Equity Research

Still Powerful

The Internet's Hidden Order

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

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

- **Despite a correction in the Internet sector over the past six months, the relationship between rank and market value of online firms continues to exhibit power law behavior.**

- **This winner-take-all, market-capitalization distribution supports the notion that competition on the Web has some unique characteristics. For example, large firms can grow as quickly as small firms and the positive reinforcing nature of the medium provides leaders with a source of competitive advantage.**

- **We would continue to focus on the companies with customers, capital, and viable long-term business models.**

Executive Summary

Despite the Internet's apparent complexity, hidden order exists. Studies demonstrate that Web pages are distributed among sites according to a universal power law. Power laws are distribution functions for measurable quantities—such as the number of earthquakes that occur at a particular intensity, or the number of cities that exist within a given range of population. The number of Web pages adheres to a remarkably stable regularity. Roughly described, power laws hold that many sites have few pages and that few sites have many pages.[1]

The distribution of Web pages is not the only online phenomenon that adheres to a power law. The distribution of visitors per site follows a power law as well: many sites have very few users, and a few sites have very many users. Although intuitive, the precise nature of these two regularities carries profound implications. Online firms that depend on the ability to attract and monetize users face significant challenges. The observed power laws suggest a low probability that newly established sites will attract a significant number of users.[2] In fact, from January 1999 to April 2000, 20 of the top 25 companies in the MediaMetrix "Top Properties" remained in the top 25.[3]

We first explored power laws and the Internet in late 1999.[4] Despite a sharp correction in Internet stock prices over the last six months, the relationship between rank and the market value of online firms is still best expressed with a power law. This winner-take-all distribution continues to support the notion that competition on the Web has unique characteristics. Not only can large firms grow as quickly as small firms, but the positive-reinforcing nature of the medium also translates into a source of sustainable competitive advantage. This report explores some of the factors that may contribute to the observed distribution.

Powerful Web

Let there be no mistake—activity on the Web is highly concentrated. To show the breakdown of Web site popularity, we ranked the number of unique site users on the horizontal axis and plotted page views on the vertical axis, using MediaMetrix data for the top 500 domains in April 2000. The result is a straight line when using a double logarithmic scale. The data show that the top 5%, or 25 sites, account for 31% of the total volume of traffic.[5] This is consistent with broader studies of the Web showing that the top sites account for a very large percentage of total traffic volume. That is, one-tenth of one percent of the total, or the top 119 sites, account for 32% of the total volume of site traffic.[6] Very few Web sites get large, and very few Web sites get a lot of user traffic.

Figure 1
Power Laws

Top 500, April 2000

Hits (000) vs. Rank
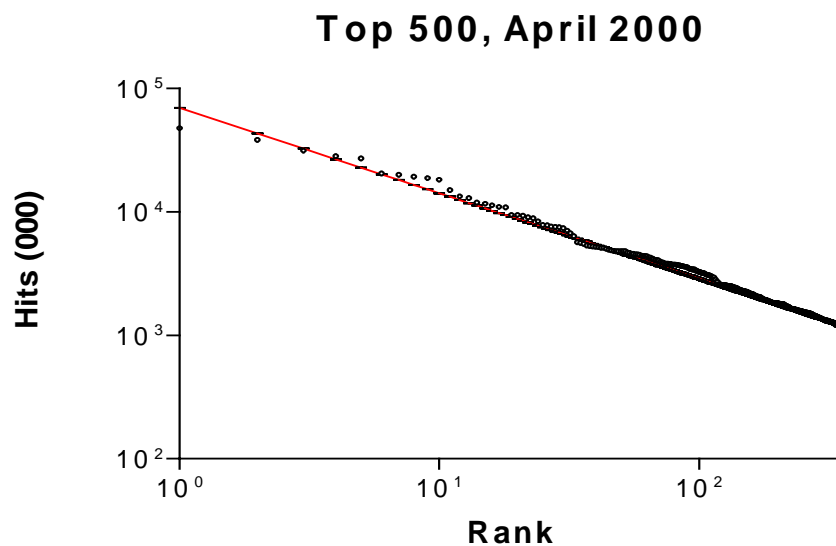
Source: MediaMetrix and CSFB analysis.

