CLASH OF THE TITANS: DOES INTERNET USE REDUCE TELEVISION VIEWING?

Stan J. Liebowitz and Alejandro Zentner*

Abstract—We examine the impact of the Internet on the leading American recreation activity: watching television. We run a panel regression using television viewing, Internet penetration, and socioeconomic variables for a large number of American cities starting before the birth of the Web. We find that the Internet's effect on television viewing varies by age group, reducing it by a moderate amount for the youngest Americans but having no impact on the viewing of the oldest Americans. We hypothesize that the overall effect is likely to increase over time as older age groups have more experience with the Internet's recreational opportunities.

I. Introduction

THE Internet is changing the way that individuals work, communicate, and recreate. As a recreational activity, it is but one in a long line of technologies—movies, radio, and television—that have altered the daily lives and behavior of ordinary citizens. Although the Internet's impact on retailing, digital distribution, supply management, and inventory enhancement have all been examined in some detail (Brown & Goolsbee, 2002; Brynjolfsson & Smith, 2000; Goolsbee, 2000; Liebowitz, 2002; Shapiro & Varian, 1999; Zentner, 2008), its impact on entertainment activities has received limited academic attention.

Yet the Internet is increasingly important for both traditional entertainment activities and novel entertainment activities such as reading blogs, joining social network sites, viewing amateur videos, and participating in online gaming worlds. The Internet has shown that it has a superb and improving capacity to entertain and inform consumers.

All of these entertainment activities share at least one common trait: they help to absorb the leisure time of the individuals engaged in these activities. The ramifications of the Internet's entertainment impact are important for many entertainment industries: live concerts, magazines and newspapers, traditional television broadcasters, cable television operators, amusement parks, travel, and many other industries that compete for the consumer's entertainment and leisure dollar. But the extent to which Internet-based entertainment activities have influenced traditional leisure activities has received scant study.

In this paper, we examine how the Internet might alter the single most important recreation activity of Americans: television viewing. Americans average between three and four hours of television viewing per day, with the average household watching television for more than eight hours per day. Although the impact of the Internet on television usage is not necessarily the same as its impact on other entertainment activities, its impact on television is likely to provide clues about the impact it will have on other off-line entertainment activities. Further, although we cannot be certain that the impact of the Internet that has occurred to this point necessarily predicts the direction of its future impact, our expectation is that the Internet is now mature enough that its previous and current impacts are harbingers of its future impact.

Our approach in this paper is to use panel data on television viewing habits in various geographic areas over time, going back to the beginning of Web browsing. Using viewing data broken down by the age of the viewer in conjunction with data on Internet use, education, and income, also broken down by the age of the viewer, we are able to measure the extent to which Internet use affects television viewing for various categories of television users.

Our findings, in short, are that the Internet has a moderate negative impact on the time younger viewers spent using television, a smaller negative impact on middle-age viewers, and virtually no impact on the time spent viewing television by the oldest cohort of individuals.

The outline of the remainder of the paper is as follows. The next section studies the possible impacts of Internet use on television viewing. The very small amount of research studying television viewing in connection with the Internet is summarized in section III. Section IV provides background metrics on television viewing and Internet use. In section V we describe our methodology and the data. Section VI lays out the model and the empirical strategy. Section VII presents the estimation results. The last section offers some conclusions.

II. Possible Impacts of Internet Use on Television Viewing

In a general manner, any time-consuming activity is a constraint on any other time-consuming activity because of the fixed amount of time in a day. To the extent that most individuals have partitions of time allocated for certain purposes, the degree of substitutability regarding the nature of the time-consuming activities should play an important role in how time spent in one activity constrains the time spent in other activities. Therefore, time spent on the Internet might not constrain the time spent on another foreground activity such as television. The reason is that television viewing, which is almost completely entertainment oriented, is not a substitute for many of the nonentertainment-oriented activities allowed by the Internet, such as conducting academic research, making airline reservations, purchasing

Received for publication August 2, 2009. Revision accepted for publication July 21, 2010.

* Liebowitz and Zentner: University of Texas at Dallas School of Management.

We would like to thank Octavian Carare for his comments, the referees and the editor Philippe Aghion who helped to improve the paper, conference participants at Telecom ParisTech ICT, and the Center for the Analysis of Property Rights and Innovation for financial support.

The Review of Economics and Statistics, February 2012, 94(1): 234–245

automobiles, or checking stock market returns. If the Internet was used only to engage in these productivity-enhancing activities, we would not expect it to substitute very much for entertainment activities such as television viewing.

Of course, the Internet also allows individuals to while away the hours watching YouTube videos or playing games in virtual worlds. These Internet-based entertainment activities are very likely going to impinge on the time spent viewing television. If individuals devoted a fixed amount of time to entertainment activities, there would be a one-to-one substitution of Internet entertainment for other forms of entertainment. Nevertheless, if, as seems likely, the relative time spent on entertainment and productivity is not fixed, then increased time spent on Internet entertainment would not come solely at the expense of other forms of entertainment.

Complications arise when the implicit quality of any block of time spent consuming entertainment changes. Such an increase in quality occurs when people switch to Internet entertainment from other forms of entertainment because the choice of Internet entertainment over alternative forms implies that they perceive those new minutes of Internet entertainment to be superior to the replaced minutes from other forms of entertainment (relative to the costs of each). Changes in the price-adjusted quality of entertainment obviously could have an impact on the time individuals spend being entertained.

On the one hand, if the quality enhancement in entertainment minutes raises the value of the marginal minute separating entertainment from productive activities, it might very well lead to an increase in time devoted to entertainment and a decrease in time devoted to productivity. On the other hand, if the entertainment quality enhancement increases the value of inframarginal entertainment minutes, this might decrease the value of the marginal entertainment minutes (since satiation comes more rapidly) and hence might decrease the time spent being entertained.

A final complication is the rapid evolution of the Internet since the birth of the Web. As connection speeds increase and the quality and variety of Internet Web sites increase, it is unclear whether individuals will increase or decrease their time spent using the Internet. Further, if early adopters of the Internet have different (entertainment) uses for it than later adopters, this will change average use as additional adopters begin to use the Internet. In spite of these uncertainties, we expect, and do find, that the percentage of the population using the Internet has increased as these Internet improvements have taken place.

This admittedly brief examination should make it clear that the expected impact of additional Internet penetration on television viewing, especially when our measurements cannot separate productive from entertainment uses of the Internet, is unclear. Nevertheless, the basic intuition that has often been applied to such a question is that Internet use will take time away from other activities, including television viewing, and this basic intuition has been supported by the limited empirical literature that has broached the subject.

III. The Literature on Television Viewing and Internet Use

The very small academic literature on this topic has provided only cursory measurements of the impact of the Internet on television viewing. One type of examination has been surveys of Internet users asking respondents about changes in their behavior brought about by the Internet. Although useful and informative, these surveys have been limited by problems typically associated with surveys: first, the ability or inability of respondents to accurately remember various details of their behavior and concerns about whether respondents are representative of the population as a whole. These surveys also tend to be wide but shallow, often asking questions on many topics but providing little detail on any particular topic.

With regard to television viewing, the *Digital Future Report* (Center for the Digital Future, 2007) found that 35.5% of Internet users said they watched less television because of the Internet (virtually none reported watching more), which would seem to put an upper limit of 35% on the diversion of television time by Internet use if these users were typical television viewers who stopped watching completely. An analysis by Nie et al. (2005) concluded that Internet users reduced their daily viewing of TV by 30 minutes. Television use for Internet users in their sample averaged 102 minutes per day (considerably less television than typical Americans watch), so this decline in viewing would imply an Internet-caused decrease in viewing of 23%.

Liebowitz (2008), while engaged in measuring the impact of file sharing on record sales, also briefly examined the impact of the Internet on radio and television use in an attempt to determine whether the entertainment impact of the Internet might be expected to have an important influence on record sales. He concluded that Internet use reduced television viewing by 10% to 15% and that Internet use reduced radio listening by a somewhat smaller amount.[1]

The small amount of prior academic research therefore, indicates that Internet use by individuals was associated with decreases in their television use. This study hopes to improve on previous work by using a much richer panel data set on television viewing for a number of cities, with a breakdown of viewing by age and sex and covering a longer time period that goes back to the beginning of the Web.

