Order Code RL30435

CRS Report for Congress

Internet and E-Commerce Statistics: What They Mean and Where to Find Them on the Web

Updated October 24, 2000

Rita Tehan
Information Research Specialist
Information Research Division



The Congressional Research Service works exclusively for the Congress, conducting research, analyzing legislation, and providing information at the request of committees, Members, and their staffs.

The Service makes such research available, without partisan bias, in many forms including studies, reports, compilations, digests, and background briefings. Upon request, CRS assists committees in analyzing legislative proposals and issues, and in assessing the possible effects of these proposals and their alternatives. The Service's senior specialists and subject analysts are also available for personal consultations in their respective fields of expertise.

Internet and E-Commerce Statistics: What They Mean and Where to Find Them on the Web

Summary

Statistics indicating Internet usage are imprecise. It is difficult to measure the scale of the Internet (or the World Wide Web), calculate the number and types of users (age, sex, race, gender, location, etc.), or forecast future growth. This report discusses the inherent complexities of estimating Internet and electronic commerce growth and describes various types of Internet statistics, discussing how to evaluate them and providing Web addresses for locating them. In addressing these topics, it is important to understand how the statistics are compiled, how they are used, and what their limitations are.

This report also provides information on the size of the Internet (number of hosts, Web sites, and online users), as well as demographic information about users (race, gender, location, etc.) and information about electronic commerce. It briefly discusses the digital divide (i.e., the discrepancy between those with easy access to the information resources of the Internet, and those without access).

Finally, the report lists selected Web sites which contain useful demographic and statistical information about the Internet and electronic commerce. This report will be updated periodically.

Contents

Difficulties in Measuring the Internet 1

Estimated Size of the Internet 3
- Number of Web Hosts 3
- Number of Web Pages 5
- Invisible Web 6

Demographics 7

Complexities of Measuring E-Commerce 8
- E-Commerce Growth Statistics 10

Digital Divide 12

Selected Web Addresses for Internet and E-Commerce Statistics 14
- E-Commerce (General) 14
- E-Commerce (Statistics) 14
- Demographics 14
- Digital Divide 14

Internet and E-Commerce Statistics: What They Mean and Where to Find Them on the Web

Difficulties in Measuring the Internet

Statistics indicating Internet usage are imprecise. It is difficult to measure the scale of the Internet (or the World Wide Web), calculate the number and types of users (age, sex, race, gender, location, etc.), or forecast future growth. In addressing these topics, it is important to understand how the statistics are compiled, how they are used, and what their limitations are.

One problem in measuring the Internet is that many domain names are unused. An individual or organization might buy one or more domain names with the intention of building a Web site; other individuals or companies buy hundreds or thousands of domain names in the hope of reselling the name. These domains can be found with search engines or Web crawlers, but their content is nonexistent or negligible.

Another reason it is difficult to count the number of Web sites is that some sites are merely synonyms for other sites. In other words, many domain names point to the exact same site. For example, *barnesandnoble.com* and *bn.com* both point to the same site. And finally, some sites are mirror sites, which are exact duplicates of the original site on another server. Usually these are created to reduce network traffic, ensure better availability of the Web site, or make the site download more quickly for users close to the mirror site (i.e., in another part of the world from the original site).

Estimates for the number of people using the Internet differ greatly. The difficulty in measuring Internet usage is partly due to the fact that analysts use different survey methods and different definitions of "Internet access." For example, some companies begin counting Net surfers at age 2, while others begin at 16 or 18. Some researchers include users who have been on the Web only within the past month, while others include people who have never used the Internet.[1] In addition, definitions of "active users" varies from one market research firm to another. Some companies count Internet users over 15 years old who surf the Web at least once every 2 weeks for any amount of time. Other companies count casual surfers or e-mail browsers in their surveys. To compare forecasts, estimates need to be adjusted for differing definitions of Internet use and population figures.

---

[1] Lake, David. Spotlight: How Big Is the U.S. Net Population? *The Standard*, November 29, 1999. [http://www.thestandard.com/metrics/display/0,2149,1071,00.html].

