Supplementary Information
Assessing the attraction of cities on venture capital from a scaling law perspective

Ruiqi Li, Lingyun Lu, Tianyu Cui, Weiwei Gu, Shaodong Ma, Gang Xu and H. Eugene Stanley

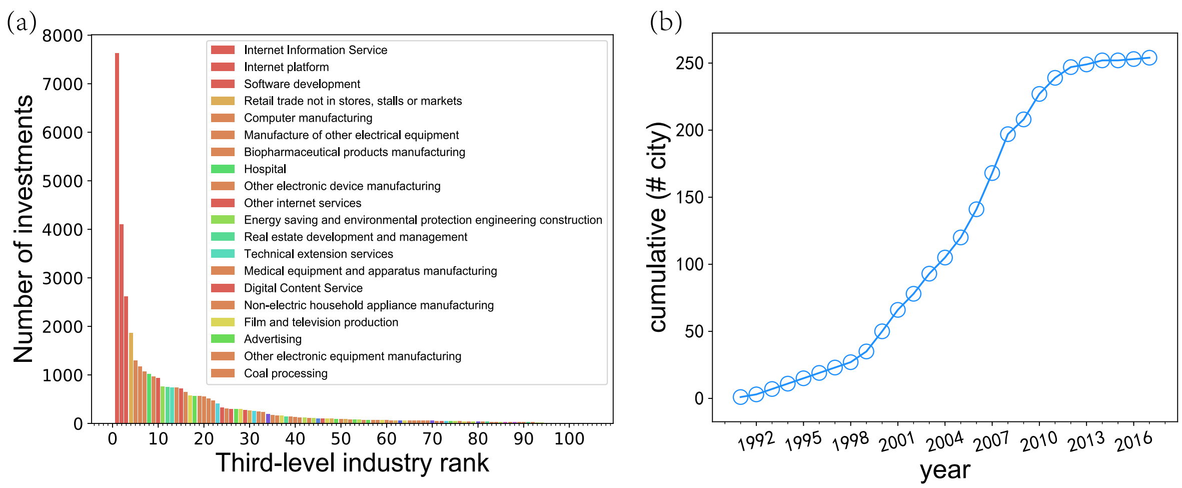

Figure 1: (a) The number of investments of each industry at the third-level of Industrial Classification for National Economic Activities of China (ICNEAC). The top 20 industries are labeled in the figure. (b) The cumulative number of cities that receive VC investments over time. Before the year 2000, the year that China made an important agreement on preparing to join the WTO, each year there are only a few (roughly four) new cities reached by VC activities; while after 2000, every year there are around fifteen new cities that receive VC investments.

