Size and growth dynamics of online stores: A case of China's Taobao.com

Article *in* Electronic Commerce Research and Applications · May 2016

Impact Factor: 1.48 · DOI: 10.1016/j.elerap.2016.04.005

4 authors, including:

Victor W. K. Chan

Rensselaer Polytechnic Institute

Xuefei Nancy Deng

California State University, Dominguez Hills

Electronic Commerce Research and Applications 17 (2016) 161–172

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Size and growth dynamics of online stores: A case of China's Taobao.com

Baojun Gao [a], Wai Kin (Victor) Chan [b,*], Lei Chi [c], Xuefei (Nancy) Deng [d]

[a] School of Economics and Management, Wuhan University, Wuhan, Hubei Province 430072, China
[b] Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, United States
[c] EmblemHealth, 55 Water Street, New York, NY 10041, United States
[d] College of Business Administration and Public Policy, California State University, Dominguez Hills, Carson, CA 90747, United States

ARTICLE INFO

Article history:
Received 11 March 2015
Received in revised form 5 April 2016
Accepted 9 April 2016

Keywords:
Consumer-to-consumer
Dynamic panel data model
Firm growth
Gibrat's law
System GMM
Taobao

ABSTRACT

This study examines Gibrat's law regarding size–growth relationships in the consumer-to-consumer (C2C) online marketplace. Using dynamic panel data models, we analyze 21,948 e-merchants from 14 industries on Taobao.com. The data analysis shows that Gibrat's law holds for large and mature stores when their size and age exceed certain threshold, but it generally fails to apply to stores whose size and age are below certain threshold. For those small stores, they grow faster than large ones in the C2C e-commerce. Results of the study provide insights into the competitive dynamics and industry structure of the C2C online marketplace.

1. Introduction

Internet-based third-party platforms where many small businesses and individuals sell goods and services provide the foundation of the online marketplace. Because the online marketplace has grown so rapidly, its growth dynamics are enigmatic.

On one hand, e-merchants must build customer trust if they are to survive and grow in a marketplace where transactions are usually among strangers. Sellers' reputations have a strong, positive impact on consumer trust and subsequent purchase intentions, price premiums, and sales volumes (Ba and Pavlou, 2002; Bente et al., 2012; Kim et al., 2008; Utz et al., 2012; Zhang, 2006). In online markets, early movers and big players tend to enjoy reputational advantages (Ou and Davison, 2009). Thus, older and larger e-merchants should generate more sales.

On the other hand, compared with traditional retail, online markets present lower entry barriers and higher profitability (McKinsey, 2013). Small businesses and individuals can launch online storefronts quickly with minimal start-up costs, but competition is intense. Taobao.com is an example of an online market that uses innovative communication technologies and electronic payments such as Wangwang (instant messaging) and Alipay (online payment) to increase transparency, enhance buyer decisions (Holsapple et al., 2014), facilitate buyer/seller trust, and expand subsequent transactions and consumer loyalty (Chen et al., 2009; Ou and Davison, 2009; Ou et al., 2014). In such a hyper-competitive environment, competitive advantages can be transient. Thus, new entrants constantly counteract any size and age advantages.

Those contradictions regarding size–growth relationships motivated us to examine the growth dynamics of stores in the consumer-to-consumer (C2C) online marketplace and to seek answers to three research questions: (1) Does size affect growth for e-merchants? (2) Do older and larger e-merchants perform better and grow faster than younger, smaller e-merchants? (3) What drives e-merchant growth in the online markets?

Gibrat's law states that firms' growth rates are independent of their initial sizes (Gibrat, 1931). It is prominent in economic models explaining growth dynamics and size distributions. Gibrat's law has been tested in manufacturing and service industries by numerous empirical studies (for reviews see Audretsch et al., 2004; Lotti et al., 2003; Santarelli et al., 2006). Most recent studies in the manufacturing context have suggested the "stylized facts" of reverse size–growth relationship, and Gibrat's law holds only for mature and large firms (Fotopoulos and Giotopoulos, 2010; Geroski,

* Corresponding author at: CII 5015 ISE Dept., RPI, 110 8th St., Troy, NY 12180, United States.

E-mail addresses: gaobj@whu.edu.cn (B. Gao), chanw@rpi.edu (W.K.(Victor) Chan), lchi@emblemhealth.com (L. Chi), ndeng@csudh.edu (X.(Nancy) Deng).

http://dx.doi.org/10.1016/j.elerap.2016.04.005
1567-4223/

1995; Goddard et al., 2002; Sutton, 1997). The small services industry rather than the manufacturing industry is now thought to be more subject to Gibrat's law (Audretsch et al., 2004; Giotopoulos and Fotopoulos, 2010; Teruel-Carrizosa, 2010). However, online stores are substantially different from traditional manufacturing and service industries in that they present much lower entry barriers, lower operation costs, and higher competition. Although scattered research has addressed the topic, online e-merchants may present different size distribution and growth dynamics.

We used GMM (Generalized Methods of Moments) dynamic estimator[1] (Arellano and Bover, 1995; Blundell and Bond, 1998) to examine size–growth relationships of stores in Taobao.com, the largest C2C online marketplace comprising 96.4% of online C2C transactions in China. We collected data for 21,948 Taobao online stores across 14 industries from March 2012 to September 2012. Using dynamic panel data models to test Gibrat's law, our results suggest that the law generally tends to fail: Taobao stores show negative size–growth relationships. However, the growth patterns differ by store sizes and ages. More specifically, Gibrat's law fails to apply when sizes and ages are below certain thresholds because smaller and younger stores tend to grow faster than larger and older ones. In contrast, Gibrat's law tends to hold when store sizes and ages are above certain thresholds because their growth rates are independent of sizes.

This study makes two major contributions. First, theoretically, the findings provide new insights into the size–growth relationship of online e-merchants and advance theoretical understandings of market concentration and competition in the online environment. We find a minimum efficient scale (MES) in online markets; that is, the minimum level of size threshold required for establishing consumers' trust and for surviving. Our findings enrich understandings of Gibrat's law in online markets: below the threshold, the law tends to fail; above the threshold, it tends to apply. Second, methodologically, this study diverges from prior e-commerce studies (e.g., Li et al., 2008; Lin et al., 2006) by applying dynamic panel GMM estimator to mitigate estimation bias (Nickell, 1981) and control for problems of unobserved individual fixed effects, growth persistence, and endogeneity (Maças Nunes and Serrasqueiro 2009).

For the rest of this paper, Section 2 reviews literature and theoretical background regarding online markets and Gibrat's law. Sections 3 and 4 discuss our research method and dynamic panel data models for testing Gibrat's law in online markets. Sections 5 and 6 discuss results and findings. Section 7 concludes with limitations and directions for future research.

2. Theoretical background

2.1. Consumer-to-consumer electronic market

This study integrates findings and theories from the literature regarding C2C e-commerce and market structure to examine size–growth relationships of e-merchants in online markets. Prior C2C e-commerce studies have provided insights into the mechanisms of consumer trust building, which drive the e-merchants' growth dynamics in the online market.

First, seller reputation is essential for building consumer trust in C2C markets (Ba and Pavlou, 2002; Kim et al., 2008; Utz et al., 2012). Online reputation systems significantly enhance online sales performance in terms of purchase intentions, price premiums, and sales volumes (Ba and Pavlou, 2002; Bente et al., 2012; Zhang, 2006). However, sellers can use bogus buyers to manipulate their reputations through fake positive feedback. You et al. (2011) analyzed a dataset from Taobao.com and found that transaction indicators (e.g., price, frequency, comment, and connectedness) and individual indicators (e.g., reputation and age) are good sources for identifying collusive traders.

Second, prior studies have examined customer attraction and retention in the C2C market from different perspectives. Chen and colleagues (2007) drew on a customer relationship management (CRM) framework and adopted a social network perspective to study the C2C market in China. They concluded that the C2C market is also a social environment and that online vendors can enhance customer loyalty by managing social relationships. Moreover, Lu et al. (2010) viewed C2C sellers and buyers as members of a virtual community, and empirically tested factors contributing to trust among them, which is essential for platform providers. Chen et al. (2009) studied a sample from Chinese C2C Web sites and found that mutual trust among members can boost their trust of a platform provider so that they become loyal buyers and sellers.

2.2. Market structure and Gibrat's law

Economists have long studied market structure, size distribution, and growth dynamics in manufacturing firms. "Gibrat's rule of proportionate growth," or Gibrat's law (1931) explaining that initial size does not determine growth rate has inspired extensive studies on firm growth and market structure.

Gibrat's initial finding has inspired extensive studies on firm growth and market structure. In general, growth dynamics reveal events and trajectories such as new entrants and growth, decline, and exits of incumbent firms, which all contribute to the size distributions of firms (De Wit, 2005). Proportionate growth assumes that size distribution approaches a lognormal distribution, one of the testable assumptions for Gibrat's law (Ganugi et al., 2005; Reichstein and Jensen, 2005). Nevertheless, the lognormal size distribution is an insufficient but necessary condition for Gibrat's law (Ganugi et al., 2005). A formal test of the law requires investigating the relationship between size and growth. Firm size distributions are often right skewed, making lognormal distribution a candidate for firm size distributions (Coad, 2009). Other distributions such as power law can provide better fits at least for the upper tail of empirical data (Axtell, 2001; Cirillo and Hüsler, 2009; Gao et al., 2015; Hart and Oulton, 1997).