A logarithmic plot of rank versus frequency that is a straight line with slope near unity (as seen in Figure 1) illustrates Zipf's law. This regularity has been observed in diverse realms, including the magnitude and frequency of earthquakes, rank and city sizes, and the distribution of words in a language. Extremities are apparent when Zipf distributions are plotted on a linear scale. A few elements score very high, a medium number of elements score in the middle of the road, and an enormous number of elements score very low. (The long tail of the distribution hugs the x-axis.)

Power laws are unusual—they represent special situations. Gaussian, or normal, distributions are much more common in nature than power law distributions. For example, the velocity of gas molecules in a room, the length of new-born babies, and the distribution of market capitalizations for the broader stock market all follow Gaussian distributions. One common trait of Gaussian systems is independence. To use gas molecules as an illustration, the velocity of any given molecule will not explicitly depend on the velocity of the other molecules. Even when the molecules collide, there is only a small degree of interaction.

In contrast, power laws occur when there is a high degree of interaction. The canonical example is a forest fire. The probability that a given tree in the forest

will burn is highly dependent on whether other trees in the forest are burning and the location of the tree relative to its neighbors. The underlying "network" position of the trees determines whether or not the blaze dies out, is confined to a small area, or becomes a conflagration. This complexity gives rise to power law distributions in the sizes of forest fires. The concentration of hits on a Web site also translates into a concentration of market value, so the most frequently used sites generally have the best opportunity to create viable business models. It's interesting that the valuation of Internet companies still conforms to a power law, despite a pronounced downturn in the segment over the last six months. This can be seen in Figure 2.

Figure 2
Internet Market Capitalization Power Law
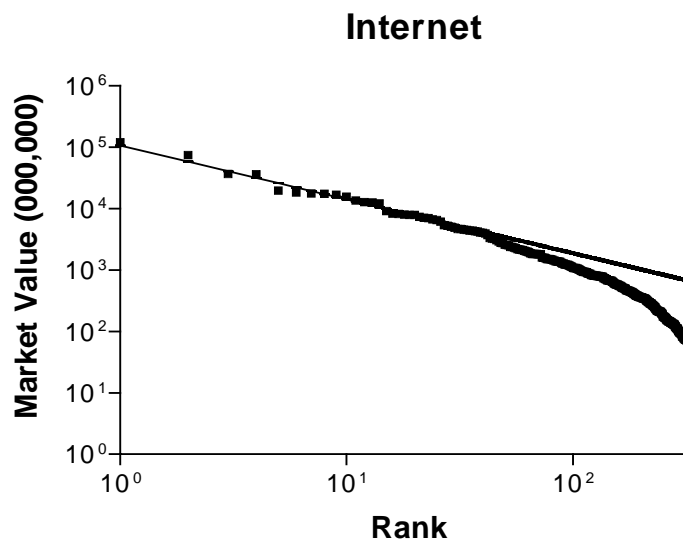
Source: CSFB analysis.

This analysis provides compelling evidence that the market continues to value Internet companies correctly on a relative basis. It is important to note that these data remain mute on the issue of absolute valuations. Contrary to popular perception, there is some discernible order to Internet valuations. So the question becomes, Is this order unique to Internet stocks? Do other industries exhibit the same characteristics, where a disproportionate share of the total market capitalization resides in a handful of firms? We analyzed 46 sectors from 18 different industry classifications in an attempt to answer this question. This broad survey revealed that 9 sectors exhibited strong power law characteristics. They are displayed in the Appendix.

The Strong Get Stronger

While no pure Internet brand is reinforced by the benefit of a physical presence, many particular characteristics of the online medium allow those companies that are already ahead to stay ahead. One example noted in "Absolute Power" is that new sites tend to link to already-popular sites in a bid for increased traffic. Such a reinforcing link structure is reconfirmed in a number of subsequent studies of the Web.

