IT&SOCIETY, VOLUME 1, ISSUE 4, SPRING 2003, PP. 1-13
http://www.ITandSociety.org

IS THE DIGITAL DIVIDE REALLY CLOSING? A CRITIQUE OF INEQUALITY MEASUREMENT IN *A NATION ONLINE*

STEVEN P. MARTIN

ABSTRACT

According to the U.S. Department of Commerce Report "A Nation Online: How Americans are Expanding their Use of the Internet," computer ownership and Internet use are rapidly becoming more equally distributed across households in the United States. The authors of "A Nation Online" use two statistical arguments to support this claim: 1) annual rates of increase for computer and Internet use are increasing most quickly for poor households, and 2) "Gini" coefficients for inequality of computer use are decreasing. These analyses critique these arguments and show that patterns that the authors attribute to decreasing inequality are instead explained by two factors: 1) computer and Internet use is increasing, and 2) households with higher incomes began using computers and the Internet earlier than households with lower incomes.

Reanalyzing these same data using odds ratios indicates that computer ownership and Internet use may actually be spreading less quickly among poorer households than among richer households. If current trends continue, poor households will eventually have the nearly universal levels of computer and Internet use currently seen among richer households, but this "catching-up" could take two decades.

Steven P. Martin is an Assistant Professor in the Department of Sociology at the University of Maryland. smartin@socy.umd.edu

I want to thank John Robinson and Alan Neustadtl for their invaluable advice and creative suggestions for this study.

Is the "digital divide" going away? The United States has shifted from a Clinton administration strongly focused on a digital divide to a Bush administration largely dismissive of it. This changed political environment has led to numerous reassessments of the evidence for and against a gap between the "haves" and the "have-nots" in terms of computer and Internet access (Compaigne 2001; Cooper 2002) and the extent to which the digital divide is a temporary phenomenon that will fix itself (Samuelson 2002).

In its 2002 report, *A Nation Online: How Americans are Expanding Their Use of the Internet* (U.S. Department of Commerce 2002), the U.S. Department of Commerce examined levels and trends in inequality of computer use across various groups of Americans. The findings were emphatic and reassuring; Computer and Internet use are increasing most rapidly among the poor and other disadvantaged groups, and the digital divide is closing quickly.

This study challenges the sanguine assessment of *A Nation Online* with respect to inequality in computer and Internet access, arguing that the key findings in *A Nation Online* depend on two types of statistical analyses—estimates of relative rates and "Gini" indices designed by the authors. Such analyses are not necessarily wrong, but they are clearly misleading when applied to trends in inequality, because they are inherently asymmetrical. When applied to questions of who owns computers or uses the Internet, these analyses consistently and automatically show that inequality is *decreasing*. However, when one reframes the analyses in terms of who *does not* own computers, they show with equal certainty that inequality is rapidly *increasing*. The authors then emphasize the results that indicate a decrease in inequality and downplay the equally valid results that indicate an increase in inequality.

One alternative, appropriate measure that is inherently symmetrical and invariant with respect to the definition of the outcome variable is the *odds ratio* (to be described later.) The data in *A Nation Online* are reevaluated using odds ratios to measure trends in inequality in computer and Internet use. While not clearly pessimistic, the results are certainly not as optimistic as those published in *A Nation Online*, and they indicate that the closing of the digital divide is far from a foregone conclusion.

DATA

The data for this analysis come from a series of Current Population Surveys (CPS) on computer and Internet use. Many of these data are summarized in *A Nation Online*, with additional analyses run on original CPS data provided by the Bureau of Labor Statistics Website (http://ferret.bls.census.gov/) and on detailed tabulations from *Current Population Reports* from the original data (U.S. Census Bureau 1985, 1990, 1994, 1998). Data from the calendar years 1984, 1989, 1993, 1997, 1998, and 2001 were included.