Extensive empirical research has tested Gibrat's law (for reviews, see Audretsch et al., 2004; Lotti et al., 2003; Santarelli et al., 2006). Tschoegl (1983) indicated that Gibrat's law is valid only when firm size does not determine growth rates, individual firms do not grow persistently from one period to the next, and firm size does not determine growth variability. Most empirical research tested only the first two conditions by using econometric analysis. Some econophysics studies focused on the relationship between size and growth variances. For example, Stanley et al. (1996) found that the relationship between size and variance of growth rate approximately follows the power law. Riccaboni et al. (2008) further indicated that the size–variance relationship undergoes a slow crossover rather than a single well-defined event.

Most empirical investigations of Gibrat's Law are in the manufacturing context, but the results are mixed. Some studies validated Gibrat's law or found positive size–growth relationships (Del Monte and Papagni, 2003; Simon and Bonini, 1958; Tschoegl, 1983). Some studies rejected Gibrat's law and showed negative size–growth relationships (Almus, 2000; Evans, 1987; Fotopoulos and Giotopoulos, 2010; Hall, 1987; Hart and Oulton, 1996). Other studies, such as Geroski (1995), Sutton (1997),

[1] We thank two anonymous referees for the suggestion to use dynamic panel data models to investigate size–growth relationships in Taobao stores. Prior studies used Chesher's (1979) cross-sectional function form (e.g., Almus, 2000; Audretsch et al., 2004; Bentzen et al., 2006; Ganugi et al., 2005; Giotopoulos and Fotopoulos, 2010; Petrunia, 2008).

Goddard et al. (2002), and Fotopoulos and Giotopoulos (2010), provide a contingency view based on the industry's minimum efficient scale (MES), defined as the smallest output that a manufacturing firm can produce in order to minimize its long run average costs. According to those scholars, Gibrat's law tends to hold for mature and large firms whose size is larger than the industry's MES, but the law does not hold for firms smaller than the industry's MES: for those small firms, they grow faster than large ones.

Economies of scale offer one explanation for the negative size–growth relationship; that is, firms must reach the MES to survive (Audretsch et al., 2004; Geroski, 1995; Sutton, 1997). Smaller firms must grow faster than larger firms do to achieve the MES as soon as possible or else incur higher costs and failure. Therefore, only surviving firms remained in the data sample, they showed faster growth, along with greater failure (Audretsch et al., 2004; Strotman, 2007).

A negative age–growth relationship is another factor affecting firm growth (Angelini and Generale, 2008; Choi, 2010; Liu et al., 1999; Lotti et al., 2003; Serrasqueiro et al., 2010). Younger firms grow faster partly because of the MES and pressure for survival (Choi, 2010; Liu et al., 1999; Lotti et al., 2003). Early in a firm's life cycle, rapid growth is essential for attaining an optimal level of efficiency; later, growth rate diminishes (Serrasqueiro et al., 2010).

Gibrat's law has been tested less frequently in the service industry. Some studies rejected it (Choi, 2010; Oliveira and Fortunato, 2008), but more validated it (Audretsch et al., 2004; Giotopoulos and Fotopoulos, 2010; Teruel-Carrizosa, 2010). In particular, a study of small Dutch service firms revealed no size–growth relationship perhaps because the service sector has lower entry barriers than the manufacturing industry (Audretsch et al., 2004). Greek ICT intensive service industries showed Gibrat's law in action, but non-ICT service industries did not (Giotopoulos and Fotopoulos, 2010). A study of manufacturing and service industries in Spain indicated that small firms in the service industries grow more slowly mainly because service industries have lower MES and diminished incentives to grow (Teruel-Carrizosa, 2010).

When Gibrat's law was applied to the eBay context and eBay seller reputation was a size proxy, seller reputation followed a log-normal distribution and smaller sellers grew faster, rejecting Gibrat's law (Lin et al., 2006). A study of both China's eBay and Taobao showed a negative relationship between growth and seller size (Li et al., 2008). In studies of the size–growth relationship, Lin et al. (2006) and Li et al. (2008) regressed the log size at time t on log size at $t-1$. Neither study considered growth persistence (Chesher, 1979) or individual fixed effects.

Those previous studies offered useful but static insights into the size–growth relationship of e-merchants in online markets. We extend the literature by providing a dynamic view of the size–growth relationship for capturing C2C e-merchant behavior and growth dynamics. We detail our data and research methods in the next two sections.

3. Data

3.1. Taobao store size measurement

Economics researchers often measure firm size by number of employees, total assets, net assets, sales revenue, and enterprise value (see Coad, 2009; Gao et al., 2015; Segarra and Teruel, 2012). Unfortunately, Taobao does not require stores to release such measures, so they are publicly unavailable. To make the study replicable and comparable, we must use publicly available data as a size proxy.

Online seller reputation, a publicly available proxy for store size (Lin et al. 2006, 2007) can be used to indicate size and business capacity of eBay online auction traders (Lin et al., 2006). Sales volume may be a more appropriate size measure for eBay and Taobao stores because it is difficult to obtain traditional size indicators such as revenue and assets from anonymous online transactions, particularly online C2C markets (Li et al., 2008). Consistent with Li et al. (2008), we also used sales volume rather than reputation as the size proxy of Taobao stores.

On Taobao, buyers and sellers rate each other after buyers confirm receiving the order. Ratings can be positive, neutral, or negative. By subtracting the number of negative ratings from the number of positive ratings, we obtained the reputation score of a Taobao store. As such, reputation reflects a Taobao store's cumulative operation since its inception rather than its current operational status. Consequently, older stores tend to have higher reputation scores.

By default, if buyers fail to rate sellers within a given period, the Taobao trading system assigns a positive seller rating. Therefore, the system counts all orders. The store's sales volume equals the total ratings, positive or negative. Moreover, most ratings are positive on Taobao. In our sample, the average ratio of positive ratings is 0.99, with both ratio median and mean over 99%. With such a high percentage of positive ratings, the seller reputation is approximately equal to the cumulative sales volume, which increases monotonically. Other firm size measures, such as number of employees, total assets, net assets, and sales, are not monotonic, but can increase or decrease depending on operational status.

Sales volume is a more appropriate size proxy for two reasons. First, sales volume may vary daily, monthly, and quarterly according to types of products, quality of service, and promotion efforts. Whereas reputation reflects cumulative operations, sales volume better captures operational status within a given period. Second, monthly sales volume data is publicly available, so future research can replicate our results. Consequently, we chose monthly sales volume to measure the size of Taobao stores, while acknowledging that the measure is imperfect. If other size proxies were available, they would have provided a more holistic measure.

3.2. Sampling and data collection

As of December 31, 2010, 5.8 million Taobao stores were registered and assigned unique integer IDs. We used the unique ID to sample store data. However, the largest store ID could be over 90 million and the ID distribution over the range of integers from 1 to 100 million is unknown to the public. Taobao did not assign most integers within this range, so a simple random sampling would have only a 5.8% probability of returning an actual store ID (5.8 million/100 million).

To increase sampling efficiency, we developed a two-phase sampling procedure. The first phase was to estimate the distribution of actual store IDs over the range of integers from 1 to 100,000. Specifically, we divided the whole range into 1000 equal-width intervals, and then randomly generated 100 integers from each interval. We validated whether a generated integer was an actual store ID by checking whether the webpage corresponding to the ID was accessible. Next, we obtained the estimated percentage of actual IDs within each interval. Based on the estimated proportion of valid IDs, we started the second phase to draw a sample from each interval with the sample size proportional to the percentage of valid IDs in that interval. As a result, we were able to draw more samples from high-proportion intervals and fewer samples from low-proportion intervals.

To ensure that our two-phase sampling procedure produced a good representative sample of the population, we compared the distribution of our sample with Fan et al.'s (2016) dataset that provided percentage distribution for stores in groups of different sizes. By conducting a chi-squared test that treats Fan et al.'s (2016)

Table 1
Descriptive statistics of Taobao stores' sales by industry.[a]

Ind. Code	N. Store	Median	Mean	Max	Std. dev.	Skew	Kurtosis	JB-test	Gini
1	6653	40	248	38,458	1233	0.32	−0.09	115.2	0.83
2	2751	28	172	24,665	750	0.39	−0.09	70.2	0.82
3	628	35	351	34,907	2170	0.51	0.45	32.8	0.89
4	726	16	137	7,114	548	0.64	0.07	50.0	0.85
5	1138	44	291	23,221	1206	0.26	−0.38	19.3	0.82
6	1186	33	227	56,073	1854	0.33	0.02	21.8	0.85
7	550	28	187	18,338	967	0.39	−0.14	14.1	0.83
8	1490	45	308	22,890	1207	0.35	−0.18	31.9	0.83
9	1465	47	311	23,839	1225	0.32	−0.23	28.6	0.83
10	1261	20	175	29,283	1004	0.50	−0.09	53.8	0.85
11	1060	27	98	14,210	468	0.19	−0.41	13.8	0.74
12	1540	15	230	44,997	1827	0.83	0.79	217.5	0.91
13	297	50	210	3,126	385	−0.09	−0.63	5.0	0.72
14	944	29	208	9,813	675	0.38	−0.38	27.8	0.82
All	21,948	32	228	56,073	1220	0.38	−0.12	546.9	0.84

Notes: (a) Data are average sales of stores from March to September 2012. Columns 3–6 show sales volume of Taobao stores; Columns 7–9 show the logarithm of sales volumes.
(b) The sample includes 21,948 Taobao stores; 21,689 stores provided industry information.
(c) The column JB-test reports the Jarque–Bera test statistics. All the JB-tests are significant at the 0.001 level for the 14 industries and for the whole sample.
(d) The 14 industries are: 1 = Apparel, Shoes, & Bags; 2 = 3C Digitals & Computers; 3 = Jewelry & Accessories; 4 = Leisure & Antiques; 5 = Baby & Children; 6 = Sports, Fitness, & Outdoors; 7 = Auto & Parts Supplies; 8 = Beauty; 9 = Home Essentials; 10 = Living & Services; 11 = Home Improvement; 12 = Video Games; 13 = Movies, Music, & Books; 14 = Food & Health.

distribution as the true distribution, we obtained a chi-square statistic 3.1 with a p-value 0.99, suggesting no significant difference between the datasets.