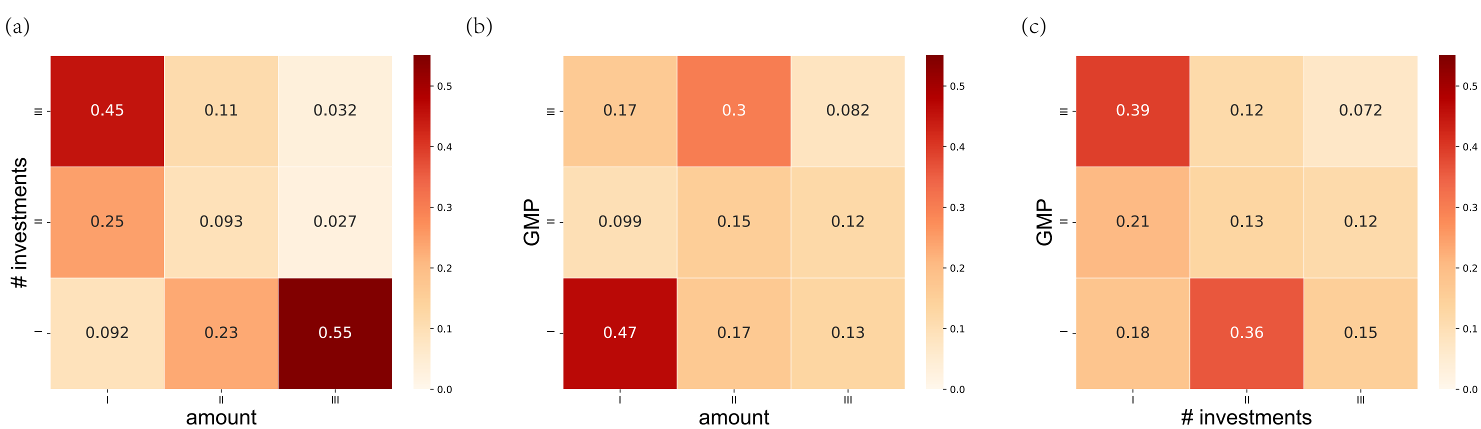

Figure 2: The Jaccard Index, which is indicated by the value in each square, between groups obtained from the standard hierarchical clustering algorithm based on the distance matrix D_{ij}. We can observe that the uprising groups on the amount of investment (group I of "amount" in (a)) and the uprising groups on the number of investments (group III of "# investment" in (a)) have a large fraction of overlapping (with a Jaccard Index equals 0.55). In comparison, the group classification results on GMP have less overlapping with the ones obtained from the investment amount and the number of investments (see (b) and (c), respectively).

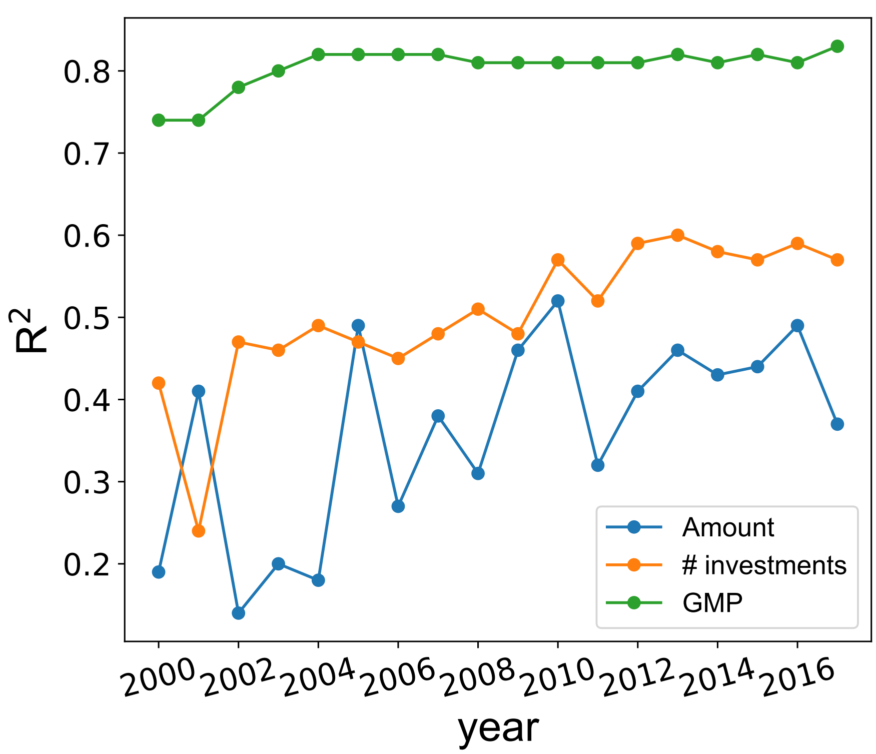

Figure 3: The R^2 of OLS regressions in the Fig. 4(a) in the main text.