IV. Background Metrics

Measurement of media use is not an easy task. Take, for example, the case of Internet use. Surveys can ask respondents how much time they spend using the Internet. Alternatively, firms engaged in measuring Internet use (for example, ComScore Media/Metrix) can install in computers

[1] A related but different question is the study of the extent to which watching television shows on the Internet (such as YouTube) displaces conventional television viewing. Waldfogel (2009) studies this question and finds evidence of modest substitution between the two outlets.

Table 1.—Pew Survey on Internet Use, December 2007

Age Group	(1) % of Population Using At All	(2) % of Users Using Yesterday	(3) % of Population Using Yesterday
18–29	92%	77%	71%
30–49	85	77	65
50–64	72	66	48
65+	37	56	21

Table 2.—Monthly Time Spent on Activity among Users and among the Entire Population (Q1 2009, Nielsen, H:M)

Age	(1) Internet	(2) Television[a]	(3) Internet Use, Entire Population
18–24	14:19	122:34	13:10
25–34	31:37	156:21	27:58
35–44	42:35	163:11	36:11
45–54	39:27	185:22	30:58
55–64	35:49	204:15	25:47
65+	28:34	215:20	10:34

[a]The television viewing values include time shifting.

of participating respondents metering programs that examine whether the Web is being accessed. One complication with metering programs is that the computer might be used and a Web browser might be open, but the real activity taking place might be the use of a word processor in another window. To avoid these problems, the monitoring programs normally focus on only the active window, but even here measurement is imprecise. A computer operator might walk away from the computer for several hours (or all night), and this might cause severe problems with the measurement system. To solve this problem, some companies with programs monitoring computer use, such as ComScore, stop counting after several minutes of inactivity. However, this creates another problem if the computer is being used for passive entertainment. If a user watches a long video or listens to Internet radio, there is no apparent activity (such as keystrokes), so comScore will seriously undercount time spent on these activities. This type of measuring algorithm will possibly miss many hours of Internet entertainment activities.

These difficulties are similar to problems that arise in measuring television viewing since the TV can be turned on but no one is watching. The latest generation of people meters, devices attached to the TV that monitor its use, try to ensure that the individual is in the room with the television on when counting viewing time, although the movement to electronic meters has been extremely contentious for many years.[2] For most American cities, however, television viewing is still measured by surveys (diaries).

A survey by the Pew Internet and American Life Project asks about Internet use (at home or at work) and breaks down the answers by age group. Table 1 provides some results where users are asked first if they use the Internet "at all" (column 1) and then, if they do use the Internet, whether they used the Internet the previous day (column 2). Column 3 is the product of column 1 and column 2 and represents the share of the total population, by age group, that used the Internet the previous day. Because frequency of use is probably related to total time using the Internet, this right-most column is probably most closely related to total use.

All three columns show a drop-off in Internet use for older age groups, but the overall drop-off is most pronounced in column 3—the share of the overall population using the Internet on the prior day. The two younger age groups are not much different from each other, but there is a moderately large drop-off in the over-50 age group and a severe drop-off in the over-65 category.

Another Internet measurement company is Nielsen Online, a sibling of the U.S. television monitoring giant, Nielsen Media. Nielsen Online has broken down the data on monthly Internet use and monthly television use by age group, as shown in table 2.[3] The Internet numbers are based on use of Internet users only and include both work and home use. The first column of data provides time spent using the Internet, and the second column provides time spent viewing television. It is clear that television viewing is still a much more heavily used activity than is the Internet, particularly given that the television universe consists of a larger percentage of individuals than the Internet universe.[4] Television use increases monotonically as individuals get older.

The Internet use numbers in table 2 (column 1), although not directly comparable to the Pew numbers in table 1, appear to tell a different story about intensity of use than the Pew numbers and seem inconsistent with our intuition. Time spent on the Internet (measured only for Internet users) does not vary monotonically with age in column 1 of table 2, unlike the numbers in table 1. According to table 2, Internet users between 35 and 44 years of age (and even 65) spend three (two) times as much time online as do 18- to 24-year olds, although this latter group might have been thought to be the prime Internet use group. One might argue that including work-based Internet use together with home-based Internet use is distorting these results, but many individuals over age 65 are likely to be retired and yet they too show considerably greater Internet use than young individuals (and the Pew numbers, which also include use of the Internet at work, show a more intuitive result). Although some of this puzzle is explained by the different populations on which these number are built (as we show momentarily), it is also possi-

[2] For discussion of the people meter controversy, see Snyder (1991) and Weiss (2001).

[3] See table 3 at http://it.nielsen.com/site/documents/A2M2_3Screens_1Q09_FINAL.pdf.

[4] Note that if these numbers are to be believed, Internet use is rarely above 20% of TV use. Because these data include Internet use at work, which is not likely a substitute for television viewing, these numbers indicate that the Internet's negative impact on TV viewing would be less than 20% if each minute of Internet use came from time otherwise spent viewing television.

ble that the problems measuring Internet use are at least in part responsible for this counterintuitive result.

We can roughly attempt to adjust the values in column 1 of table 2, based only on Internet users, to represent values based on the entire population by applying the share of the population using the Internet values from column 1 of table 1 to the Internet hours used in column 1 of table 2.[5] The result of this calculation is found in column 3 of table 2, labeled "Internet Use, Entire Population." Here we find a more reasonable result that those over age 65 have the least Internet use, although the fairly low Internet use reported by Nielsen for the 18–24 age group is still surprising.

The 2003 Census data that we use in our analysis do not provide statistics on use intensity, but they do provide Internet penetration rates by age group. Although we use different age groupings in our main analysis, if we arrange the Census data to match the Pew age groups we get the following penetration rates: 54% (18–29), 58% (30–49), 51% (50–64), and 23% (65+). These numbers provide a pattern very similar to that found in the Pew numbers (although the overall values are lower, which is not surprising given that Internet penetration was lower in 2003 than in 2007). The two youngest age groups have the highest Internet penetration, and the values are similar to each other (although the ordering is reversed from the Pew ordering). There is a drop-off to the third age group and then a much larger drop-off for the oldest age group of those over 65 years of age, similar to the Pew data.

Our own interpretation of the data, influenced by our intuition, is that younger individuals are more likely to have higher Internet intensity and use than older individuals, as indicated by the Pew and Census data, and we will retain this assumption throughout the paper. Independent of the question of use by age, all data sources seem to indicate that many individuals use the Internet and that those using the Internet use it frequently.

V. Methodology and Data

Our econometric investigation requires the merging of several data sets. First, Nielsen Media Research (not to be confused with Nielsen NetRatings) sells data on the amount of television viewing in DMAs (designated market areas, which are constructed to measure television viewing and advertising markets) subdivided into age groups and gender. We procured this data covering the period 1990 through 2006.

Second, the U.S. Census, as part of its Current Population Survey (CPS) undertaken for the Bureau of Labor Statistics, conducts surveys on Internet and computer use. These surveys, conducted six times between 1994 and 2003, provide information on the penetration of home Internet; the type of Internet connection; household income; Metropolitan Statistical Area (MSA) of residence; the age, sex, race, and education of the respondent; as well as a host of other variables. This information is based on responses from approximately 130,000 individuals. Because we have data on each of these individuals, we can create variables by MSA (DMA), such as Internet penetration or education level, for subsets of the population based on age. Unfortunately, the size of the sample in smaller MSAs is sometimes insufficient to allow great confidence in the measurements for various slices of demographic data.

Combining these data sets is not a trivial task.[6] In highly populated metropolitan areas, television-based DMAs are almost always larger than Census-based MSAs, and the television-advertising market, measured by the Nielsen DMA, often covers many Census-Based MSAs. DMAs also include rural areas, whereas MSAs do not, so that the Census data based on MSAs leave out a portion of the DMA population. This basic methodology of combining Census data with Nielsen data was undertaken in Liebowitz (2008) and originally came from Boorstin (2004).

