# Predicting Eligibility Rates for Rare Populations in RDD Screening Surveys 

Donald P. Camburn, Abt Associates Inc.

Robert A. Wright, National Center for Health Statistics
Donald Camburn, Abt Associates Inc., 101 North Wacker Drive, Suite 400
Chicago, IL 60606, (312) 621-2071, donald_camburn@abtassoc.com

KEY WORDS: RDD, eligibility, surveys

## 1. Background

Accurately predicting eligibility rates for rare populations is a critical step in designing random-digit-dial (RDD) surveys. If the eligibility rate encountered during data collection is higher than predicted, survey costs may increase with no appreciable gain in the precision of the estimates. If the actual eligibility rate is lower than the predicted rate, increasing the sample size to achieve a fixed number of completed interviews may be necessary. This may increase survey costs and delay the data collection schedule. If, because the eligibility rate was overestimated originally, the number of interviews completed is fewer than called for in the survey design, the reliability of estimates derived from the data may suffer.

In surveys of rare populations even very small differences between the predicted and observed eligibility rates can have enormous implications. If the predicted eligibility rate is $5 \%$ and the observed rate is $4.5 \%$, the $0.5 \%$ absolute difference represents a $10 \%$ shortfall in comparison to the expected rate.

This paper is an outgrowth of the National Immunization Survey (NIS) data collection. This large RDD CATI survey, described more fully in the next section, is conducting telephone screening interviews from 1994 through 1997 to identify those households with children 19-35 months of age. The survey design, sample frame design, and cost estimates assumed an eligibility rate of $5 \%$. As data collection for the NIS has progressed, however, the observed eligibility rates have been below predicted levels, thereby reducing survey response rates and increasing survey production time and costs. During 1994 the observed eligibility rate stood at $4.1 \%$ versus the predicted $5.0 \%$, a shortfall of $18 \%$.

As part of an ongoing series of research efforts designed to identify possible sources of the lower than expected eligibility rate, consideration was given to the possibility that the number of unique telephone lines per household is lower in telephone households eligible for the NIS than in the general population. If such a differential exists, it could result in substantial misestimation of the eligibility rate in RDD surveys that screen for rare populations, since RDD sample designs typically assume that the number of telephone lines is uniformly distributed across all household types.

This assumption is reinforced when data from household sources (such as the decennial Census, the Current Population Surveys, or the National Health Interview Survey) are used to calculate estimated eligibility rates for telephone surveys. Eligibility rates calculated from household data, when applied to telephone surveys, implicitly assume that the number of telephone lines per household is the same in both eligible and ineligible sample units. The probability of selection for sample units in telephone surveys is directly related to the number of telephone lines with voice capability that access the household. Therefore, predicted eligibility rates derived from household-level survey data will result in over- or underestimation of the number of eligible households identified in screening surveys of rare populations, if the number of telephone lines per household in the rare population of interest differs substantially from the general population.

This paper uses data from the 1994 NIS, the 1990 Census, two recent CBS News surveys, and other sources to show that fewer telephone lines are found in households with young children and telephones than in the general population of households with telephones. The implications for eligibility rates in the NIS are examined using data at the national level and for specific subregions of the country. The results show that using household-level data to predict eligibility rates in the design of the RDD sample resulted in an overestimation of the eligibility rate. The results of this analysis, and the effects they have for RDD sample designs that screen for rare populations, are discussed.

## 2. An Overview of the National Immunization Survey

The prevention of disease through vaccination has been a significant success for the nation's public health system. During the 20th century, the United States has seen the incidence of childhood diseases such as measles, pertussis, and diphtheria fall by more than $98 \%$, with much of this improvement attributable to nationwide vaccination programs (Table 2-1). ${ }^{1,2}$ In 1979 the Surgeon General of the U.S. Public Health Service established a series of national health goals for 1990. One goal was that $90 \%$ of all children should have completed their basic vaccination series by age two. ${ }^{3}$ These goals were expanded in 1991 with the publication of Healthy People 2000, which again called for increased vaccination levels, especially among the very young and other high-risk groups. ${ }^{4}$

Table 2-1. Reported Cases of Vaccine-Preventable Childhood Disease in the U.S.

