

Internet a

Donna L. Hoffman, William D. Kalsbeek,
and Thomas P. Novak

TAPPING THE ENORMOUS POTENTIAL OF THE INTERNET AS A COMMERCIAL MEDIUM and market is proving to be challenging. Computer-mediated environments, like the World-Wide Web, are not well understood and possess unique characteristics distinguishing them in significant ways from traditional, physical targets of opportunity [11, 15]. Electronic commerce—as research area, business and investment opportunity, and entire new industry—is in its infancy. There is a great deal of uncertainty in this complex commercial

environment and little solid information on which to base critical business decisions.

A variety of studies have suggested that the demographic characteristics of individuals are highly correlated with their use of the Internet. For example, previous research suggests that Internet use is related to gender, education, income, race, occupation, and even geographic region and that males are generally more likely than females to use the Net, peo-

ple with higher socioeconomic status are more likely to use the Net, and whites are more likely to use the Net [21, 22]. Recent research also suggests that the demographics of Internet use are shifting, so the Internet appears to be going more “mainstream” in its demographic makeup, and that this trend is likely to continue as the Internet moves toward critical mass as a commercial medium [9].

Our objectives in this article are twofold:

These baseline numbers and marketing insights for commercial development reveal how to estimate Internet use in all its ambiguity using the CommerceNet/Nielsen Internet Demographic Survey as a case study.

and Web Use in the U.S.

- Provide interested parties—consumers, market researchers, Internet-based business interests, investors, the media, and policymakers—the accurate baseline numbers and marketing insights necessary for decision making and stimulating inquiry in this complex and rapidly changing environment.
- Demonstrate the difficulties and questions inherent in estimating a complex moving target.



Other Studies of Internet Use

Several studies have been performed to date attempting to address the issue of who is on the Internet in the U.S. and what they are doing there. For a good summary of many of the studies, see the CyberAtlas at www.cyberatlas.com [7].

At first glance, different studies seem to show disparate results that are difficult to reconcile. For example, the CommerceNet/Nielsen Internet Demographic Survey reported that 37 million people 16 years and older in the U.S. and Canada have access to the Internet; 24 million of these people use the Inter-

net; 18 million use the Web; and 2.5 million of these Web users have bought products through the Internet [3, 4].

In contrast, a number of market research firms and publishers have reached different conclusions.

For example, O'Reilly & Associates found 5.8 million adults use the Internet [20]; FIND/SVP estimated that 8.4 million adults and 1.1 million children use the Internet [8]; and Times Mirror estimated that 25 million adult Americans are online [21].

Closer examination suggests that, apart from methodological flaws or survey bias, the differences are reconciled—at least in part—by understanding what is meant by Internet “use.”

O'Reilly, with the lowest estimate, uses a restrictive definition of Internet use, defining an Internet user as an individual 18 or older with direct access to the Internet, as well as access to email and other Internet applications. The O'Reilly count excludes individuals whose only access to the Internet is through an online service.

We estimate that 28.8 million people in the U.S. who are 16 years and older have access to the Internet, 16.4 million use the Internet, 11.5 million use the Web, and 1.51 million have used the Web to purchase something.

FIND/SVP, with a somewhat larger estimate than O'Reilly, uses a similar but less restrictive definition than O'Reilly, because its definition of an adult Internet user is an individual 18 years or older who uses at least one Internet application (in addition to email) and includes individuals accessing the Internet through commercial online services.

Times Mirror, with the largest estimate, uses the most liberal definition of online use, defining an "online" user as an individual 18 or older using a computer from home, work, or school to connect with computer bulletin boards, commercial online services, or other computers through the Internet.

Deriving Valid Estimates of Internet Use

Our goal is to obtain accurate baseline estimates of Internet use in the U.S. among people 16 years and older from a survey sample performed in August 1995. We set up the survey in the following ways:

- The weighted sample is representative of the U.S. population on several key variables known to affect Internet use.
- The estimates of Internet use developed are based on conceptually valid and logically consistent definitions of Internet access and use for the U.S. Internet market.
- Internet market segments derived from these data provide important insights into the nature of the Internet experience and facilitate the study of its evolution over time.

To demonstrate the difficulties in achieving such goals, consider the CommerceNet/Nielsen Internet Demographics Survey (CNIDS) [3, 4]. The CNIDS was based on an unrestricted random-digit dial sample frame [1] of telephone numbers in the U.S. and Canada. While publicly available reports [3] provide results for the combined U.S. and Canadian population ages 16 and older to facilitate comparison with U.S. Census data, our analyses of the CNIDS data are based on only the 3,785 usable responses from the U.S. sample 16 years and older.