Using this sampling approach, we collected a data sample of 70,000 Taobao stores in April 2011. By March 2012, only 55,147 Taobao store websites were accessible; the other stores appeared to have exited the market. We scraped the data of the surviving 55,147 stores on a monthly basis from March 2012 to September 2012, and generated the raw dataset for this study.

It's common for some Taobao stores to have no sales in some months and then generates sales in subsequent months. This sales pattern differentiates C2C stores from traditional brick-and-mortar firms. Among the 55,147 stores scraped monthly from March to September in 2012, only 16,344 had sales in all months of the period, and the remaining 38,803 reported zero sales in at least one month during the period. In addition, we treated the 12,647 stores that had zero sales in all months as non-active, although Taobao listed them as "alive". They could not have survived if they had gone offline given the high operational costs in the traditional market.

The last step is to clean the data so that it meets the requirements of econometric methodology. First, following a common data-cleaning procedure (e.g., Li et al., 2008; Oliveira and Fortunato, 2006a), we excluded observations with missing values or zero sales to avoid taking the logarithm of zeros. Second, we kept only firms with at least five consecutive non-zero sales periods after removing the observations, because the system GMM estimator requires at least four consecutive periods, and we need one extra period to calculate growth rate (Bond et al., 2001). The data-cleaning process produced a final sample of 21,948 stores during the seven-month period from March to September 2012.

3.3. Summary statistics

To investigate the size distribution and size–growth relationship of online stores, we analyzed the data by industries, which is a common practice because different industries may yield different results (Audretsch et al., 2004; Bentzen et al., 2006; Goddard et al., 2006; Lotti et al., 2003; Lotti and Santarelli, 2004; Petrunia, 2008; Teruel-Carrizosa, 2010; Zhang et al., 2009; Liu, 2016). Moreover, separating the data sample by industry allowed us to see whether Gibrat's law depends on industry for validity.

Table 1, which presents the summary statistics of sales volume, shows that the mean values of sales are always substantially larger than median values. This is not surprising given that we expect a skewed firm size distribution. In addition, the skewness measures of logarithmic sales of all industries except Industry 13 are positive, indicating that the logarithmic sales are still right skewed. This suggests that firm size distribution is more skewed than a lognormal distribution in most industries.

If Gibrat's law holds, the size distribution should be approximately lognormal. Based on the results of Jarque–Bera (JB) test in the second-last column of Table 1, we rejected the log-normality hypotheses at the 0.01 significance level for the fourteen Taobao industries and the whole sample. However, as log-normality is a result of proportionate growth, rejecting the lognormal hypothesis is insufficient to reject Gibrat's law (Sutton, 1997). We also computed the Gini coefficient to measure unequal sales distribution. As the last column of Table 1 shows, all Gini coefficients are greater than 0.7, and most are greater than 0.8, indicating that sales vary significantly among stores within the same industry. That is, a small number of big stores generated most sales.

As firms grow, the distribution of log firm size moves toward the right side (i.e., right tail), and the degree of skewness decreases significantly (Angelini and Generale, 2008; Cabral and Mata, 2003). To investigate the size distribution of different age groups, we split the data into two subsamples based on the median age of an industry. Table 2 shows some key statistics for the size distribution of the younger and older groups: the mean and standard deviation of the older groups are generally larger than those values of the younger group, with a few exceptions. However, the skewness of the older group is not consistently smaller than that of the younger group. On the contrary, among the fourteen industries, the skewness measures of the eight industries in the older group are larger than the skewness measures of the same industries in the younger group. This result differs from that of offline firms (Angelini and Generale, 2008; Cabral and Mata, 2003) as the skewness of those firms tends to diminish in age (Angelini and Generale, 2008).

As to the normality test, our data did not provide sufficient evidence to reject the null hypothesis for four industries in the younger group and three industries in the older group. Even when we found sufficient evidence to reject the null hypothesis, some JB-tests were significant only at the 0.05 significance level, but

Table 2
Key statistics of younger and older groups.

	Stores younger than industry age median				Stores older than industry age median			
Ind_Code	Mean	Std. dev.	Skew	JB-test	Mean	Std. dev.	Skew	JB-test
1	141	467	0.15	20.1***	370	1734	0.26	47.3***
2	138	553	0.26	16.1***	209	920	0.38	32.4***
3	252	1476	0.46	10.4**	489	2875	0.52	15.5***
4	128	367	0.22	12.9**	231	1063	0.54	28.6***
5	268	1155	0.12	0.7	424	1622	0.19	4.5
6	106	277	0.20	8.8*	332	2628	0.26	6.3*
7	143	704	0.32	4.5	235	1212	0.23	3.4
8	166	670	0.21	5.7	481	1822	0.22	11.2**
9	233	919	0.30	13.1**	334	920	0.22	14.2**
10	116	425	0.36	16.2***	251	1414	0.52	26.7***
11	82	168	0.10	6.4*	119	670	0.22	7.9*
12	211	1353	0.94	149.0***	195	1602	0.64	49.8***
13	237	427	−0.20	2.6	153	294	−0.02	1.6
14	142	456	0.32	9.5**	275	872	0.31	13.3**

Notes: ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 level, respectively. Refer to the notes of Table 1 for corresponding names of the 14 industries.

Table 3
Descriptive statistics of store size, growth, and age.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std.dev
Size	0	2.30	3.56	3.62	4.90	11.25	1059
Growth	−4.21	−0.44	−0.02	−0.03	0.36	4.67	0.86
Age	11	24	36	40	52	113	20.56

when both age groups were combined, the null hypotheses were rejected at the 0.001 significance level.

Table 3 reports the summary statistics of the three key variables—size, growth, and age—used in the econometric analysis for the whole sample. Size is the logarithmic sales; growth is the dependent variable measured by the difference of logarithmic sales in two consecutive periods; age is a variable measured by the number of months since the store's registration.

4. Model specification

Gibrat's law states that the growth rate of each firm in an industry is unrelated to its current size. According to Goddard et al. (2002), the model of firm growth and size is:

$$Growth_{it} = \alpha_i + \delta_t + (\beta - 1)Size_{it-1} + u_{it}; \quad u_{it} = \rho u_{it-1} + \varepsilon_{it} \tag{1}$$

where t and i indicate time period and firm, respectively. $Size_{it-1}$ is the natural log size for firm i at time $t-1$ and $Growth_{it}$ is the growth rate of firm i at time t measured by the difference between $Size_{it}$ and $Size_{it-1}$. δ_t is time effects, and α_i is the unobserved time-invariant individual (firm) specific effects, which allows for the heterogeneity across firms. Parameter β determines the relationship between size and growth. ρ captures serial correlation in u_{it}—the disturbance term of the growth equation. ε_{it} is a random disturbance assumed to be normal, independent, and identically distributed (IID) with $E(\varepsilon_{it}) = 0$ and $\text{var}(\varepsilon_{it}) > 0$.

In Eq. (1), the size–growth relationship can be tested based on the estimate of β and the following hypotheses: the null hypothesis (H_0: $\beta = 1$) states that the probability distribution of growth rates is the same for all firm classes. If $\beta \geqslant 1$, $\alpha_i = 0$ for all i, it implies an explosive growth path: firms tend to grow faster as they get larger. Such a pattern is conceivable for a limited time, but presumably could not continue indefinitely. The variance of the cross-sectional firm size distribution and the level of concentration both increase over time. $\beta = 1$ implies that the growth is unrelated to firm size. In this case, Gibrat's law holds, and the mean and variance of growth are independent of size. If $\beta < 1$, firm size is mean-reverting and small firms tend to grow faster. In this case, α_i can be considered IID with $E(\alpha_i) = 0$ and $\text{var}(\alpha_i) \geqslant 0$. If $\text{var}(\alpha_i) = 0$, individual firm effects are homogenous (all firms tend to revert toward the same mean size); and if $\text{Var}(\alpha_i) > 0$, they are heterogeneous; that is, the mean size is firm-specific.

Even if β is close to unity, Gibrat's law will not hold if the error terms are serially correlated (i.e., $\rho \neq 0$ in Eq. (1)) (Chesher, 1979). Serial correlation in growth rates can be ascribed to the persistence of chance factors that contribute to abnormal growth. A positive value of ρ suggests that an above average growth rate persists from one period to the next, while a negative value of ρ indicates below average to above average growth rates in consecutive periods.