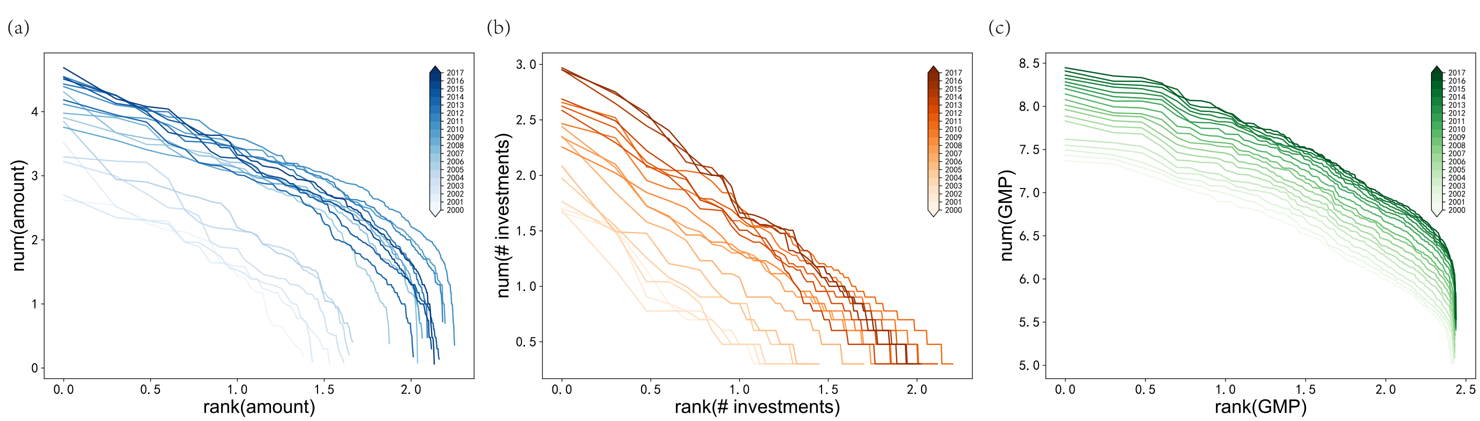

Figure 4: The Zipfian distribution on (a) the total amount of investment, (b) the number of investments, and (c) GMP from the year 2000 to 2017.