The Internet presents a unique problem for surveying. For example, college students are prime Web users, but there is really no economical way to measure students' usage of the Internet. Another example of a demographic which is hard to capture is workers who use the Internet on the job. Many businesses will not allow measurement software to be installed on their sites. In addition, Web sites with a large overseas audience cannot always separate the U.S. users from the international ones.

At the heart of the issue is the methodology used to collect responses from individual users. Since there is no central registry of all Internet users, completing a census or attempting to contact every user of the Internet is neither practical nor financially feasible. Therefore, Internet surveys attempt to answer questions about all users by selecting a subset to participate in the survey. This process is called sampling. The following discussion of survey methodologies is excerpted from the Georgia Institute of Technology's *GVU's World Wide Web User Survey Background Information* Web page.[2]

There are two types of sampling, random and non-probabilistic. Random sampling creates a sample using a random process for selecting members from the entire population. Since each person has an equal chance of being selected for the sample, results obtained from measuring the sample can be generalized to the entire population. Non-probabilistic sampling is not a pure random selection process, and can introduce bias into the sampling selection process because, for example, there is a desire for convenience or expediency. With non-probabilistic sampling, it is difficult to guarantee that certain portions of the population were not excluded from the sample, since elements do not have an equal chance of being selected.

Since Internet users are spread out all over the world, it becomes quite difficult to select users from the entire population at random. To simplify the problem, most surveys of the Internet focus on a particular region of users, which is typically the United States, though surveys of European, Asian, and Oceanic users have also been conducted. Still, the question becomes how to contact users and get them to participate. The traditional methodology is to use random digit dialing (RDD). While this ensures that the phone numbers and thus the users are selected at random, it potentially suffers from other problems as well, notably, self-selection.

Self-selection occurs when the entities in the sample are given a choice to participate. If a set of members in the sample decides not to participate, it reduces the ability of the results to be generalized to the entire population. This decrease in the confidence of the survey occurs since the group that decided not to participate may differ in some manner from the group that participated. It is important to note that self-selection occurs in nearly all surveys of people. Thus, conventional means of surveying Internet usage are subject to error.

Most of the statistics gathered during the early days of the Internet only concerned the number of hosts connected to the Internet or the amount of traffic

---

[2] GVU's WWW User Survey Background Information. Graphics, Visualization & Usability Center, Georgia Institute of Technology. [http://www.cc.gatech.edu/gvu/user_surveys/].

flowing over the backbones. Such statistics were usually collected by large universities or government agencies on behalf of the research and scientific community, who were the largest users of the Internet at the time. This changed in 1991 when the National Science Foundation lifted its restrictions on the commercial use of the Internet. More businesses began to realize the commercial opportunities of the Internet, and the demand for an accurate accounting of the Internet's population increased.

Estimated Size of the Internet

Number of Web Hosts

The number of Web hosts (computers with domain names) keeps on growing, according to Alexa Research's Internet archiving project. In May 1999, Alexa counted 2.5 million hosts. In September 1999, the number had risen to 3.4 million. That is a 31% growth rate in just 4 months, or 125% per year.[3] Every two months, Alexa crawls the entire Web and counts the number of unique top-level pages. Whatever page shows up at www.example.com is considered a top-level page. Alexa then counts these pages, *removing duplicates* for an estimate of total unique Web hosts.

The Internet connects more than 72 million host computers in 247 countries.[4] The Internet is now growing at a rate of about 40% to 50% annually (for machines physically connected to the Internet), according to data from the Internet Domain Survey, the longest-running survey of Internet hosts. Such exponential growth has led to the expansion of the Internet from 562 connected host computers in 1983 to 72.3 million such computers in January 2000.[5] At any time from 1983 through 1996, half of the Internet's historical growth had occurred in the preceding 12 to 14 months.[6]

Another way to think about growth in Internet access is to compare it to other technologies from the past. It took 38 years for the telephone to penetrate 30% of U.S. households. Television took 17 years to become that available. Personal

---

[3] Grow Web, Grow! *Internet Trends Report: 1999 Review*. Alexa Research, February 1, 2000. [http://www.alexaresearch.com/top/report_4q99.cfm]

[4] International E-mail Accessibility Based on International Standard ISO 3166 Codes. Released June 1, 1999, at: [http://www.nsrc.org/codes/]. This document is a guide to country codes, showing which countries have access to the Internet or general e-mail services.