| Disease | Maximum Number <br> of Cases (year) | 1994 <br> (provisional) | \% Change |
| :--- | ---: | ---: | ---: |
| Diphtheria | $206,939(1921)$ | 1 | -99.9 |
| Measles | $894,134(1941)$ | 814 | -99.9 |
| Mumps | $152,209(1968)$ | 957 | -99.4 |
| Pertussis | $265,269(1934)$ | 2,203 | -99.2 |
| Poliomyelitis-paralytic | $21,269(1952)$ | 0 | -100.0 |
| Rubella | $57,686(1969)$ | 204 | -99.6 |
| Congenital Rubella Syndrome | $20,000(1964)$ | 2 | -99.9 |
| Tetanus | 1,560 deaths (1923) | 22 | -98.6 |

A series of federal, state, and local programs have been mounted to raise vaccination levels among young children. Among these programs is the Childhood Immunization Initiative (CII) established by President Clinton to create a long-term national vaccination system. The CII established a goal for 1996 of increasing vaccination levels for 2-year-old children to at least $90 \%$ for one dose of measles-mumps-rubella vaccine (MMR) and at least three doses each of diphtheria and tetanus toxoids and pertussis vaccine (DTP), oral poliovirus vaccine, and Haemophilus influenzae type b vaccine (Hib). ${ }^{5}$ In addition, CII established a goal for 1996 to increase vaccination levels for 2-year-old children to at least $70 \%$ for three or more doses of hepatitis B (Hep B) vaccine. ${ }^{6}$ During the first quarter of 1994, national vaccination levels for children 19-35
months were $89.6 \%$ for measles containing vaccine, $25.5 \%$ for three or more doses of Hep B, $70.6 \%$ for three or more doses of Hib, $87.0 \%$ for three or more doses of DTP, and $76.0 \%$ for three or more doses of polio vaccine. ${ }^{7}$

From 1959 to 1985, national vaccination levels in the U.S. were monitored by the United States Immunization Service. Beginning in the late 1980's, discussions between the National Immunization Program (NIP) and the National Center for Health Statistics (NCHS) led to the inclusion of vaccination questions on the 1991 Child Health Supplement to the National Health Interview Survey (NHIS). These questions have been included as the Immunization Supplement to the NHIS since 1992. Commencing with the 1993 data collection, the Immunization Supplement to the NHIS has provided quarterly estimates of vaccination coverage at the national level.

The National Immunization Program continues to work throughout the 50 states, the District of Columbia and the U.S. Territories and Commonwealths to raise early childhood vaccination coverage levels. To provide current baseline estimates of vaccination levels for children 19 through 35 months of age and to monitor change in these levels, the NIS is being conducted in 78 of these Immunization Action Plan (IAP) areas, consisting of the 50 states, the District of Columbia and 27 metropolitan areas. Beginning with the second quarter of 1994 and continuing through the fourth quarter of 1997, the NIS data collection effort will conduct independent quarterly surveys in each of the 78 IAP areas. The 78 IAP areas taken together will provide complete coverage of the 50 states and the District of Columbia.

## 3. NIS Design.

The NIS uses a two-phase sample design. For the first-phase, a quarterly random sample of telephone numbers for each IAP area is drawn, and a screening questionnaire is administered to identify households with one or more children 19 to 35 months of age. The first-phase sample design uses list-assisted random-digit-dialing (RDD) methods to screen the large number of households required to obtain a sample of age-eligible children efficiently. Interviews are conducted using computer-assisted telephone (CATI) data collection methods.

The second-phase is a sample of health care providers of vaccinations of children in
households surveyed during the first-phase. The second-phase is included in the design to supply the data needed to improve the accuracy of the vaccination coverage estimates. The secondphase is necessary because research showed that parents tend to underestimate the number of doses received for multiple-dose vaccines and to overestimate coverage for single-dose vaccines. Household reports of vaccinations often contain errors and the two-phase sample design permits parental reports to be combined with information from health care providers' records, yielding a more accurate estimate of vaccination coverage levels than either source alone.

In the first-phase, screening for eligible households is conducted through brief CATI interviews. When an eligible household is found, a full vaccination interview is conducted with the adult in the household who is the most knowledgeable about the vaccination histories of all age-eligible children. Data on five types of vaccinations, including dates and/or numbers of vaccination events, are obtained for each child in the target age range living within the household. To implement the second-phase provider component for the NIS, provider names, addresses, and consent to contact those providers are obtained concurrently with the first-phase CATI vaccination interview. Subsequently, a questionnaire is mailed to providers that collected vaccination information for the identified children.