The most comprehensive study of the Web to date reveals some interesting insights into its macroscopic structure. Researchers at IBM's Almaden Research Center in San Jose, California, examined over 200 million pages and 1.5 billion links (roughly one-fifth of the estimated current Web). They found that the Web's connectivity resembles a bow tie. On the left side of the tie is what they dub the "newbies." This group of sites consists of 43 million pages (or 21% of the total) that are relatively new. These sites link to the center, or what is called the "Strongly Connected Component" (SCC). However, these established players in the center—with over 56 million pages—do not link back to the newbies. Sites in the SCC consist of portals and many of the top corporate sites in the MediaMetrix data. On the right side of the bow are the "introverts." This group includes a lot of lesser-known e-commerce sites. Although the SCC is linked to these introverts, the introverts do not provide links back to the strongly connected center.[7]

Most search engines rely on so-called "crawlers." Crawlers work by indexing a Web page, jumping to other pages linked to it, then indexing those, and so on. The Almaden study contradicts earlier suggestions that any two Web pages are connected by a relatively small number of hyperlinks. In order to give wide breadth of coverage in the future, search engines will have to crawl from a greater diversity of starting points than they do today.[8] Until this occurs, they will continue to channel traffic to the most popular and entrenched Web sites—without any reciprocity for the newbies. With only a fraction of the growing Web appearing in most search queries, traffic patterns to the existing strongly connected sites are positively reinforced.

Advertising Spiral

So how does this translate into economic value? One example of the value concentration is advertising. Advertisers aren't taking any chances with the "new" medium: about 95% of Web advertising is spent on 1% of all ad-supported Web sites. And while the *absolute* number of Web sites receiving online ad revenues continues to grow, the *percentage* share of ad dollars being spent at the top 50 sites continues to rise.

Table 1
Publishers' Share of Total Web Ad Revenues

	1998	1999	2000
Top 10 Sites	72%	74%	76%
Top 25 Sites	85%	87%	89%
Top 50 Sites	92%	95%	95%

Source: eMarketer estimates.

Although this profile may change as advertisers shift their dollars away from portals and toward more highly targeted vertical Web sites, advertisers continue to gravitate toward high-traffic sites.

Web Competition

Models that try to explain the characteristics of electronic markets—where a few sites thrive while most are driven to the verge of extinction—conclude that sudden transitions occur. As the competition between Web sites heats up, a profound shift can take place from an environment where many sites simultaneously thrive to a winner-take-most market where a handful of sites attract most users. This is attributable to a nonlinear interaction among sites, which effectively reduces the growth rate of some, owing to extreme competitive pressure from others.[9]

With firms spending enormous sums to sustain growth, it is easy to see that it is possible to reach a critical spending point where funding dries up (in tandem with the business fundamentals). This phenomenon can occur even with only two competing sites, and is comparable to the "principle of mutual exclusion" in ecology. Two predators of the same prey cannot coexist in equilibrium when competitive predation is very strong.[10]

Once the fixed development cost of operating a Web site is spent, it is relatively inexpensive to scale capacity to meet increased demand. So as Web sites proliferate, focus on aggregate demand becomes the order of the day. Unfortunately, many Web company "predators" are competing for the attention and resources of the same consumer "prey."

Conclusion

Figure 3
Trailing Venture Capital Money Flows
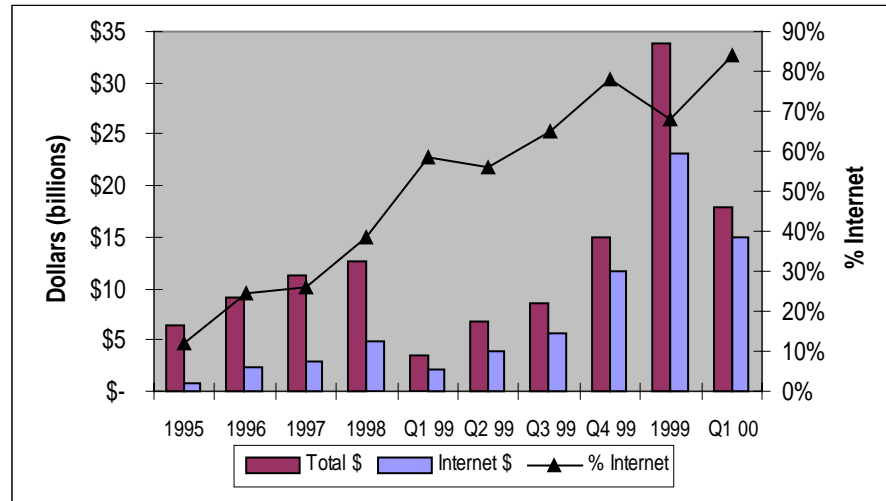
Source: VentureOne Corp. and CSFB analysis.