[5] Number of Internet Hosts. Network Wizards, January 1999 at: [http://www.nw.com/zone/host-count-history].

[6] *Standard & Poor's Industry Surveys*. Computers: Consumer Services and the Internet. September 30, 1999. p. 6.

computers took 13 years. Once the Internet became popular because of the World Wide Web, it took less than 7 years to reach a 30% penetration level.[7]

Although the number of people using the Internet can only be estimated, the number of host computers can be counted fairly accurately. A host is a computer server hooked to the Internet. The growth of Internet hosts is shown below:

Year	Number of Internet Hosts
1969	4 hosts
04/1971	23 hosts
08/1981	213 hosts
08/1983	562 hosts
12/1987	28,174 hosts
07/1988	33,000 hosts
07/1989	130,000 hosts
10/1990	313,000 hosts
07/1991	535,000 hosts
07/1992	992,000 hosts
07/1993	1,776,000 hosts
07/1994	2,217,000 hosts
07/1995	6,642,000 hosts
01/1996	9,472,000 hosts
01/1997	17,753,266 hosts
01/1998	29,670,000 hosts
01/1999	43,230,000 hosts
06/1999	56,218,000 hosts
1/2000	72,398,092[8] hosts

Packet traffic, a measure of the amount of data flowing over the network, continues to increase exponentially. Traffic and capacity of the Internet grew at rates of about 100% per year in the early 1990s. There was then a brief period of explosive growth in 1995 and 1996. During those two years, traffic grew by a factor of about 100, which is about 1,000% a year. In 1997, traffic growth slowed to about 100% per year.[9]

---

[7] *State of the Internet: USIC's Report on Use & Threats in 1999.* U.S. Internet Council, April 1999 at: [http://www.usic.org/].

[8] Internet statistics are compiled by Mark Lottor of Network Wizards. The Internet Domain Survey attempts to discover every host on the Internet by doing a complete search of the Domain Name System (DNS). It is sponsored by the Internet Software Consortium, whose technical operations are subcontracted to Network Wizards. Survey results are available from Network Wizards at: [http://www.isc.org/ds/WWW-200007/index.html].

[9] *The Size and Growth Rate of the Internet.* First Monday, October 5, 1998, at: [http://www.firstmonday.dk/issues/issue3_10/coffman/index.html].

Number of Web Pages

In July 2000, Cyveillance, an Internet consulting company, estimated that there were 2.1 billion unique, publicly available pages on the Internet.[10] Cyveillance states that the Internet grows by 7.3 million pages each day, which means that it will double in size by early 2001. The study also found that of the more than 350 million links considered, about 10½% generated broken link error messages.

In May 2000, researchers at IBM, Compaq, and AltaVista conducted a study which argues against the widely held impression that the entire Internet is highly connected.[11] The study looked at roughly 200 million Web pages and the five billion links to and from each page. On the basis of their analysis, the researchers set out a "bow tie theory" of Web structure. The World Wide Web is fundamentally divided into four large regions, each containing approximately the same number of pages.

The researchers found that four distinct regions make up approximately 90% of the Web (the bow tie), with approximately 10% of the Web completely disconnected from the entire bow tie. The "strongly-connected core" (the knot of the bow tie) contains about one-third of all Web sites. These include portal sites like Yahoo, large corporate sites like Microsoft, and popular news and entertainment destinations. Web surfers can easily travel between these sites via hyperlinks; consequently, this large "connected core" is at the heart of the Web.

This surprising pattern became apparent almost immediately. "About half the time, we'd follow all the links from a page and the whole thing would peter out fairly quickly," according to Andrew Tomkins, a researcher at IBM's Almaden Research Center. "The other half of the time, the list of links would grow and grow and eventually we'd find 100 million other pages – half of the whole universe."[12]

In January 2000, researchers at NEC Research Institute and Inktomi completed a study that estimated that the Web has more than one billion unique pages.[13] Interestingly, although Inktomi has crawled more than a billion pages on the Web, Inktomi's chief scientist commented at a search engine conference that "[i]t was difficult to find 500 million legitimate pages after culling duplicates and spam. We

---

[10] Size of Net Will Double Within Year. *eMarketer*, July 11, 2000. [http://www.emarketer.com/estats/20000713_size.html?ref=wn] .
See also: Internet Exceeds 2 Billion Pages. Cyveillance press release, July 10, 2000, at: [http://www.cyveillance.com/newsroom/pressr/000710.asp].

