		Local Regression	
	1st order	2nd order	bw=15	bw=30	1st order	2nd order	bw=15	bw=30
Participants	0.287	0.796*	0.550	0.363	-61.986	92.652	124.960	-67.965
	(0.235)	(0.454)	(0.812)	(0.299)	(58.978)	(170.003)	(72.316)	(48.181)
Constant	0.971***	0.958***	0.618	1.480**	149.577*	72.218	81.596	222.343**
	(0.169)	(0.181)	(0.784)	(0.590)	(83.115)	(80.134)	(184.977)	(99.265)
First stage								
Z	0.295***	0.223***	0.192	0.319***	0.326***	0.185	-0.105	0.316***
	(0.046)	(0.067)	(0.152)	(0.101)	(0.087)	(0.118)	(0.145)	(0.151)
F-statistic	36.91	38.21	1.84	1.94	14.76	16.66	5.95	1.44
Observations	1,012	1,012	52	103	265	265	22	45
R^2	0.285	0.229	0.628	0.470	0.129		0.998	0.468
Covariates?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel C: Regression results including Generation 7 and excluding covariates

Dependent Variable	Listed in AngelList				Number of Followers in AngelList			
	Polynomial		Local Regression		Polynomial		Local Regression	
	1st order	2nd order	bw=15	bw=30	1st order	2nd order	bw=15	bw=30
Participants	0.534***	0.567	0.544	0.400*	13.160	86.506	218.867	84.700
	(0.143)	(0.544)	(0.469)	(0.224)	(34.603)	(151.604)	(196.054)	(62.665)
Constant	0.180***	0.137	0.178	0.209**	20.585	-27.223	-118.549	-24.070
	(0.045)	(0.208)	(0.209)	(0.101)	(18.094)	(100.089)	(138.964)	(41.272)
First stage								
Z	0.271***	0.102***	0.143***	0.218***	0.285***	0.091****	0.134	0.191***
	(0.029)	(0.039)	(0.075)	(0.053)	(0.051)	(0.068)	(0.108)	(0.081)
F-statistic	196.78	379.73	3.63	16.72	73.35	143.82	1.54	5.57
Observations	2,394	2,394	175	334	669	669	73	129
R^2	0.146	0.141	0.208	0.165	0.037			
Covariates?	No	No	No	No	No	No	No	No

Panel D: Regression results including Generation 7 and including covariates

Dependent Variable	Listed in AngelList				Number of Followers in AngelList			
	Polynomial		Local Regression		Polynomial		Local Regression	
	1st order	2nd order	bw=15	bw=30	1st order	2nd order	bw=15	bw=30
Participants	0.421**	0.866**	0.542	0.398	-31.795	85.211	-920.657	-40.284
	-0.191	-0.395	-1.423	-0.297	-52.309	-140.627	-5293.09	-46.999
Constant	1.104***	1.101***	1.894*	1.497**	81.549**	69.852	665.194	132.296
	-0.16	-0.167	-1.013	-0.592	-35.465	-60.585	-3780	-145.504
First stage								
Z	0.319***	0.228***	0.104	0.296***	0.329***	0.190	0.036	0.317***
	(0.042)	(0.060)	(0.155)	(0.097)	(0.074)	(0.101)	(0.208)	(0.134)
F-statistic	40.38	42.07	1.51	1.63	17.83	19.91	1.33	1.49
Observations	1,220	1,220	61	122	342	342	30	58
R^2	0.274	0.212	0.657	0.472	0.1			0.273
Covariates?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table AI. Definition of Variables

Name of variable	Definition
	Covariates
Age	Leader's age of startups before SUP.
Money Raised	Dummy variable equals 1 if startup has money raised before SUP.
Gender	Dummy variable equals 1 if startup leader is female before SUP.
Incorporated	Dummy variable equals 1 if startup is incorporated before SUP.
Continent	Category variable indicates the continent of startup belongs to.
Industry	Category variable indicates the industry of startup.
Field of leader	Category variable indicates the leader of startup degree field.
	Outcome variables
Listed in AngelList	Dummy variable equals 1 if startup is listed in AngelList.
Listed in Crunchbase	Dummy variable equals 1 if startup is listed in Crunchbase .
Listed in LinkedIn	Dummy variable equals 1 if startup is listed in LinkedIn.
Num of startups AngelList	Number of startups listd in AngelList.
Num of startups Crunchbase	Number of startups listd in Crunchbase .
Num of startups LinkedIn	Number of startups listd in LinkedIn.
Num of Followers in AngelList	Number of followers in AngelList
Num of Followers in LinkedIn	Number of followers in LinkedIn.
Num of Facebook Likes	Number of Facebook likes.
Google Search	Google search times by 07/12/2013
Alexa global rank	Global rank on Alexa
Alexa US Rank	US rank on Alexa

Num of Alexa Site Link	Number of Alexa Site Links
AngelList Funding	Capital raised via AngelList
Crunchbase Funding	Capital raised via Crunchbase
LinkedIn Company Size	Company size shows in LinkedIn