Our procedure consists of either matching a Census MSA to a DMA when a DMA contained only one MSA or combining the data from constituent Census MSAs into the DMA in which they belonged. As described in the data appendix, we try to note instances where the (summed) populations of the MSAs assigned to a DMA indicated a potentially poor match to the DMA population and therefore might not properly reflect the economic characteristics of the DMA. A variable named "coverage" measures the percentage difference between the DMA population and the populations constructed from underlying MSAs.

The Nielsen viewing data for these DMAs come from diary entries and set-top meters. For many smaller DMAs, diaries are the only method used to collect data. For larger DMAs, the diary data are merged with electronic set-top meter data when available (20,000 of these meters were in use as of 2006). Set-top meters are fairly primitive and merely report that the television is on and to which channel it is tuned; they do not indicate who, if anyone, is watching. Since 2002, more advanced meters, known as people meters, which provide information on who is watching television, have been rolled out in the largest cities, but as of

[5] The process used to create column 3 of table 2 consisted of finding the product of Pew's Internet penetration variable "% of Population Using at all" from table 1 and Nielsen's average Internet use (in hours per month) in table 2. Because the Pew and Nielsen age groupings do not exactly match up, some of the Nielsen age groups required combining information from more than one Pew age group. This was done by assuming a uniform distribution of users by age and that the average value of the group applied to individuals of all ages in the group. Then we weighted the Pew values by the number of years included in the Nielsen age category. For example, the Nielsen 25–34 category consisted of five years of the Pew 18–29 group and five years of the Pew 30–49 group; thus, the viewing time was a pure average of those two groups. When a Nielsen category was entirely contained within the Pew category, the Pew value was used.

[6] Among other issues, the Census Internet variables change over time, although we tried to use the most consistent definitions. A complete description of variables used in the analysis is described in the data appendix.

2006, only the ten largest DMAs had these people meters providing local data. Diary data are the old standby in the Nielsen stable of measuring instruments and are still used in most cities to rate television at the local level and to set local advertising prices. These ratings come during quarterly sweeps periods during which approximately half a million households participate in each of the quarterly month-long sweeps. These forms of local data collection are in contrast to a much smaller sample (fewer than 10,000 people meters) that Nielsen uses to create a national sample whose results are often trumpeted in press releases discussing national viewing habits or overnight ratings.[7]

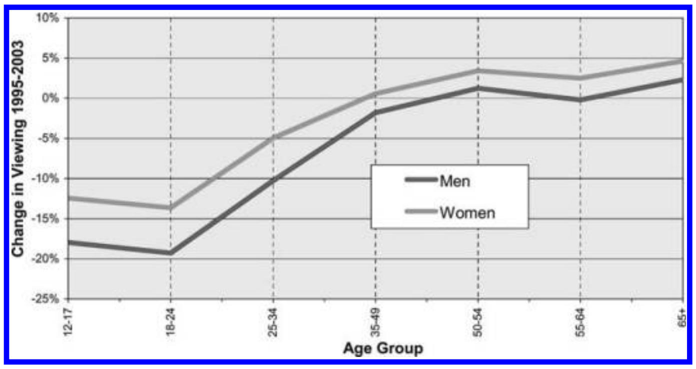
Television viewing, according to the Nielsen data, has remained remarkably constant over the seventeen-year period beginning in 1990, with an overall decline in the average hours viewed during that period of 1.4%. For the period 1995 to 2003, the one covered by our regressions, the decline was somewhat larger, though still small, at 2.2%. Nor did this small overall change mask a temporal change of any substance: the largest deviation from the mean television rating over the seventeen-year period was 6% in 2005, when the national rating dropped to 19.3 compared to the seventeen-year average of 20.5.[8] The consistency of these measurements in television viewing is somewhat surprising given the period's rapid advancement of cable and satellite television, which might be thought likely to increase viewing time, as well as other technological changes leading to alternative forms of entertainment such as the Internet.

Needless to say, with the large changes taking place in these underlying factors, the placid facade of overall viewing habits hides some underlying turbulence. Figure 1 shows the changes in television viewing by age and gender from 1995 through 2003.

Figure 1 reveals a less tranquil change in viewing habits than indicated by the overall viewing numbers. Both genders showed a consistent pattern of younger viewers decreasing their viewership, whereas older individuals were keeping their viewership constant or increasing it slightly. Men in all age groups appear to have a somewhat more negative change in viewing compared to women.

The results from figure 1 provide some preliminary clues consistent with the hypothesis that the Internet may be having a negative impact on viewing, since young people are thought to be the heaviest Internet users (and they presumably use it more for entertainment), and they show the largest falloff in television viewing. But this figure cannot tell us whether there is a link between Internet use and television viewing. To better approximate an answer to that question, we turn to multiple regression analysis, looking at the relationship between variation in Internet penetration and television viewing in cities while controlling for demographic differences and city and time fixed effects.

FIGURE 1.—VIEWING CHANGE BY AGE AND GENDER, 1995–2003



VI. Model and Empirical Strategy

We discussed in section II that the direction of the impact of Internet use on television viewing is unclear and not as straightforward as it might appear. We now lay out the empirical model used in our econometric investigation of this question.

An individual, represented by i, watches television, uses the Internet, and devotes time to work and leisure activities. Formally, on day t and in city j, individual i watches T_{ijt} hours of television, uses the Internet for I_{ijt} hours, and employs R_{ijt} hours in alternative work or leisure activities:

$$24 = T_{ijt} + I_{ijt} + R_{ijt}.$$

Summing across individuals in city j and dividing by the population we get,

$$24 = \frac{\sum_{i=1}^{N_j} T_{ijt}}{N_j} + \frac{\sum_{i=1}^{N_j} I_{ijt}}{N_j} + \frac{\sum_{i=1}^{N_j} R_{ijt}}{N_j}.$$

Rearranging terms, we write the average number of hours of television viewing in city j in terms of the total time endowment minus the average number of hours of Internet use and the average number of hours people devote to other activities:

$$\frac{\sum_{i=1}^{N_j} T_{ijt}}{N_j} = 24 - \frac{\sum_{i=1}^{N_j} I_{ijt}}{N_j} - \frac{\sum_{i=1}^{N_j} R_{ijt}}{N_j}.$$

[7] As a practical matter, there is no alternative source of local viewing data, although we note that the results of the two samples are not entirely consistent with one another. According to the National Nielsen Media statistics, time spent viewing television increased slowly but fairly smoothly for every age category from 1995 to 2005 (an increase of about 13% in total). Since the local market data generate the majority of revenue for Nielsen, are based on a fairly consistent methodology throughout the multiyear period we examine, are based on a much larger sample, and seem intuitively more plausible, we think that use of the diary data is appropriate.

[8] Ratings measure the percentage of a viewing day (defined as 18 hours in the data we use) that is taken up by television viewing, so that a rating of 20 indicates average daily viewing of 3.6 hours (20% × 18 hours).

TABLE 3.—SUMMARY STATISTICS

		1995	1997	1998	2000	2001	2003
Television Viewing (minutes per day)	Mean	215.0	210.6	208.5	206.6	201.1	206.0
	SD	21.4	23.6	23.8	28.3	27.2	29.7
Internet penetration household	Mean	NA	0.21	0.30	0.46	0.57	0.61
(fraction of the population)	SD	NA	0.08	0.09	0.10	0.09	0.09
Household income (dollars per year)	Mean	34,930	37,471	39,881	41,417	42,500	44,940
	SD	7,727	7,751	7,281	8,092	7,742	9,707
Education (ordinal one to five)	Mean	2.12	2.15	2.18	2.20	2.21	2.26
	SD	0.28	0.26	0.25	0.26	0.24	0.28
Black (percentage)	Mean	12.5%	12.7%	12.8%	13.0%	13.7%	12.2%
	SD	0.12	0.13	0.13	0.13	0.13	0.12
Hispanic (percentage)	Mean	8.7%	9.8%	9.7%	10.5%	10.9%	11.7%
	SD	0.17	0.17	0.17	0.18	0.17	0.18
Viewing ages 6–34 (minutes per day)	Mean	174.5	167.5	163.9	155.7	148.3	149.0
Viewing ages 35–54 (minutes per day)	Mean	214.9	210.7	209.0	210.1	207.6	214.3
Viewing ages 55+ (minutes per day)	Mean	317.5	313.8	312.4	316.5	308.1	316.7

The information on total television viewing and on television viewing by age groups comes from Nielsen. All the other variables are from the Current Population Survey by the Census. See the data appendix for a detailed definition of each variable.