RESULTS

Annual Growth in the Rate of Internet Use: One method the authors of *A Nation Online* use to measure inequality is the annual growth rate of Internet use. This measure is calculated by the following procedure. At two time points—t_1 and t_2— one measures the proportions $P(t_1)$ and $P(t_2)$ of persons using computers or the Internet The annual growth rate *GR* is then a function of the times and proportions:

$$GR^{(t_2-t_1)} = \frac{P(t_2)}{P(t_1)} \tag{1}$$

The problem with the annual growth rate as a measure of inequality is that it is biased toward groups with a low value of $P(t_1)$. Because poorer households had lower levels of computer ownership or computer use at earlier times, they are almost guaranteed to have higher annual growth rates.

An unsatisfactory solution to this problem would be to calculate the annual rate of *decline* for individuals and households that *do not* use the Internet or own computers. The authors of *A Nation Online* also calculate this measure and duly warn the reader about its inherent bias. At two time points—t_1 and t_2—one measures the proportions $1 - P(t_1)$ and $1 - P(t_2)$ of persons *not* using computers or the Internet. The annual rate of decline in nonuse *(DR)* is then a function of the times and proportions, where:

$$DR^{(t_2-t_1)} = \frac{[1-P(t_2)]}{[1-P(t_1)]} \tag{2}$$

The annual rate of decline has the opposite weakness as a measure of inequality; it is biased toward groups with a *high* value of $P(t_1)$. Because richer households had higher levels of computer ownership or internet use at earlier times, they are almost guaranteed to have higher annual rates of decline in the fraction still *not* using a given technology.

To overcome this problem with growth rates, researchers commonly measure trends using *odds ratios*, a measure of population proportions that is invariant with respect to whether differences are measured for the proportion of observations *inside* a category or *outside* a category (Agresti 1990). To compute odds, one evaluates the proportions twice, once *using* computers or the internet and once *not using* computers or the Internet. An odds ratio *T*, then, is a ratio of two odds calculated at different times t_1 and t_2:

$$\Theta = \frac{P(t_2)*[1-P(t_1)]}{P(t_1)*[1-P(t_2)]} \tag{3}$$

TABLE 1: THE ONLINE POPULATION. FAMILY INCOME AND INTERNET USE FROM ANY LOCATION BY INDIVIDUALS AGE 3 AND OLDER: 1998-2001

	Internet use (%)			Annual Growth in use rate	Increase in odds of use
	Dec. 1998	Sept. 2001	Difference		
Family income					
<$15,000	14	25	11	25	2.1
$15,000-24,999	18	33	15	24	2.2
$25,000-34,999	25	44	19	22	2.3
$35,000-49,999	35	57	22	20	2.5
$50,000-74,999	46	67	22	15	2.5
$75,000+	59	79	20	11	2.6

Note: At 25% annual growth rate, households with $75,000 or more in family income would have had Internet use at 108% in September 2001.
Source: Data from A Nation Online: Table 2-3, page 27.

Table 1 shows the application of these different measures for the case of Internet use, from any location, by individuals age three and older. The levels of Internet use and annual growth rates are from *A Nation Online*, with a measure of the changing *odds* of Internet use added.

Two patterns are clear in Table 1. First, individuals with the lowest family incomes have a much lower chance of using the Internet in either time period (14% in 1998, 25% in 2001) than individuals with the highest family incomes (59% in 1998, 79% in 2001). Hence, there is still clear evidence of a 'digital divide' between individuals in the highest and lowest income categories. Secondly, Internet use increased quite rapidly for individuals at all levels of family income—from an 11% increase at the lowest income levels to a 20% increase at the highest income levels. Note that the largest *differences* in Internet use were for individuals in the *middle* income categories, who passed across the 50% threshold for Internet use between 1998 and 2001. (This pattern is consistent with an "S-shaped" curve commonly seen in populations adopting a technology or experiencing some other sort of transition. It does not necessarily mean that groups temporarily in the center of the S-curve are experiencing the most rapid transitions across a longer span of time).

The increase of Internet use across all income categories is an encouraging social trend, but the authors of *America Online* argue further that poorer families are adopting Internet use *more rapidly* than richer families. They base their argument on the annual growth rates in the fourth column of Table 1. The annual growth rate is highest for the poorest income categories and more than twice as high for the lowest income category as for the richest category (25% and 11%, respectively). However, as argued above, the annual growth rate measure is biased toward groups with low initial percentages, so this striking pattern does not really tell one about inequality in the distribution of Internet use.