By assuming homogeneity between firms, that is, assuming $\text{Var}(\alpha_i) = 0$, we can re-write $\alpha_i = \alpha$ in Eq. (1), reducing it to the cross-sectional function form used in many empirical studies (Audretsch et al., 2004; Bentzen et al., 2006; Chesher, 1979; Ganugi et al., 2005; Lotti, et al. 2003; Petrunia, 2008). However, if heterogeneity is present in α_i, firm specific effects α_i is subsumed in the disturbance term. In other words, if $\text{Var}(\alpha_i) > 0$, the cross sectional ordinary least squares (OLS) estimator is inconsistent and may overestimate β whenever $\beta < 1$. This is a special case of Nickell's (1981) results.

As such, panel data model is appropriate for dealing with firm heterogeneity. Note also that Gibrat's law is equivalent to the presence of a unit-root in log firm size. Gibrat's law has been tested with panel unit root tests (Chu et al., 2008; Del Monte and Papagni, 2003; Goddard et al., 2002; Harris and Trainor, 2005; Resende, 2004). As Goddard et al. (2002) suggested, we rewrote Eq. (1) for panel estimation:

$$Growth_{it} = \alpha_i + \delta_t + (\beta - 1)Size_{it-1} + \rho Growth_{it-1} + \gamma Age_{it-1} + \eta_{it} \tag{2}$$

where $\eta_{it} = \varepsilon_{it} + \rho(1-\beta)Size_{it-2}$, so $\eta_{it} = \varepsilon_{it}$ under H_0: $\beta = 1$.

In Eq. (2), Age_{it} is the logarithmic age of Store i in month t. We incorporated age in the right side of Eq. (2) because age and growth often have a negative relation (Angelini and Generale, 2008; Choi, 2010; Liu et al., 1999; Lotti et al., 2003).

Eq. (2) is a dynamic panel data model as it contains one lagged dependent variable. To deal with the unobserved heterogeneity (α_i), if the within (demeaning) transformation is applied, the results will suffer from "Nickell bias," particularly under a small number of time periods (T) and a large number of individual units (N) because the demeaning process creates a correlation between the regressor and the error term of the model (Nickell, 1981). The correlation between the regressor and error leads to biased and inconsistent estimates. The degree of bias is of order $1/T$. For small T, this can be a big problem. The bias disappears only when T goes to infinity.

To deal with this bias, we use the system GMM estimator (Arellano and Bover, 1995; Blundell and Bond, 1998), consistent with some recent studies of Gibrat's law and firm growth (Giotopoulos and Fotopoulos, 2010; Maçãs Nunes and Serrasqueiro, 2009; Oliveira and Fortunato, 2006a,b). The system GMM estimator uses equations in first-differences, and eliminates individual specific effects by transformation. This estimator deals with endogeneity by using deeper lags of endogenous variables (including the lagged dependent variable and other endogenous explanatory variables in the right side) to instrument the earlier ones. The assumption made on variable endogeneity determines the instruments used in the model. We assumed that firm size is an endogenous variable and firm age is pre-determined. Therefore, we used the deeper lags of firm growth and size to instrument the earlier ones. We then used the Sargan test for the validity of over-identifying restrictions (Arellano and Bond, 1991) to select appropriate instruments for each estimation. Guided by the Sargan test results on a variety of specifications, we used only the second and a varying number of higher-order lags of firm growth and size as instruments in the estimations.

To ensure the validity of the results of dynamic estimators, the data must meet two conditions: (1) the restrictions, a consequence of using instrumental variables, must be valid; and (2) no second-order autocorrelations should occur in the first-differenced residuals (Arellano and Bond, 1991). We tested instrument validity by using a Sargan test of over-identifying restrictions, which has an asymptotic chi-square distribution under the null hypothesis that these moment conditions are valid. The Sargan test checks the null hypothesis of validity of the instruments used and consequent restrictions against the alternative hypothesis of non-validity of the instruments used and consequent restrictions. If the null hypothesis is rejected, the system GMM estimator results cannot be considered valid.

Testing for autocorrelation of the residuals is another diagnosis procedure in dynamic panel data estimation. As residuals of the differenced equation often have serial correlation, it is important to ensure that the differenced residuals have no significant second-order autocorrelation, because such an autocorrelation will make the second lags of endogenous variables inappropriate instruments for their current values. Hence, to ensure the appropriateness of instruments in the model, we performed the second-order autocorrelation test to check the null hypothesis (absence of second-order autocorrelation) against the alternative hypothesis (existence of second-order autocorrelation). If we reject the null hypothesis, we reject the validity of the GMM estimator system as well. When the assumption of serial independence in the original error term is satisfied, the differenced residuals are unlikely to exhibit significant second-order autocorrelation.

In summary, the GMM system gives possibly valid results only if both the Sargan test and the second-order autocorrelation test do not reject their null hypotheses.

5. Results

Table 4 shows the system GMM estimation results for the fourteen Taobao industries. March to September 2012 is the period for the panel. We used a two-step version of the system GMM (Arellano and Bover, 1995) with standard errors that are asymptotically robust to heteroskedasticity. In each estimation, we controlled for the time fixed effects and individual fixed effects.

As the upper panel of Table 4 shows, the estimated coefficients of size are all negative and the coefficients in twelve of fourteen Taobao industries are significant, indicating that smaller stores

Table 4
System GMM estimation results for the whole sample.

Dependent variable: $Growth_{it}$														
Industries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Growth_{it-1}$	0.055***	−0.027	0.022	−0.022	0.003	0.031	−0.099**	−0.024	0.040·	−0.020	−0.085**	−0.090***	−0.037	−0.043
	(0.010)	(0.016)	(0.032)	(0.036)	(0.025)	(0.024)	(0.032)	(0.026)	(0.022)	(0.032)	(0.028)	(0.026)	(0.040)	(0.038)
$Size_{it-1}$	−0.070***	−0.083**	−0.139*	−0.238***	−0.048	−0.101**	−0.032	−0.140***	−0.100***	−0.168***	−0.115**	−0.009	−0.113*	−0.106
	(0.011)	(0.030)	(0.062)	(0.061)	(0.030)	(0.031)	(0.042)	(0.035)	(0.024)	(0.039)	(0.041)	(0.039)	(0.058)	(0.067)
Age_{it-1}	0.034***	0.011	0.040	−0.019	0.015	0.059*	−0.014	0.131***	0.032	0.035	0.041·	0.021	−0.028	0.047
	(0.008)	(0.012)	(0.029)	(0.037)	(0.022)	(0.023)	(0.031)	(0.031)	(0.021)	(0.028)	(0.023)	(0.020)	(0.039)	(0.040)
N. Stores	6,653	2,751	628	726	1,138	1,186	550	1,490	1,465	1,261	1,060	1,540	297	944
N. Obs.	32,049	14,208	3,061	3,509	5,733	5,945	2,880	7,749	7,563	6,154	5,538	6,873	1,502	4,719
$Wald_{JS}$	46.804	39.015	6.153	44.529	4.424	13.59	19.017	47.556	18.864	46.772	54.42	29.028	8.22	15.385
	[0.000]	[0.000]	[0.104]	[0.000]	[0.219]	[0.004]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.042]	[0.002]
Sargan test	247.811	25.588	15.932	22.033	37.756	36.805	23.439	17.344	40.995	41.394	19.584	33.964	18.646	30.522
	[0.000]	[0.082]	[0.597]	[0.577]	[0.001]	[0.001]	[0.174]	[0.500]	[0.012]	[0.001]	[0.357]	[0.066]	[0.230]	[0.023]
$m2$	−0.193	−1.585	0.327	0.236	0.932	−1.599	−1.506	−1.508	−0.36	−0.596	−0.982	−0.002	0.217	−0.269
	[0.847]	[0.113]	[0.744]	[0.813]	[0.351]	[0.110]	[0.132]	[0.132]	[0.719]	[0.551]	[0.326]	[0.998]	[0.828]	[0.788]

Notes: (1) All estimates control for the time fixed effects and individual fixed effects. (2) Asymptotic standard errors robust to heteroskedasticity are in parentheses. The null hypothesis that each coefficient is equal to zero is tested using robust standard errors. (3) ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 level, respectively. (4) $Wald_{JS}$ is the Wald statistic of joint significance of the independent variables excluding time dummies and the constant term. *P*-values are in square brackets. (5) Sargan test is a test of the over-identifying restrictions based on the efficient two-step GMM estimator, asymptotically distributed as χ^2 under the null of instruments' validity. *p*-values are in square brackets. We assumed that *Size* is an endogenous variable and *Age* is pre-determined. Guided by the results of Sargan test on a variety of specifications, only the second and a varying number of higher-order lags of *Growth* and *Size* are used as instruments in the estimations. (6) $m2$ is tests for the second-order serial correlation in the first-differenced residuals, under the null hypothesis of no serial correlation (based on the efficient two-step GMM estimator). *p*-values are in square brackets. (7) N. Stores and N. Obs. indicate the number of Taobao stores and number of observations used in each estimation. (8) Refer to the Table 1notes for the names of the 14 industries generally grew faster during the observed period. The estimated coefficients of Age are either positive or negative, and most are not significant. With respect to serial correlation in growth rate; that is, coefficient of $Growth_{it-1}$, most of the fourteen industries have a negative coefficient, while ten of fourteen cases are not significant at the 0.05 significance level.