The average number of hours of Internet use in city j has increased significantly over the past decade. This is because more people use the Internet and because Internet users spend more hours online. We aim to discover whether this has resulted in a reduction in the average number of hours of television viewing or in a reduction in the time spent in other work or leisure activities.

We start our statistical analysis by using cross-sectional variation in television viewing and Internet use; we study if the average number of hours of television watched per day in city j in a given time, T_j,

$$T_j = \alpha^1 + \beta^1 X_j + \psi^1 I_j + u_j^1, \tag{1}$$

is correlated with the average number of hours of Internet use per day in that time, I_j. In equation (1), X_j represents other covariates for city j, and u_j^1 represents the error, which accounts for the effect on television viewing of all the unobservable alternative work and leisure activities that are included in R_j.

Cross-section regressions estimate the effect of Internet use on television viewing by comparing cities with low Internet use and high Internet use. But the level of Internet use is unlikely to be randomly assigned across cities conditional on observables characteristics. Panel data variation provides a far superior approach because it controls for time-invariant city unobservable characteristics that may be correlated with the level of Internet use. The average number of hours of television viewing per day in city j and time t, T_{jt}, is represented by

$$T_{jt} = \alpha^2 + \beta^2 X_{jt} + \psi^2 I_{jt} + \beta_t dummy_t + \beta_j dummy_j + u_{jt}^2, \tag{2}$$

where α^2 is a constant, X_{jt} is a set of variables with time variation within cities, β_t is a time t fixed effect (which is common for all DMAs), $dummy_t$ is an indicator variable equal to 1 in time t, β_j represents a city-specific and time-invariant effect, $dummy_j$ is a an indicator variable equal to 1 for city j, ψ^2 measures the effect of increasing the average number of hours of Internet use in a city on the average number of hours of television viewing in that city, and u_{jt}^2 is the error. This regression compares how television viewing in cities decays as the number of hours of Internet use increases.

We do not observe the average number of hours of Internet use at the city level, but we observe their level of Internet penetration. We propose a linear relationship between the number of hours of Internet use and Internet penetration, P_{jt}:

$$\frac{\sum_{i=1}^{N_j} I_{ijt}}{N_j} = I_{jt} = aP_{jt}. \tag{3}$$

Substituting equation (3) into equations (1) and (2), we get

$$T_j = \alpha^1 + \beta^1 X_j + \psi^1 aP_j + u_j^1 \tag{1'}$$

and

$$T_{jt} = \alpha^2 + \beta^2 X_{jt} + \psi^2 aP_{jt} + \beta_t dummy_t + \beta_j dummy_j + u_{jt}^2, \tag{2'}$$

where the term ψa indicates the impact of an increase in Internet penetration on television viewing. Equations (1′) and (2′) are the econometric models that we use in the following section.

VII. Econometric Results

Our main interest is to use regression analysis to measure the impact of the Internet on television viewing. Before we get to the regressions, however, it is useful to examine some characteristics of the data in table 3. The statistics in this table are averaged across DMAs. We use 1995, as the initial year because the Census changed geographic definitions of

TABLE 4.—YEARLY OLS REGRESSIONS ON AVERAGE MINUTES OF TELEVISION VIEWING

	1997	1998	2000	2001	2003
Internet penetration	−132.3	−68.3	−47.0	−185.1***	−203.7**
	(80.6)	(61.4)	(73.1)	(67.0)	(79.4)
Percentage black	164.3***	167.5***	160.3***	122.2***	132.7***
	(42.6)	(41.3)	(46.3)	(38.3)	(41.5)
Percentage Hispanic	84.8***	75.7***	63.3**	38.4	53.4*
	(29.2)	(26.9)	(28.6)	(24.2)	(28.3)
Median income per member of the household	0.500	−0.100	2.500	0.300*	1.100
	(2.3)	(1.7)	(1.7)	(1.7)	(1.4)
Education	33.0	19.8	−34.9	−9.5	26.7
	(32.5)	(29.1)	(31.7)	(28.2)	(34.2)
Constant	149.6***	174.5***	259.9***	282.2***	248.7***
	(48.5)	(49.0)	(46.7)	(48.8)	(61.1)
Observations	82	82	82	82	82
R^2	0.38	0.42	0.36	0.40	0.34

The regressions use DMAs with coverage ratios above 70% and are weighted by the number of observations in the CPS in the corresponding year. Robust standard errors in parentheses. *Significant at 10%, **5%, ***1%. See table 3 for units of measurement and summary statistics of each variable and the data appendix for detailed definitions.

MSAs in 1995, and for that reason we were unable to match the MSAs to DMAs in years prior to 1995.

Overall, average individual viewing (using DMAs as the unit of analysis) fell slightly from 1995 through 2003, which is consistent with the small changes in average results already reported from 1990 through 2006. Table 3 shows a decay in television viewing for the youngest age group while viewing remained fairly constant for the two older age groups.

The large increase in Internet use that we know occurred during this time period is reflected in the changes in average home penetration rate across DMAs. There is no value given for Internet penetration in 1995 since the Census did not collect that statistic in 1995, which undoubtedly reflected the newness of Internet use. Other sources, however, indicate that the 1995 figure was quite low, perhaps in the vicinity of 5% of the population.[9] The nature of Internet access was far different then than it is now, particularly in the realm of entertainment. Slow dial-up speeds (topping out at 28.8K) restricted the ability to listen to streaming music, and streaming video was all but impossible.[10] The major ISPs (AOL, CompuServe) were just in the process of opening up to the Web, allowing users to crawl over the ISP's walled gardens of self-contained content, self-contained e-mail, and self-contained limitations. Although the future was bright, it seems fair to treat this year as part of the dawn of the Internet age, at least in terms of entertainment.[11]

According to table 3, nominal household income rose, although at an uneven pace, during this period. The education variable, an ordinal construct ranging from 1 to 5 to cover education levels from less than a high school diploma to a college degree or above, indicates a slight increase in education levels over this period of time.[12] Also, there was an increase in the share of the population comprising Hispanic individuals while the share of the population comprising by black Americans was largely constant.[13] Finally, there are important differences in time spent viewing television by age group, with older individuals watching more television.

As a first step in the analysis, we ran yearly regressions. Because there are important differences among cities that the limited control variables at our disposal are unlikely to fully capture, we do not expect these pure cross-section regressions to provide the most reliable results.

In addition to the Internet penetration variable, we included variables on education, median income per member of the household, and minority status since it is well known that black Americans watch considerably more television than the population at large does.[14] Industry literature on television viewing reports that television viewing is negatively related to income, although it is not clear whether this is true when minority status is being held constant.[15] The regressions in table 4 are limited to DMAs where the coverage ratios are greater than 70% (meaning that the MSA populations for which we had data were at least 70% of the overall DMA population), since using DMAs with lower coverage ratios risks polluting observations with Internet penetration ratios and other socioeconomic data that do not match the true population values. This leads to 82 observations. These regressions were weighted by the number of sample observations used to cre-

[9] A claim that 6.7% of the population had Internet access in 1995 can be found at http://web.archive.org/web/20060209014044/www.nua.com/surveys/how_many_online/n_america.html.

[10] The 28.8 top modem speed is based on modem advertisements in issues of *PC Magazine* from late 1995.

[11] We realize that text-based computer games have been played since before PCs (*Star Trek, Spacewar!*) came into existence and do not wish to disparage the entertainment value of these games or offend anyone who might have wasted far too much time playing them. Nevertheless, we refer in the text to activities that would appeal to large segments of the population.

[12] The values: less than high school diploma (1), high school graduates (2), some college (3), associate degree (4), and bachelor's degree or higher (5).

[13] The relatively larger changes in the percentage of black population in 2001 and 2003 might be explained by sampling error since the sample on which it is based are likely to contain very few observations in DMAs with small black populations.

[14] According to recent Nielsen statistics, blacks watch 41% more television than the average, whereas individual Hispanics watch only 88% of the average amount. See Nielsen Media Research (2007).

