CREDIT SUISSE | FIRST BOSTON

CREDIT SUISSE FIRST BOSTON CORPORATION

Equity Research

Americas | U.S./Investment Strategy | December 21, 1999

Absolute Power

The Internet's Hidden Order

Michael Mauboussin
1 212 325 3108
michael.mauboussin@csfb.com

Alexander Schay
1 212 325 4466
alexander.schay@csfb.com

Stephen Kawaja
1 212 325 3922
stephen.kawaja@csfb.com

- On-line firms are subject to an even more ruthless statistical distribution than the bell curve—the power law.
- Popular wisdom suggests that every firm with dot.com in its name is commanding a huge valuation. The data do not support this view.
- Power laws seen in the volume of traffic, as well as the size of Web sites, both strongly support the view that on-line markets are winner-take-all.
- The stock market is not dumb. Far from encountering systematic overvaluation, the market capitalization of Internet firms in our index followed a power law distribution.

Executive Summary

The dot.com phenomenon has been widely derided as a "bubble." Some savvy market watchers have noted that many industries that have had a profound impact on the economy initially spawned a multitude of competitors. But these industries, which include airlines and automobiles, ultimately witnessed little value creation. Unlike the automobile and airline industries of yesteryear—in which firms competed for product market share—most on-line firms are competing for neural mind share. *However, their plight will be the same. Just like many of their "real world" brethren, a significant number of on-line firms will either fail or be absorbed by other competitors in the coming years.*

In Lake Wobegon, Minnesota, all the children are said to be above average.[1] (See notes at end of text.) While refraining from comment about the strength or beauty of the parents, our data show that the on-line offspring of venture capital firms are now eagerly snapped up by the public markets and are subject to an even more ruthless statistical distribution than the bell curve—the power law.

Power laws are distribution functions for measurable quantities, like how many earthquakes there are of a particular size or the number of cities within a given range of populations. Nascent scientific study of the Internet has revealed that it is subject to the same statistical regularities that govern these other "natural" phenomena, and as we will see, these regularities have important implications for the future of dot-com hopefuls.

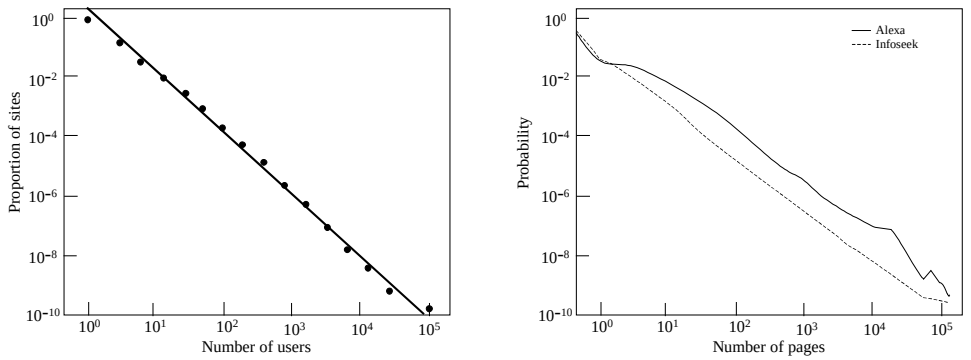
Powerful Web

Despite being convinced of the disruptive nature of the on-line medium[2], we believe that value creation for most Internet companies will be restricted owing to the difficulty of achieving sufficient scale. This begs some important questions. Who has the advantage in achieving this scale? Does the Internet, as a self-organizing system, have any interesting or unexpected characteristics?

Underneath the Internet's apparent complexity and confusion, there exists hidden order. Web pages are the building blocks of the World Wide Web, and every page is identified by a unique URL (uniform resource locator). Studies of the new medium have demonstrated that Web pages are distributed among sites according to a universal power law: many sites have very few pages, and very few sites have many pages.[3]

The right side of Figure 1 shows the distribution of site sizes for two "crawls," one conducted by Alexa and the other by Infoseek, covering a large number of sites. Both data sets illustrate a power law. That is, on a log-log scale the distribution of the number of pages per site appears as a straight line.[4] The existence of a power law in the growth of the Web carries some profound implications. One important suggestion is that the expected number of sites of a particular size can be estimated by extrapolating the power law to any large value of "n." For example, if data were collected from 250,000 sites, the probability of finding a site with a million pages would be 10^{-4}, or one in ten thousand.[5] It seems that size is hard to come by.