In the lower panel of Table 4, the first two rows, "*N.* Stores" and "*N.* Obs." indicate the number of e-merchants (Taobao stores) and number of observations in each estimation. The next two rows are the results of Wald joint test ($Wald_{JS}$), which tests the joint significance of the estimated coefficients with a null hypothesis that all coefficients, except for the coefficients for time dummies, are equal to zero jointly. According to the *p*-value shown in squared brackets, we rejected the null hypothesis at the 0.001 significance level, except for Industries 3 and 5.

The validity of the dynamic estimator depends on the validity of the instrument set measured by the Sargan test. The first and second rows of the Sargan test in Table 4 are the chi-square statistics and *p*-values, respectively. As the Sargan tests are significant at the 0.05 significance level for Industries 1, 5, 6, 9, 10, and 14, the null hypotheses (i.e., the instruments are valid) are rejected.

The consistency of the GMM estimator also requires the absence of second-order serial correlation in the residuals, testable by using the *m*2 statistics. In Table 4, the numbers in square brackets below the *m*2 statistics are the *p*-values of the test. All the *m*2 statistics are insignificant, which shows no significant evidence of second-order serial correlation for the fourteen cases. Based on the results of the Sargan tests and second-order autocorrelation tests, we concluded that the results of the system GMM dynamic estimator are valid for only nine of fourteen industries, except for Industries 1, 5, 6, 9, 10, and 14.

As discussed in Section 2, Gibrat's law might hold only for large and mature firms but not for small and young firms (Audretsch et al., 2004; Fotopoulos and Giotopoulos, 2010; Geroski, 1995; Sutton, 1997). To examine whether this result applies to our dataset, we separated the Taobao stores into groups based on size and age. We then tested whether the size–growth relationship varies across different size and age groups. Specifically, for the size quantity, we divided the sample into larger and smaller groups based on the size median of each industry. Similarly, for age, we split the data into two groups based on the median age of each industry, as was done previously (Audretsch, 1995; Comanor and Wilson, 1967; Daunfeldt and Elert, 2013; Sutton, 1991).

Tables 5 and 6 report the results for the smaller size subsample and the larger size subsample, respectively. Overall, larger and smaller stores have quite different size–growth relationships.

As Table 5 shows, for the smaller size subsample, the size coefficients are all negative. Moreover, the coefficients are all significant except for Industry 3, which indicates that the smaller stores generally grow faster, consistent with the results for the whole sample shown in Table 4. With regard to the Sargan and second-order serial correlation tests, the null hypothesis of the Sargan test is accepted for ten of fourteen industries, except Industries 1, 2, 6, and 10, while the null hypothesis of the second-order serial correlation test is accepted in all cases. This indicated instrument validity for ten industries but not for four industries.

Larger stores, however, showed quite different results. As Table 6 shows, the Sargan and second-order serial correlation tests support the validity of the instruments for eleven of fourteen industries, except Industries 1, 2, and 5. All coefficients of size remain negative, but are not significant at the 0.05 significance level for eight of fourteen industries, indicating that their growth tends to be *independent* of size. The Wald joint test further supports this result. The *p*-values of the Wald test for Industries 3, 7, 8, 12, 13, and 14 are all larger than 0.1. For these six industries, one cannot reject the null hypothesis that the coefficients of size, growth,

Table 5
System GMM estimation results for store size smaller than the industry median.

Industries	Dependent variable: $Growth_{it}$													
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Growth_{it-1}$	−0.035*	−0.087**	−0.032	−0.055	−0.091**	−0.044	−0.111	−0.002	−0.077*	0.024	−0.102*	−0.174***	−0.021	−0.043
	(0.018)	(0.026)	(0.059)	(0.065)	(0.035)	(0.039)	(0.069)	(0.039)	(0.031)	(0.046)	(0.045)	(0.040)	(0.080)	(0.041)
$Size_{it-1}$	−0.109**	−0.211***	−0.249	−0.589***	−0.173*	−0.234**	−0.326**	−0.369***	−0.167**	−0.526***	−0.279**	−0.232*	−0.435**	−0.440**
	(0.038)	(0.062)	(0.144)	(0.147)	(0.076)	(0.086)	(0.125)	(0.080)	(0.063)	(0.108)	(0.089)	(0.103)	(0.161)	(0.144)
Age_{it-1}	0.003	0.019	−0.003	0.001	0.004	0.032	0.010	0.064*	0.011	−0.017	−0.008	0.018	0.042	−0.005
	(0.011)	(0.020)	(0.039)	(0.051)	(0.030)	(0.030)	(0.052)	(0.031)	(0.026)	(0.036)	(0.040)	(0.025)	(0.093)	(0.041)
N. Stores	3,326	1,372	314	363	569	593	274	745	732	627	529	770	148	470
N. Obs.	14,481	6,490	1,352	1,520	2,600	2,706	1,295	3,500	3,456	2,726	2,505	3,093	677	2,109
$Wald_{JS}$	59.642	132.964	14.693	66.991	42.775	46.195	59.002	63.881	45.097	62.775	73.677	117.864	20.889	38.042
	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Sargan test	69.217	27.79	17.624	20.289	29.218	25.566	20.439	21.681	28.784	39.686	24.517	17.359	32.475	22.552
	[0.000]	[0.052]	[0.283]	[0.317]	[0.212]	[0.043]	[0.309]	[0.300]	[0.228]	[0.005]	[0.139]	[0.298]	[0.116]	[0.311]
m2	−0.732	−1.111	0.491	0.091	−1.034	−1.351	−1.386	−1.206	−1.024	0.626	−0.509	−0.333	−0.757	−1.892
	[0.464]	[0.266]	[0.624]	[0.927]	[0.301]	[0.177]	[0.166]	[0.228]	[0.306]	[0.532]	[0.611]	[0.739]	[0.449]	[0.059]

Notes: as in Table 4.

Table 6
System GMM estimation results for store size larger than the industry median.

	Dependent variable: Growth$_{it}$													
Industries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Growth$_{it-1}$	0.201***	0.071***	0.087·	−0.029	0.253***	0.176***	0.086*	0.005	0.200***	0.028	−0.007	0.033	0.048	0.041
	(0.012)	(0.017)	(0.046)	(0.041)	(0.038)	(0.031)	(0.043)	(0.034)	(0.031)	(0.028)	(0.043)	(0.034)	(0.057)	(0.041)
Size$_{it-1}$	−0.252***	−0.067*	−0.112·	−0.082	−0.300***	−0.164***	−0.166·	−0.048	−0.217***	−0.158***	−0.118·	−0.045	−0.055	−0.098·
	(0.021)	(0.030)	(0.057)	(0.069)	(0.060)	(0.034)	(0.101)	(0.041)	(0.039)	(0.040)	(0.066)	(0.053)	(0.080)	(0.057)
Age$_{it-1}$	0.128***	−0.015	0.046	0.032	0.146**	0.041	0.053	0.039	0.038	0.020	0.063*	0.023	−0.038	0.061
	(0.017)	(0.014)	(0.038)	(0.026)	(0.050)	(0.026)	(0.051)	(0.036)	(0.036)	(0.024)	(0.028)	(0.023)	(0.053)	(0.043)
N. Stores	3,323	1,375	314	363	569	593	274	745	731	630	529	770	148	471
N. Obs.	17,544	7,697	1,709	1,989	3,133	3,239	1,573	4,249	4,095	3,408	3,021	3,780	819	2,598
Wald$_{JS}$	294.844	25.421	5.09	9.195	46.183	37.121	4.727	2.16	54.959	17.859	9.523	1.323	1.105	3.157
	[0.000]	[0.000]	[0.165]	[0.027]	[0.000]	[0.000]	[0.193]	[0.540]	[0.000]	[0.000]	[0.023]	[0.724]	[0.776]	[0.368]
Sargan test	144.111	39.148	21.285	22.333	60.398	25.44	16.835	24.976	23.71	17.484	13.465	28.711	21.5	32.14
	[0.000]	[0.002]	[0.321]	[0.099]	[0.000]	[0.085]	[0.329]	[0.126]	[0.165]	[0.490]	[0.763]	[0.231]	[0.255]	[0.124]
m2	0.851	−2.09	−0.39	0.276	3.384	−0.91	−0.612	−0.087	−1.424	−1.231	−0.737	−0.309	1.413	0.381
	[0.395]	[0.037]	[0.697]	[0.783]	[0.001]	[0.363]	[0.540]	[0.931]	[0.154]	[0.218]	[0.461]	[0.758]	[0.158]	[0.703]

Notes: as in Table 4.

Table 7
System GMM estimation results for store age younger than the industry median.