The CNIDS data has two serious problems that must be rectified before accurate

estimates of Internet use can be derived:

- **Not adequately weighted.** The CNIDS sample was not adequately weighted by Nielsen for the U.S. population [13]. Table 1 compares CNIDS, weighted only by gender, and the 1995 U.S. Census Current Population Survey (CPS) with respect to marginal percentages for age, education, and gender, three variables known to affect Internet use. The marginal percentages are the frequencies in each demographic category, expressed as a percentage of the total. For example, the percentage of respondents with a high school education or less is 42.8% in the CNIDS weighted sample but 54.2% in the U.S. population. Table 1 shows that the weighting developed by Nielsen overrepresents CNIDS respondents who are ages 16–24, have a college degree or higher, or are women and underrepresents respondents who are ages 55 and older, have a high school education or less, or are men. The general pattern is that the groups overrepresented by CNIDS contain respondents more likely to access and use the Internet and the Web, while the underrepresented groups contain respondents less likely to do so. The only exception is the relatively small tendency for CNIDS to overrepresent women, who are less likely to use the Internet and the Web than men. Overall, the effect of this demographic imbalance in the CNIDS sample is to inflate the Nielsen estimates of Internet and Web access and use.

Table 1. Weighted marginal percentages for U.S. respondents ages 16 and older

		1995 CNIDS (%)	1995 Census CPS (%)
Age:	16–24	18.3	16.3
	25–34	20.9	20.8
	35–44	22.8	21.3
	45–54	16.5	15.4
	55+	21.5	26.1
Education:	High school or less	42.5	54.2
	Tech/some college	28.3	25.6
	College degree or more	28.9	20.2
Gender:	Men	47.0	48.1
	Women	53.0	51.9
Estimated total number of respondents:		199 million	199 million

Measure:	North America		U.S. only		Estimate corrected for respondent inconsistency ¹ :	
	(a)	(b)	(c)	(d)	(e)	
	Nielsen weights ²	Nielsen weights ³	Project 2000 weights	Nielsen weights	Project 2000 weights	
1. Do you personally have access to the Internet?	36.8 million	33.2 million	28.8 million	— ⁴	—	
2. Have you accessed the Internet at home, at work, or at school within the past three months?	24.0 million	22.0 million	18.4 million	19.7 million	16.4 million	
3. When connected to the Internet, do you have access to the Web ⁵ (in past three months)	18.2 million	16.9 million	14.1 million	14.0 million	11.5 million	
4. Thinking about the times you have accessed the Web, have you ever used it to purchase a product or service?	2.5 million ⁶	1.97 million	1.58 million	1.85 million	1.51 million	
<p>¹ To correct for respondent inconsistency, items asked subsequently in the survey were used to eliminate respondents who later indicated they did not have access to or use the Internet or the Web. Thus, the corrected estimate is more conservative and is based only on those respondents consistently stating in two or more questions that they actually used the Internet or the Web.</p> <p>² Source: CommerceNet/Nielsen [2].</p> <p>³ We multiplied the original Nielsen weights by a constant so that the sum of weights equaled the estimated number of U.S. respondents ages 16+ in 1995. This was done to ensure comparability of population estimates obtained using the Nielsen and Project 2000 weightings.</p> <p>⁴ No items were available to correct this estimate for inconsistency.</p> <p>⁵ This estimate is a composite variable constructed from a positive response to one of four questions related to Web access.</p> <p>⁶ The number reported in the CNIDS Executive Summary [2, 3] is incorrect. The correct number for the North American sample, using the original Nielsen weights, is 2.08 million.</p>						

Table 2. Estimates of Internet and Web access and use based on Nielsen and revised Project 2000 weights, ages 16 and older

• **No check for logical consistency.** The publicly reported CNIDS estimates of Internet and Web use (see Table 2) did not incorporate a check for logical consistency of responses. The CNIDS survey instrument contains more than 40 items pertaining to individual access and use of the Internet and online services. Multiple items concerning Internet and Web access and use were present in the survey instrument, allowing a check on individual consistency in responses to all questions related to a topic. The practical effect of the distributional differences on key demographic variables is to produce large percentage differences in the estimates of Internet and Web access and use. Although the differences in these single-variable marginals between CNIDS and CPS in Table 1 are relatively modest, the joint distribu-

tions (not shown in the interests of space) reveal much larger differences. When the adjustment for weighting on the demographic variables is combined with the corrections for inconsistency, the percentage differences in the various estimates become dramatic.