Figure 1
Power Laws



Source: Bernardo Huberman, Lada Adamic; Xerox Palo Alto Research Center.

The largest sites also get the most traffic. Researchers at Xerox PARC found that for all the sites in their sample, as well as sites in specific categories of the sample, the distribution of *visitors* per site followed a universal power law.[6] As was the case with the number of pages per site, the results for the number of visitors per site indicate that many sites have very few users, and a few sites have very many users. This deceptively simple assertion belies a powerful statistical truth. Very few companies have met with success on-line, and we believe fewer still will succeed on-line in the future. Such a disproportionate distribution of user volume among sites is characteristic of winner-take-all markets. A select few reap all the rewards, and latecomers confront serious difficulties cracking the "lineup." The percentage of total volume accounted for by the top sites (measured by unique users) is displayed below.

Table 1
Percentage Volume by User

% Sites	All sites	Adult Sites	Educational Sites
0.1%	32.36%	1.4%	2.81%
1%	55.63%	15.83%	23.76%
5%	74.81%	41.75%	59.50%
10%	82.26%	59.29%	74.48%
50%	94.92%	90.76%	96.88%

Source: Xerox PARC, data from December 1, 1997.

.

The top 119 sites, representing one-tenth of one percent of the total, captured an enormous 32.36% of user volume, while the top 1% of sites garnered more than half of the total user volume. Although the ability to attract users does not ipso facto lead to monetization of traffic, much less profitability, it is a necessary precondition. With Internet market share beholden to these strong regularities, the implications for competition—and investors sizing up the competition—are enormous.

Models that account for the observed power law behavior indicate a low probability that newly established sites will attract a significant number of users.[7]

Taking Web sites and ranking them on the horizontal axis according to popularity—as determined by the number of unique users—and then placing pageviews on the vertical axis, results in a straight line when using a double logarithmic scale.[8]

This so-called Zipf distribution[9] is a power law and has been observed in the distribution of words used in a natural language. Zipf-like distributions have also been seen in many other realms, including earthquakes, city sizes, and the variation of cotton prices. Extremities become apparent when Zipf distributions are plotted on a linear scale. (See the bottom of Figure 3.) A few elements score very high, a medium number of elements get middle-of-the-road scores, and an enormous number of elements score very low (the long tail of the distribution which hugs the x-axis).

Knowledge Is Power

We have observed that success begets success ever since the dawn of recorded history.[10] Thrust into the modern context of massively interconnected networks, this hoary truism takes on a powerful new form. We now outline some of the reasons power laws are persistent on the Internet. On-line, popularity is determined by a collective process that has been called "social search."[11] That is, referrals from friends, relatives, and acquaintances drive growth and traffic.

Search Costs

With only a Web address to differentiate itself from the crowd, an on-line firm naturally flirts with obscurity. The importance of awareness can be traced back to the high search costs associated with locating companies in Internet markets. These costs stem from the sheer volume of competing information available. A virtual avalanche of marketing studies have supported the notion that consumers reach a point of diminishing return when confronted with ever-increasing choice.[12]

For example, in the book business over the last 15 years, consumers have been confronted with a greater selection than in any other period in history. Ironically, despite the abundance of offerings, the share of book sales represented by the 30 top-selling hard cover books in America nearly doubled from 1986 to 1996.[13] The brand name author as "safe haven" is mirrored in the on-line distribution data, as Internet users flock to the same sites in order to reduce search costs.