	Dependent variable: Growth$_{it}$													
Industries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Growth$_{it-1}$	0.058***	0.012	0.051	0.003	0.040	0.070*	−0.093*	−0.020	0.050·	−0.013	−0.061	−0.159***	0.097·	−0.092*
	(0.014)	(0.024)	(0.043)	(0.047)	(0.037)	(0.033)	(0.041)	(0.039)	(0.029)	(0.040)	(0.038)	(0.035)	(0.053)	(0.037)
Size$_{it-1}$	−0.069***	−0.137**	−0.102	−0.228***	−0.129***	−0.144**	−0.044	−0.158**	−0.071*	−0.143**	−0.136**	0.079	−0.139*	−0.055
	(0.017)	(0.043)	(0.065)	(0.061)	(0.036)	(0.048)	(0.057)	(0.056)	(0.028)	(0.053)	(0.047)	(0.060)	(0.059)	(0.058)
Age$_{it-1}$	0.002	0.001	0.062	−0.020	−0.057	0.054	−0.099	0.119*	0.076·	0.079	0.061	0.037	0.020	0.004
	(0.021)	(0.034)	(0.068)	(0.079)	(0.056)	(0.059)	(0.074)	(0.059)	(0.040)	(0.056)	(0.062)	(0.063)	(0.085)	(0.058)
N. Stores	3,326	1,372	314	363	569	593	275	745	731	630	530	770	148	472
N. Obs.	15,638	6,989	1,496	1,759	2,824	2,862	1,442	3,782	3,710	3,000	2,758	3,327	741	2,345
Wald$_{JS}$	20.972	20.07	2.559	25.502	15.833	10.169	9.739	25.305	9.068	17.92	27.064	28.512	6.319	16.678
	[0.000]	[0.000]	[0.465]	[0.000]	[0.001]	[0.017]	[0.021]	[0.000]	[0.028]	[0.000]	[0.000]	[0.000]	[0.097]	[0.001]
Sargan test	80.828	18.879	13.059	10.471	25.986	31.142	21.638	18.265	23.538	32.431	20.143	30.36	15.232	33.092
	[0.000]	[0.219]	[0.788]	[0.915]	[0.131]	[0.008]	[0.248]	[0.438]	[0.488]	[0.020]	[0.325]	[0.173]	[0.646]	[0.102]
m2	−0.841	−0.836	1.5	0.904	0.1	−0.998	−1.367	−1.308	0.452	0.913	0.044	0.102	−0.75	−1.382
	[0.400]	[0.403]	[0.134]	[0.366]	[0.920]	[0.318]	[0.171]	[0.191]	[0.651]	[0.361]	[0.965]	[0.919]	[0.453]	[0.167]

Notes: as in Table 4.

Table 8
System GMM estimation results for store age older than the industry median.

	Dependent variable: Growth$_{it}$													
Industries	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Growth_{it-1}$	0.055^{***}	$-0.057^{\cdot}$	-0.025	-0.136^{**}	0.019	0.007	$-0.072^{\cdot}$	-0.066^{*}	0.042	-0.005	-0.072^{*}	-0.077^{*}	-0.168^{**}	-0.037
	(0.014)	(0.030)	(0.044)	(0.045)	(0.030)	(0.034)	(0.043)	(0.032)	(0.033)	(0.031)	(0.034)	(0.031)	(0.061)	(0.040)
$Size_{it-1}$	-0.079^{***}	-0.043	-0.098	-0.039	$-0.067^{\cdot}$	-0.077^{**}	-0.036	$-0.060^{\cdot}$	-0.139^{***}	-0.141^{**}	$-0.107^{\cdot}$	-0.045	-0.053	-0.113
	(0.016)	(0.052)	(0.091)	(0.079)	(0.034)	(0.029)	(0.057)	(0.033)	(0.038)	(0.043)	(0.059)	(0.041)	(0.107)	(0.071)
Age_{it-1}	0.049^{**}	0.027	-0.065	0.043	0.052	0.071	0.063	$0.091^{\cdot}$	0.085	0.043	0.034	0.010	-0.087	0.024
	(0.018)	(0.033)	(0.076)	(0.062)	(0.053)	(0.045)	(0.050)	(0.049)	(0.053)	(0.060)	(0.061)	(0.045)	(0.087)	(0.075)
N. Stores	3,326	1,375	314	363	569	593	275	745	732	630	530	770	148	472
N. Obs.	16,407	7,195	1,565	1,750	2,909	3,083	1,438	3,967	3,841	3,152	2,780	3,546	755	2,374
$Wald_{JS}$	27.723	22.163	6.615	25.602	4.259	11.054	9.624	19.946	14.808	16.401	24.285	15.94	15.584	9.142
	[0.000]	[0.000]	[0.085]	[0.000]	[0.235]	[0.011]	[0.022]	[0.000]	[0.002]	[0.001]	[0.000]	[0.001]	[0.001]	[0.027]
Sargan test	160.997	24.163	22.808	21.117	38.417	23.635	14.081	24.511	41.402	40.267	15.641	23.586	17.275	27.133
	[0.000]	[0.150]	[0.246]	[0.274]	[0.003]	[0.167]	[0.724]	[0.139]	[0.003]	[0.005]	[0.739]	[0.485]	[0.504]	[0.298]
$m2$	0.645	-1.159	1.229	-1.472	0.58	-1.052	-1.109	-1.108	-0.66	-1.541	-1.274	-0.532	0.634	-1.286
	[0.519]	[0.246]	[0.219]	[0.141]	[0.562]	[0.293]	[0.268]	[0.268]	[0.509]	[0.123]	[0.203]	[0.595]	[0.526]	[0.198]

Notes: as in Table 4.

and age are equal to zero jointly; namely, the joint hypotheses of $\beta = 1$ and $\rho = 0$ are not rejected. Thus, Gibrat's law still holds for these six industries. The coefficients of size are not significant for Industries 4 and 11, indicating that their growth rates are independent of size. When we excluded Industries 1, 2, and 5 for failing the instrument validity test and examined the remaining eleven industries, the growth was independent of size for eight of the eleven industries. In other words, as stores become larger, they were more likely to adhere to Gibrat's law. Thus, Gibrat's law has a higher possibility of holding for larger stores, which is consistent with previous studies (Evans, 1987; Hall, 1987; Hart and Oulton, 1996).

Similarly, the size–growth relationships of the two age groups also differ significantly. Table 7 shows the results for the subsample of the younger group. As Table 7 shows, the null hypothesis of the Sargan test is rejected for Industries 1, 6, and 10. The *p*-value of the Wald conjoint test is nearly 0.1 for Industry 13, indicating that Gibrat's law holds for this industry. The size coefficients of fourteen industries are all negative, and ten are significant. Excluding Industries 1, 6, and 10, which failed in the instrument validity test and instead examining the remaining eleven industries, six have a negative and significant size coefficient and an insignificant result for the Wald conjoint test. That is, the younger stores tend to grow faster in six of eleven industries for the younger subsample.

Table 8 shows the results for the stores older than the industry age median. The older group shows overwhelmingly different results. The size coefficients are all negative, but ten are not significant at the 0.05 significance level. The null hypothesis of the Sargan test is rejected for Industries 1, 5, 8, and 9. When they were excluded, nine of the remaining ten industries have a negative but insignificant size coefficient, which indicates that their growth is independent of their size. Once again, as stores get older, they show smaller deviations from Gibrat's law. In other words, Gibrat's law has a higher possibility of holding for older stores.

The results of these four subsamples led us to conclude that store growth tends to be independent of store size when size or age reaches a certain threshold. In this study, we specified the respective median values for each industry as the thresholds of age and size for that industry.

6. Discussion

Our finding that smaller stores grow faster than larger ones suggests that Gibrat's law does not generally hold in the fourteen industries on Taobao. Furthermore, Gibrat's law is conditioned on store age and size. For stores whose size and age are above certain threshold, store growth is more likely to be independent of size. For stores whose size and age are below certain threshold, store growth is negatively associated with size.

These results are consistent with findings in the manufacturing industry (Fotopoulos and Giotopoulos, 2010; Geroski, 1995; Goddard et al., 2002; Sutton, 1997). Smaller firms must grow faster to reach the industry minimum efficient scale (MES), or else face higher costs and higher probability of failure. Therefore, in the sample of surviving firms, smaller firms tend to grow faster (Geroski, 1995; Sutton, 1997).

Likewise, in C2C online markets, consumer trust is essential to e-merchant growth and survival. Seller reputation strongly and positively affects consumer trust and sales performance (Ba and Pavlou, 2002; Bente et al., 2012; Kim et al., 2008; Utz et al., 2012; Zhang, 2006). To survive, sellers must build their reputations quickly. As mentioned, nearly 100% of all Taobao customer feedback was positive. As such, reputation increases almost as rapidly as sales. Increasing sales volume is the best strategy for new stores with smaller scales to overcome trust-building disadvantages.

Higher sales volumes could inspire *herding behavior*—in which individuals follow the crowd even if they know other choices (Banerjee, 1992). When consumers have little knowledge about the sellers, they tend to follow the decisions of others under information asymmetry. Sales volume indicates that many previous customers trusted the company. Therefore, younger and smaller stores must obtain high sales volumes if they are to establish their presence and build their reputations in the C2C online market.

We argue that the C2C market is like manufacturing in having a minimum efficient scale (MES). However, the manufacturing scale focuses on minimizing long run costs, but the C2C scale focuses on minimum reputation levels for establishing, sustaining, or growing consumer trust. Smaller and younger stores have a lower reputation level than the industry MES, and are thus under great pressure to grow faster to reach the industry MES.

The findings imply that smaller stores are more likely to fail, which is consistent with prior findings that survival depends on size for manufacturing firms (Audretsch et al., 2004). In our sample, 6298 of 55,147 Taobao stores exited[2] Taobao from March to September 2012. To examine the relationship between store failure and size, we further analyzed store sales during the last six months prior to store exit. Based on the six-month sales volume, stores were categorized as having sales of zero, 1–10, 11–100, 101–1000, and over 1000 units.

[2] If a store leaves Taobao, their webpage is no longer accessible (Wang et al., 2013). Therefore, we can pinpoint stores that exited by checking their homepages.