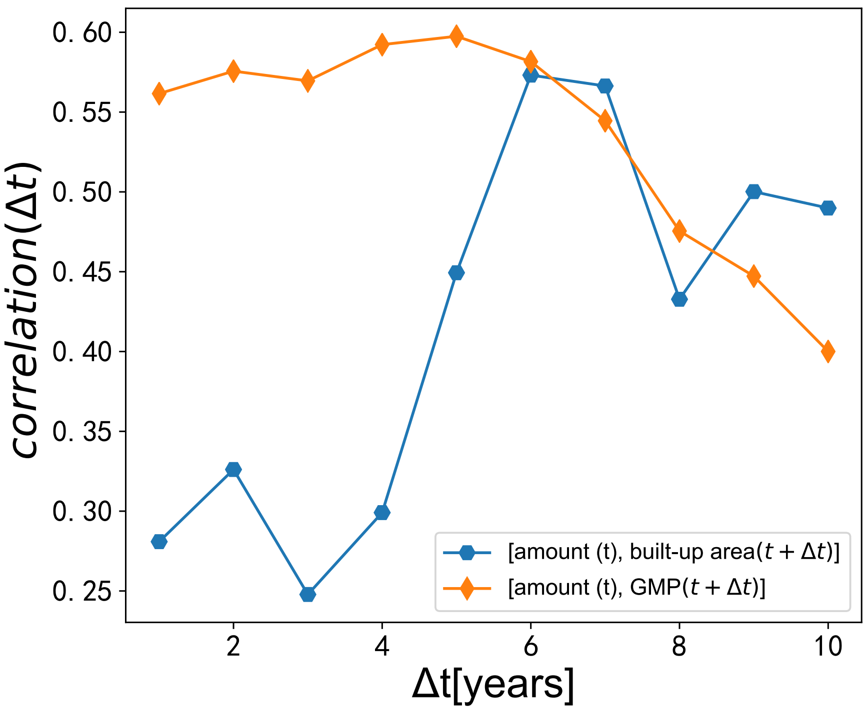

Figure 5: The lagged temporal autocorrelation between investment amount and GMP, built-up area of cities, where the lagged temporal autocorrelation of each city is calculated as follows: $C_i(\Delta t) = \sum_t [Amt_i(t) GMP_i(t + \Delta t)] / (\sqrt{\sum_t Amt_i(t)^2} \sqrt{\sum_t GMP_i(t + \Delta t)^2})$. Then we average over all cities and obtain the above figure. As we can see that when $\Delta t = [4, 6]$ years, the correlation between VC investments and GMP is among the highest, which indicates that the venture capital investments may need 4 to 6 years to have a stronger impact on the GMP of cities. While, for the built-up area, it peaks when $\Delta t = [6, 7]$ years, which corresponds to a longer cycle.

	Size	Mean	S.D.	Max	Min
Investment amount	37309	62.59454	429.4949	29327.26	0
GMP	4968	84231.02	208083.9	3063299	1038.47
Population	4968	108.3727	197.2505	2425.68	4

Table 1: Descriptive statistics of the data used in the paper, which is from the year 2000 to 2017. Data before the year 2000 is neglected. The raw data on venture capital investments is purchased from the SiMuTong dataset of Zero2IPO Group (www.pedata.cn). The population and GMP data of prefecture-level cities in our study are collected from the China City Statistical Yearbook. Since the spatial resolution of the location of the company is limited, the city refers to the administrative definition. The unit of Investment amount and GMP are in RMB/Million, the unit of Population is 10,000. The Sample Size of GMP and Population refers to the number of cities that have ever appeared in the scaling analysis in our paper.

