DOCUMENT RESUME

ED 446 154 TM 031 973

AUTHOR Henke, Robin R.; Kaufman, Phillip; Broughman, Stephen P.; Chandler, Kathryn

TITLE Issues Related to Estimating the Home-Schooled Population in the United States with National Household Survey Data. Technical Report.

INSTITUTION National Center for Education Statistics (ED), Washington, DC.; MPR Associates, Berkeley, CA.

REPORT NO NCES-2000-311

PUB DATE 2000-10-00

NOTE 107p.

AVAILABLE FROM ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free).

PUB TYPE Reports - Evaluative (142)

EDRS PRICE MF01/PC05 Plus Postage.

DESCRIPTORS Census Figures; *Estimation (Mathematics); *Home Schooling; National Surveys; Research Design; *Research Methodology; Tables (Data)

IDENTIFIERS *Current Population Survey; *National Household Education Survey

ABSTRACT

Estimates of the number and proportions of students who are home schooled vary. This report explores differences in survey design and execution that may have contributed to the discrepancies between the estimates for the October 1994 Current Population Survey Education Supplement (CPS:Oct94) and the National Household Education Survey, Parent and Family Involvement/Civic Involvement component of 1996 (NHES:96). Data for the CPS:Oct94 came from household sampling that includes two interviews and six telephone follow-ups for each family surveyed. The NHES:96 was a random-digit dial telephone survey. Differences in sample coverage and response rates are discussed. Missing data are not likely to affect the CPS:Oct94 estimate greatly, but the effect of the lower NHES:96 household response rate cannot be estimated. Observational errors can be introduced by data collection procedures, survey instruments, and respondents. The impact of data processing errors should also be considered, but, in fact, the potential direction and magnitude of estimate differences could not be predicted for most of the methodological differences between the two surveys. The study raises research questions that should be examined to ensure accuracy in estimating the numbers of home schooled students. Appendixes contain technical notes and standard error tables. (Contains 22 tables, 7 figures and 30 references.) (SLD)

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report

October 2000

Issues Related to Estimating the Home-Schooled Population in the United States With National Household Survey Data

U.S. Department of Education
Office of Educational Research and Improvement

NCES 2000-311

Customer Survey of OERI Publication Users

To help us improve future editions of this publication and give you better customer service, we would appreciate your comments on this survey form. Please check the appropriate box(es) below for each question. Responses will be kept completely confidential. You may return the survey by mail or FAX. It can be folded and taped closed to allow mailing to the address listed on the reverse side of this form, or it can be returned by FAX to 202–219–1321. **Many thanks for your customer feedback—it is very important to us!**

1a. Name of publication

1b. Publication number

1c. Author name

2. How did you receive a copy of this publication?

- ❑ Bought it
- ❑ Borrowed it
- ❑ Mailing list membership
- ❑ Telephone request
- ❑ Internet request
- ❑ Other (please describe)

3. Was this publication easy to get?

❑ Very ❑ Somewhat ❑ Not at all

4. How did you find out about this and other OERI publications? (Check all that apply.)

- ❑ Conferences
- ❑ Journal articles
- ❑ Teacher/educator
- ❑ Professional associations
- ❑ Internet (WWW)
- ❑ Publication announcement
- ❑ Received in mail
- ❑ OERI staff contact

5. For what purposes did you use this OERI publication? (Check all that apply.)

- ❑ Planning
- ❑ Policy or legislation
- ❑ Administrative decisions
- ❑ Teaching, class material
- ❑ Research/analysis
- ❑ General information
- ❑ Writing news articles, TV or radio material
- ❑ Marketing, sales, or promotion
- ❑ Other (please describe)

6. Did the publication help you accomplish whatever you needed it for?

❑ Yes ❑ No ❑ Partially

7. What is your occupation?

❑ Parent ❑ Teacher ❑ Administrator
❑ Librarian ❑ Researcher ❑ Statistician
❑ Journalist/writer ❑ Policy Analyst ❑ Student
❑ Program Planner
❑ Other (please specify)