The fifth column of Table 1 shows an unbiased measure of trends over time—that is, the increase in odds of Internet use from t_1 to t_2. By this measure, the increase in odds was substantial across all income categories, especially considering that the interviews were fewer than three years apart. However, there is a clear pattern indicating that individuals from the poorest income category are adopting Internet use *a bit more slowly* than individuals from the richest income category (the odds ratios are 2.1 and 2.6, respectively).

As described above, an unacceptable, alternative way to measure trends in inequality is to use the annual rate of decrease in the proportions of individuals *not* using the Internet. However, the offline population is a significant policy concern, and the authorr of *A Nation Online* took time to reassess the data about Internet use in terms of the population *not* using the Internet. Table 2 shows their results.

The first three columns of Table 2 show the same data as in Table 1, but the percentages have been subtracted from 100%, and the signs of the differences have all changed. Note in the fourth column that the annual decline in the non-use rate indicates a much more rapid decline in the offline population among the richest households than among the poorest households (22% and 5%, respectively). While this finding suggests that Internet non-users are disappearing rapidly from higher income households, it is a biased measure of trends in inequality of Internet use—just as the measure of annual growth rates in Table 1 is a biased measure. The authors of *A Nation Online* recognized this problem and strongly warned readers not to interpret the annual decline in the non-use rate as a measure of trends in inequality (U.S. Department of Commerce 2002, page 75). However, the authors did *not* include warnings about the similarly biased annual growth rates, but instead they touted the results from the annual growth rates as one of the key findings of the report (U.S. Department of Commerce 2002, page 1, page 11).

Perhaps the most unbiased measure of trends in inequality of Internet nonuse is the odds ratio, as reported in the fifth column of Table 2. The change in odds of computer nonuse was large for all income categories and largest for the richest households (2.6), indicating a promising overall trend but also a slight increase in inequality. Note that these odds ratios are identical to the corresponding ones in Table 1, a necessary result because the odds ratios, unlike growth rates, are invariant with respect to how trends are defined.

Gini coefficients for trends in inequality of computer ownership: The authors of *A Nation Online* used another measure to argue that the digital divide was decreasing over time; a version of the *Gini* coefficient commonly used in studies of time trends in income inequality. In its standard form, the Gini coefficient is calculated by ordering the individuals in a population from highest income to lowest along the *x*-axis, then summing the cumulative distribution of

TABLE 2: THE OFFLINE POPULATION. FAMILY INCOME AND INTERNET "NON-USE" FROM ANY LOCATION BY INDIVIDUALS AGE 3 AND OLDER: 1998-2001

	Not using Internet (%)			Annual Decline in Non-use rate	Decrease in odds of non-use
	Dec. 1998	Sept. 2001	Difference		
Family income					
<$15,000	86	75	-11	5	2.1
$15,000-24,999	82	67	-15	7	2.2
$25,000-34,999	75	56	-19	10	2.3
$35,000-49,999	65	43	-22	14	2.5
$50,000-74,999	54	33	-22	17	2.5
$75,000+	41	21	-20	22	2.6

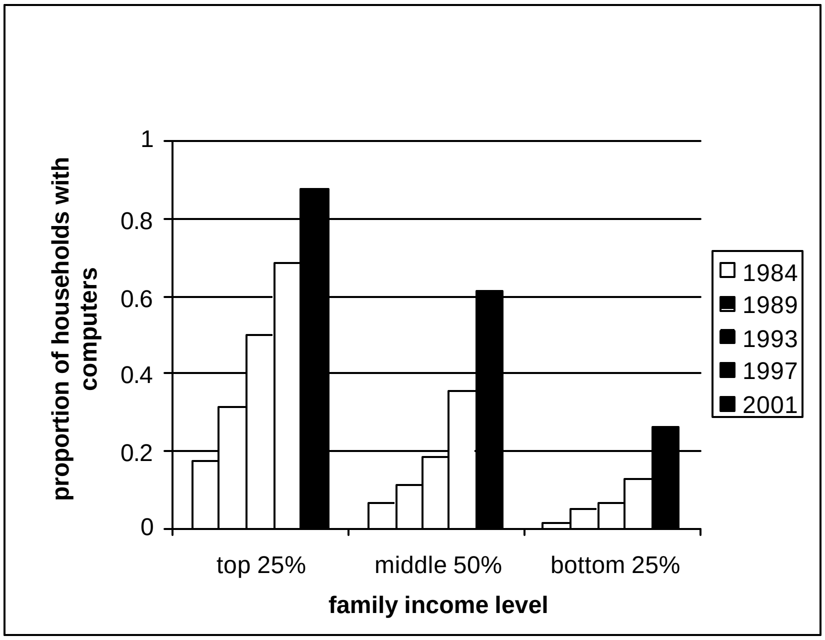
Source: Data from A Nation Online: Table 8-2, page 82.