Differentiation on-line is becoming more and more challenging as the number of entrants to the market increases daily. In an environment in which record-breaking sums are being allocated for traditional advertising, one recent article (*Wall Street Journal*) quipped, "Americans are being dot-commed to death." The fallout from this overwhelming assault is that many consumers are just not paying attention. After all, it's hard to remember whether that ad was for Pets.com, Petstore.com, PETsMART.com, or Petopia.com.[14]

Links with the Past

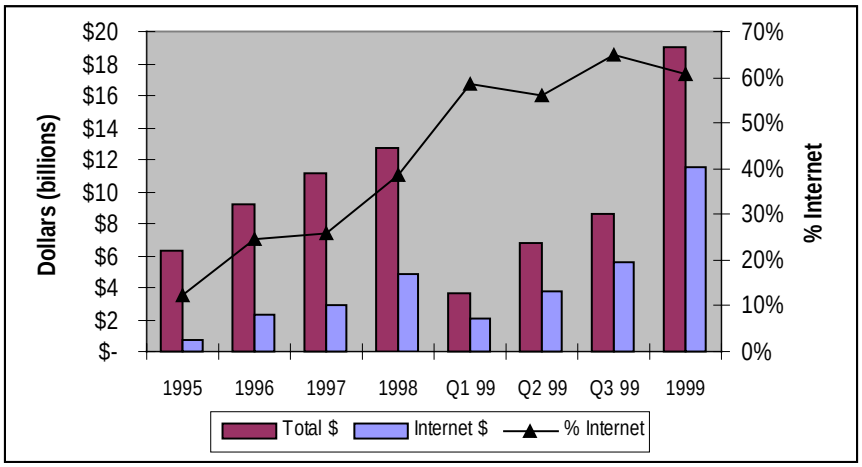
New sites tend to link to already popular sites in a bid for increased traffic. This results in a rich-get-richer phenomenon, whereby already popular sites get even more popular. Studies conducted at the University of Notre Dame have found that the proportion of Web pages with many outside links—directing traffic to that

particular page—is much higher than would be expected in a random network.[15] If people had linked their pages with others at random, pages with at least 1,000 links would be virtually nonexistent. However, this is not the case. A recent survey conducted by Forrester Research showed that retailers with on-line distribution deals generated 36% of their site traffic and 33% of their sales through links with firms that rank among the top-50 in volume.[16]

Studying the links into and out from the 325,729 pages on Notre Dame's domain "nd.edu," physicists discovered that the pages follow a power law distribution. (Other domains saw similar results.) The number of pages with n links was proportional to $1/n^2$. For example, pages with 20 links were one quarter as frequent as pages with 10 links. This is strong evidence in support of the Web as a self-organizing system.