Weighting the Sample to Make It Representative

Poststratification adjustment [16, 19] is a commonly used methodology for reweighting data from a sample survey to ensure the weighted joint distribution of the sample with respect to key demographic variables (e.g., the cross-classification of gender, age, and education) matches the joint distribution from the best available external source. As the CNIDS was conducted in August 1995, the March 1995 CPS provides the best available estimate of the joint distribution of gen-

	High Web Use	Low Web Use	Never Use Web
High Internet Use	Segment 1 4.94 million	Segment 2 4.55 million	Segment 3 3.39 million
Low Internet Use		Segment 4 2.06 million	Segment 5 1.51 million
Never Use Internet			Segment 6 182.5 million

- Segment definitions:**
- 1) Use the Internet once a week or more and check top two boxes of a fourpoint scale of Web use frequency.
 - 2) Use the Internet once a week or more and check bottom two boxes of a fourpoint scale of Web use frequency.
 - 3) Use the Internet once a week or more and do not use the Web.
 - 4) Use the Internet a couple times a month or less and use the Web.
 - 5) Use the Internet a couple times a month or less and do not use the Web.
 - 6) Never use the Internet or the Web.

Table 3. Internet and Web use segments

ment, we reviewed which correlated demographics in the CNIDS interview—age, gender, education, and income—would also pass the consistency test and would be comparable in the CNIDS and the 1995 Census data.

We excluded income, not because we thought it failed to qualify as a legitimate predictor of Internet use, but for two other reasons:

- Because the notorious “softness” of income measures in surveys in general makes income a weak predictor, and
- Because of the relatively high nonresponse rate for income in the U.S. CNIDS sample in particular.

der, age, and education in the U.S. at the time the CNIDS data was collected.

Nielsen used a poststratification adjustment procedure, but poststratified the sample to compensate only for gender imbalances. Since characteristics other than gender are likely to be useful for adjust-

Further, education is a reasonable proxy for socioeconomic status and is highly correlated with Internet access and use. Thus, we added age and education (actually the highest level of education attained) to the list of adjustment variables for three reasons:

- These are potentially useful adjustment variables;

		High Internet Use			Low Internet Use		
		(1) High Web Use	(2) Low Web Use	(3) No Web Use	(4) Low Web Use	(5) No Web Use	(6) Non- User
Age:	16–24	22.1	23.6	32.7	20.4	34.6	15.5
	25–34	32.6	28.1	25.1	27.8	25.1	20.1
	35–44	26.9	24.1	19.0	27.7	18.1	21.1
	45–54	15.7	18.0	15.3	17.0	16.4	15.3
	55+	2.7	6.3	8.0	7.1	5.9	28.0
		100%	100%	100%	100%	100%	100%
Education:	High school or less	13.5	15.1	26.6	17.8	20.4	57.5
	Tech/some college	30.7	27.5	31.1	24.4	40.1	25.2
	College degree +	55.8	57.4	42.3	57.8	39.5	17.3
		100%	100%	100%	100%	100%	100%
Gender:	Men	78.0	72.0	55.7	62.8	44.2	46.5
	Women	22.0	28.0	44.3	37.2	55.8	53.5
		100%	100%	100%	100%	100%	100%
Children:	Children under 17	51.0	45.9	46.9	52.4	44.6	41.1
	No children under 17	49.0	54.1	53.1	47.6	55.4	58.9
		100%	100%	100%	100%	100%	100%
Segment size in millions:		4.94	4.55	3.39	2.06	1.51	182.5

Table 4. Demographic characteristics of Internet and Web use segments and nonuse segments

		High Internet Use			Low Internet Use		
		(1) High Web Use	(2) Low Web Use	(3) No Web Use	(4) Low Web Use	(5) No Web Use	(6) Non- User
Access type:	Direct only	26.5	19.9	19.1	15.1	10.3	0
	Direct and online service	48.3	39.9	30.0	33.1	28.3	0
	Online service only	21.5	36.0	46.8	45.4	51.9	7.8
	Don't know/none	1.7	4.3	4.1	6.4	9.5	92.2
		100%	100%	100%	100%	100%	100%
Home Computers Owned by Household:	None	7.1	6.8	12.9	6.2	19.2	61.6
	None/plan to buy	3.2	6.6	7.0	4.8	4.4	6.0
	1	43.7	40.4	50.0	45.9	43.8	25.6
	2	26.4	26.8	21.4	28.8	22.2	4.7
		19.6	19.4	8.7	14.3	10.4	2.0
		100%	100%	100%	100%	100%	100%
Own a modem:	Yes	82.2	80.0	68.1	73.6	64.6	114.5
	No	17.8	20.0	31.9	26.4	35.4	85.5
		100%	100%	100%	100%	100%	100%
How long a computer user:	Less than 1 year	2.1	5.1	6.4	10.1	14.4	10.1
	1-3 years	9.3	10.9	21.4	7.1	13.2	9.5
	3-5 years	8.1	9.7	15.7	18.3	7.8	6.7
	5+ years	77.5	70.1	48.7	61.5	58.4	20.6
	No current	3.1	4.2	7.8	3.0	6.2	53.1
Home/work access		100%	100%	100%	100%	100%	100%
Segment size in millions:		4.94	4.55	3.39	2.06	15.1	182.5

Table 5. Computer use behavior of Internet and Web use segments

- Population data for U.S. civilian, noninstitutionalized populations—the population as defined for the survey—is readily available; and
- Questions on these two items were included in the study interview.