their incomes along the *y*-axis. In a perfectly equal society, everyone contributes the same amount to the cumulative income distribution, so the x-y graph is a perfectly straight diagonal. In a perfectly unequal society, one person has all the income, so the graph hugs the x axis at y = 0%, then shoots up to y = 100% at the very last x-observation. Graphs for true populations are somewhere in the middle, and the Gini index measures the departure of the observed distribution from the perfectly equal distribution, on a scale of 0 to 1. Gini indices for US income have risen from about .41 to about .46 in recent decades (U.S. Department of Commerce 2002, page 89).

The authors of *A Nation Online* attempted to adapt the Gini coefficient to study trends in the inequality of computer ownership and Internet use. However, instead of measuring the ordered distribution of one variable against the cumulative distribution of that same variable, they measured the ordered distribution of one variable (such as income) against the cumulative distribution of a different variable (such as whether a household owns a computer). Unfortunately, as shown below, this measure is subject to exactly the same sort of bias as the measures of annual growth rates and annual rates of decline. To be specific, when poor households have fewer computers than rich households, any increase in computer ownership will cause the Gini coefficient to decline. After replicating the Gini coefficients reported by the authors, the same data and estimation procedure were used to calculate a second Gini coefficient for inequality in households without computers. This second Gini is just as valid conceptually as the first, but it produces exactly the opposite trends.

Descriptively, Figure 1 shows the proportion of U.S. households with computers, by year and by reported family income level. In Figure 1, it is clear that the households in the top quarter of the income distribution were most

FIGURE 1: PROPORTION OF U.S. HOUSEHOLDS WITH COMPUTERS, BY YEAR AND FAMILY INCOME LEVEL



Source: U.S. Current Population Surveys for 1984-2001.

likely to have a computer in any year, and that the highest income households had the largest absolute increase in the proportion of households with a computer—from 18% in 1984 to 88% in 2001. However, it is also clear that computer ownership increased dramatically at all income levels. Indeed, the increase at the bottom quarter of the income distribution was from 2% in 1984 to 26% in 2001, a relative thirteen-fold increase. (Obviously, the high-income households could not have had such a relative increase, because the proportion with a computer would have greatly exceeded 100%).

The question, then, is whether these patterns are consistent with an increase or a decrease in inequality of computer ownership as shown in Figure 1 and Table 3.

The first row of figures in Table 3 shows the Gini coefficients for computer inequality calculated by the authors of *A Nation Online* and intended to measure the change in inequality shown graphically in Figure 1. By this measure, the inequality in computer ownership plummeted by nearly half, from a Gini of .44 in 1984 to .23 in 2001. Translated into practical terms, this means that in 1984, the few computers in households were mostly in the hands of the richest families, while in 2001, the computers were distributed more equitably.

TABLE 3: "GINI"-STYLE COEFFICIENTS FOR INEQUALITY IN THE DISTRIBUTION OF HOUSEHOLDS WITH COMPUTERS AND WITHOUT COMPUTERS, BY FAMILY INCOME LEVEL 1984-2001

	1984	1989	1993	1997	2001
Gini based on distribution of households *with* computers	.44	.40	.39	.33	.23
Gini based on distribution of households *without* computers	.04	.07	.12	.20	.33

Source: Calculated from Current Population Surveys for 1984-2001.