These surveys are done quarterly in each of the 78 IAP areas, concurrently, to obtain detailed vaccination data on children 19 through 35 months of age. This makes it possible to combine four consecutive quarters of survey data to provide annualized estimates of the coverage rates for DTP, Polio, MMR, Hib, and Hep B with the degree of precision sufficient for analytic purposes within each of the 78 IAP areas. Also, by using the same methodology in each IAP area, the NIS can produce rates of vaccination coverage that are comparable among IAP areas with the precision of these estimates achieving a coefficient of variation of no more than $5 \%$ for an estimated proportion of .50 .

## 4. Methods

The need to collect vaccination data independently for each of the 78 IAP areas quarterly, combined with the small proportion of households in the United States that contain a child 19 through 35 months of age, requires a large initial sample of telephone numbers. For the survey
year 1994, consisting of three quarters, the total sample size was $1,177,140$ randomly generated telephone numbers. The NIS design was based on the estimate that, on average, identifying one household with one or more age-eligible children will be necessary to complete screening interviews with 20 households (i.e., an eligibility rate that averages 5\% across all U. S. households with telephones). Table 4-1 compares the expected and observed sample sizes for the NIS at several key stages of the sample implementation. These data show the impact of the lower than expected eligibility rates on the number of sample cases required to approximate design goals for the number of completed cases. Despite these increases, the number of eligible households fell short of the design goal of 30,282 by $13.4 \%$,

Table 4-1. Comparison of Expected and Observed Sample Size Characteristics for the 1994 National Immunization Survey

| 1994 National Immunization Survey |  |  |  |
| :--- | ---: | ---: | ---: |
| Sample Stage | Expected | Observed | \% Change, <br> Expected to <br> Observed |
| Total CATI Sample Size | 961,345 | $1,177,140$ | $+22.4 \%$ |
| Working Residential Numbers | 628,078 | 712,956 | $+13.5 \%$ |
| Completed Screening Interviews | 605,647 | 643,328 | $+6.2 \%$ |
| Households Screened, Not Eligible | 575,365 | 617,113 | $+7.3 \%$ |
| Households Screened, Eligible | 30,282 | 26,215 | $-13.4 \%$ |
| Eligibility Rate | 5.0 | 4.1 | $-18.0 \%$ |

Regional data. The number of data sources that permit examination of the relationship between the number of telephone lines per household and expected eligibility rates is limited. After repeated contacts with Ameritech, the local service provider for Illinois, Indiana, Michigan, Ohio, and Wisconsin, data were obtained from Ameritech sources that showed the percentage distribution of multiple line households in this five-state region. These data show that $89 \%$ of households in this region have a single telephone line, $10 \%$ have two lines, and $1 \%$ have three or more telephone lines. Data from the NIS can be used to calculate the distribution of telephone lines in households that contain a child 19 through 35 months of age. These data show that only $2.4 \%$ of households with children 19-35 months of age have two or more telephone lines. This is
in contrast to the $11.0 \%$ level observed for all households.
By combining data for these five states from the NIS, the Census PUMS data, and data from the Statistical Abstract, estimating eligibility rates based on the number of telephone lines connected to households is possible, rather than the number of households. If differences exist between households with age-eligible children and other households in the number of unique voice-capable telephone lines (i.e., telephone numbers) connected to them, estimates calculated using the number of telephone lines may vary from eligibility estimates calculated based on household eligibility. Table 4-2 shows the results of these calculations. The key numbers in the table are:
(a) Ameritech data on the proportion of households with multiple telephone lines;
(b) NIS data (weighted) on the number of telephone lines in eligible households in the IAPs that comprise the five-state region;
(c) Census data (PUMS) on the estimated eligibility rate at the household level (5.5\%), for these five states;
(d) Census data on the total number of households in the five-state region.

The estimated distribution column, under the Households with Telephones heading, in the All Households row, shows the Ameritech data. Applying these rates to the total number of households with telephones for the region $(15,266,630)$ yields the distribution of households shown in the Number column. Applying the Census-based household eligibility rate of 5.5\% (Household Based Predicted Eligibility Rate) yields the total number of eligible households in the region $(834,050)$.

At this stage, the NIS rate for multiple line households can be used to form the distribution of eligible households by the number of telephone lines connected to those households. Next, the same distribution can be formed for ineligible households by subtraction, both for the estimated distribution and for the number of households.