[11] Altavista, Compaq, and IBM Researchers Create World's Largest, Most Accurate Picture of the Web. IBM Research Almaden News press release, May 11, 2000. [http://www.almaden.ibm.com/almaden/webmap_release.html].

[12] Study Reveals Web as Loosely Woven. *New York Times*, May 18, 2000. [http://www.nytimes.com/library/tech/00/05/circuits/articles/18webb.html].

[13] Web Surpasses One Billion Documents. Inktomi press release, January 18, 2000. [http://www.inktomi.com/new/press/billion.html].

found 445 million, but had to go digging to get the index to 500 million."[14] A number of facts have emerged from the study:

Number of documents in Inktomi database: over one billion
Number of servers discovered: 6,409,521
Number of mirrors (identical Web sites) in servers discovered:1,457,946
Number of sites (total servers minus mirrors): 4,951,247
Number of good sites (reachable over 10-day period): 4,217,324
Number of bad sites (unreachable): 733,923

Top level domains	Percentage
.com	54.68%
.net	7.82%
.org	4.35%
.gov	1.15%
.mil	0.17%

Percentage of documents in English: 86.55%
Percentage of documents in French: 2.36%

Researchers at the Online Computer Library Center (OCLC) are conducting a study to describe the structure, size, usage, and content of the Web. In September 1999, the Web Characterization Project published the results of a study which determined that there were 3.6 million Web sites on the World Wide Web. In addition, the largest 25,000 sites account for approximately 50% of content on the Web. Out of the 3.6 million total sites, 2.2. million offer publicly accessible content with about 300 million Web pages. This represents a considerable increase from 1997, when there were 800,000 sites. Approximately 400 million pages are restricted, requiring either fees or authorization. The report further notes that about 2% of the public sites, or 42,000, contain sexually explicit material. The mean average Web site contains 129 pages, which is a 13% jump over 1998's estimate of 114 pages.[15]

Invisible Web

In addition, it is necessary to account for the "invisible Web" (databases within Web sites). According to an August 2000 study by BrightPlanet, an Internet content company, the World Wide Web is 400 to 550 times bigger than previously

[14] Sherman, Chris. 'Old Economy' Info Retrieval Clashes with 'New Economy' Web Upstarts at the Fifth Annual Search Engine Conference. *Information Today*, April 24, 2000. [http://www.infotoday.com/newsbreaks/nb000424-2.htm].

[15] OCLC Research Project Measures Scope of the Web. OCLC press release, September 8, 1999. [http://www.oclc.org/oclc/press/19990908a.htm].

estimated.[16] According to this study, the Web consists of hundreds of billions of documents hidden in searchable databases unretrievable by conventional search engines – what it refers to as the "deep Web." The deep Web contains 7,500 terabytes of information, compared to 19 terabytes of information on the surface Web. A *single* terabyte of storage could hold each of the following: 300 million pages of text, 100,000 medical x-rays, or 250 movies.[17]

Search engines rely on technology that generally identifies "static" pages, rather than the "dynamic" information stored in databases. Deep Web content resides in searchable databases, the results from which can only be discovered by a direct query. Without the directed query, the database does not publish the result. Thus, while the content is there, it is skipped over by traditional search engines which cannot probe beneath the surface. Examples of Web sites with "dynamic" databases are: THOMAS (legislative information), PubMed and Medline (medical information), SEC corporate filings, Yellow Pages, classifieds, shopping/auction sites, library catalogs, etc. BrightPlanet has developed a software called "LexiBot" which searches not only pages indexed by traditional search engines, but delves into Internet databases as well.

Demographics

Various research and consulting firms have estimated the number of U.S. Internet users to be between 123 and 144 million in 2000.[18] These figures do not include military computers, which for security reasons are invisible to other users. Many hosts support multiple users, and hosts in some organizations support hundreds or thousands of users.

About 144 million people—52% of the U.S. population—have home access to the Internet, according to an August 2000 survey by Nielsen/Net Ratings.[19] The figures represent an increase of 35% over the 104.6 million people the company reported online at home in July 1999.