Note that it would have been appropriate to also use race, but Nielsen did not collect it.

Full details of our weighting procedure using a trivariate poststratification adjustment are provided in [10]. The outcome of this adjustment is a set of new weights, which we refer to as the Project 2000 weights. The CNIDS raw data, when weighted using the Project 2000 weights, has three properties:

- Marginal distributions of gender, age, and education match exactly the 1995 CPS percentages reported in the right column of Table 1;
- The joint distribution of gender, age, and education matches exactly the corresponding 1995 CPS percentages; and
- The total number of U.S. respondents ages 16 and older is 199 million, the same number estimated by the 1995 CPS.

Note that our weighting analysis may still be biased upward (that is, our estimates might be reduced even further) because we use only three demographic vari-

ables (e.g., race and region are also known to be related to Internet use) and three broad categories of education. Finer category distinctions might produce even larger differences in the joint distributions.

ONE ADDITIONAL FACTOR THAT MAY CONTRIBUTE to the statistical problems with the CNIDS weighted sample, besides a deficient weighting scheme, is the cooperation rate, calculated as the percentage of eligible individuals who agreed to participate in the survey. Low cooperation rates may make the sample nonrepresentative, especially if those declining to participate are different from sample respondents in ways that affect the responses to the variables under study.

For example, if Internet users were more likely to participate in the survey, perhaps because of an intrinsic interest in high-technology surveys, and non-Internet users were less likely to participate due to lack of interest in such surveys, this sampling bias could cause inflation in the estimates. The cooperation rate per se does not reveal the source of the bias, but low rates may indicate the presence of bias. The sample cooperation rate for the CNIDS was not publicly released but is reported in the proprietary Final Report [4], available for a fee from the CommerceNet Consortium or from Nielsen Media Research.

Our adjustment procedure is also intended to par-

		High Internet Use			Low Internet Use		
		(1) High Web Use	(2) Low Web Use	(3) No Web Use	(4) Low Web Use	(5) No Web Use	(6) Non- User
Men	Communication	36.5% ¹	30.4%	21.7%	6.8%	4.2%	26.3%
	Download software	15.2%	8.3%	1.4%	2.5%	10.3%	8.9%
	Interactive discussion	8.5%	12.1%	2.2%	3.8%	4.2%	7.6%
	Non-Interact discussion	19.2%	9.3%	14.6%	8.7%	7.8%	13.3%
	Use another computer	16.0%	6.9%	4.3%	4.2%	3.9%	9.0%
	Real time audio or video	3.6%	0.2%	3.3%	1.7%	0%	2.1%
Segment size (men) in millions		3.85	3.28	1.89	1.29	0.66	10.97
Women	Communication	40.9%	27.9%	41.7%	16.3%	6.7%	27.9%
	Download software	9.7%	5.5%	4.0%	5.1%	4.5%	5.5%
	Interactive discussion	6.0%	11.5%	9.4%	2.8%	4.4%	7.1%
	Non-Interact discussion	23.5%	7.0%	10.1%	14.2%	9.6%	11.9%
	Use another computer	4.7%	5.8%	7.0%	0.70%	0.62%	4.2%
	Real time audio or video	4.6%	2.7%	1.7%	6.5%	0%	2.8%
Segment size (women) in millions		1.09	1.27	1.50	0.77	0.84	5.47

Table 6. Activities frequently performed by Internet and Web use segments

¹Percentage of segment that frequently performs each activity (selects top category on five-point rating scale)

tially correct for the biasing effects of low cooperation rates.

Correction for Respondent Inconsistency

The CNIDS estimates of Internet and Web use are logically inconsistent because individuals saying they used the Internet in one question only to later say they did not use the Internet were included in counts of use. Our approach to Internet estimation is conservative. We correct for individuals' inconsistency in their responses by using their responses to all questions related to a topic. Our corrections for inconsistency removed respondents providing conflicting information as to whether they actually used the Internet or the Web. The logic of our corrections for respondent inconsistency includes:

- **Internet access.** The first measure attempts to estimate an individual's access to the Internet yet provides no definition of what is meant by such access. "Do you personally have access to the Internet, either directly through an Internet service provider or an online service like America Online or by any other means?" This is a broad measure of access, subject to interpretation by the individual. As such, it provides insight into market potential, but it should not be taken as estimating the number of

individuals with actual access to the Internet in the U.S. There were no items available in the survey to correct this measure for inconsistency. As the Internet access measure includes potential as well as actual use, from a marketing perspective this measure may be analogous to aided awareness in an advertising context, tapping whether individuals have ever heard of the Internet as much as whether they actually have access to it. Thus, its interpretation must be viewed with caution.