Balance of Power

Figure 2
Trailing Venture Capital Money Flows

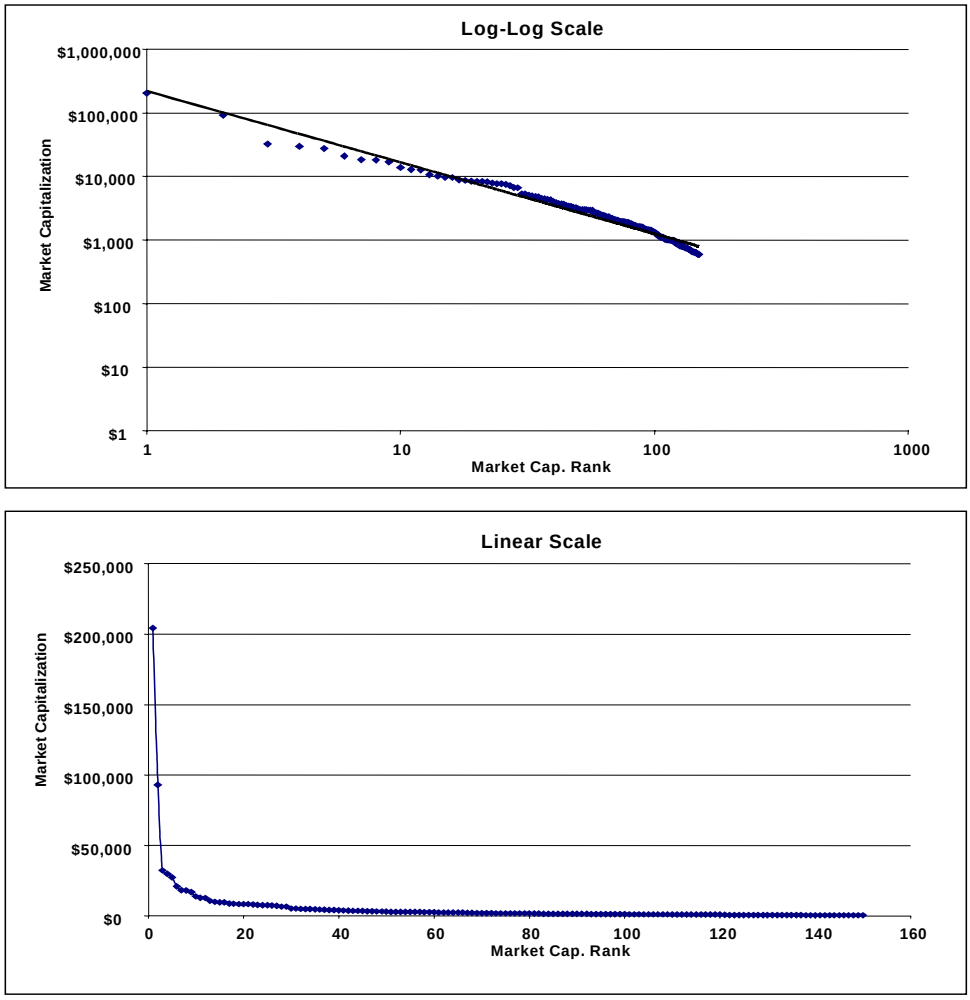


Source: VentureOne Corporation and CSFB analysis.

Over the last three years venture capital firms have put a staggering amount of money to work. In the third quarter of 1999 alone, a record shattering $8.6 billion flowed into the coffers of over 750 venture-backed firms. Almost two-thirds of these investment dollars were earmarked for Internet companies, reflecting an upward spiral in the overall percentage of venture capital dollars allocated to on-line firms. In 1995 roughly 12% of total venture capital outlays went to on-line companies. For the first three quarters of 1999, that percentage skyrocketed to 60%, nearly tripling the total investment dollars distributed in all of 1995. (See Figure 2.)

With evidence mounting to support the idea that the on-line medium is a winner-take-all market, it would seem that many venture capital firms have placed substantial amounts of investment dollars in harms way. Pundits would observe that the "bubble-like" valuations awarded in the public markets could continue to bail out the venture capital firms. Dubious of market inefficiency on such a grand scale, we plotted the market capitalization of an index of pure play dot.com companies. The results are stunning. Far from encountering systematic overvaluation, we observed a power law distribution in the market capitalization of the largest 150 Internet firms in our index.[17]

Figure 3
Dot.Com Market Capitalization Power Law
market capitalization in millions



Source: CSFB analysis.

Conclusion

The market is not dumb. The percentage of total market capitalization contributed by the top on-line firms is remarkably similar to the percentage of total volume accounted for by the top sites in the individual segments in Table 1. The top 5% of on-line firms in terms of market capitalization accounted for half of the total market capitalization of the group, while the top 10% accounted for 61% of the total, and the top 50% accounted for 90%.

Popular wisdom suggests that every firm with dot.com in its name is commanding a huge valuation. The data do not support this view. Our findings do show that the market capitalization of the Internet is dominated by a handful of leaders. This is consistent with the winner-take-all outcomes implied by the power law distributions.

It is important to underscore that this analysis provides compelling evidence that the market is valuing Internet companies relatively correctly. These data provide no evidence, however, that the market is valuing the sector correctly on an absolute basis (either overvaluation or undervaluation).

In sum, very few Web sites get large, and very few Web sites get a lot of user traffic. In addition, very few Internet firms account for a disproportionate share of the total market value of the group. We think this will continue to hold true for the future of the medium.

[1]During the introduction to every broadcast of "A Prairie Home Companion," Garrison Keillor observes that in the mythical town of Lake Wobegon, Minnesota, "all the women are strong, the men are good looking, and the children above average."

[2]See *The Retailer's Dilemma,* Credit Suisse First Boston, October 25, 1999.

[3]"*Growth Dynamics of the World-Wide Web.*" Huberman, B.A., Adamic, L. *Nature* Vol. 401, September 9, 1999.

[4]What does this mean? Mathematically, these straight lines are called power laws since they show that some quantity *N* can be expressed as some power of another quantity *s*:

$N(s) = s^{-\tau}$ *How Nature Works* Bak, Per (Copernicus 1996) p. 27.