The market research firm Media Metrix said a survey of 55,000 Internet users found low-income users, especially those from households earning less than $25,000 a year, grew nearly 50% in the 12 months ending in June 2000.[20] While the lower-

---

[16] The Deep Web: Surfacing Hidden Value. Bright Planet, July 2000. [http://www.completeplanet.com/help/help_deepwebFAQs.asp].

[17] The Life Cycle of Government Information: Challenges of Electronic Innovation. 1995 FLICC Forum on Federal Information Policies, Library of Congress, March 24, 1995. [http://lcweb.loc.gov/flicc/forum95.html].

[18] How Many Online. Nua Internet Surveys. Regularly updated. [http://www.nua.ie/surveys/how_many_online/n_america.html].

[19] Chidi, George. More Than Half of U.S. Homes Online. *PC World.com*, August 24, 2000. [http://www.pcworld.com/shared/printable_articles/0,1440,18210,00.html].

[20] The Dollar Divide: Household Income and Web Usage. Media Metrix, August 2000. [http://www.mediametrix.com/home.jsp?language=us].

(continued...)

income Internet users grew more than 49% to 7.5 million users, they still represent only 10% of the total online population.

Web surfers earning more than $75,000 make up nearly one-third of total users. Americans spend more hours online than their counterparts in Australia, Britain, Canada, France, Germany, and Japan.[21] American home users spend an average of more than 13 hours per month online, based on measurement figures by Media Metrix in May 2000. The United States has more Web users than the next 15 countries combined. But only three of the next six biggest Web-using nations—Australia, Britain, and Canada—have predominantly English-speaking populations.

Web use became balanced between the sexes for the first time with 31.1 million men and 30.2 million women online in April 2000, according to Media Metrix.[22] The fastest growing group among Web users is women over 55. In 2000, their numbers increased by 98.1% over 1999.

Complexities of Measuring E-Commerce

A new activity emerged to fill the need to gather e-commerce statistics: gathering and selling strategic and statistical information about the Internet. Internet organizations, such as the Internet Society and the International Telecommunications Union, began to compile information on the size and growth of the Internet. Traditional research and polling firms such as Nielsen and Gallup also entered the field, along with new types of companies, such as Forrester Research, Gartner Group, Jupiter Communications, International Data Corporation (IDC), and Zona Research, which provide estimates of the impact of the Internet on consumers and business.

Web traffic measurement may seem tedious, but with the number of dollars at stake, it is very important to businesses. Advertisers, who are spending a lot of money, are very interested in knowing how many eyes, and whose, are viewing their ads. Federal and state governments need demographic information also, to assist them in distributing resources and financial assistance.

Another factor must be considered in evaluating Internet statistics. More and more businesses have decided that the Internet is the key to success and are aware of their acute need for e-commerce guidance. With increasing frequency, yet another e-commerce research firm declares that it has all the answers for e-commerce strategies and schemes. "No one knows how much real insight online consulting firms

---

[20] (...continued)
See also: Press release [http://www.mediametrix.com/press/releases/20000821.jsp].

[21] Chea, Terence. Americans Are No. 1 in Time Online. *Washington Post*, August 3, 2000. p. E4.

[22] Austen, Ian. Studies Reveal a Rush of Older Women to the Web. *New York Times*, June 29, 2000. [http://www.nytimes.com/library/tech/00/06/circuits/articles/29numb.html].

provide, but what would-be Internet player can afford not to subscribe? Planning for the future is hard in a mature industry; it's nearly impossible in one still teething."[23]

Forrester Research in Cambridge, MA, and Jupiter Communications in New York City are two of the largest Internet research firms. They are so-called syndicated research firms, which means that they publish a wide range of reports with high subscription fees to a small, targeted audience of corporate executives. For approximately $20,000, a company can buy a subscription for one of eight subjects from Jupiter Communications. Then every month for a year, it receives a 32-page report filled with analysis and advice, survey data, and industry forecasts, all on the impact of e-commerce.