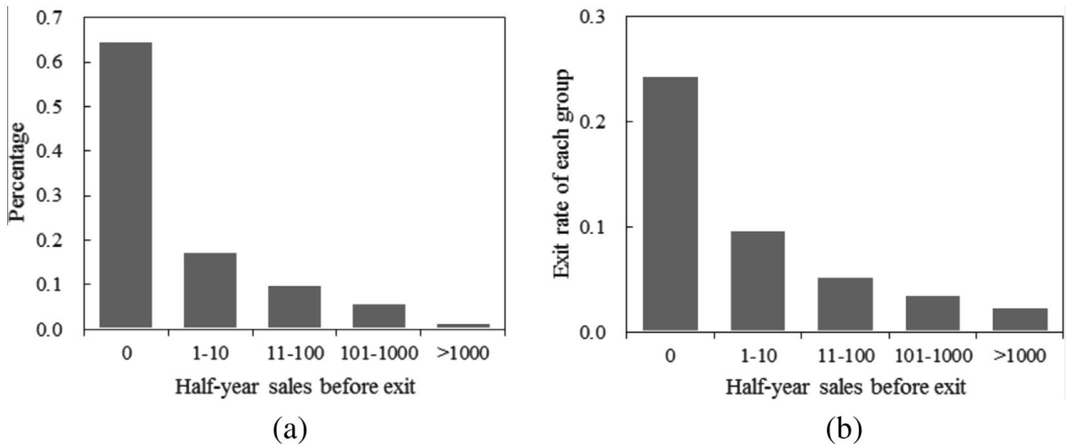

Fig. 1. Analysis of exit rates of Taobao stores.

Fig. 1(a) shows the distribution of exited stores by store size in terms of sales. During the last six months prior to exiting, about 65% had zero sales, followed by 17.5% with 1–10 unit sales; only 5% generated sales more than 100 units. Fig. 1(b) compares the exit rates of Taobao stores by size groups. We measured the store exit rate by the percentage of exited stores within each group. As size increased (i.e., more units of sales), the exit rate decreased. For example, stores with zero sales for six consecutive months experienced a 25% exit rate, compared with the 4% exit rate for stores with sales of more than 100 units.

Thus, smaller stores are more likely to exit the C2C online market. To increase their chances of survival, they should focus on expanding their businesses quickly to reach the industry MES. Once they reach the threshold of the industry MES, they are more likely to obtain customer trust. After that, factors other than size will be more important for growth. After growing beyond a certain survival threshold, growth depends less on size. Therefore, store growth was likely to be independent of size for larger or older stores, validating Gibrat's law.

For young and small Taobao newcomers to survive in a hypercompetitive online market, they must rapidly enhance their reputations, perhaps through product and process strategies for promoting reputation and increasing sales (Porter, 1985). In Taobao's hypercompetitive environment, upfront heavy marketing spending and aggressive price discounting may not provide sustainable advantage, as evidenced in the growth of Internet companies before the dot-com bubble (Eisenmann, 2006). Instead, innovative products and services combined with competitive pricing and enhanced product quality and services (e.g., delivery and post-sale service) could become the key differentiators for building good reputation and obtaining more sustainable competitive advantage. Furthermore, e-merchants must identify unique resources they can access to sustain their seller reputation. For industries with low barriers to entry and imitation, firms can sustain competitive advantage by acquiring less obvious and difficult to imitate resources, such as access to customers (Makadok, 1998).

7. Conclusions, limitations, and future research

This study employs dynamic panel data models to examine the relationship between size and growth of Taobao stores and to test how well Gibrat's law predicts the size–growth relationship in online stores. Our analysis suggests that Gibrat's law is conditioned on the age and size of online stores; the law tends to hold for stores whose size and age exceed certain threshold. For these larger and olders e-merchants, growth tends to be independent of their size, perhaps because older and larger e-merchants have exploited their advantages as first- or early-movers (Lieberman and Montgomery, 1988). However, Gibrat's law does not apply to the younger and smaller stores in our sample: they tend to grow faster. These results suggest that relatively young and small newcomer e-merchants on Taobao may be pressured to build their reputations and grow fast to survive in a hypercompetitive online sales market, while larger and older e-merchants struggle to sustain their early entry competitive advantages.

In this study, we focused on C2C e-commerce on a third-party platform and examined particularly the relationship between online store size and growth. Consequently, the results do not necessarily generalize to other types of e-commerce, such as business-to-business (B2B) or business-to-consumer (B2C). Therefore, we make no claims about the generality of Gibrat's law on e-commerce. Instead, we tested Gibrat's law on the Chinese C2C online market, a dynamic and hypercompetitive e-commerce environment featuring extremely low barriers to entry and a highly fragmented market.

This study has several limitations. First, other proxies might be better than sales volume for measuring store size. If traditional measures of firm size such as assets and revenues had been available, they would have provided a holistic measure for store size. However, sales volume reflects the amount of transactions and can potentially influence store survival, thus making it an appropriate size proxy for our study. Future research may adopt multiple size proxies and test the robustness of empirical results. Second, resource constraints and information restrictions necessitated using monthly Taobao store data, although economics studies more commonly use annual data. Nevertheless, the data provide initial evidence to support our hypotheses. Future studies will generate further insights by employing annual data over a longer period.

Our study offers three potential extensions. First, researchers might apply our methods to examine e-merchants in different cultural contexts such as the United States, Europe, or other Asian countries. Our results provide good insights into the dynamics and structure of e-merchants in China's online market. Yet, China's online market is unique in its cultural context and its predominant individual sellers and microbusinesses. Second, future research may examine the underlying factors that cause Gibrat's law to hold or fail. Results will provide insights into competitive growth strategies for e-merchants and policymakers. Last, industries often evolve from early stage to maturity and shakeout (Klepper, 1996), so it would be interesting to predict whether the online market on a third-party platform such as Taobao is likely to follow the traditional evolutionary path. We hope that our in-depth investigation of the size–growth relationship in the largest Chinese C2C online marketplace will help e-merchants and business policymakers better understand and cope with online market challenges.