[5]Using: $P(n_2) = P(n_1) \times (n_2 / n_1)^{-\beta}$. The expected number of sites of size n_2 in a crawl of N sites would be $NP(n_2)$ "*Growth Dynamics of the World-Wide Web.*"

[6]"The Nature of Markets in the World Wide Web." Huberman, B.A., Adamic, L. from the proceedings of *Computing In Economics and Finance*, May 6, 1999.

[7]ibid. p.3. Note that the observed power law behavior was derived from usage logs for all Web sites, not usage logs for commercial Web sites alone.

[8]S. Glassman, *Computer Networks ISDN Systems* No. 27, 165 (1994).

[9]Zipf, G.K. *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge, Massachusetts, 1949).

[10]Matthew 25:29, "For everyone who has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him."

[11]Bernardo Huberman quoted in The *New York Times* article *"Not a Great Equalizer After All?"* by John Markoff.

[12]Cristol, S. M., Sealey, P. "*Replacement Marketing: Antidote for Consumer Stress.*" available at http://haas.berkely.edu/~wba268/replacement.htm.

[13]Malcolm Gladwell, *"The Science of the Sleeper" The New Yorker*, Annals of Marketing, October 4, 1999.

[14]The *Wall Street Journal*, *"A New Spin for Web Firms: Not.Com"* Suein L. Hwang. On-line video tape and DVD purveyor BigStar Entertainment Inc. has commandeered a fleet of trucks to help it with its on-line effort but not to haul any of its goods. Rather, BigStar has emblazoned its name across other firms trucks in order to create a sense of "real-world permanence." November 11, 1999.

[15]Reka, A.,Hawoong, J.,Barabasi, A. "Diameter of the World-Wide Web." Nature, Vol. 401, September 1999.

[16]Including affiliate deals and pacts with vertical sites, these distribution agreements accounted for 52% of traffic, and 50% of sales. "*The Parting of the Portal Seas*" Forrester Research, Charlene Li, December 1999.

[17]We combined the Bloomberg U.S. Internet Index (196 members) with our own database of pure-play "dot.coms," resulting in a 400 company index. The lowest ranked 162 firms by market capitalization (a $100 million market cap cutoff) accounted for one half of one percent of the total market capitalization. The top 1% of companies in the index accounted for 40% of the total market capitalization, while the top 5% and 10% accounted for 63% and 77%, respectively. Finally, the top 50% of firms in the index accounted for 99% of the total. The median market capitalization of the index was $277 million.

CREDIT SUISSE | FIRST BOSTON

Americas

Eleven Madison Avenue
New York, NY 10010, U.S.A.
1 212 325 2000

Atlanta	1 404 656 9500	Mexico City	1 525 202 6000
Baltimore	1 410 223 3000	Pasadena	1 626 395 5100
Boston	1 617 556 5500	Philadelphia	1 215 851 1000
Buenos Aires	1 541 394 3100	San Francisco	1 415 836 7600
Chicago	1 312 750 3000	São Paulo	55 11 822 4862
Houston	1 713 220 6700	Toronto	1 416 352 4500

Europe

One Cabot Square
London E14 4QJ, England
44 171 888 8888

Amsterdam	31 20 575 4890	Paris	33 1 40 76 8888
Budapest	36 1 202 2188	Prague	42 2 248 10937
Frankfurt	49 69 75380	Vienna	43 1 512 3023
Geneva	41 22 707 0130	Warsaw	48 22 695 0050
Madrid	34 91 532 0303	Zug	41 41 726 1020
Milan	39 02 7702 1	Zurich	41 1 335 7711
Moscow	7 501 967 8200		

Asia Pacific

Three Exchange Square
8 Connaught Place
Central, Hong Kong
852 2101 6000

Auckland	64 9 302 5500	Shanghai	86 21 6219 0808
Beijing	86 10 6410 6611	Singapore	65 538 6322
Melbourne	61 3 9 280 1666	Sydney	61 2 9394 4400
Mumbai	91 22 284 6888	Taipei	8862 2718 5919
Osaka	81 6 243 0789	Tokyo	81 3 5404 9000
Seoul	82 2 3707 3700	Wellington	64 4 474 4400