Investment by venture capital firms has been unprecedented in the last two years. In the fourth quarter of 1999 alone, $14.9 billion was spent on new business opportunities. This *quarterly* total was higher than any *annual* amount in the history of U.S. venture capital. Picking up where 1999 left off, venture-backed firms raised $17.8 billion in the first quarter of the new millennium. About 84% went to Internet-related ventures. We register our skepticism today, as we did six months ago, about the ability of most online firms to survive the unforgiving rule of the power law.

Since the beginning of the second quarter there has been a marked change in perception of Internet businesses. Investors are now focusing on a new acronym: P2P—the path to profitability. With some high-profile firms at 52-week lows and others closing their doors altogether, investors must remember the inherent advantage of today's market leaders. They have the customers. They have the capital. And they benefit directly from the self-reinforcing nature of the Web.

[1] This was initially reported in *"Growth Dynamics of the World-Wide Web."* Huberman, B.A., Adamic, L. *Nature* Vol. 401, September 9, 1999. Since then the results have been reconfirmed in *"Graph Structure in the Web."* Broder, Andrei, et. al. IBM Almaden Research Center, San Jose, CA., June 2000.

[2] ibid. Huberman. It's important to note that the observed power law behavior was derived from usage logs for all web sites, not usage logs for commercial web sites alone. Corporate strategy can influence the ability to crack the line-up.

[3] Six properties were absorbed by competitors that remained in the top 25 in April. The balance of the firms remained in the top 100.

[4] Michael Mauboussin, Alexander Schay, Stephen Kawaja, *"Absolute Power,"* Credit Suisse First Boston Equity Research, December 21, 1999.

[5] Although we could find no direct statistical correlation between site traffic and market capitalization, 14 of the top 25 companies in site volume were also in the top 25 in terms of market value.

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

[7] *"Graph Structure in the Web."* Broder, Andrei, et. al. IBM Almaden Research Center, San Jose, CA., June 2000.

[8] *"Souped Up Search Engines." Nature* Vol. 405, May 11, 2000 *www.nature.com*

[9] "Competitive Dynamics of Web Sites." Maurer, Sebastian M., Huberman, Bernardo A. Xerox Palo Alto Research, March 17, 2000.

[10] Ibid.

Appendix

Complements

We analyzed 46 sectors from 18 different industries for possible power law trends. Data are from FactSet industry classifications. This broad survey shows that 9 of these industries show strong power law characteristics. (See Table 2.) In every industry studied we found that the largest companies best follow a power law. We show the relative size of this subset of companies in the first column. The second column gives a measure of the extent to which each industry is dominated by its largest companies, and the third column quantifies how well the companies fit a power law. (Smaller numbers closer to zero correspond to better fits.) Here we employ the standard statistical practice of taking the square root of the mean squared residual as a goodness of fit test.

Table 2
Power Law Sectors

Industry	Portion Following Power Law	Market Share Held By Companies in Top 5%	RMS of Residuals in Power Law Fit
Internet	47%	64%	0.152
Biotech	46%	64%	0.094
Savings & Loan	45%	66%	0.091
Telecom Equipment	39%	86%	0.209
Specialty Chemicals	35%	40%	0.121
Semiconductors	28%	64%	0.137
Restaurants	19%	85%	0.214
Software	17%	85%	0.126
Apparel	16%	67%	0.127

Source: CSFB analysis.

To further illustrate the procedure used to establish the results in Table 2, we present the Internet industry results in greater detail.