		city model				person model
		lognormal		Gaussian		
	database	$\delta = 2$ (OLS)	$\delta \in [1,3]$	$\delta = 1$	$\delta \in [1,2]$	
2000	Amount	$\mathbf{0.86} \pm 0.63 \rightarrow^*$	$1.00 \pm 0.68 \rightarrow^*$	$1.30 \pm 0.60 \circ^*$	$1.30 \pm 0.73 \rightarrow^*$	$1.31 \pm 0.81 \nearrow$
	GDP	$\mathbf{1.06} \pm 0.09 \nearrow$	$1.06 \pm 0.09 \nearrow$	$1.16 \pm 0.11 \nearrow$	$1.09 \pm 0.12 \rightarrow$	$1.11 \pm 0.10 \nearrow$
	Times	$\mathbf{0.82} \pm 0.43 \rightarrow^*$	$0.69 \pm 0.41 \rightarrow^*$	$1.60 \pm 0.59 \nearrow$	$0.99 \pm 0.82 \rightarrow$	$1.57 \pm 0.74 \nearrow$
2001	Amount	$\mathbf{1.28} \pm 0.57 \rightarrow^*$	$1.36 \pm 0.64 \rightarrow^*$	$2.00 \pm 0.97 \nearrow$	$2.00 \pm 1.25 \circ$	$2.00 \pm 1.16 \nearrow$
	GDP	$\mathbf{1.10} \pm 0.07 \circ$	$1.10 \pm 0.09 \nearrow$	$1.22 \pm 0.23 \nearrow$	$1.10 \pm 0.23 \rightarrow$	$1.09 \pm 0.10 \nearrow$
	Times	$0.55 \pm 0.27 \circ^*$	$\mathbf{0.72} \pm 0.31 \rightarrow$	$1.01 \pm 0.65 \rightarrow$	$1.01 \pm 0.53 \rightarrow$	$1.03 \pm 0.60 \rightarrow$
2002	Amount	$\mathbf{0.60} \pm 0.45 \rightarrow^*$	$0.82 \pm 0.54 \rightarrow^*$	$1.11 \pm 0.76 \rightarrow$	$1.11 \pm 0.70 \rightarrow$	$1.17 \pm 0.79 \nearrow$
	GDP	$\mathbf{1.13} \pm 0.06 \circ$	$1.13 \pm 0.09 \nearrow$	$1.25 \pm 0.21 \nearrow$	$1.13 \pm 0.25 \circ$	$1.10 \pm 0.10 \nearrow$
	Times	$0.69 \pm 0.27 \circ^*$	$\mathbf{0.71} \pm 0.24 \circ$	$1.27 \pm 0.61 \circ$	$0.86 \pm 0.44 \rightarrow$	$1.22 \pm 0.59 \circ$
2003	Amount	$0.81 \pm 0.49 \rightarrow^*$	$\mathbf{1.18} \pm 0.63 \rightarrow^*$	$2.00 \pm 0.77 \nearrow$	$1.35 \pm 0.61 \circ$	$1.36 \pm 0.87 \nearrow$
	GDP	$\mathbf{1.13} \pm 0.07 \nearrow$	$1.13 \pm 0.07 \nearrow$	$1.19 \pm 0.09 \nearrow$	$1.13 \pm 0.10 \circ$	$1.12 \pm 0.09 \nearrow$
	Times	$0.67 \pm 0.28 \circ^*$	$\mathbf{0.65} \pm 0.20 \searrow^*$	$1.49 \pm 0.45 \nearrow$	$0.96 \pm 0.26 \rightarrow$	$1.34 \pm 0.42 \nearrow$
2004	Amount	$\mathbf{0.78} \pm 0.52 \rightarrow^*$	$0.92 \pm 0.58 \rightarrow^*$	$2.00 \pm 0.61 \nearrow$	$1.12 \pm 0.92 \circ$	$1.97 \pm 1.27 \nearrow$
	GDP	$\mathbf{1.13} \pm 0.06 \nearrow$	$1.12 \pm 0.06 \nearrow$	$1.14 \pm 0.08 \nearrow$	$1.11 \pm 0.08 \circ$	$1.09 \pm 0.08 \nearrow$
	Times	$0.79 \pm 0.32 \rightarrow^*$	$\mathbf{0.71} \pm 0.22 \circ$	$1.48 \pm 0.44 \nearrow$	$1.48 \pm 0.81 \rightarrow$	$1.56 \pm 0.54 \nearrow$
2005	Amount	$\mathbf{1.29} \pm 0.50 \rightarrow^*$	$1.32 \pm 0.54 \rightarrow^*$	$2.00 \pm 0.44 \nearrow$	$1.59 \pm 0.66 \nearrow$	$2.00 \pm 0.83 \nearrow$