- **Internet use within the past three months.** The second measure "Have you accessed the Internet at home, at work, or at school within the past three months?" taps Internet use in the past three months but does not define the parameters of use, other than to specify that use must have occurred within the past three months. Thus, this measure also leaves it to the individual to interpret what is meant by "accessing" the Internet in the past three months. This measure was corrected by eliminating individuals giving conflicting responses to questions dealing with reported access, or actual use, of the Internet.
- **Web use within the past three months.** This measure is more complex. Web use within the past three months was reported by CommerceNet/Nielsen [3] as specifying people who "have used

In all cases, the Project 2000 weights produce lower estimates than the Nielsen weights.

the WWW in the past three months.” However, the uncorrected measure is defined as follows: (1) people who used the Web the last time they accessed the Internet; (2) people who, if they did not use the Web the last time, use the Internet at least sometimes to access (i.e., use) the Web; or (3) people who, if they never use the Internet to access the Web, have access to (i.e., can potentially use) the Web. Thus, the uncorrected definition mixes actual use and potential use of the Web. Our corrected estimate removes both the potential users as well as inconsistent actual users by excluding respondents stating they have never accessed the Web.

- **Web purchase.** The question “Thinking about the times you have accessed the Web, have you ever used the Web to purchase a product or service?” applies a similar correction to eliminate individuals saying they have never accessed the Web from the group of people saying they have used the Web to purchase a product or service.

Results

We estimate that 28.8 million people in the U.S. who are 16 years and older have access to the Internet, 16.4 million use the Internet, 11.5 million use the

Web, and 1.51 million have used the Web to purchase something. Our emphasis is not on the absolute numbers per se but on the breakdowns that are of interest from a commercial perspective. These and other key results (see Tables 2–7) are discussed in this section. Here we highlight several of our key findings from the analysis that demonstrate the benefit of such information for decision making:

Aggregate estimates of Internet and Web access and use. To figure the effect of reweighting the sample, Table 2 reports estimates of Internet and Web access and use based on both the Nielsen and Project 2000 weights for the four measures of Internet use that were publicly released. Column (a) in Table 2 shows the original estimates released by the CommerceNet Consortium and Nielsen Media Research. These projections were derived from the sample of North Americans 16 years and older and were weighted only by gender.

The effect of reweighting the CNIDS raw data is shown by comparing columns (b) and (c) in Table 2. The CNIDS estimates for the U.S., using Nielsen’s weights, are shown in column (b) of Table 2. The corresponding estimates using the Project 2000 weights appear in column (c) in Table 2. We can also

examine the effect of reweighting by studying columns (d) and (e) in Table 2 to compare the consistency-corrected estimates of Internet and Web access and use using the Nielsen and Project 2000 weights, respectively.

In all cases, the Project 2000 weights produce lower estimates than the Nielsen weights. For example, column (b) shows that CNIDS estimates 16.9 million Americans use the Web, while the Project 2000 estimate in column (c) is that 14.1 million Americans use the Web.

To correct for

Table 7. Activities performed on the Web by Web use segments

		High Internet Use		Low Internet Use	
		(1) High Web Use	(2) Low Web Use	(4) Low Web Use	(1, 2, 4) All Web Use
Men	Very skilled at navigating Web ¹	15.6%	1.5%	2.2%	8.1%
	Navigating Web is very challenging	8.9%	8.7%	20.9%	10.6%
	Search for product/service info ²	72.2%	48.6%	42.5%	58.4%
	Made purchase based on Web info	28.5%	16.9%	6.7%	20.6%
	Search for company/org. Info	71.6%	56.3%	45.6%	61.7%
	Search for other info on Web	85.8%	66.2%	59.7%	74.2%
	Browse/explore	94.5%	86.5%	83.6%	89.7%
	See what's new at favorite web sites	75.7%	56.8%	45.7%	63.7%
	Ever used Web for business	56.6%	43.7%	36.6%	48.5%
	Segment size (men) in millions		3.85	3.28	1.29
Women	Very skilled at navigating Web	12.1%	0.9%	1.6%	5.0%
	Navigating Web is very challenging	10.3%	15.2%	27.1%	19.5%
	Search for product/service info	55.7%	33.1%	31.2%	40.5%
	Made purchase based on Web info	17.7%	7.8%	5.2%	10.6%
	Search for company/org. Info	66.5%	45.3%	25.7%	47.9%
	Search for other info on Web	82.8%	65.7%	50.7%	68.0%
	Browse/explore	90.6%	82.2%	74.5%	83.2%
	See what's new at favorite web sites	67.5%	44.1%	25.6%	47.7%
	Ever used Web for business	46.3%	29.7%	30.6%	35.7%
	Segment size (women) in millions		1.09	1.27	0.77

¹Percentage selecting top category of five-point rating scale reported for the skill and challenge items

²Percentage checking "yes" category of yes/no scale reported for the remaining seven items.