Figure 4
Internet Services Industry
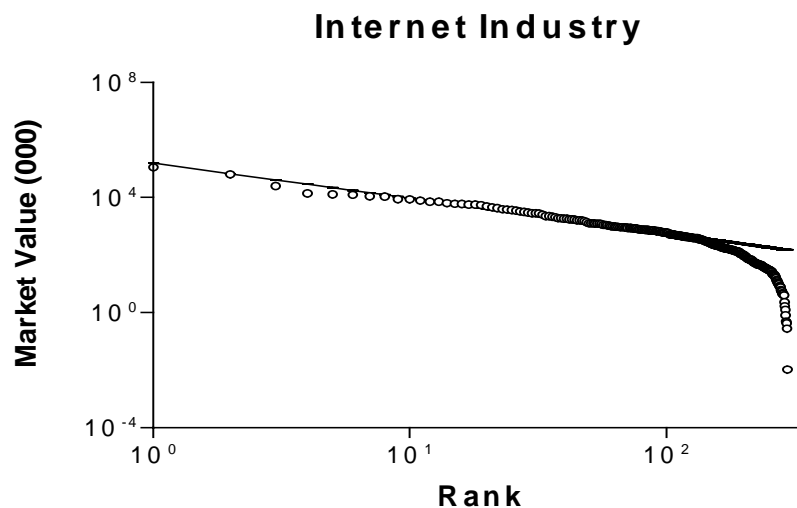
Source: CSFB analysis.

Figure 4 displays the market value (in thousands) of each of the 297 companies in the Internet industry. As a function of rank, the top 140 companies (47%) closely follow the fit line. Figure 4 demonstrates how the fit line was chosen. Each point is the result of a power law fit to a portion of the total industry. When more than 140 companies are included in the fit, the RMS values begin to systematically increase, which indicates divergence from power law behavior. Therefore the fit line in Figure 5 is chosen to best fit the top 140 companies.

Figure 5
Internet Services Industry

Best Fit at (140, 0.152)
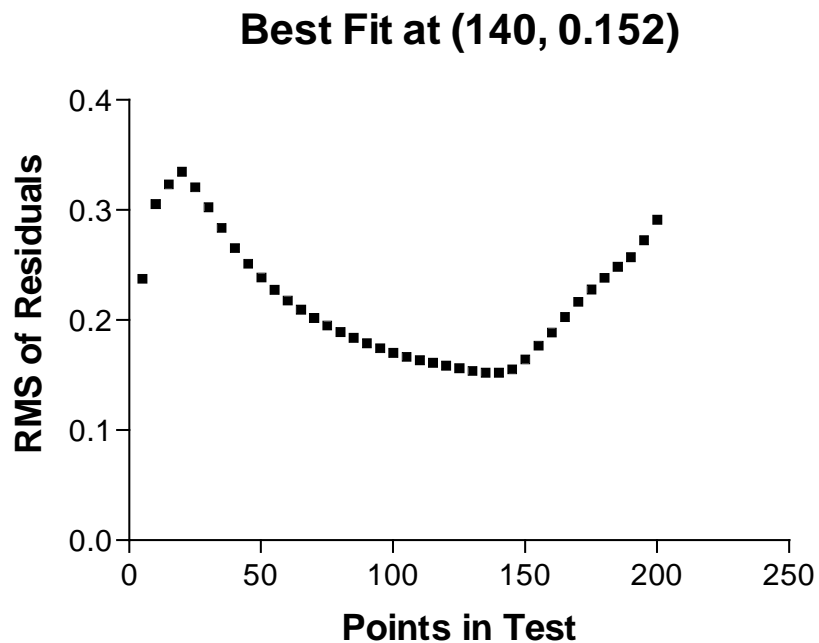
Source: CSFB analysis.