8. How could this OERI publication (or other OERI publications) better meet your needs? (Check all that apply.)

- ❑ More important topics in education
- ❑ More timely release of data
- ❑ More text introductions to each section
- ❑ More research statistics
- ❑ Shorter reports (less than 10 pages)
- ❑ Other (please describe)

9. Overall, how satisfied are you with this product?	Very Satisfied ❑	Satisfied ❑	Dissatisfied ❑
a. Comprehensiveness of information	❑	❑	❑
b. Clarity of writing (readability, interpretability)	❑	❑	❑
c. Clarity of presentation (e.g., tables, charts)	❑	❑	❑
d. Timeliness of information	❑	❑	❑
e. Accuracy of information	❑	❑	❑
f. Clarity of technical notes	❑	❑	❑
g. Usefulness of resources and bibliography	❑	❑	❑
h. Organization	❑	❑	❑
i. Length	❑	❑	❑
j. Format	❑	❑	❑

PAPERWORK BURDEN STATEMENT
Office of Education Research and Improvement (OERI)
Publication Customer Survey

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 1880-0529. The time required to complete this information collection is estimated to average 10 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have any comments concerning the accuracy of the time estimate(s), suggestions for improving this form, or concerns regarding the status of your individual submission of this form, write directly to: **Information, Room 204, Media and Information Services, OERI, U.S. Department of Education, 555 New Jersey Avenue NW, Washington, DC 20208-5570.**

OERI Publication Customer Survey
Media and Information Services
U.S. Department of Education
Washington, DC 20202

Official Business
Penalty for Private Use, $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST-CLASS MAIL PERMIT NO. 012935 WASHINGTON DC

POSTAGE WILL BE PAID BY U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
Mail Code: 5570
400 Maryland Avenue, S.W.
Washington, D.C. 20277—2935

Fold on line—TAPE CLOSED—DO NOT STAPLE

10. Do you have any suggestions regarding the content or format of future editions of this publication or other comments?

NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report October 2000

Issues Related to Estimating the Home-Schooled Population in the United States With National Household Survey Data

Robin R. Henke
Phillip Kaufman
MPR Associates, Inc.

Stephen P. Broughman
Kathryn Chandler
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement NCES 2000-311

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
C. Kent McGuire
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue NW
Washington, DC 20208–5652

October 2000

The NCES World Wide Web Home Page is: ***http://nces.ed.gov***
The NCES World Wide Web Electronic Catalog is: ***http://nces.ed.gov/pubsearch/index.asp***

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *Issues Related to Estimating the Home-Schooled Population in the United States With National Household Survey Data,* NCES 2000–311, by Robin R. Henke and Phillip Kaufman. Project Officer: Stephen P. Broughman. Washington, DC: 2000.

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

or call toll free 1–887–4ED–Pubs.