Extrapolating from the household data, the right-hand three columns of Table 4-2 show the number of telephone lines in the region. For example, if $13,587,301$ households have a

Table 4-2. Comparison of Predicted Eligibility Rates for Five Midwestern States
Using Census Household Data, NIS Data, and Ameritech Data

| Household Type | Number of Telephone Lines in Household | Households with Telephones |  | Household Based Predicted Eligibility Rate | Telephone Lines |  | Telephone Based Predicted Eligibility Rate |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Estimated Distribution | Number |  | Estimated Distribution | $\begin{gathered} \hline \text { Number of } \\ \text { Telephone Lines } \end{gathered}$ |  |
| All Households | One | 89.0\% ${ }^{\text {b }}$ | 13,587,301 |  | 79.4\% | 13,587,301 |  |
|  | Two | 10.0\% | 1,526,663 |  | 17.9\% | 3,053,326 |  |
|  | Three or More | 1.0\% | 152,666 |  | 2.7\% | 457,999 |  |
|  | All | 100.0\% | 15,266,630 ${ }^{\text {c, } 8}$ | N/A | 100.0\% | 17,098,626 | N/A |
| Eligible Households | One | 97.6\% ${ }^{\text {d }}$ | 814,295 |  | 95.3\% | 814,295 |  |
|  | Two | 2.3\% | 19,374 |  | 4.5\% | 38,748 |  |
|  | Three or More | 0.1\% | 381 |  | 0.2\% | 1,143 |  |
|  | All Eligible | 100.0\% | 834,050 | 5.5\% ${ }^{\text {e }}$ | 100.0\% | 854,186 | 5.0\% |
| Ineligible Households | One | 88.5\% ${ }^{\text {I }}$ | 12,773,006 |  | 78.6\% | 12,773,006 |  |
|  | Two | 10.4\% | 1,507,289 |  | 18.6\% | 3,014,578 |  |
|  | Three or More | 1.1\% | 152,285 |  | 2.8\% | 456,856 |  |
|  | All Ineligible | 100.0\% | 14,432,580 | 94.5\% | 100.0\% | 16,244,440 | 95.0\% |

${ }^{\mathrm{b}}$ Data on the distribution of telephone lines for all households were obtained from Ameritech corporate data. Ameritech is the local telephone service provider for the five states included in this analysis: Illinois, Indiana, Michigan, Ohio, and Wisconsin.
${ }^{\text {C }}$ Estimates of the number of households with telephones for these five states are from 1993 Census data on the total number of households (Statistical Abstract of the Unites States, 1994), adjusted for a telephone ownership rate of $95.5 \%$ ( 1990 Census PUMS data).
${ }^{d}$ Estimates of the distribution of the number of telephone lines for eligible households are calculated from weighted NIS data for all IAP areas, using the combined data for four quarters of NIS data collection beginning in the second quarter of 1994 through the first quarter of 1995.
${ }^{\mathrm{e}}$ The 0.055 estimate of the eligibility rate among households with telephones was calculated from the 1990 Census PUMS data file for Illinois, Indiana, Michigan, Ohio, and Wisconsin
${ }^{\mathrm{f}}$ Calculated from Ameritech data in comb ination with 1990 Census PUMS estimate of the household ineligibility rate of 0.945 for Illinois, Indiana, Michigan, Ohio, and Wisconsin.
single telephone line (see the All Households row and the Number column under the Households with Telephones heading), this implies that these households contain a total of 13,587,301 telephone lines. Similarly, the 1,526,663 households with two telephone lines imply that these households contain 3,053,326 telephone lines ( $2 \times 1,526,663$ ). Through this process the total number of telephone lines in the region can be calculated $(17,098,626)$ and the distribution of age -eligible children by telephone lines can be calculated. This extrapolation from the household data to telephone line data was carried out for the balance of the rows in Table 4-2, yielding the following results:
(a) the total number of telephone lines in the region is $17,098,626$;
(b) 854,186 of these reach households with children 19 through 35 months of age, or 5.0\% of all telephone lines in the region.

These data show, for the East North Central Census region, that the proportion of households, with telephones, that would be encountered in an RDD design is $10 \%$ lower than would be expected if the eligibility rate is calculated using household data. This is due to the much lower level ( $11 \%$ versus $2.4 \%$ ) of multiple telephone lines in households that also contain a child 19-35 months of age.