Figure 6
Power Law Segments

Internet

Biological Technology

Savings and Loans
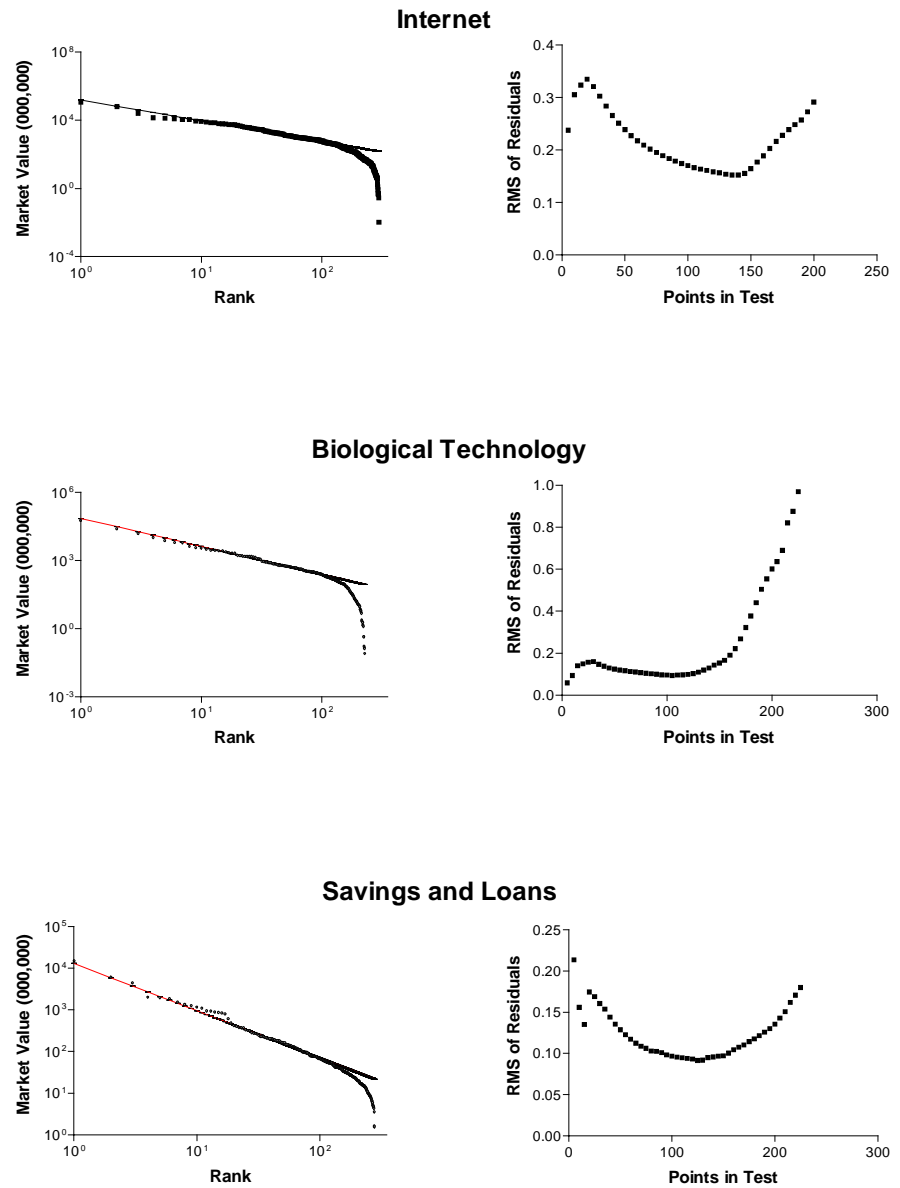
Source: CSFB analysis.