Studies purporting to estimate the size of the Internet and its potential as a commercial market must demonstrate the weights used to produce estimates are adjusted according to the best available figures on population composition.

respondent inconsistency, we examined its effect on the estimates of Internet and Web access and use by comparing columns (b) and (d) and columns (c) and (e) in Table 4. The first comparison, (b) vs. (d), shows the effect on the CNIDS estimates, using the Nielsen weights when they are consistency-corrected. The table shows that the estimates are reduced when corrected for consistency. The Project 2000 estimates, comparing (c) and (e), are similarly reduced when corrected for inconsistency. For example, the Project 2000 estimate of Internet use is 18.4 million before correction and is reduced to 16.4 million after.

In sum, because the numbers in column (e) of Table 2 are the best available estimates of Internet and Web access and use for the U.S. in 1995, we use them as the base for our subsequent analyses.

Market segments based on frequency of use. Table 3 presents five market segments based on the stated Internet and Web use of 16.4 million Internet users and includes a sixth segment of almost 183 million nonusers for comparison. We constructed these Internet and Web use segments by cross-classifying frequency categories for the two services according to the definitions in Table 3. We sought segments that would be sizable, logically consistent, and useful for business planning purposes.

Segment 1 comprises the “High Internet/High Web” users who represent 30% of all Internet users 16 years and older in the U.S. The 4.94 million individuals in this segment use the Internet once a week or more, and they say they use the Web frequently. “High Internet/Low Web” users, who use the Internet once a week or more but do not use the Web frequently, represent 28% (4.55 million) of all Internet users.

“High Internet/No Web” users are those 21% (3.39 million) who use the Internet at least once a week but never use the Web. “Low Internet/Low Web” users make up our fourth segment. These 2.06 million individuals (13% of all Internet users 16 years and older in the U.S.) use the Internet no more than a couple of times a month and use the Web infrequently. Our fifth (and final) segment defines “Low Internet/No Web” users. These infrequent users, 9% (1.51 million) of Internet users (total 101% due to rounding), are those who use the Internet no more

than a couple of times a month and do not use the Web. The sixth segment consists of the estimated 182.5 million individuals who never use the Internet or the Web.

Note that the variable used to define Web use, based on a four-point scale, is not as clearly definable as the variable used to define Internet use, based on specific itemized categories of frequency. This definition of Web use limits somewhat the interpretability of the five use segments. Future surveys should use consistent scales when measuring aspects of usage for different online applications.

Table 4 describes the five Internet and Web use segments and the one nonuse segment in terms of respondent demographic characteristics. Space does not permit an exhaustive analysis of the patterns apparent in this data. However, we observe in general that the demographics of Internet use depend on the particular use segment, so we must be careful about drawing aggregate conclusions. For example, women are 44.3% of the High Internet/No Web segment but only 22% of the High Internet/High Web segment, while the nonuser segment is more than 50% female. And while individuals 55 years and older make up 7.1% of the Low Internet/Low Web use segment, only 2.7% of individuals in this group are frequent Internet and Web users (the High Internet/High Web segment).

Table 5 shows the relationship between computer use behavior and the use segments. Not surprisingly, there are strong relationships between type of access, computer and modem ownership, length of time of computer use, and segments of Internet and Web use. Frequent Internet users who also use the Web (High Internet/High Web and High Internet/Low Web) are more likely to have both a direct connection to the Internet and access to an online commercial service rather than only an online service, while infrequent Internet users (Low Internet/Low Web and Low Internet/No Web) and the High Internet/No Web segment are more likely to have only an online service connection. The latter type of access may play a role in these segments’ infrequent use, perhaps because of the generally higher costs of using online services rather than direct access providers to access the Internet (use-based pricing vs. flat fee). We also notice that Low Internet users (segments 4 and 5) are more likely to be novice comput-

er users (less than one year).

Table 6 lists computer-related activities performed frequently by the different use segments and is further broken down by gender. For example, women are much more likely to use the Internet for communication (e.g., email) than men—an effect especially strong for the High Internet/No Web and Low Internet/Low Web segments. Additionally, High Internet/Low Web women users are much more likely to use real-time audio or video than men in this segment. In contrast, High Internet/High Web male users are more likely to download software and use a second computer than females in this same segment.