The second row of figures in Table 3 shows the comparable Gini indices based on the distribution of households without computers. By this measure, inequality of computer ownership increased dramatically from .04 in 1984 to .33 in 2001. In practical terms, this means that in 1984, few people of any income owned household computers, and rich and poor households alike were pretty much alike in not having computers. By 2001, many households had computers, and most of the remaining households without computers were concentrated at the lowest income levels. Of course, neither Gini index is telling the whole story.

The odds ratio can again be used to obtain an invariant measure of changing inequality in the distribution of households with computers.

Table 4 shows the change in the odds that a household will have a computer for the household data shown in Figure 1. Table 4 shows that the odds of a household owning a computer increased at every income level and across every time period; this should be taken as quite encouraging news.

Trends by family income are more mixed. In different time intervals, the Top 25%, Middle 50%, or Bottom 25% of the income distribution each showed the largest increase in the odds of owning a computer. However, the highest income category showed the most consistent and largest overall increases in the odds of having a computer (odds ratio = 33.7 for 1984 to 2001), whereas the lowest income category showed the least consistent, smallest overall increases (odds ratio = 20.6 for 1984 to 2001).

One way to interpret these trends is to assume that the future will be like the past in that Internet use and computer ownership will continue to spread among all income groups. Such an assumption is consistent with a "normalization" model of technology diffusion, as compared to a "stratification" model of technology diffusion (c.f. Norris 2001). Table 5 presents the results from this predictive exercise for the case of households having a computer. The values for 2001 are the most recent observed data, with 88 percent of the highest-income households having a computer, while only 26 percent of the lowest-income households have a computer. If the odds ratios continue to

TABLE 4: FAMILY INCOME AND THE ODDS OF A U.S. HOUSEHOLD HAVING A COMPUTER. 1984-2001

Family Income:	Change in odds				
	1984-1989	1989-1993	1993-1997	1997-2001	1984-2001
Top 25%	2.2	2.1	2.2	3.3	33.7
Middle 50%	1.8	1.8	2.4	2.9	22.2
Bottom 25%	2.9	1.4	2.1	2.4	20.6

Source: U.S. Current Population Surveys for 1984-2001.

TABLE 5: FAMILY INCOME AND THE PREDICTED PROBABILITY OF A U.S. HOUSEHOLD HAVING A COMPUTER FOR FUTURE YEARS

Family Income	Change in Odds of having a computer		Proportion with a computer		
			observed	predicted	
	1984-2001	Annual	2001	2005	2010
Top 25%	33.7	1.23	.88	.94	.98
Middle 50%	22.2	1.20	.61	.76	.89
Bottom 25%	20.6	1.19	.26	.42	.64

increase at the same annual rates as in the past, then by 2010 a full 98 percent of the highest income households will have a computer and 64 percent of the lowest-income households will have a computer. Such a pattern would represent a decrease in *absolute* levels of inequality, in that the difference between high- and low-income households would be smaller in 2010 than in 2001. However, the persistent *relative* inequality in technology diffusion means that there could be a significant proportion of poorer households without a computer for more than a decade. A comparable exercise for internet use would predict a more rapid uptake, with nearly 80 percent of the poorest individuals using the internet by 2010. However, such a prediction has a very high degree of uncertainty, as it is based on a much shorter time interval of observed data.

In summary, the reassessment of the data in *A Nation Online* shows no evidence for a decrease in the unequal diffusion of computer ownership or Internet use over time. The evidence for such a decrease comes from measures that are clearly predestined to show such a decrease. Both an informal assessment of trends and a formal comparison by odds-ratios indicate that computer ownership and use of the Internet are increasing for individuals of all income levels, but the increases are generally more pronounced at the highest income levels. This increasing inequality may not prevent the lowest income groups from reaching high or even universal levels of Internet and computer access, but it could clearly increase the time lag between high-income and low-income groups.

DISCUSSION

Given that Internet use and computer ownership are spreading across all segments of American society, does it matter if the spread is slower among the lowest income groups? One possibility is that the increase of Internet use and computer ownership will stall for the most disadvantaged groups in the United States, as depicted in Figure 2. This outcome appears unlikely, because past trends have shown a consistent increase across all measures of computer use, all groups, and all time periods. However, computer ownership and Internet use were still quite low for the lowest-income echelons in 2001, so it will take a decade or more until a large majority of poor individuals use the Internet and/or have computers at home. It is certainly premature to conclude that past patterns of increase will automatically persist into the future.