	GDP	$\mathbf{1.12} \pm 0.06 \nearrow$	$1.11 \pm 0.06 \nearrow$	$1.13 \pm 0.08 \nearrow$	$1.10 \pm 0.06 \circ$	$1.10 \pm 0.04 \nearrow$
	Times	$0.81 \pm 0.29 \rightarrow^*$	$\mathbf{0.68} \pm 0.32 \circ$	$1.65 \pm 0.58 \nearrow$	$0.99 \pm 0.65 \rightarrow$	$1.55 \pm 0.59 \nearrow$
2006	Amount	$\mathbf{1.05} \pm 0.43 \rightarrow^*$	$1.04 \pm 0.57 \rightarrow^*$	$2.00 \pm 1.53 \nearrow$	$1.14 \pm 0.93 \circ$	$1.83 \pm 1.34 \nearrow$
	GDP	$\mathbf{1.16} \pm 0.06 \nearrow$	$1.15 \pm 0.07 \nearrow$	$1.18 \pm 0.06 \nearrow$	$1.15 \pm 0.07 \nearrow$	$1.12 \pm 0.06 \nearrow$
	Times	$0.86 \pm 0.29 \rightarrow^*$	$\mathbf{0.83} \pm 0.23 \rightarrow$	$1.67 \pm 0.58 \nearrow$	$1.45 \pm 0.55 \rightarrow$	$1.64 \pm 0.68 \nearrow$
2007	Amount	$\mathbf{1.19} \pm 0.29 \rightarrow^*$	$1.11 \pm 0.33 \rightarrow^*$	$1.41 \pm 0.68 \nearrow$	$1.41 \pm 0.55 \circ$	$1.57 \pm 0.56 \nearrow$
	GDP	$\mathbf{1.17} \pm 0.06 \nearrow$	$1.15 \pm 0.07 \nearrow$	$1.17 \pm 0.07 \nearrow$	$1.15 \pm 0.08 \nearrow$	$1.12 \pm 0.05 \nearrow$
	Times	$0.86 \pm 0.23 \rightarrow^*$	$\mathbf{0.86} \pm 0.17 \rightarrow^*$	$1.68 \pm 0.39 \nearrow$	$1.06 \pm 0.25 \rightarrow$	$1.47 \pm 0.44 \nearrow$
2008	Amount	$\mathbf{0.92} \pm 0.28 \rightarrow^*$	$0.89 \pm 0.34 \rightarrow^*$	$1.59 \pm 0.76 \nearrow$	$0.91 \pm 0.70 \rightarrow$	$1.45 \pm 0.61 \nearrow$
	GDP	$\mathbf{1.16} \pm 0.06 \nearrow$	$1.15 \pm 0.07 \nearrow^*$	$1.16 \pm 0.07 \nearrow$	$1.15 \pm 0.08 \nearrow$	$1.10 \pm 0.05 \nearrow$
	Times	$0.90 \pm 0.19 \rightarrow^*$	$\mathbf{0.89} \pm 0.19 \rightarrow$	$1.73 \pm 0.36 \nearrow$	$1.09 \pm 0.27 \circ$	$1.47 \pm 0.43 \nearrow$
2009	Amount	$\mathbf{1.21} \pm 0.24 \rightarrow^*$	$1.13 \pm 0.33 \rightarrow^*$	$2.00 \pm 1.05 \nearrow$	$1.21 \pm 0.57 \nearrow$	$1.81 \pm 1.03 \nearrow$
	GDP	$\mathbf{1.16} \pm 0.06 \nearrow^*$	$1.15 \pm 0.06 \nearrow^*$	$1.17 \pm 0.08 \nearrow$	$1.15 \pm 0.08 \nearrow$	$1.11 \pm 0.05 \nearrow$
	Times	$0.88 \pm 0.21 \rightarrow^*$	$\mathbf{0.87} \pm 0.18 \rightarrow^*$	$1.64 \pm 0.44 \nearrow$	$1.05 \pm 0.19 \rightarrow$	$1.36 \pm 0.41 \nearrow$
2010	Amount	$\mathbf{1.34} \pm 0.19 \circ^*$	$1.32 \pm 0.24 \circ^*$	$2.00 \pm 0.60 \nearrow$	$1.37 \pm 0.35 \nearrow$	$1.50 \pm 0.63 \nearrow$
	GDP	$\mathbf{1.15} \pm 0.06 \nearrow$	$1.14 \pm 0.07 \nearrow$	$1.13 \pm 0.11 \nearrow$	$1.14 \pm 0.08 \nearrow$	$1.06 \pm 0.07 \nearrow$
	Times	$\mathbf{1.08} \pm 0.15 \rightarrow$	$1.10 \pm 0.16 \rightarrow$	$1.67 \pm 0.39 \nearrow$	$1.17 \pm 0.15 \nearrow$	$1.36 \pm 0.32 \nearrow$