National level estimates using Census data. Unfortunately, the Ameritech data do not exist at the national level. Other sources of information on the distribution of telephone lines across households were required to continue the analysis. A convenient source of data on the distribution of telephone lines in households is a large national telephone survey, since to calculate we ights for the probability of selection it is necessary that the survey questionnaire asks questions about the number of voice-capable telephone lines in the household. Data from two CBS News/New York Times National Surveys were used to calculated this dis tribution, one conducted between August 5 and 9, 1995, the other conducted between September 18 and 22, 1995. Using weighted data from these surveys it is possible to calculate the national distribution of households with one, two, and three or more telep hone lines. The results show that nationally, $8.4 \%$ of all households have two or more telephone lines. ${ }^{9}$

Table 4-3 contains the results of analysis that parallels the analysis shown in Table $4-2$,

Table 4-3. Comparison of Predicted Eligibility Rates for U.S. Households
Using 1993 Census Household Data

| Household Type | Number ofTelephone Lines in Household | Households with Telephones |  | HouseholdBasedPredictedEligibilityRate | Telephone Lines |  | TelephoneBasedPredictedEligibility Rate |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Estimated Distribution | Number ${ }^{\text {a10 }}$ |  | Estimated Distribution | Number of Telephone Lines |  |
| All Households | One | $91.6 \%{ }^{\text {b }}$ | 83,721,136 |  | 83.6\% | 83,721,136 |  |
|  | Two | 7.1\% | 6,497,023 |  | 13.0\% | 12,994,047 |  |
|  | Three or More | 1.3\% | 1,160,509 |  | 3.4\% | 3,481,527 |  |
|  | All | 100.0\% | 91,378,688 | N/A | 100.0\% | 100,196,710 | N/A |
| Eligible Households | One | 97.2\% ${ }^{\text {c }}$ | 4,900,739 |  | 94.5\% | 4,900,739 |  |
|  | Two | 2.7\% | 136,645 |  | 5.3\% | 273,290 |  |
|  | Three or More | 0.1\% | 4,568 |  | . $02 \%$ | 13,703 |  |
|  | All | 100.0\% | 5,041,951 | 5.5\% ${ }^{\text {d,11 }}$ | 100.0\% | 5,187,732 | 5.2\% |
| Ineligible Households | One | 91.3\% ${ }^{\text {c }}$ | 78,820,397 |  | 83.0\% | 78,820,397 |  |
|  | Two | 7.4\% | 6,360,378 |  | 13.4\% | 12,720,757 |  |
|  | Three or More | 1.3\% | 1,155,942 |  | 3.6\% | 3,467,825 |  |
|  | All | 100.0\% | 86,336,717 | 94.5\% | 100.0\% | 95,008,979 | 94.8\% |

${ }^{\text {a }}$ Estimates of the number of households in the U.S. with telephones are based on 1993 Census data (Stati stical Abstract of the United States, 1994), adjusted for a telephone ownership rate of $94.8 \%$ nationally ( 1990 Census PUMS data).
${ }^{\mathrm{b}}$ Data on the distribution of telephone lines for all households are from CBS News National Telephone Surveys conducted during the fall of 1995. Data obtained through the Roper Center.
${ }^{c}$ Estimates of the distribution of the number of telephone lines for eligible households are calculated from weighted NIS data for all IAP areas, using the combined data for four quarters of NIS data collection beginning in the second quarter of 1994 through the first quarter of 1995.
${ }^{d}$ The 0.055 estimated eligibility rate for households with telephones was calculated from 1990 Census PUMS data for the U.S.
${ }^{e}$ Calculated from CBS News data in combin ation with the 1990 Census PUMS estimated ineligibility rate of 0.945 for the United States for households with telephones.
but at the national level. The CBS News/New York Times data were used as the basis of the distribution of multiple telephone line households. NIS data for all 78 IAP areas, weighted appropriately, were used to estimate the distribution of multiple telephone lines in households that contain a child 19 through 35 months of age. Census data from the PUMS files were used to calculate the household based predicted eligibility rate. Data from the 1994 Statistical Abstract of the United States were the source of information on the number of households.