A more likely scenario is that Internet use and computer ownership will become nearly universal for lower-income as well as higher-income individuals. If this occurs, Americans need to decide what an acceptable lag time is for the disadvantaged groups in society. Figures 3, 4, and 5 illustrate possible trajectories for technology diffusion. In Figure 3, disadvantaged groups experience a lag in technology diffusion, but the odds ratios for diffusion are the same across groups, so the disadvantaged groups trail the advantaged groups by the same number of years throughout the diffusion process. Figure 4 represents an optimistic scenario in which the odds of technology use increase most quickly for the disadvantaged group, so group differences diminish over time, and inequality persists for a shorter duration. Figure 5 represents the scenario most consistent with trends so far; odds of technology use increase for all groups, but more slowly for the disadvantaged group. Inequality in technology use still disappears in this scenario, but the inequality persists for several additional years.

Why should lag time matter if all groups eventually end up on equal footing? The answer to that question depends on the relationship between technology use and other social and economic outcomes. So long as technological inequalities persist, they may exacerbate other forms of social and economic inequality. If these forms of social and economic inequality are automatically ameliorated when the technological inequality disappears, then a long lag in diffusion innovation will not have permanent effects on society. If, however, these forms of social and economic inequality become "locked in," then a return to relative technological equality will not suffice to undo the effects of past technological inequality. In that case, it is important to make the lag in technology diffusion as brief as possible by actively engaging the digital divide as long as it persists.

Figure 2: Diffusion pattern with 5-year lag, stratified diffusion

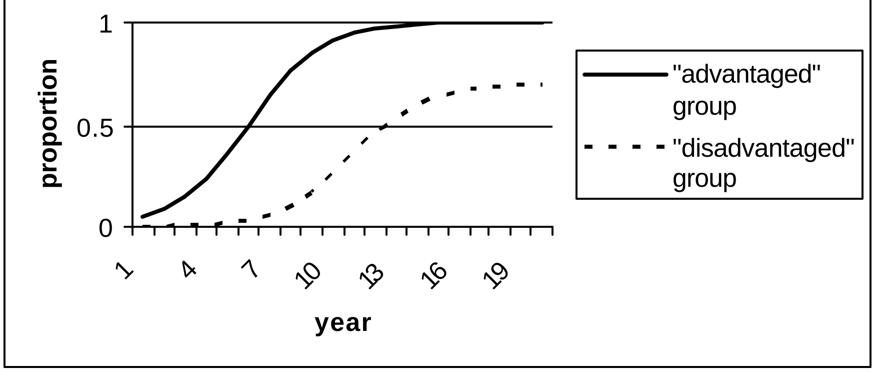
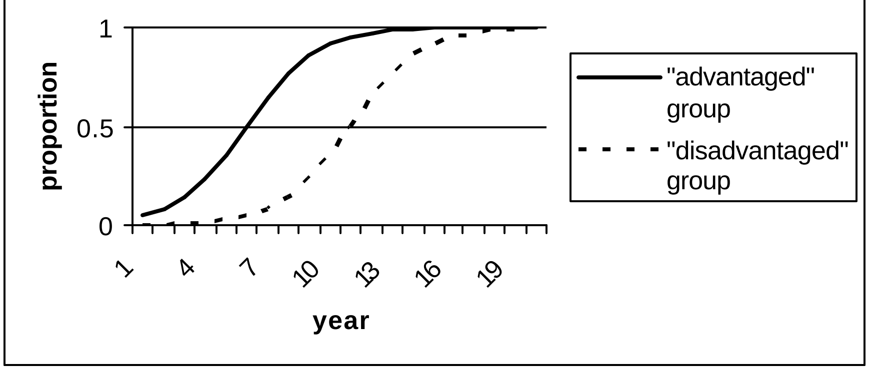