[15] Nielsen reports that households watch less television as income rises. See Nielsen Media Research (2007).

TABLE 5.—FIXED EFFECTS REGRESSIONS ON AVERAGE MINUTES OF TELEVISION VIEWING, 1995, 1997, 1998, 2000, 2001, 2003

	All Ages		Ages 6–34		Ages 35–54		Ages 55+	
	cov>0.7	all DMAs	cov>0.7	all DMAs	cov>0.7	all DMAs	cov>0.7	all DMAs
Internet penetration	−31.6**	−22.0**	−32.8***	−23.6***	−18.7	−11.5	−6.1	−2.7
	(13.4)	(9.9)	(11.1)	(7.9)	(12.9)	(9.4)	(20.7)	(14.8)
Percentage black	−19.1	−11.2	−0.8	3.5	−19.0	−8.4	−10.7	−11.1
	(27.8)	(17.2)	(16.9)	(10.4)	(20.0)	(13.5)	(31.2)	(21.3)
Percentage Hispanic	−15.8	2.4	−17.2	−2.4	−0.2	5.7	3.2	17.8
	(25.1)	(19.6)	(15.8)	(13.0)	(22.9)	(16.0)	(32.2)	(24.0)
Median income per member of the household	−0.200	0.000	0.300	0.400	−0.400	−0.100	0.300	0.400
	(0.5)	(0.3)	(0.4)	(0.3)	(0.4)	(0.3)	(0.5)	(0.3)
Education	5.1	5.2	1.3	0.8	6.7	5.8*	−8.7	−6.6
	(7.4)	(4.9)	(4.1)	(2.6)	(5.2)	(3.3)	(6.4)	(4.0)
Constant	246.9***	228.7***	188.6***	175.7***	232.9***	220.1***	356.1***	344.9***
	(20.2)	(13.5)	(11.8)	(8.8)	(16.7)	(10.5)	(17.9)	(11.0)
Observations	492	852	492	852	492	852	492	852
R^2	0.92	0.92	0.93	0.93	0.90	0.90	0.84	0.86

The regressions include years and DMA fixed effects and are weighted by the number of observations in the CPS in the year 2000. Robust standard errors in parentheses. *Significant at 10%, **5%, ***1%. See table 3 for units of measurement and summary statistics of each variable and the data appendix for detailed definitions.

ate the Census (MSA)-based variables in each DMA since we were warned by representatives of the Census, and simple logic implies, that information from small MSAs (particularly for subgroups of the population) is less reliable than is the case for large MSAs.[16]

The results in table 4 indicate that the Internet penetration coefficient is negative in all years, although it is statistically significant only in 2001 and 2003. The coefficients are very large in these two years, as well as in 1997. The coefficient for 2003 indicates, for example, that cities would have experienced a decrease in average television viewing time of 20.4 minutes for each 10 percentage point increase in Internet penetration rate. Given that Internet penetration rates average 61% in 2003, an extrapolation of the coefficient from 0 to the 2003 level would be expected to lead to a reduction in viewing that is more than half (~56%) of the level of viewing in 1995. It seems rather far-fetched to believe that the Internet has led to this size of decline while viewing has remained relatively stable, since this would imply that a very large and heretofore uncommented on countervailing factor must have been at work. A countervailing factor of this size seems unlikely to have escaped discussion in the television literature. Fortunately, the more reliable fixed-effects regression will have more reasonable measures of the impact of the Internet on viewing.

The most consistently significant variables in these cross-section regressions are the minority group variables. The size of the black population coefficient implies that daily television viewing in a city that was 100% black would be about 150 minutes more than in a city with no black population. Of course, this is an out-of-sample extrapolation since a city with a 100% black population does not exist in the data. The size of the impact, though, overstates somewhat the average extra television viewing of blacks, which is about 130 minutes, but, of course, the regression coefficient picks up other city characteristics that are found in cities with a higher share of blacks. Cities with higher shares of Hispanics also appear to have greater average television viewing even though unconditional statistics on viewing indicate that the average Hispanic watches somewhat less television than nonminority viewers.

Income has a very slight hint of having a positive impact on viewing in the later years, but education has no consistent impact on viewership.

We turn now to fixed-effects regressions using the panel data. Table 5 reports the results by pairs of columns, with the first column of each pair consisting of regressions based on just those DMAs with coverage above 70% and the second column based on the entire set of DMAs. The first pair of columns, labeled "All ages," performs fixed-effects regressions for all individuals regardless of age. The later pairs of columns perform these regressions based only on subgroups of individuals within certain age categories.

The all-ages regressions indicate that Internet penetration has a negative and statistically significant impact on television viewing. The coefficient indicates a decline of 32 minutes of television viewing per day when an Internet subscription exists (a 22-minute decline when all DMAs are included). This is approximately a 15% (10%) decline in viewing from the 1995 level. Since the Internet penetration rate is actually considerably smaller than 100% (it was estimated to be 72% in 2008),[17] extrapolating the measured coefficient to current circumstances (2008 data) implies that television viewing is currently lower by approximately 11% (7%) from what it might have been if the Internet did not exist.[18]

[16] Statistics at the MSA level are computed using the number of observations in the sample for each MSA. Means statistics for small MSAs are computed with fewer observations, and according to the Census, they provide less information. Wooldridge (2005) analyzes an analogous application.

[17] See http://www.internetworldstats.com/am/us.htm.

[18] This calculation and analogous calculations that follow may be slightly overstated because they assume that the initial (1995) Internet penetration rate is 0.

The 2003 Internet penetration rate of 61% would imply a viewing decline of about 19 minutes due to the Internet (using the results for DMAs with coverage greater than 70%), which is somewhat greater than the actual decline that did occur from 1995 to 2003 (9 minutes according to table 3). Since the standard error is about 13 minutes, we cannot have great confidence that the decline due to the Internet is larger than the decline that occurred. In other words, it is far from certain that other factors raised television viewing by the 10-minute differential between the predicted Internet impact and the actual impact.

As we saw in figure 1, the change in television viewing over the past decade appears far different for young people than it is for older individuals, and we now turn to regressions examining the impact of the Internet on the television viewing habits of individuals in age-based cohorts. The last three pairs of columns in table 5 separate the sample into three age categories and present the fixed-effects regressions, one pair at a time, for each age group.

In these regressions the variables *Internet penetration, education, percentage black, percentage Hispanic*, and *Income* per household member are calculated separately for each age group in each DMA. Thus, our age group regressions allow us to compare how television viewing in cities changes as the rate of Internet penetration (or education or minority status or income) for that particular age group changes. The statistical penalty for separating the data by age group is a reduction in the number of underlying observations in the Census data on which the measurement of the DMA-level variables is based, making any DMA measurements of these variables less precise, even though the number of observations (DMAs) remains unchanged.

Since *Internet penetration* is defined for each age group, the interpretation of the coefficient on this variable indicates the extent to which an increase in Internet adoption for a given age group reduces television viewing for that group. There are several possible reasons that the impact of the Internet on television viewing may differ by age groups. For example, if younger individuals use the Internet with different intensity than older users (as found in table 1), or younger individuals use it for different purposes than older individuals do (for productivity versus entertainment, for example), then the impact of Internet penetration on television viewing could differ by age group.[19]

The results of these regressions are in conformance with general expectations. Individuals in the age group from 6 to 34 years experience the largest negative impact on television viewing due to Internet use. Note as well from table 3 that throughout the period, these young individuals watched considerably less television than did older individuals. Thus, this coefficient represents a larger percentage decline in viewing for the youngest age group than for the overall population (18.8% versus 14.7%, at the point estimates assuming 100% Internet penetration), and the difference is greater than might be thought from looking only at the coefficients.

For individuals between the ages of 35 and 54, the results (a 19-minute decline in viewing due to an Internet penetration rate of 100%) imply an 8.7% decline in television viewing due to Internet use, although the result is not statistically significant. For individuals older than age 55, there is a virtually imperceptible decline in television viewing (1.9%), which is very far from statistical significance. This last result is consistent with a view that older individuals use the Internet less intensively and are less likely to use it for entertainment.