Table 2: comparisons between regression results with five models in the framework of sophisticated maximum likelihood estimation [1]. Each item denotes the scaling exponent $\beta \pm b$, where the error bars b were computed with bootstrapping with replacement. The value obtained through ordinary least-squares (OLS) fitting in log-scale coincides with the lognormal model with fixed fluctuations reported in the first column. The asterisk indicates that the model is compatible with the data, which is identified by a p-value smaller than 0.05. The best model is highlighted in bold, when none of the models is accepted, then it is judged according to the Bayesian Information Criterion (BIC). BIC$= -2\ln\mathcal{L} + k\ln N$, where $\mathcal{L}$ is the maximum-likelihood of the model, k is the number of free parameters, and N the number of observations. If the difference Δ BIC between the BIC of each model with the same model with a fixed $\beta = 1$ is below 0, the model is linear ($\rightarrow$), between zero and six is inconclusive (open circle $\bigcirc$) and higher than six (strong evidence) is super-linear ($\nearrow$)/sublinear ($\searrow$). The models were also compared between each other using the respective BICs within the same noise model (grey background has lower BIC) and between all others (bold text indicates the model with the lowest BIC). See more details of the MLE regression in Ref.[1].

		city model				person model
		lognormal		Gaussian		
	database	$\delta = 2$	$\delta \in [1,3]$	$\delta = 1$	$\delta \in [1,2]$	
2011	Amount	**1.12** $\pm 0.23 \rightarrow$	$1.09 \pm 0.27 \rightarrow$	$1.89 \pm 0.77 \nearrow$	$1.08 \pm 0.37 \rightarrow$	$1.49 \pm 0.58 \nearrow$
	GDP	**1.15** $\pm 0.06 \nearrow^*$	$1.14 \pm 0.07 \nearrow^*$	$1.12 \pm 0.08 \nearrow$	$1.14 \pm 0.09 \nearrow$	$1.06 \pm 0.06 \nearrow$
	Times	$1.03 \pm 0.17 \rightarrow^*$	**1.05** $\pm 0.16 \rightarrow^*$	$1.77 \pm 0.41 \nearrow$	$1.19 \pm 0.15 \nearrow$	$1.43 \pm 0.32 \nearrow$
2012	Amount	**1.16** $\pm 0.26 \rightarrow^*$	$1.13 \pm 0.24 \rightarrow^*$	$1.41 \pm 0.63 \nearrow$	$1.41 \pm 0.37 \nearrow$	$1.46 \pm 0.27 \nearrow$
	GDP	**1.16** $\pm 0.06 \nearrow^*$	$1.14 \pm 0.07 \nearrow^*$	$1.13 \pm 0.11 \nearrow$	$1.14 \pm 0.08 \nearrow$	$1.06 \pm 0.08 \nearrow$
	Times	$1.06 \pm 0.16 \rightarrow^*$	**1.08** $\pm 0.16 \rightarrow^*$	$1.90 \pm 0.32 \nearrow$	$1.28 \pm 0.15 \nearrow$	$1.50 \pm 0.34 \nearrow$
2013	Amount	**1.28** $\pm 0.27 \circ^*$	$1.22 \pm 0.27 \rightarrow^*$	$1.71 \pm 0.46 \nearrow$	$1.53 \pm 0.42 \circ$	$1.62 \pm 0.53 \nearrow$
	GDP	**1.15** $\pm 0.06 \nearrow^*$	$1.13 \pm 0.07 \nearrow^*$	$1.12 \pm 0.10 \nearrow$	$1.14 \pm 0.09 \nearrow$	$1.06 \pm 0.07 \nearrow$
	Times	$1.12 \pm 0.17 \rightarrow$	**1.10** $\pm 0.20 \rightarrow^*$	$1.97 \pm 0.37 \nearrow$	$1.26 \pm 0.18 \nearrow$	$1.59 \pm 0.38 \nearrow$