Content Contact:
Stephen P. Broughman
(202) 502–7315

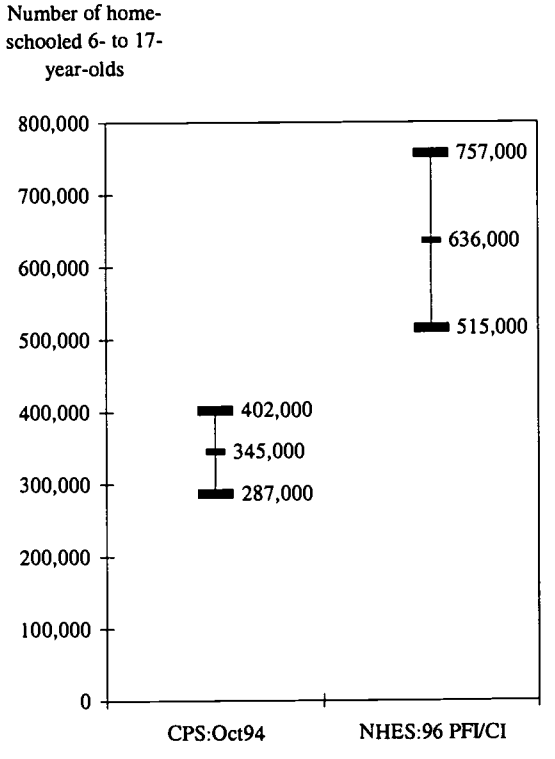
Executive Summary

Home schooling in the United States has become a topic of interest to education policymakers, administrators, and the general public. Currently, published estimates of the number of children who are home schooled vary by hundreds of thousands of children and are of uncertain reliability. Informed discussions of home-schooling policy are compromised without accurate estimates of how many children are educated at home and whether the proportion of children who are so educated is changing.

Estimates of the number and proportion of students who were home schooled derived from two sets of national survey data from the mid-1990s—the October 1994 Current Population Survey Education Supplement (CPS:Oct94); and the 1996 National Household Education Survey, Parent and Family Involvement/Civic Involvement component (NHES:96, PFI/CI)—also vary. The point estimates of the number of children ages 6 to 17 who were home schooled ranged from 345,000 in CPS:Oct94 to 636,000 in NHES:96 (figure A). Taking estimated sampling variance into account, the 95-percent confidence interval around the CPS:Oct94 point estimate ranges from 287,000 to 402,000, and the 95-percent confidence interval from 515,000 to 757,000 around the NHES:96 point estimate. According to CPS:Oct94, 0.8 percent of children were home schooled, and according to NHES:96, 1.4 percent of children were home schooled.

Although the differences between these surveys' estimates may reflect growth in the number and proportion of students who are home schooled, it seems unlikely that the number of home-schooled children nearly doubled in less than two years (Lines 1998; Ray 1999). This report explores differences in survey design and execution that may have contributed to these two different estimates.

Figure A—CPS:Oct94 and NHES:96 point estimates and their 95-percent confidence intervals of number of home-schooled 6- to 17-year-olds



SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement/Civic Involvement Component (NHES:96, PFI/CI).

The report is based on the premise that for any given year, there is some "true" number of homeschooled children in the population. Point estimates derived from the CPS and NHES depart from this true value by some amount of error. Errors in surveys include errors in nonobservation, errors in observation, and data processing errors (Groves 1991). After describing the data sources, this report examines each type of error.

Data Sources

For decades the Census Bureau has conducted the Current Population Survey (CPS) each month on behalf of the Bureau of Labor Statistics in order to study labor force participation and unemployment. The CPS includes a set of basic labor force and demographic questions that are repeated each month and a supplement, whose topic varies from month to month. Each October's supplement focuses on participation in education programs for civilians age 3 and older, and in 1994 the October supplement included questions related to home schooling.

CPS samples households using addresses from the most recent Decennial Census and updates to it as the sampling frame. Each sampled household is part of CPS for 8 months. In its first- and fifth-month interviews, the household's interview is conducted in person: a Census Bureau interviewer visits the home and conducts the interview with a laptop computer. With the household's permission, the remaining six interviews are conducted by telephone. Interviewers attempt to speak with the most knowledgeable person in the household, although any household member 15 years old or older may serve as the respondent. Respondents answer questions regarding all household members.

The National Center for Education Statistics (NCES) has conducted the National Household Education Survey (NHES) five times since the first administration in 1991. NCES uses NHES to collect data on education issues on which households, rather than education institutions, are best able to provide data. Each time the NHES is fielded, a Screener interview is used to determine whether the household includes members who are eligible for either of two extended topical interviews. In 1996, one of these interviews—the Parent and Family Involvement/Civic Involvement (PFI/CI) component—of the NHES included questions on children's schooling, including home schooling. The PFI/CI component sampled children from age 3 through 12th grade, with a maximum age of 20.

NHES is a random-digit-dial (RDD) telephone survey, i.e., it samples households via telephone numbers. Interviewers in telephone centers use computer-assisted telephone interviewing (CATI) to conduct interviews from January through April of the administration year. Interviewers ask to speak with a household member at least 18 years old, who responds to the Screener questions. In 1996, once the interviewer determined through the Screener that a child in the household was eligible for the PFI/CI, the interviewer asked to speak with the parent or guardian who knew the most about the sampled child's care and education.