"Stepping boldly into the answer vacuum, the companies [Forrester and Jupiter] have transformed themselves into prediction factories, supplying the world with a steady stream of airy, context-free, yet reassuringly precise projections ...."[24] These firms don't simply pull forecasts out of thin air; they strive to guess methodically: the analysts interview advertisers and executives at top Web sites, review annual reports, adjust overly-optimistic figures, assemble historical research comparing ad spending with consumer research, and estimate spending for online advertising. Forrester's chief Internet advertising analyst says, "The interesting thing about projections is that they come out looking very exact. But really, it's just your opinion expressed numerically."[25]

One observer says the models used by Forrester or Jupiter "have little resemblance to statistical techniques, like regression analysis or time-series analysis, used by traditional market researchers or industrial forecasters to determine next year's worldwide consumption of, say, gasoline or Coca-Cola. That, of course, is because enterprises that forecast the consumption of resources or consumer items can draw on decades of historical data."[26] Using such techniques would not now work for Forrester or Jupiter because the Internet e-commerce environment has existed for only 3 to 4 years for most businesses.

There is much debate over which Web measurement company's methods are more accurate. The research firms disagree about sampling methods and panel selection, since differing methods of identifying and soliciting the survey participants result in different ratings.

Yet, in the business of assessing the Internet's future, there seems to be little accountability for incorrect forecasts. Many industry executives "admit that they view

---

[23] Roth, Daniel. My, What Big Internet Numbers You Have! *Fortune*, March 15, 1999. p. 114-20.

[24] Frederick, Jim. $6 Billion on Online Holiday Sales by the End of This Month! $24 billion in Internet Ads by 2003! 2.3 Trillion E-biz Predictions by 2010! *New York Times Magazine*, December 19, 1999. p. 70-73.

[25] Roth, p. 120.

[26] Frederick, p. 70.

the companies and the media froth they generate as necessary evils ...."[27] Businesses need the Internet research consultants' estimates to write business plans and attract investors, but, at this time, the estimates are not based on historically valid measurement models.

Another controversy centers around how to measure Web use in the workplace, which is estimated to account for 35% to 45% of Internet use. Companies are notoriously reluctant to place Web measurement software on their workers' computers.[28] Privacy and the protection of proprietary business information would most likely have to be resolved before this could become a common measurement tool.

E-Commerce Growth Statistics

With all the caveats discussed above, below is a sampling of estimates of the size and growth of e-commerce. (For a selected list of Web sites for e-commerce statistics, see the *Selected Web Addresses for Internet and E-Commerce Statistics*, below.) Recent e-commerce statistics include:

- According to ActivMedia, e-commerce activity for the year 2000 will reach $132 billion worldwide, more than double the $58 billion reported in 1999.[29] This figure includes revenues from business-to-business, business-to-consumer, and Internet services. Online retail sales for 1999 were estimated at $66 billion.[30]

- Worldwide Christmas/holiday sales are forecast to be worth $19.5 billion in 2000, according to the Gartner Group.[31] North America will account for more than half of online sales in this period this year, with a 55% share. Europe and Asia/Pacific are projected to grow 96% and 91%, respectively, with a combined market share of 36%. 1999 online holiday sales hit $7 billion (customers who shopped between November 1 and December 31, 1999).[32]

---

[27] Ibid., p. 73.

[28] O'Leary, Mike. Web Measurers Wrestle with Methodologies, Each Other. *Online*, May/June 1999. p. 106.

[29] Real Numbers Behind 'Net Profits 2000. ActivMedia, June 2000. [http://www.activmediaresearch.com/real_numbers_2000.html]

[30] ActivMedia Research LLC, January 12, 2000. [http://www.activmediaresearch.com].

[31] Gartner Forecasts Worldwide Online Holiday Sales to Top $19 Billion in 2000. Gartner Group press release, September 5, 2000. [http://gartner6.gartnerweb.com/public/static/aboutgg/pressrel/pr20000905a.html].

[32] Online Holiday Sales Hit $7 Billion, Consumer Satisfaction Rising. Jupiter Communications press release, January 13, 2000. [http://www.jup.com/company/pressrelease.jsp?doc=pr000113].

- A July 2000 survey of online consumer shopping in 27 countries, *Global eCommerce Report*, found that 10% of worldwide Internet users shop online (defined as users who have bought or ordered online in the last month).[33] The United States, Japan, and Norway have the highest percentage of users who shop online (27%, 20%, and 19%, respectively).