2014	Amount	**1.41** $\pm 0.38 \circ^*$	$1.62 \pm 0.43 \circ^*$	$2.00 \pm 0.41 \nearrow$	$1.59 \pm 0.48 \nearrow$	$1.97 \pm 0.71 \nearrow$
	GDP	**1.15** $\pm 0.05 \nearrow^*$	$1.13 \pm 0.07 \nearrow^*$	$1.12 \pm 0.09 \nearrow$	$1.13 \pm 0.08 \nearrow$	$1.07 \pm 0.06 \nearrow$
	Times	$1.17 \pm 0.22 \rightarrow^*$	**1.11** $\pm 0.28 \rightarrow$	$2.00 \pm 0.15 \nearrow$	$1.40 \pm 0.25 \nearrow$	$1.81 \pm 0.44 \nearrow$
2015	Amount	**1.53** $\pm 0.32 \nearrow^*$	$1.65 \pm 0.32 \nearrow^*$	$2.00 \pm 0.00 \nearrow$	$1.71 \pm 0.34 \nearrow$	$2.00 \pm 0.41 \nearrow$
	GDP	**1.17** $\pm 0.07 \nearrow^*$	$1.14 \pm 0.07 \nearrow^*$	$1.14 \pm 0.09 \nearrow$	$1.15 \pm 0.08 \nearrow$	$1.08 \pm 0.07 \nearrow$
	Times	$1.22 \pm 0.24 \circ^*$	**1.24** $\pm 0.20 \circ^*$	$2.00 \pm 0.08 \nearrow$	$1.44 \pm 0.12 \nearrow$	$1.79 \pm 0.41 \nearrow$
2016	Amount	**1.55** $\pm 0.26 \nearrow^*$	$1.51 \pm 0.29 \circ^*$	$2.00 \pm 0.07 \nearrow$	$1.53 \pm 0.32 \nearrow$	$1.90 \pm 0.38 \nearrow$
	GDP	**1.17** $\pm 0.05 \nearrow$	$1.15 \pm 0.06 \nearrow$	$1.14 \pm 0.07 \nearrow$	$1.15 \pm 0.08 \nearrow$	$1.10 \pm 0.06 \nearrow$
	Times	$1.27 \pm 0.21 \circ^*$	**1.23** $\pm 0.22 \circ^*$	$2.00 \pm 0.09 \nearrow$	$1.51 \pm 0.21 \nearrow$	$1.91 \pm 0.42 \nearrow$
2017	Amount	**1.28** $\pm 0.30 \rightarrow^*$	$1.23 \pm 0.33 \rightarrow^*$	$1.84 \pm 0.56 \nearrow$	$1.84 \pm 0.59 \nearrow$	$1.92 \pm 0.60 \nearrow$
	GDP	**1.19** $\pm 0.05 \nearrow^*$	$1.17 \pm 0.05 \nearrow^*$	$1.18 \pm 0.08 \nearrow$	$1.17 \pm 0.08 \nearrow$	$1.13 \pm 0.08 \nearrow$
	Times	**1.24** $\pm 0.23 \circ^*$	$1.22 \pm 0.22 \circ^*$	$2.00 \pm 0.12 \nearrow$	$1.37 \pm 0.37 \nearrow$	$1.88 \pm 0.39 \nearrow$

Table 2 (continued).

	(1)	(2)
β	0.9765^{***} (0.0000)	0.9889^{***} (0.0002)
γ	0.0621^{***} (0.0000)	0.1270^{***} (0.0000)

Table 3: From the perspective of econometrics, in order to test if there are significant differences between large cities and other cities, or before and after the year 2013, we introduce an interaction term in the regression model: $\ln(amount_{it}) = \alpha + \beta \times \ln(P_{it}) + \gamma \times I_{it} \times \ln(P_{it}) + \varepsilon_{it}$, where $amount_{it}$ is the total amount of VC investment of city i in the year t, P_{it} is the population size of city i in the year t. In Model 1, the dummy variable I_{it} indicates if a city is a large city or not, which is determined in line with the definition in Fig. 4 in the main text. In Model 2, I_{it} indicates if it is after the year 2013. The results indicate that there are significant differences between large cities and other cities, as well as before the year 2013 and after the year 2013.

References

[1] Leitao JC, Miotto JM, Gerlach M, Altmann EG. 2016 Is this scaling nonlinear?. *Royal Society open science* **3**, 150649.