Figure 3: Diffusion pattern with 5-year lag, no inequality of diffusion



The late 1990s were a time of rapid increase in Internet and computer use across all levels of American society. However, as these results indicate, that increase was more pronounced among individuals and families at the highest income levels. Some inequality in technological diffusion may not matter much if the rapid increases of the late 1990s continue, but these increases occurred amid a robust increase in wealth and income, and amid a

Figure 4: Diffusion Pattern with 5-year lag, decreasing inequality of diffusion

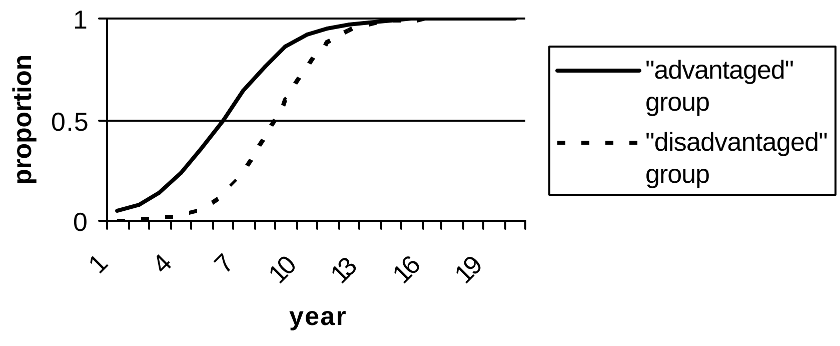
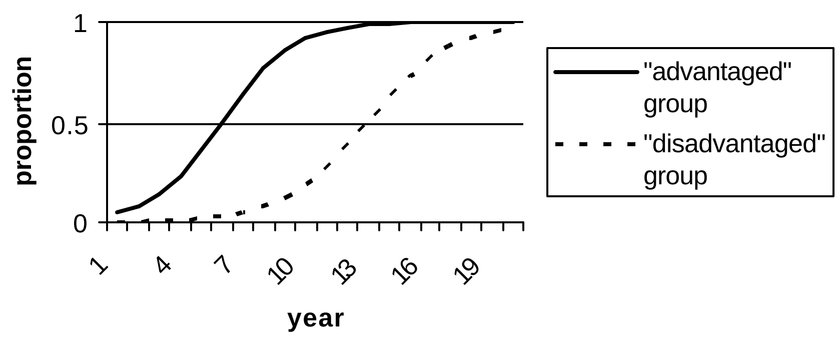


Figure 5: Diffusion pattern with 5-year lag, increasing inequality of diffusion



strong political commitment to address inequalities in computer access. As the political will and the economic boom dissipate, one may see only modest increases in the proportion of poorer households with computers or Internet access. If that occurs, the digital divide could easily persist for a generation or longer.

REFERENCES

Agresti, A. 1990. *Categorical Data Analysis*. New York: John Wiley and Sons.

Compaigne, B. 2001. Declare the War Won. p. 315-336. In Compaigne, Benjamin M. (ed.) *The Digital Divide: Facing a Crisis or Creating a Myth?* Cambridge, MA: MIT Press Sourcebooks.

Cooper, M. 2002. Does the Digital Divide Still Exist? Bush Administration Shrugs, but Evidence Says "Yes." *Consumers Union*. Available at: http://www.consumerfed.org/DigitalDivideReport20020530.pdf Accessed on 4/24/03.

Norris, P. 2001. *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge, MA: Cambridge University Press.

Samuelson, R. 2002. Debunking the Digital Divide. *Washington Post* Op-Ed Piece, March 20.

U.S. Census Bureau. 1985. Computer Use in the U.S. 1984, *Current Population Reports P23-155*.

U.S. Census Bureau. 1990. Computer Use in the U.S. 1989, *Current Population Reports P23-171*.

U.S. Census Bureau. 1994. Computer Use in the U.S. 1993, *Current Population Reports*.

U.S. Census Bureau. 1998. Computer Use in the U.S. 1997, *Current Population Reports P20-522*. Available at: http://www.census.gov/population/computer/report97/

U.S. Department of Commerce, Economics and Statistics Administration. 2002. *A Nation Online: How Americans are Expanding Their Use of the Internet*. Washington DC: U.S. Department of Commerce. Data from the September 2001 Current Population Survey are available at: http://ferret.bls.census.gov/