**Figure 7
Power Law Segments**

Telecommunication Equipment

Specialty Chemicals

Semiconductors
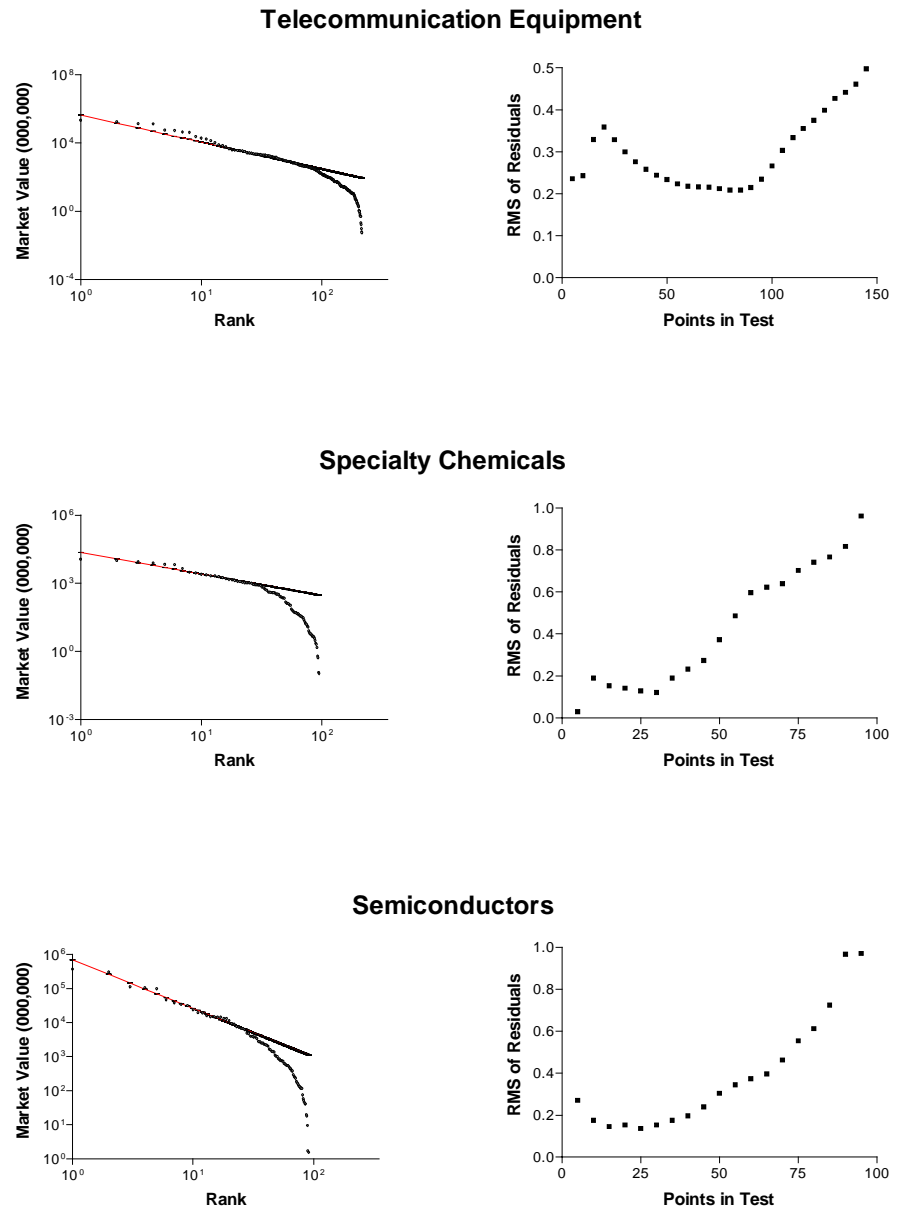
Source: CSFB analysis.

Figure 8
Power Law Segments

Restaurants
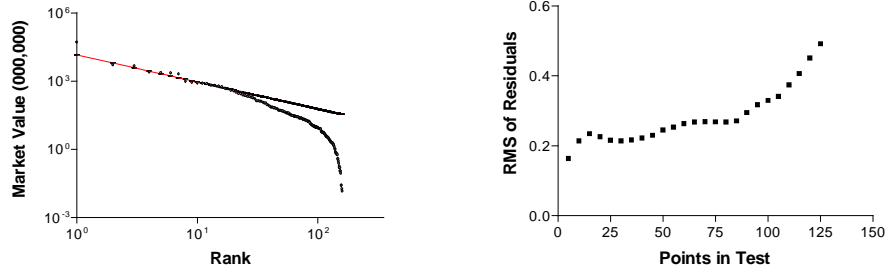
Software
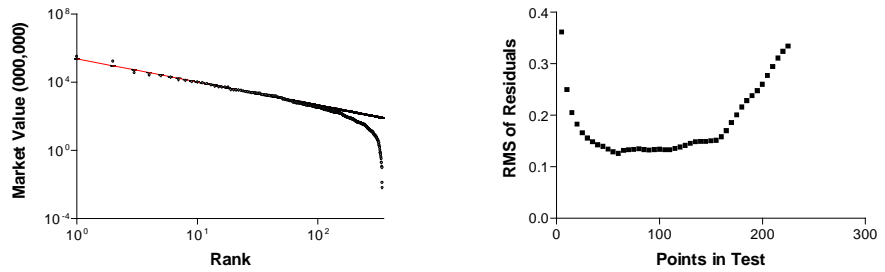
Apparel

Source: CSFB analysis.