In August 2000, the U.S. Department of Commerce released total e-commerce sales figures for the second quarter of 2000.[34] The Department estimated that U.S. retail e-commerce sales for second quarter 2000, not adjusted for seasonal, holiday, and trading-day differences, was $5.518 billion, an increase of 5.3% from the revised first quarter 2000 level. The first quarter estimate was revised from $5.260 billion to $5.240 billion. The e-commerce sales report provides a simple tally of Internet retail sales, based on a survey of roughly 2,000 Web merchants. One analyst concludes, "That is a far cry from the monthly retail reports released by the Census Bureau, which break down sales totals of traditional retailers by categories like shoes, liquor, and furniture."[35] Census Bureau officials hope to achieve a similar level of detail with e-commerce reports, but they must overcome obstacles which have slowed the effort.

First is the issue of how to categorize retailers who appear, disappear, or change their products with dizzying regularity. For instance, Lee Price, chief economist for the Economics and Statistics Administration, said, "It's not just a question of taxonomy. It's one of evolving taxonomy. Amazon used to just sell books. Now they sell a much more varied selection. You have to figure out how to capture that."[36] Also, businesses are still developing interactions between brick and mortar establishments and their e-commerce equivalents (for example, Wal-Mart's physical stores compete with its Internet presence).

Second, projecting data from 2,000 respondents so it represents the entire universe of e-commerce spending also presents analytical dilemmas. For example, analysts must weigh the data from those who did not respond, just as they weigh the results of those who responded. The Commerce Department cannot compel companies to respond. Finally, there is the hurdle of actually collecting the data. Some electronic retailers do not want to be bothered, and others want assurances that the information they send will be safe and protected.

Industry analysts and executives are hopeful that Commerce Department figures will provide more reliable information than is now available. This will be important

---

[33] 10% of Worldwide Net Users Buy Online. *eMarketer*, July 31, 2000. [http://www.emarketer.com/estats/20000731_tns.html?ref=wn].

[34] Retail E-Commerce Sales in Second Quarter 2000 Increased 5.3 Percent from First Quarter 2000, Census Bureau Reports. Bureau of the Census press release, August 31, 2000. [http://www.census.gov/mrts/www/current.html]

[35] Tedeschi, Bob. Government Figures Will Shed Little Light on Holiday Online Sales. New York Times Cybertimes, January 10, 2000. [http://www.nytimes.com/library/tech/00/01/cyber/commerce/10commerce.html].

[36] Ibid.

for seeing long-term trends, but they will not provide immediate data on Internet sales. According to Jack Staff, chief economist with Zona Research, "It'll take at least three years for the government data to be highly usable. But there's a whole segment of the Internet industry that's devoted to the numbers, and it'll be fundamentally changed once better numbers come along. And that's as it should be."[37]

Digital Divide

More Americans than ever have access to telephones, computers, and the Internet. At the same time, however, according to the U.S. Department of Commerce's recent report, *Falling Through the Net: Defining the Digital Divide*, there is still a significant "digital divide" separating American information "haves" and "have nots."[38] Indeed, in many instances, the digital divide has widened in the last year. As the Commerce Department report (issued by its National Telecommunications and Information Administration) states, there is a persistent discrepancy between the information rich (who frequently include whites, Asians/Pacific Islanders, those with higher incomes, those more educated, and dual-parent households) and the information poor (many of whom are younger, with lower incomes and education levels, certain minorities, and those in rural areas or central cities).

Households with annual incomes of $75,000 and above are more than 20 times as likely to have Internet access as households at the lowest income levels. Households that identified themselves as being black or Hispanic are only 40% as likely as white households to be online. The differences cannot be explained by income alone. More than a third of white families earning between $15,000 and $35,000 per year owned computers, but only one in five black families at the same income level did.

Regardless of income, Americans living in rural areas are, in general, lagging behind in Internet access, according to the Commerce Department's *Falling Through the Net* report. Indeed, at the lowest income levels, those in urban areas are more than twice as likely to have Internet access than those earning the same income in rural areas. Community access centers, such as schools, libraries, and other public locations, play an important role. The 1998 data demonstrate that community access centers are particularly well used by those groups who lack access at home or at work. These same groups (which often include those with lower incomes and education levels, certain minorities, and the unemployed) are also using the Internet at higher rates to search for jobs or take courses.