References

Almus, M., 2000. Testing "Gibrat's law" for young firms-empirical results for West Germany. Small Bus. Econ. 15 (1), 1–12.
Angelini, P., Generale, A., 2008. On the evolution of firm size distributions. Am. Econ. Rev. 98, 426–438.
Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58 (2), 277–297.
Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. J. Econometrics 68 (1), 29–51.
Audretsch, D.B., 1995. Innovation, growth and survival. In: Int. J. Ind. Organiz. 13 (4), 441–457.
Audretsch, D.B., Klomp, L., Santarelli, E., Thurik, A.R., 2004. Gibrat's law: are the services different? Rev. Ind. Organ. 24 (3), 301–324.
Axtell, R.L., 2001. Zipf distribution of US firm sizes. Science 293, 1818–1820.
Banerjee, A.V., 1992. A simple model of herd behavior. Quart. J. Econ. 107 (3), 797–817.
Ba, S., Pavlou, P.A., 2002. Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior. MIS Quarterly 26 (3), 243–268.
Bente, G., Baptist, O., Leuschner, H., 2012. To buy or not to buy: Influence of seller photos and reputation on buyer trust and purchase behavior. Int. J. Hum. Comput. Stud. 70 (1), 1–13.
Bentzen, J., Madsen, E.S., Smith, V., 2006. The growth opportunities for SMC? J. Bus. Econ. Manage. 7 (3), 139–145.
Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. J. Econometrics 87 (1), 115–143.
Bond, S.R., Hoeffler, A., Temple, J., 2001. GMM Estimation of Empirical Growth Models. Discussion Paper No. 2048, Centre for Economic Policy Research.
Cabral, L.M.B., Mata, J., 2003. On the evolution of the firm size distribution: facts and theory. Am. Econ. Rev. 93, 1075–1090.
Chen, J., Zhang, C., Yuan, Y., Huang, L., 2007. Understanding the emerging C2C electronic market in China: an experience-seeking social marketplace. Electron. Markets 17 (2), 86–100.
Chen, J., Zhang, C., Xu, Y., 2009. The role of mutual trust in building members' loyalty to a C2C platform provider. Int. J. Electron. Commerce 14, 147–171.
Choi, B.P., 2010. The US property and liability insurance industry: firm growth, size, and age. Risk Manage. Insurance Rev. 13 (2), 207–224.
Chesher, A., 1979. Testing the law of proportionate effect. J. Ind. Econ. 27, 403–411.
Chu, H.P., Sher, P.J., Yeh, M.L., 2008. Revisiting Gibrat's law using panel SURADF tests. Appl. Econ. Lett. 15 (2), 137–143.
Cirillo, P., Hüsler, J., 2009. On the upper tail of Italian firms' size distribution. Physica A 388 (8), 1546–1554.
Coad, A., 2009. The Growth of Firms: A Survey of Theories and Empirical Evidence. Edward Elgar Publishing.
Comanor, W.S., Wilson, T.A., 1967. Advertising, market structure and performance. Rev. Econ. Stat. 49 (4), 423–440.
Daunfeldt, S.O., Elert, N., 2013. When is Gibrat's law a law? Small Bus. Econ. 41 (1), 133–147.
De Wit, G., 2005. Firm size distributions: an overview of steady-state distributions resulting from firm dynamics models. Int. J. Ind. Organ. 23 (5), 423–450.
Del Monte, A., Papagni, E., 2003. R&D and the growth of firms: empirical analysis of a panel of Italian firms. Res. Policy 32 (6), 1003–1014.
Eisenmann, T.R., 2006. Internet companies' growth strategies: determinants of investment intensity and long-term performance. Strategic Manage. J. 27 (12), 1183–1204.
Evans, D.S., 1987. The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. J. Ind. Econ. 35 (4), 567–581.
Fan, Y., Ju, J., Xiao, M., 2016. Reputation premium and reputation management: Evidence from the largest e-commerce platform in china. Int. J. Ind. Organ., forthcoming
Fotopoulos, G., Giotopoulos, I., 2010. Gibrat's law and persistence of growth in Greek manufacturing. Small Bus. Econ. 35 (2), 191–202.
Ganugi, P., Grossi, L., Gozzi, G., 2005. Testing Gibrat's law in Italian macro-regions: analysis on a panel of mechanical companies. Stat. Methods Appl. 14 (1), 101–126.
Gao, B., Chan, W.K.V., Li, H., 2015. On the increasing inequality in size distribution of China's listed companies. China Econ. Rev. 36 (12), 25–41.
Geroski, P.A., 1995. What do we know about entry? Int. J. Ind. Organ. 13 (4), 421–440.
Gibrat, R., 1931. Les Inegalites Economiques. Librairie du Recueil Sirey, Paris.
Giotopoulos, I., Fotopoulos, G., 2010. Intra-industry growth dynamics in the Greek services sector: firm-level estimates for ICT-producing, ICT-using, and non-ICT industries. Rev. Ind. Organ. 36 (1), 59–74.
Goddard, J., Wilson, J., Blandon, P., 2002. Panel tests of Gibrat's law for Japanese manufacturing. Int. J. Ind. Organ. 20 (3), 415–433.
Goddard, J., McMillan, D., Wilson, J.O., 2006. Do firm sizes and profit rates converge? Evidence on Gibrat's Law and the persistence of profits in the long run. Appl. Econ. 38 (3), 267–278.
Hall, B.H., 1987. The relationship between firm size and firm growth in the U.S. manufacturing sector. J. Ind. Econ. 35 (4), 583–606.
Harris, R., Trainor, M., 2005. Plant-level analysis using the ARD: another look at Gibrat's law. Scottish J. Political Econ. 52 (3), 492–518.
Hart, P.E., Oulton, N., 1996. The growth and size of firms. Econ. J. 106 (3), 1242–1252.
Hart, P.E., Oulton, N., 1997. Zipf and the size distribution of firms. Appl. Econ. Lett. 4 (4), 205–206.
Holsapple, C., Lee-Post, A., Pakath, R., 2014. A unified foundation for business analytics. Decis. Support Syst. 64, 130–141.
Kim, D.J., Ferrin, D.L., Rao, H.R., 2008. A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. Decis. Support Syst. 44 (2), 544–564.
Klepper, S., 1996. Entry, exit, growth, and innovation over the product life cycle. Am. Econ. Rev. 86 (3), 562–583.
Li, D., Li, J., Lin, Z., 2008. Online consumer-to-consumer market in China – a comparative study of Taobao and eBay. Electron. Commerce Res. Appl. 7 (1), 55–67.
Lieberman, M.B., Montgomery, D.B., 1988. First-mover advantages. Strategic Manage. J. 9 (1), 41–58.
Lin, Z., Janamanchi, B., Li, D., Huang, W., 2006. Reputation distribution and consumer-to-consumer online auction market structure: an exploratory study. Decis. Support Syst. 41 (2), 435–448.
Lin, Z., Li, D., Huang, W., 2007. Traders beware an examination of the distribution of eBay sellers' online reputation. Int. J. Electron. Bus. 5 (5), 499–517.
Liu, J.T., Tsou, M.W., Hammitt, J.K., 1999. Do small plants grow faster? Evidence from the Taiwan electronics industry. Econ. Lett. 65 (1), 121–129.
Liu, S., 2016. How the user liaison's understanding of development processes moderates the effects of user-related and project management risks on IT project performance. Inf. Manage. 53 (1), 122–134.
Lotti, F., Santarelli, E., Vivarelli, M., 2003. Does Gibrat's Law hold among young, small firms? J. Evol. Econ. 13 (3), 213–235.
Lotti, F., Santarelli, E., 2004. Industry dynamics and the distribution of firm sizes: a nonparametric approach. South Econ. J. 70 (3), 443–466.
Lu, Y., Zhao, L., Wang, B., 2010. From virtual community members to C2C e-commerce buyers: trust in virtual communities and its effect on consumers' purchase intention. Electron. Commerce Res. Appl. 9 (4), 346–360.
Maçãs Nunes, P.J., Serrasqueiro, Z.M., 2009. Gibrat's law: empirical test of Portuguese service industries using dynamic estimators. Serv. Ind. J. 29 (2), 219–233.
Makadok, R., 1998. Can first-mover and early-mover advantages be sustained in an industry with low barriers to entry/imitation. Strategic Manage. J. 19 (7), 683–696.
McKinsey Global Institute, 2013. China's E-Tail Revolution: Online Shopping as a Catalyst for Growth. Available at <http://www.mckinsey.com/global-themes/asia-pacific/china-e-tailing>. Last accessed March 14, 2016.
Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49 (6), 1417–1426.
Oliveira, B., Fortunato, A., 2006a. Testing Gibrat's Law: empirical evidence from a panel of Portuguese manufacturing firms. Int. J. Econ. Bus. 13 (1), 65–81.
Oliveira, B., Fortunato, A., 2006b. Firm growth and liquidity constraints: a dynamic analysis. Small Bus. Econ. 27 (2–3), 139–156.
Oliveira, B., Fortunato, A., 2008. The dynamics of the growth of firms: evidence from the services sector. Empirica 35 (3), 293–312.
Ou, C.X., Davison, R.M., 2009. Why eBay lost to Taobao in China: the glocal advantage. Commun. ACM 52 (1), 145–148.
Ou, C.X., Pavlou, P.A., Davison, R., 2014. Swift Guanxi in online marketplaces: the role of computer-mediated communication technologies. MIS Quarterly 38 (1), 209–230.
Petrunia, R., 2008. Does Gibrat's law hold? evidence from Canadian retail and manufacturing firms. Small Bus. Econ. 30 (2), 201–214.
Porter, M., 1985. Competitive Advantage: Creating and Sustaining Superior Performance. Free Press, New York.
Reichstein, T., Jensen, M.B., 2005. Firm size and firm growth rate distributions—the case of Denmark. Ind. Corporate Change 14 (6), 1145–1166.
Resende, M., 2004. Gibrat's law and the growth of cities in Brazil: a panel data investigation. Urban Stud. 41 (8), 1537–1549.
Riccaboni, M., Pammolli, F., Buldyrev, S.V., Ponta, L., Stanley, H.E., 2008. The size variance relationship of business firm growth rates. In: Proceedings of the National Academy of Sciences, vol. 105. No. 50, pp. 19595–19600.
Santarelli, E., Klomp, L., Thurik, A.R., 2006. Gibrat's Law: an overview of the empirical literature. In: Santarelli, E. (Ed.), Entrepreneurship, Growth, and Innovation. Springer, US, pp. 41–73.
Segarra, A., Teruel, M., 2012. An appraisal of firm size distribution: does sample size matter? J. Econ. Behav. Organ. 82 (1), 314–328.
Serrasqueiro, Z., Nunes, P.M., Leitão, J., Armada, M., 2010. Are there non-linearities between SME growth and its determinants? A quantile approach. Ind. Corporate Change 19 (4), 1071–1108.
Stanley, M.H., Amaral, L.A., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Stanley, H.E., 1996. Scaling behavior in the growth of companies. Nature 379 (6568), 804–806.
Simon, H.A., Bonini, C.P., 1958. The size distribution of business firms. Am. Econ. Rev. 48 (4), 607–617.

Strotman, H., 2007. Entrepreneurial survival. Small Bus. Econ. 28 (1), 87–104.
Sutton, J., 1991. Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration. MIT Press, Cambridge, MA.
Sutton, J., 1997. Gibrat's Legacy. J. Econ. Lit. 35 (1), 40–59.
Teruel-Carrizosa, M., 2010. Gibrat's law and the learning process. Small Bus. Econ. 34 (4), 355–373.
Tschoegl, A.E., 1983. Size, growth and transnationality among the world's largest banks. J. Bus. 56, 187–201.
Utz, S., Kerkhof, P., Bos, J., 2012. Consumers rule: how consumer reviews influence perceived trustworthiness of online stores. Electron. Commerce Res. Appl. 11 (1), 49–58.
Wang, Y., Wang, S., Fang, Y., Chau, P.Y., 2013. Store survival in online marketplace: an empirical investigation. Decis. Support Syst. 56, 482–493.
You, W., Liu, L., Xia, M., Lv, C., 2011. Reputation inflation detection in a Chinese C2C market. Electron. Commerce Res. Appl. 10 (5), 510–519.
Zhang, J., 2006. The roles of players and reputation: evidence from eBay online auctions. Decis. Support Syst. 42 (3), 1800–1818.
Zhang, L., Liu, H., Hu, E., Yue, M., 2009. The relationship between size and growth: the case of Chinese listed companies. Appl. Econ. Lett. 16 (18), 1849–1853.