Table 7 shows a series of consumer-oriented Web activities by the different use segments, again broken down by gender. Notice that male users are more likely to search for product information, make a purchase based on Web information, and use the Web for business than female users. Perhaps not surprisingly, males report being more skilled at navigating the Web than females do, and females report being more likely to find navigation a challenge than males do. The comparison of gender differences in Tables 6 and 7 is fascinating, warranting further study.

The business value of the market segments presented in Tables 3–7 will ultimately derive from how well we understand the people in those segments.

Thus, we have presented in this article a simple analysis relating the Internet and Web use segments to the demographic characteristics of the individuals in those segments and to their computer use and consumer behaviors. These tables are only a first step in the direction of understanding the demographics of Internet use and the associated consumer behavior issues. Knowledge of such relationships not only contributes to our understanding of the Internet experience in the U.S. but is also strategically advantageous as the Internet continues its explosive commercial development.

What It Means

At this stage in the development of the Internet as a commercial medium, individuals and firms must cooperate and share information. The results we have presented here concerning Internet use among people 16 and older in the U.S. represent one step toward creating the fundamental research necessary to develop the market effectively and productively. In the emerging commercial market, many more such efforts are necessary, including development and evaluation of Web measurement standards, best business practices, further survey research, and models for measuring consumer demand in a revolutionary new medium. Achieving critical mass of the Internet

as a commercial medium, at both the firm and consumer levels, depends on such research.

We recommend that anyone evaluating the results of Internet research insist on a higher standard for such research—a standard that not only advances our understanding of the Internet experience but contributes to the population's education on this important social phenomenon. Such a standard includes the following guidelines:

- **Demand an open methodological standard.** The Internet as a medium and as an emerging market is characterized by technological complexity, rapid growth, lack of standards, unstable customer preferences, and intense competition. The Internet market is characterized by uncertainty. One way to reduce the uncertainty and improve the value of research as an aid to decision making in this environment is to adhere to an open methodological standard in the dissemination of such research. The result would be increased understanding of the numbers themselves and increased confidence in their value as adjuncts to the decision-making process. It would also permit easier evaluation and comparison of the relative merits—and limitations—of different studies. If we cannot determine how a particular result was arrived at, how can we be sure of its reliability and validity?
- **Use the best possible adjustments to sample weights.** Studies purporting to estimate the size of the Internet and its potential as a commercial market must demonstrate the weights used to produce estimates are adjusted according to the best available figures on population composition. A related issue is that researchers should employ standard demographic descriptors to ensure comparability across different surveys. In addition to age, gender, and education, region, marital status, and race are important variables known to influence Internet use that can be included to further verify the projectability of the weighted sample.
- **Report the cooperation rate.** Any survey that purports to generalize to a population should also report its cooperation rate. Low cooperation rates may suggest the presence of nonresponse bias and the need for effective adjustments to the weights and could affect the final estimates.
- **Know what is meant by “use.”** As we have shown in this article, measuring consumer use of the Internet is difficult, owing in part to the complexity of the medium, consumers' lack of understanding of the medium, and the conceptual difficulties involved in determining what it means to “use” the Internet. In order to understand an estimate of Internet use, we must understand the measures



that were used to derive the estimate—possible only if we examine the survey instrument used to collect the data.

If information on the demographic composition of the weighted sample, the cooperation rate, and the survey instrument itself is not presented along with the results, ask for it. If the provider of the information declines to provide such information, no conclusions regarding the validity of the projections should be drawn.

Conclusion

Open standards—important because they foster innovation [2]—contribute to our confidence in the results of Internet research and have the additional value of educating the industry about these important issues. All participants in the fledgling Web-based commerce industry work to raise the bar even as the industry itself, and the theory and practice of Internet measurement, evolves. Demanding a higher standard will yield benefits to the entire Web-based commerce industry [12, 14]—not the least of which is dissemination of actionable information.

An open methodological standard, easily addressed in the normal course of doing business, is appropriate, regardless of whether the research in question is proprietary in whole or in part. Such disclosure, rather than compromising the commercial value of, say, a proprietary study, actually serves as a competitive advantage to firms willing and able to demonstrate the high quality of their research.