We can now put in perspective the change in television viewing over time that was illustrated in figure 1 and table 3. To do this we use the Internet penetration rates for each age group in 2003. For the group of individuals between 6 and 34 years of age, the 26-minute decline in television viewing is not fully explained by the growth of the Internet's impact on television viewing (17.4 minutes), and it is possible that other factors, such as video games, might explain the difference.[20]

The middle-age group's Internet coefficient of 19 minutes implies a decrease in viewing due to the Internet of 11 minutes at 2003 penetration rates (57%). Average television viewing for this group did not change from 1995 to 2003, as seen in table 3. Although at typical levels of statistical significance, we cannot reject a hypothesis that there was no Internet impact on viewing, nevertheless, television viewing did not fall as much as predicted by the Internet penetration changes. Thus, it is possible that some factors left out of the analysis (e.g., the number of channels) might have a small positive influence.

Finally, the Internet does not appear to have any impact on the television viewing habits of individuals older than 55 years of age. Table 3 indicates that this group has had fairly constant television viewing, with the implication that the net impact of other factors that might have been thought to increase viewing, such as the availability of a larger number of cable networks, did not have an impact on their viewing.

Because our empirical design includes city and year fixed effects, identification derives from idiosyncratic variation in Internet connectedness within cities and not from national aggregate variation in Internet penetration across time. Although this should help limit many potential confounding factors, there is an important potential concern in the analysis that some confounding unobservable factors, correlated with television viewing and the increase in Internet penetration across cities and across time, could create a spurious correlation between city-level Internet adoption trends and city-level trends in television viewing.

For example, the increase in the time young individuals devote to playing video games might have a pattern across

[19] Younger individuals, for example, are more likely to use the Internet to illegally download music or videos than is the case for older individuals. See Liebowitz (2006) or Waldfogel (2009).

[20] The Internet penetration rate for 6–34 year olds is 53%. Multiplying the coefficient for this age group (−32.8) by the rate of Internet penetration (53%) results in an explained decay of approximately 17.4 minutes.

cities and time similar to the increase in Internet penetration. Correspondingly, the use of cell phones has greatly increased during this period, as has the viewing of DVDs, and it is possible that their growth pattern is also similar to that of the Internet. There also may be other activities with these characteristics that we have not enumerated here.

If we had no information about these other activities, we would have to acknowledge that our analysis and results are likely to reflect the joint impact of various new digital forms of entertainment on television viewing and not just, or even mainly, the impact of the Internet. Nevertheless, the basic findings, which are still likely to be of interest, would require only a slightly altered understanding. These findings would represent the impact of this larger class of new technologies on television viewing. Further, the borders between these various activities and technologies sometimes get blurry. For example, when an Xbox user hooks up to the Internet (Xbox live) to play games online or to purchase items, should that be viewed as an Internet activity or a video game console activity? The correct answer is unclear.[21]

We are not without some information about these other technologies, however. We do not believe, for example, that cell phones are likely to have strong impacts on television viewing since the major change in telephone talk time they might engender comes from their ability to allow calls to be made when the user is away from home and such mobile calls are not a replacement for television viewing. Nor are cell phones likely to increase the time talking at home since local cell phone calls are metered, whereas local calls on landlines are not (although the existence of cell phones does allow individuals at home to speak to people who are outdoors, possibly increasing home-based talk time somewhat). Finally, some cell phone uses (checking e-mail, checking the weather, sending text messages, and so forth) are probably best categorized as forms of Internet use, and thus any accidental inclusion of the influence of cell phone use is not as harmful to our results as it might seem.

Prerecorded movies (DVDs and VHS) also seem to be unlikely candidates to greatly influence television viewing since, unlike the large increase in Internet penetration, there was not much growth in the use of prerecorded movies during the period of our study (DVDs largely replaced VHS tapes, and while sales increased, rentals fell).[22] If the viewing of prerecorded movies did not change very much during our period of measurement, then the viewing of prerecorded movies is unlikely to have altered television viewing during this period and thus is unlikely to have been a confounding factor. Nonetheless, while the lack of growth in prerecorded movie viewing decreases the likelihood that movies could be an important confounding factor, we are not claiming it is dispositive.

Video games, however, especially those played on consoles, are a particular concern since they are clearly a substitute for television viewing, grew rapidly (in terms of revenue) during the period of our study, and also take up a good deal of (particularly) young people's time when they might otherwise be watching television. We do not have a panel of city-level information on video games penetration to incorporate into the econometric analysis and thus are unable to measure to what extent video games influence our measurement of the Internet's impact on viewing.

Nonetheless, we were able to obtain some limited data on video game penetration beginning in 2006 for the ten largest DMAs.[23] We would intuitively expect that if video game use was going to confound our results, it would cause us to overstate the impact of the Internet because we might credit the Internet for decreases in television viewing that were in fact caused by video games. If this expectation were correct, it would imply that cities with high Internet penetration might also tend to have high video game penetration.[24] Surprisingly, the correlations between 2003 (or earlier) Internet penetration and 2006 video game console penetration were negative, and although they were not small, they also were not statistically significant.[25] With nonmatching years and only ten DMAs, one cannot form any but the most hesitant conclusions. Nevertheless, this very limited evidence indicates that the cumulative changes in video game penetration across cities appear to be dissimilar to the cumulative changes in Internet penetration that took place across cities, reducing the likelihood that video game consoles are an important confounding factor that would make the Internet's negative impact appear stronger than it is.[26]

[21] Note, however, that viewing television over the Internet is not a problem for our regression analysis since our data end before the advent of YouTube or other sites making television available.

[22] This is documented in Liebowitz (2006), where the revenue decline in rentals largely, but not completely, offsets the revenue increase in video purchases. However, since rentals are less expensive than purchases, meaning that more movies can be watched per dollar of rental (we do not believe that repeated viewing of owned movies would compensate for this), the decline in rental revenue likely understates the decrease in rental viewing relative to viewing from movie purchases, causing total viewing of prerecorded movies to be close to flat.

[23] These data were purchased from Nielsen, which did not have video game data from earlier years.

[24] This potential relationship is complicated by the fact that video games already had penetrated many homes prior to the advent of home Internet access. These earlier video game consoles, however, appealed mainly to children, and the market had flattened out prior to the arrival of 32-bit machines in 1995, mainly the Sony Playstation, which revved up the market and broadened the user base up through 2002, at which point the market flattened again. Thus, video games went through a substantial growth spurt during the period of our regression analysis, and if the pattern of growth in these video game consoles was similar to the growth of the Internet, then it would be a confounding factor in our measurements. But if the growth patterns were similar, Internet penetration and video game penetration are likely to have a positive correlation because the cumulative penetration of these newer consoles into households would be related to cumulative Internet penetration rates.

[25] The correlations between these variables ran from −42% to −9% for 2006 video game penetration and 1997, 1998, 2000, 2001, and 2003 Internet penetration.

[26] This of course does not imply that video games do not have an impact on television viewing. For example, figure 1 shows a decay in television viewing for both males and females but a larger decline for males. Males are more intensive video games players than females are, and this fact may explain the larger decay in the television viewing for males.

Therefore, although our results could be viewed as the joint impact of Internet use, video gaming, and perhaps some other factors, we lean toward a view that the coefficients on the Internet variable are likely to represent largely the impact of the Internet.

VIII. Conclusions

The Internet arrived with great fanfare and was pronounced by many to be a regime changer. Although it has not quite lived up to its original hype, it has progressed over the years to become a formidable source of news and entertainment and is in the process of remaking or eliminating many information and entertainment industries. Our focus in this paper has been on the Internet's impact on television viewing.

Does the Internet reduce television viewing? The answer appears to be yes, although it depends on how the Internet is being used and who is using it. We have found that the Internet has reduced television viewing for individuals with Internet connections. Nevertheless, the size of the impact on television viewing, although substantial, may be smaller than might have expected by some analysts who view the Internet as all consuming.

The strongest result was found for the cohort of individuals who have grown up since the personal computer was developed. The Internet has had a somewhat smaller negative impact on the television viewing habits of individuals in middle age and a minimal impact on the television viewing of the oldest individuals. These results are not surprising or counterintuitive. Middle-age users have lower Internet intensities and are likely to be somewhat more interested in the Internet's ability to provide information than entertainment. The oldest age group has far less Internet intensity than the other groups, so it is not surprising that the impact on television viewing for members of that group is so small.