The estimated eligibility rate for the U.S. calculated from 1990 Census data is $5.5 \%$, s imilar to the rate shown above for the five-state East North Central region. Based on the distribution of the number of telephone lines in eligible and ineligible households, these data predict a telephone -based eligibility rate of $5.2 \%$ (Table $4-3$, right-hand most column.) This is $5.5 \%$ lower than the level predicted from the household-level data.

National-level estimates using NHIS data. Because the 1990 Census data are several years old, they may not accurately predict estimated eligibility rates for a survey to be conducted several years later. More recent data are available from the NHIS from 1993. The NHIS is a national area probability survey conducted by the National Center for Health Statistics. The data used for this analysis are from the 1993 NHIS, the most recent year for which complete annual data are available. Using the NHIS data, repeating the previous analysis is possible (Table 4-3). The results of this additional analysis are shown in Table 4-4.

An eligibility rate of $4.9 \%$ was calculated from the NHIS data for households with telephones and children 19 through 35 months of age. This rate is much lower than the $5.5 \%$ household eligibility rate calculated from national Census data and demonstrates an additional problem in predicting eligibility rates in surveys of rare populations: the estimated eligibility rate may vary greatly, depending on the data on which the calculations are based.

Following previously described algorithms, the NHIS eligibility rate is recalculated based on the CBS News/New York Times survey data for the distribution of the number of telephone lines in the general population. These data are used with the NIS information on the number of telephone lines in eligible households and Census data on the number of ho useholds in the U.S. to estimate an eligibility rate for telephone numbers that reach households with children 19 through 35 months of age. This rate, shown in the right-hand most column of Table 4-4 is calculated at 4.6\%. This is very much closer to the eligibility rate observed for the NIS of $4.1 \%$.

Table 4-4. Comparison of Predicted Eligibility Rates for U.S. Households Using 1993 National Health Interview Survey Eligibility Rates

| Household Type | Number ofTelephone Lines inHousehold Household | Households with Telephones |  | Household Based Predicted Eligibility Rate | Telephone Lines |  | TelephoneBasedPredictedEligibility Rate |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Estimated Distribution | Number ${ }^{\text {a,12 }}$ |  | Estimated Distribution | Number of Telephone Lines |  |
| All Households | One | $91.6 \%{ }^{\text {b }}$ | 83,721,136 |  | 83.6\% | 83,721,136 |  |
|  | Two | 7.1\% | 6,497,023 |  | 13.0\% | 12,994,047 |  |
|  | Three or More | 1.3\% | 1,160,509 |  | 3.4\% | 3,481,527 |  |
|  | All | 100.0\% | 91,378,688 | N/A | 100.0\% | 100,196,170 | N/A |
| Eligible Households | One | 97.2\% ${ }^{\text {c }}$ | 4,348,916 |  | 94.5\% | 4,348,916 |  |
|  | Two | 2.7\% | 121,259 |  | 5.3\% | 242,517 |  |
|  | Three or More | 0.1\% | 4,053 |  | . $02 \%$ | 12,160 |  |
|  | All | 100.0\% | 4,474,228 | 4.9\% ${ }^{\text {d,13 }}$ | 100.0\% | 4,603,593 | 4.6\% |
| Ineligible Households | One | 91.3\% ${ }^{\text {e }}$ | 79,372,219 |  | 83.0\% | 79,372,219 |  |
|  | Two | 7.4\% | 6,375,765 |  | 13.4\% | 12,751,530 |  |
|  | Three or More | 1.3\% | 1,156,456 |  | 3.6\% | 3,469,367 |  |
|  | All | 100.0\% | 86,904,440 | 95.1\% | 100.0\% | 95,593,116 | 95.4\% |

${ }^{\text {a }}$ Estimates of the number of households with telephones for the U.S. are based on 1993 Census data (Statistical Abstract of the United States, 1994), adjusted for a telephone ownership rate (nationally) of 04.8\% (1990 Census PUMS data).
${ }^{\mathrm{b}}$ Data on the distribution of telephone lines for all households are from CBS News National Telephone Surveys conducted during the fall of 1995.
Data obtained through the Roper Center.
${ }^{\mathrm{c}}$ Estimates of the distribution of the number of telephone lines for eligible households are calculated from weighted NIS data for all IAP areas, using the combined data for four quarters of NIS data collection beginning in the second quarter of 1994 through the first quarter of 1995.
${ }^{d}$ The 0.049 estimated eligib ility rate for households with telephones was calculated from 1993 National Health Interview Survey data, weighted with the basic household weight adjustment by segment and special household sampling factors.
${ }^{\mathrm{e}}$ Calculated from CBS News data in combination with 1993 NHIS estimated ineligibility rate for telephone householdsof 0.9448 for the U.S. total.