---

[37] Ibid.

[38] U.S. National Telecommunications and Information Administration. *Falling Through the Net: Defining the Digital Divide*, July 8, 1999, at: [http://www.ntia.doc.gov/ntiahome/fttn99/contents.html].

Some commentators have expressed skepticism about the existence of a digital divide. The Commerce Department report relied on results from a 1998 survey, and given the fast pace at which Americans buy computers and Internet access accounts, failure to use more current numbers skews the results. Skeptics believe that the divide is economic rather than racial.[39] A study by the Joint Center for Political and Economic Studies showed that 11% of African American households with incomes under $15,000 reported using the Internet at home or at work, while 83% of blacks with incomes over $90,000 used the Internet.[40]

President Clinton proposed federal subsidies to narrow the digital divide and help millions of low-income families go online. The FY2001 budget proposal calls for spending:

- $2 billion over 10 years in tax incentives to encourage private sector donation of computers, sponsorship of community technology centers, and technology training for workers.
- $150 million to help train all new teachers entering the workforce to use technology effectively.
- $100 million to create 1,000 Community Technology Centers in low-income urban and rural neighborhoods.
- $50 million for a public/private partnership to expand home access to computers and the Internet for low-income families.
- $45 million to promote innovative applications of information and communications technology for underserved communities.
- $25 million to accelerate private sector deployment of broadband networks in underserved urban and rural communities.
- $10 million to prepare Native Americans for careers in information technology and other technical fields.[41]

In addition, the Commerce Department's Technology Opportunities Program is offering $12.5 million in grants to fund projects that will help close the digital divide. The funds will be made available to local governments and non-profit organizations that develop projects that provide opportunities for technologically underserved

---

[39] Hubbard, Lee. A Disingenuous Divide. *Intellectual Capital.com*, February 10, 2000. [http://www.intellectualcapital.com/issues/issue345/item8150.asp].

[40] National Opinion Poll Factsheet, Internet Use, 1998 and 1999. [http://www.jointcenter.org/selpaper/inetfcts.htm].

[41] The Clinton-Gore Administration: From Digital Divide to Digital Opportunity. White House press release, February 2, 2000. [http://www.pub.whitehouse.gov/uri-res/I2R?urn:pdi://oma.eop.gov.us/2000/2/2/4.text.1]. Another summary is available at the Digital Divide Network at: [http://www.digitaldividenetwork.org/clinton_budget.adp].

communities to use the Internet in homes, education, and business. The program, under its previous name, Telecommunications and Information Infrastructure Assistance Program, has made $135 million in matching grants since 1994.[42]

Selected Web Addresses for Internet and E-Commerce Statistics

Sites listed are generally stable and timely with Internet and e-commerce statistics:

E-Commerce (General)

- U.S. Government Electronic Commerce Policy [http://www.ecommerce.gov/]
- *Digital Economy 2000* (report from the U.S. Department of Commerce, June 2000) [http://www.esa.doc.gov/de2k.htm]
- Center for Research in Electronic Commerce (University of Texas, Austin) [http://cism.bus.utexas.edu/]
- University of California E-conomy Project [http://e-conomy.berkeley.edu/]

E-Commerce (Statistics)

- eMarketer [http://www.emarketer.com/estats/]
- Internet Economy Indicators [http://www.internetindicators.com/facts.html]

Demographics

- Headcount (Internet usage around the world) [http://www.headcount.com/count/datafind.htm?choice=country]
- Nua Internet Surveys—*How Many Online* [http://www.nua.ie/surveys/how_many_online/index.html]
  General demographic information [http://www.nua.ie/surveys/]
- Internet Domain Survey (Network Wizards) [http://www.isc.org/ds]

Digital Divide

- Digital Divide (*Education Week*) [http://www.edweek.org/context/topics/issuespage.cfm?id=48]

---

[42] For more information, see the U.S. Department of Commerce. Technology Opportunities Program Overview Web page at: [http://www.ntia.doc.gov/otiahome/top/index.html].

- Digital Divide Network [http://www.digitaldividenetwork.org/] Click on Grants & Funding for information on government, corporate, and private funding.