Although the generational differences that we have measured may continue, we think it more likely that the results from the younger generations will spread into the older generations as time goes by since the knowledge of how to use the Internet is not likely to disappear as individuals age. If this diminution in generational differences takes place, the future of television viewing appears less bright than would be the case if the only factor changing were Internet penetration rates, since those rates do not have all that far left to go. Television executives should be concerned about whether these viewing habits will migrate in this way over time and what that portends for the future of their industry.

Obviously a great deal of additional work is called for if we are to be in a position to predict the future impact of the Internet (and other new media) on traditional sources of entertainment. This paper is but a small first step in that direction.

REFERENCES

Boorstin, E., "Music Sales in the Age of File Sharing," Princeton University mimeograph (2004).
Brown, J., and A. Goolsbee, "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy* 110:3 (2002), 481–507.
Brynjolfsson, E., and M. Smith, "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* 46:4 (2000), 563–585.
Center for the Digital Future, *The Digital Future Report* (Los Angeles: USC Annenberg School, 2007).
Goolsbee, A., "In a World without Borders: The Impact of Taxes on Internet Commerce," *Quarterly Journal of Economics* 115:2 (2000), 561–576.
Liebowitz, S., *Rethinking the Network Economy: The Real Forces That Drive the Digital Marketplace* (New York: Amacom, 2002).
——— "File-Sharing: Creative Destruction or just Plain Destruction?" *Journal of Law and Economics* 49 (2006), 1–28.
——— "Testing File-Sharing's Impact on Music Album Sales in Cities," *Management Science* 54:4 (2008), 852–859.
Nie, N., A. Simpser, I. Stepanikova, and L. Zheng, "Ten Years after the Birth of the Internet, How Do Americans Use the Internet in Their Daily Lives?" (Palo Alto, CA: Stanford Center for the Quantitative Study of Society, 2005).
Nielsen Media Research, *Report on Television* (New York: Report on Television, 2007).
Shapiro C., and H. Varian, *Information Rules: A Strategic Guide to the Network Economy* (Boston: Harvard Business School Press, 1999).
Snyder, A., "Trouble in Nielsenland," *Newsday* (April 21, 1991).
Waldfogel, J., "Lost on the Web," *Information Economics and Policy* 21:2 (2009), 158–168.
Weiss, M., "Trying to Clean Up Sweeps; Are Any Alternatives Ready for Prime Time?" American *Demographics* (May 1, 2001), http://findarticles.com/p/articles/mi_m4021/is_2001_May_1/ai_74303123/?tag=content;col1.
Wooldridge, J., *Introductory Econometrics: A Modern Approach* (Mason, OH: Thompson South-Western, 2005).
Zentner, A., "Online Sales, Internet Use, File Sharing, and the Decline of Retail Music Specialty Stores," *Information Economics and Policy* 20:3 (2008), 288–300.

DATA APPENDIX

Television Viewing, Internet, and Demographics

We acquired data on television viewing by DMA for the years 1990 to 2006 from Nielsen Media.[27] For each DMA and each year, the data measure average television viewing per day based on the sweeps periods (in February, May, July, and November). Daily television viewing is measured from 7:00 a.m. to 1:00 a.m. The data on viewing are contained in a variable that Nielsen labels PUT (Persons Using Television), which is the percentage of persons using television at any given time. To convert this measure into average minutes of viewing per day, we multiply the PUTs by the number of minutes measured per day (1,080 minutes or 18 hours, 7:00 a.m. to 1:00 a.m.). The viewing data are further subdivided into sixteen age and gender groups: children, 2–5; children, 6–11; females, 12–17; males, 12–17; females, 18–24; males, 18–24; females, 25–34; males, 25–34; females, 35–49; males, 35–49; females, 50–54; males, 50–54; females, 55–64; males, 55–64; females, 65 plus; and males, 65 plus.

Nielsen collects television viewing data using two measurement tools: diaries and meters. The older measurement tool is the diary. With diaries, each member of randomly selected households' records information for every channel tuned in for five or more minutes in a quarter-hour. Each year Nielsen processes about 2 million from households during the sweeps periods.

[27] There are 212 DMAs in the data. In 1999, Nielsen condensed three DMAs into a single DMA: Tuscaloosa, Anniston, and Birmingham, Alabama. Information for Juneau, Alaska, is available beginning in 1998. There was no measurement for New Orleans in 2006.

Nielsen also uses two different types of meters when collecting viewing information. Set meters track household tuning but do not provide demographic information about who watches television, whereas people meters collect information on viewing by individual members within a household. In some markets and in some years, diaries alone are used. In other markets, diary results are mixed with metered results. In yet other markets, people meters alone are used.

Nielsen has historically created two distinct samples for different purposes: a national sample for quick overnight information about overall viewing and local samples (as used in our study), which provided far more detailed demographic information about the nature of who was watching, by metropolitan area. People meters are currently employed in both samples, although for much of our time period, people meters were not used at all in the local sample. Although Nielsen started using people meters to collect data from the national sample in 1987, it started to use this technology to collect information from local markets in only 2002 (Boston was the first city to use people meters for the local sample). Still by 2006, only ten of the top local markets were being serviced using people meters.[28] By 2006, a combination of meters and diaries was employed for the next 46 largest local markets and diaries only for the rest of the markets. The data set we acquired (that uses the local sample) was thus collected using diaries for diary-only markets, using set-top boxes and diaries for cities using that combination, and was based on people meters for the small number of markets and years where this system was employed. For each DMA and each year, Nielsen also provided population counts by age and gender breaks and counts of blacks and Hispanic households, although we did not use the Nielsen minority data in our regressions since we had more informative information about minorities (by age group) in the Census data.[29]

We also use data from special surveys included as part of the Current Population Survey by the Census on Internet use, demographics, and household income.[30] Although the CPS is a monthly survey of households and the primary source of information on the labor force characteristics of the U.S. population, the special Internet and computer use surveys (supplements) include variables not found in the ordinary CPS. These special surveys were conducted irregularly for the years 1997, 1998, 2000, 2001, and 2003 and tend to have about 130,000 respondents.[31] For each respondent, these surveys provide information on the age, sex, race, education, household income, and geographical information on the MSA of residence. By aggregating the raw data, we were able to create age group variables for share of blacks, share of Hispanics, education levels, income per household member, and Internet penetration rates.

We combined the data from Nielsen with the data from the CPS by comparing the geographical definitions of MSAs and DMAs. DMAs cover the entire U.S. territory and thus cover more territory and more inhabitants than MSAs, which cover only metropolitan areas. Some DMAs have more (sometimes considerably more) than one MSA. We define the variable coverage as a measurement of the extent to which the Census data from all the MSAs within a DMA can be considered to be representative of the entire DMA. This variable is defined using the difference between the sum of the population in all the MSAs within a DMA for year 1998 and the population in the DMA provided by Nielsen for that same year.[32] Occasionally an MSA is split between two or more DMAs.[33] Although we have information on television viewing for 210 DMAs, not all of the DMAs contain an MSA.

Over the years, the CPS has slightly changed the questions aimed at measuring Internet penetration in the Internet and computer use supplements. For 1997, we measured Internet penetration combining two variables (PESCU12A and PESCCU4H)—one asking individuals ages 15 and older and the other asking individuals 3 to 14 whether they use the Internet at home. We use household identifiers to create a household-level variable measuring Internet penetration if any one of the household members uses the Internet at home. For 1998 and 2000, we used the variable HESIU3 to measure Internet use in the household and PR11 to measure individual-level Internet use. Finally, for 2001 and 2003, we used the variable HESINT1 to measure household-level Internet use and the variable PRNET2 to measure individual-level Internet use. Individual-level Internet use is employed to compute Internet use by age groups. In all cases, we weight each observation by the weighting variable provided by the special CPS, PWSSWGT.