## 5. Summary

Despite the limitations of the data used in these analyses, they show the difficulty of estimating eligibility rates in any survey of rare populations, but especially for telephone surveys that screen for rare populations. The analyses presented in this paper demonstrate that for telephone surveys the distribution of multiple telephone lines in the eligible and ineligible segments of the population can results in a lower than expected eligibility rate. While these results are for households that contain young children, the results may apply to other segments of the population that share similar socio economic and demographic characteristics (for example, low income).

These results suggest that the true eligibility rate for the NIS may range between $4.1 \%$ (the observed NIS rate) and $5.5 \%$ (the Census estimated rate). Because of these widely disparate estimates, it is important that at the design stage the implications of error in estimates of the eligibility rate are taken into consideration. In the case of the NIS, the estimated eligibility rate of $5.0 \%$ was used, lower than the Census rate, but potentially higher than the rate adjusted for the distribution of multiple telephone line households calculated from the NHIS data.

In addition, recent news releases in the past 12 to 18 months suggest that an explosive growth may have occurred recently in the number of households that have multiple telephone lines . A recent article in the New York Times concluded that "the percentage of households with more than one line has nearly doubled this year [1995], to 16 percent." ${ }^{14}$ In part, this growth partially has been the result of the recent increase in sales of home computers with modems. Some portion this growth, however, is attributable to an increase in the proportion of households with multiple voice -capable telephone line. Given recent trends, this proportion is very likely to continue to increase into the ne ar future.

This increase, as noted in the New York Times article, is concentrated in more affluent households. This outcome, if true, will increase the likelihood that the estimated eligibility rates for telephone surveys of rare populations will deviate significantly from the observed eligibility rates if appropriate steps are not taken to account for this differential in the design phase of the survey.

## 6. References and Notes

1. Hinman, A.R. 1990. "Immunizations in the United States." Pediatrics. 1064-1066.
2. United States Department of Health and Human Services, Public Health Service. Morbidity and Mortality Weekly Report, October 7, 1994. Volume 43, No. 39.
3. United States Department of Health and Human Services, Public Health Service. 1980. Promoting Health, Preventing Disease: Objectives for the Nation.
4. United States Department of Health and Human Services, Public Health Service. 1991. Healthy People 2000: National Disease Prevention and Health Promotion Objectives for the Nation.
5. United States Department of Health and Human Services, Public Health Service. Morbidity and Mortality Weekly Report, October 7, 1994. Volume 43, No. 39.
6. United States Department of Health and Human Services, Public Health Service. Morbidity and Mortality Weekly Report, October 7, 1994. Volume 43, No. 39.
7. United States Department of Health and Human Services, Public Health Service. Morbidity and Mortality Weekly Report, March 3, 1995. Volume 44, No. 8
8. U.S. Bureau of the Census, Statistical Abstract of the United States: 1994 (114 ${ }^{\text {th }}$ edition.) Washington, DC, 1994.
9. The authors wish to thank Michael Butterworth of CBS News for directing them to the CBS/News/New York Times surveys as a source of these data. In addition we also wish to thank George Pettinico of the Roper Center for Public Opinion Research for assisting in the analysis of these data.
10. U.S. Bureau of the Census, Statistical Abstract of the United States: 1994 (114 ${ }^{\text {th }}$ edition.) Washington, DC, 1994.
11. Estimates of the household eligibility rate are based on analysis of data from the United States Census Bureau of the Census. Census of Population and Housing, 1990: Public Use Microdata Samples U.S. Washington, DC, 1992.
12. U.S. Bureau of the Census, Statistical Abstract of the United States: 1994 (114 ${ }^{\text {th }}$ edition.) Washington, DC, 1994.
13. Estimates of the household eligibility rate are based on analysis of data from the United States Census Bureau of the Census. Census of Population and Housing, 1990: Public Use Microdata Samples U.S. Washington, DC, 1992.
14. "Multiple Family Phone Lines, a Post -Postware U.S. Trend." New York Times, December 26, 1995. This article cited several sources, including research conducted by IDC/Link Resources and local telephone service providers, such as Nynex Corporation.