Many of the strategic issues surrounding the commercialization of the Internet have been clouded in hype and misinformation. The time has come to raise the bar on the reporting standards for publicly released studies of who is on the Internet. Further development of the Internet, arguably one of the most exciting commercial markets in the history of modern society, depends on accurate and valid information that, at its core, helps reduce the uncertainty we face in this complex and emerging market. □

References

1. Bagozzi, P. In *Principles of Marketing Research*. Basil Blackwell, Cambridge, Mass., 1994, pp. 116–161.
2. Brynjolfsson, E., and Kemerer, C.F. Network externalities in microcomputer software: An econometric analysis of the spreadsheet market. *Management Science* (1996).
3. CommerceNet/Nielsen. The CommerceNet/Nielsen Internet Demographic Survey, Executive Summary. CommerceNet Consortium/Nielsen Media Research, Palo Alto, Calif., Oct. 1995. See www.commerce.net/information/surveys/exec_sum.html.
4. CommerceNet/Nielsen. CommerceNet/Nielsen Internet Demographic Survey, vols. 1 and 2. CommerceNet Consortium/Nielsen Media Research, Palo Alto, Calif., Nov. 1995.
5. CommerceNet/Nielsen. CommerceNet/Nielsen Internet Demographic Survey, Frequently Asked Questions. CommerceNet Consortium/Nielsen Media Research, Palo Alto,

- Calif., Dec. 1995. See www.commerce.net/information/surveys/cnet-qa3.html
6. Current Population Survey. Conducted by the Bureau of the Census for the Bureau of Labor Statistics and Technical Documentation. Prepared by Administrative and Publications Services Division, Microdata Access Branch. Bureau of the Census, Washington, D.C., March 1995.
7. CyberAtlas. CyberAtlas: Your Guide to Cyberspace Research. I/PRO, San Francisco, 1996. See www.cyberatlas.com.
8. FIND/SVP. The American Internet User Survey: New Survey Highlights, 1996. See etrg.findsvp.com/features/newinet.html.
9. Hoffman, D.L. Cyberspace to Congress: The Net is mainstream—and it votes! *MicroTimes* 148 (Mar. 4, 1996). See www.microtimes.com/prof.html.
10. Hoffman, D.L., Kalsbeek, W.D., and Novak, T.P. Internet use in the United States: 1995 baseline estimates and preliminary market segments. Project 2000 Working Paper, Apr. 12, 1996. Vanderbilt University, Nashville, Tenn. See www2000.ogsm.vanderbilt.edu/baseline/1995.Internet.estimates.html.
11. Hoffman, D.L., and Novak, T.P. Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing*, 60, 3 (July 1996), 50–68.
12. Hoffman, D.L., and Novak, T.P. A new marketing paradigm for electronic commerce. *The Information Society* (forthcoming).
13. Hoffman, D.L., and Novak, T.P. The CommerceNet/Nielsen Internet Demographics Survey: Is it representative? Project 2000 Note, Dec. 12, 1995. See www2000.ogsm.vanderbilt.edu/surveys/cn.questions.html.
14. Hoffman, D.L., and Novak, T.P. Wanted: Net census. *Wired* 2, 11 (Nov. 1994), 93–94.
15. Hoffman, D.L., Novak, T.P., and Chatterjee, P. Commercial scenarios for the Web: Opportunities and challenges. *J. Computer-Mediated Commun.* 1, 3 (Dec. 1995). See shum.huji.ac.il/jcmc/vol1/issue3/hoffman.html.
16. Kalton, G. *Elementary Survey Sampling*. Sage, New York, 1983. Section 10.
17. Kalton, G. Compensating for missing survey data. Res. Rep. Survey Research Center, Institute for Social Research, Univ. of Michigan, Ann Arbor, Mich., 1983.
18. Kish, L. *Survey Sampling*. Wiley, New York, 1983.
19. Lessler, J.T., and Kalsbeek, W.D. *Nonsampling Error in Surveys*. Wiley, New York, 1992.
20. O'Reilly/Trish. *Defining the Internet Opportunity*. O'Reilly & Associates/Trish Information Services, Sebastopol, Calif., 1995.
21. Times Mirror. *Technology in the American Household: Americans Going Online*. Times Mirror Center for the People and the Press, Washington, D.C. (now the Pew Research Center). Oct. 16, 1995.
22. Yankelovich Partners. *The Yankelovich Cybercitizen Report*. Yankelovich Partners, Atlanta, 1995.

DONNA L. HOFFMAN (hoffman@colette.ogsm.vanderbilt.edu) is an associate professor in the Owen Graduate School of Management at Vanderbilt University and co-director of Project 2000, a sponsored research center (see www2000.ogsm.vanderbilt.edu).

WILLIAM D. KALSBECK (kalsbeek.sru@mhs.unc.edu) is an associate professor in the Survey Research Unit in the Department of Biostatistics at the University of North Carolina at Chapel Hill.

THOMAS P. NOVAK (novak@moe.ogsm.vanderbilt.edu) is an associate professor in the Owen Graduate School of Management at Vanderbilt University and co-director of Project 2000.

Permission to make digital/hard copy of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.