CREDIT SUISSE | **FIRST BOSTON**
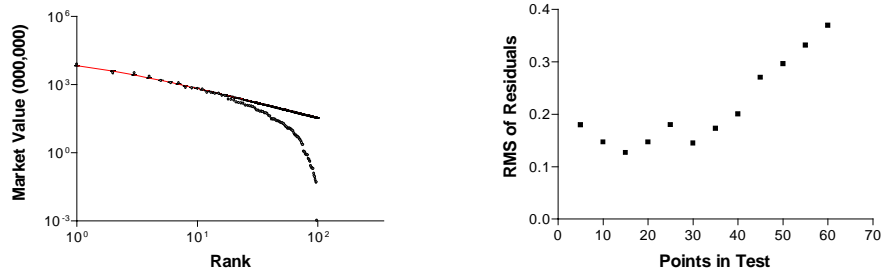
AMSTERDAM 31 20 5754 890	**LONDON**................ 44 20 7888 8888	**SAN FRANCISCO** ... 1 415 836 7600
ATLANTA.................1 404 656 9500	**MADRID** 34 91 423 16 00	**SÃO PAULO** 55 11 3841 6000
AUCKLAND 64 9 302 5500	**MELBOURNE** 61 3 9280 1666	**SEOUL** 82 2 3707 3700
BALTIMORE 1 410 223 3000	**MEXICO** 52 5 283 89 00	**SHANGHAI**............ 86 21 6881 8418
BEIJING 86 10 6410 6611	**MILAN** 39 02 7702 1	**SINGAPORE** 65 538 6322
BOSTON 1 617 556 5500	**MOSCOW** 7 501 967 8200	**SYDNEY** 61 2 8205 4400
BUDAPEST 36 1 202 2188	**MUMBAI**................... 91 22 230 6333	**TAIPEI** 886 2 2715 6388
BUENOS AIRES 54 11 4394 3100	**NEW YORK**.............. 1 212 325 2000	**TOKYO** 81 3 5404 9000
CHICAGO.................1 312 750 3000	**PALO ALTO**............. 1 650 614 5000	**TORONTO**............... 1 416 352 4500
FRANKFURT..............49 69 75 38 0	**PARIS** 33 1 40 76 8888	**VIENNA** 43 1 512 3023
GENEVA..................41 22 394 70 00	**PASADENA** 1 626 395 5100	**WARSAW**............... 48 22 695 0050
HOUSTON................1 713 220 6700	**PHILADELPHIA**....... 1 215 851 1000	**WELLINGTON**........... 64 4 474 4400
HONG KONG 852 2101 6000	**PRAGUE**................ 420 2 210 83111	**ZUG** 41 41 727 97 00
		ZURICH 41 1 333 55 55

NI2571.doc

[1] This was initially reported in *"Growth Dynamics of the World-Wide Web."* Huberman, B.A., Adamic, L. *Nature* Vol. 401, September 9, 1999. Since then the results have been reconfirmed in *"Graph Structure in the Web."* Broder, Andrei et. al. IBM Almaden Research Center, San Jose, CA., June 2000.

[2] ibid. Huberman. It's important to note that the observed power law behavior was derived from usage logs for all Web sites, not usage logs for commercial Web sites alone. Corporate strategy can influence the ability to crack the line-up.

[3] Six properties were absorbed by competitors that remained in the top 25 in April. The balance of the firms remained in the top 100.

[4] Michael Mauboussin, Alexander Schay, Stephen Kawaja, *"Absolute Power,"* Credit Suisse First Boston Equity Research, December 21, 1999.

[5] Although we could find no direct statistical correlation between site traffic and market capitalization, 14 of the top 25 companies in terms of site volume were also in the top 25 in terms of market value.

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

[7] *"Graph Structure in the Web."* Broder, Andrei et. al. IBM Almaden Research Center, San Jose, CA., June 2000

[8] *"Souped Up Search Engines."* Nature Vol. 405, May 11, 2000 *www.nature.com*

[9] "Competitive Dynamics of Web Sites." Maurer, Sebastian M., Huberman, Bernardo A. Xerox Palo Alto Research, March 17, 2000.

[10] Ibid