We measure educational attainment using a variable from the CPS with five categories (PREDUCA5): less than high school diploma (1), high school graduates (2), some college (3), associate degree (4), and bachelor's degree or higher (5).[34]

Household income is measured in nominal dollars per year. We created this variable using the variable HUFAMINC from the CPS. In answering this question, respondents can choose from fourteen income brackets for the years 1995, 1997, 1998, 2000, and 2001 and sixteen income brackets for 2003. We used the middle dollar value of the income bracket and $100,000 for the top bracket ($75,000 and over) for the years 1997, 1998, 2000, and 2001 and a value of $175,000 for the top bracket in 2003 ($150,000 and over). We computed two variables: average household income and median income per member of the household. The choice of median income per household member versus average household income had only a minor impact on the results. We report only the results using median income measurements. Because the top income is an open-ended bracket for which we chose a single value, our average measurements are less likely to be affected by large outliers than if actual income amounts had been included. Median income for each DMA is based on the respondents' prorated share of family income over all respondents in a city. Computing income per member of the household allows us to create separate age-group-based median income variables.

For minorities, we used the variables PRHSPNON for Hispanics and PERACE for blacks for the years 1995, 1997, 1998, 2000, and 2001. For 2003, we used PEHSPNON for Hispanics and PTDTRACE for blacks. PRDTRACE has many categories for combination of races; we recoded to black if the answer was "black only."[35] Using the age of each individual, we constructed minorities variables (blacks and Hispanics) broken down by age groups. These variables are used in the regressions by age breaks.

[28] The pattern of the trend in television viewing and the trends by age groups are similar when excluding the top ten markets that started measuring audience by using local people meters. Television viewing levels, on the other hand, are significantly lower when excluding the largest markets using local people meters. This indicates that television viewing is higher in the top ten markets, which is not surprising given the larger number of local stations in these large local markets. By 2003, there were 5,000 households in the national people meters sample. The effective sample size of the national sample increased to about 10,000 households in 2005 with the integration of the local people meters samples into the national sample.

[29] The data from Nielsen provide counts of total, black, and Hispanic households. It is possible to calculate the proportion of black and Hispanic households for each DMA and each year. This information could be used instead of the Census information in the regressions, which combine all age groups into one whole. The coefficients on the Internet variables employing information on minorities from Nielsen are largely unchanged (-31.6 in table 5 versus -30.4 when using DMAs with coverage above 70% and -22.0 in table 5 versus -20.9 when using all DMAs). Using the data on minorities' shares from Nielsen, the coefficients on the panel regressions suggest that an increase in the share of blacks (and to a smaller extent also Hispanics) increases television viewing, but the standard errors are large. The information from Nielsen does not contain information on minorities broken down by age groups. Therefore, the regressions by age group cannot rely on the minority data from Nielsen.

[30] These data are made easily accessible with the Census's DataFerrett tool found at http://dataferrett.census.gov/.

[31] The CPS also conducted Internet and computer use supplements for years 1994 and 2007. The geographical definitions of the MSAs for these two years are different from the definitions the Census used for 1997, 1998, 2000, 2001, and 2003.

[32] In a few instances, the population of the sum of the MSAs is higher than the population of the DMA. This is because for a few MSAs, the CPS defines the strata to be larger than the MSA. The regressions using DMAs with a coverage rate higher than 0.7 drop six DMAs because their population as measured by Nielsen is more than 30% lower than the sum of the populations of all the MSAs within them.

[33] Only sixteen MSAs are in more than one DMA. Dropping them from the regressions does not change the results significantly. We do not use the MSA Merced, California, in our analysis because of incomplete information.

[34] This question is applicable to individuals 15 years and older.

[35] Other categories for partial black ethnicity have very few observations.

This article has been cited by:

1. Avi Goldfarb, Catherine Tucker. 2019. Digital Economics. *Journal of Economic Literature* **57**:1, 3. [Crossref]
2. Sung Wook Ji. 2018. The Internet and Changes in the Media Industry: A 5-Year Cross-National Examination of Media Industries for 51 Countries. *Journalism & Mass Communication Quarterly* **5**, 107769901880791. [Crossref]
3. Ai-Phuong Hoang, Robert J. Kauffman. 2018. Content Sampling, Household Informedness, and the Consumption of Digital Information Goods. *Journal of Management Information Systems* **35**:2, 575-609. [Crossref]
4. Tim Prostka, Edlira Shehu, Michel Clement. 2018. Cannibalisation effects in the early market stage of e-books – an analysis of the German book market. *Journal of Media Business Studies* **15**:2, 108-126. [Crossref]
5. Qian Zhang, Chao Wu, Han Qiao, Shouyang Wang. 2017. No advertising, but more sponsorship?. *Chinese Management Studies* **11**:1, 90-106. [Crossref]
6. J. Adam Avant, Kyongseok Kim, Jameson L. Hayes. 2017. Thirty Years of Advertising Research in Leading Communication and Marketing Journals: Learning From the Parent Disciplines. *Journal of Current Issues & Research in Advertising* **38**:1, 44-64. [Crossref]
7. Elena Raptou, Georgios Papastefanou, Konstadinos Mattas. 2017. Investigating the influence of eating habits, body weight and television programme preferences on television viewing time and domestic computer usage. *Perspectives in Public Health* **137**:1, 59-72. [Crossref]
8. Daegon Cho, Michael D. Smith, Alejandro Zentner. 2016. Internet adoption and the survival of print newspapers: A country-level examination. *Information Economics and Policy* **37**, 13-19. [Crossref]
9. Nakil Sung, Eunkyoung Kwack. 2016. IPTV's videos on demand for television programs, their usage patterns, and inter-channel relationship in Korea. *Telecommunications Policy* **40**:10-11, 1064-1076. [Crossref]
10. Stan J. Liebowitz, Alejandro Zentner. 2016. The internet as a celestial TiVo: What can we learn from cable television adoption?. *Journal of Cultural Economics* **40**:3, 285-308. [Crossref]
11. Shinjae Jang, Minsoo Park. 2016. Do New Media Substitute for Old Media?: A Panel Analysis of Daily Media Use. *Journal of Media Economics* **29**:2, 73-91. [Crossref]
12. Yizao Liu, Rigoberto A. Lopez. 2016. The impact of social media conversations on consumer brand choices. *Marketing Letters* **27**:1, 1-13. [Crossref]
13. Ai Phuong Hoang, Robert J. Kauffman. Experience Me! The Impact of Content Sampling Strategies on the Marketing of Digital Entertainment Goods 4750-4759. [Crossref]
14. Victor M. Bennett, Robert Seamans, Feng Zhu. 2015. Cannibalization and option value effects of secondary markets: Evidence from the US concert industry. *Strategic Management Journal* **36**:11, 1599-1614. [Crossref]
15. Andriana Bellou. 2015. The impact of Internet diffusion on marriage rates: evidence from the broadband market. *Journal of Population Economics* **28**:2, 265-297. [Crossref]
16. Jin Li, Zhiling Guo, Robert J. Kauffman. Recovering Household Preferences for Digital Entertainment 4276-4284. [Crossref]
17. Oliver Falck, Robert Gold, Stephan Heblich. 2014. E-lections: Voting Behavior and the Internet. *American Economic Review* **104**:7, 2238-2265. [Crossref]
18. Ambarish Chandra, Ulrich Kaiser. 2014. Targeted Advertising in Magazine Markets and the Advent of the Internet. *Management Science* **60**:7, 1829-1843. [Crossref]
19. Matthew Gentzkow. 2014. Trading Dollars for Dollars: The Price of Attention Online and Offline. *American Economic Review* **104**:5, 481-488. [Crossref]
20. Robert Seamans, Feng Zhu. 2014. Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers. *Management Science* **60**:2, 476-493. [Crossref]
21. Brett Danaher, Michael D. Smith, Rahul Telang. 2014. Piracy and Copyright Enforcement Mechanisms. *Innovation Policy and the Economy* **14**, 25-61. [Crossref]
22. Satoshi Kitamura. 2013. The Relationship Between Use of the Internet and Traditional Information Sources. *SAGE Open* **3**:2, 215824401348969. [Crossref]
23. Michael R. Ward. 2012. Learning to surf: Spillovers in the adoption of the Internet. *Technological Forecasting and Social Change* **79**:8, 1474-1483. [Crossref]