Impact of Nonobservational Errors

Errors of nonobservation occur when members of the target population are excluded from the sampling frame or when sampled members of the population fail to participate in the survey or some part thereof. This report discusses both of these sources of nonobservation error: sample coverage and nonresponse.

Sample Coverage

Both the CPS and NHES sampling frames undercover some groups within the U.S. noninstitutionalized population, although each undercovers different segments. The Census Bureau estimates that CPS undercovers between 7 and 13 percent of infants through 19-year-olds in the population. Among children, males, blacks, and older children are more likely than females, nonblacks, and younger children to be missed. Sampling weights adjust for undercoverage with respect to these demographic characteristics, but to the extent that undercovered groups home school at rates different from the general population, these weights may not eliminate error in estimates related to home schooling. However, because home schooling is a rare event and the rates of undercoverage are low, even if the relatively small undercovered groups were home schooled at rates considerably higher or lower than the general population, the error in the estimates would be small.

NHES has two primary sources of undercoverage: the exclusion of nontelephone households and the exclusion of some residential telephone numbers due to the particular method of random digit dialing used to sample households. CPS:Oct94 data indicate that approximately 6 percent of households did not have telephones. Sampling weights adjust NHES estimates to population controls derived from the Census, and therefore adjust for the undercoverage of households without telephones. CPS:Oct94 data indicate that children in nontelephone households were home schooled at the same rate as children in telephone households, and therefore there is no evidence of error due to the exclusion of nontelephone households.

To reduce costs, the NHES uses the list-assisted method of random digit dialing, and studies of the list-assisted method indicate that 3 to 4 percent of residential telephone numbers are excluded from the sampling frame when this method is used. It is not possible to determine empirically whether children in these households are more or less likely to be home schooled than are children in included households. However, the rate of home schooling is generally low and the proportion of excluded households is small. Therefore, even if the rate of home schooling were considerably different among excluded households compared with included households, the potential error in the estimated number and percentage of home-schooled children would be small.

Although there is some potential for error in the studies' sampling frames, neither of the studies' sample designs appears to be biased. Both studies sample randomly from households within their frames and oversample some minority groups to collect sufficient data for reliable estimates concerning those groups. NHES:96 PFI/CI randomly sampled children within households, depending on the number of children who were eligible to participate within a household.

Response Rates

Response rates were calculated at three levels—household, supplement or extended interview, and item—for each survey. The CPS:Oct94 household response rate (94 percent) is considerably higher than the NHES:96 Screener response rate (70 percent). The low household response rate in the NHES allows for the possibility that homeschooling families, who may not wish to be identified or involved in government-related research (Kaseman and Kaseman 1991), may have participated at a lower rate than other families. However, because families with children in grades K–12 make up approximately 30 percent of the population of households in the United States, approxi-

mately 9 percent, rather than the entire 30 percent, of nonresponding households might include children in the desired age/grade range who were home schooled. At the second level, supplement in CPS and PFI/CI interview in NHES, the CPS again had a higher response rate than the NHES (97 percent compared with 89 percent, respectively).

In both surveys the item response rates were high for items used in these analyses. Among the items that identify home-schooled children in CPS:Oct94, all of the items had response rates of at least 92 percent, and nearly all relevant items in NHES:96 had item completion rates approaching 100 percent. It appears, therefore, that families who participated in the surveys were not unwilling to discuss home schooling.

However, because missing data for many items were not imputed in the CPS:Oct94 data set, some cases had to be excluded from the CPS analyses because it was not possible to determine whether they met the criteria that defined the sample or whether they were home schooled. The excluded cases represented about 2 million of the 46 million 6- to 17-year-olds in the United States. If the excluded children were home schooled at the same rate as children who were included, approximately 30,000 additional children would be home schooled. However, the characteristics of excluded children, especially age, suggest that excluded children may well be home schooled at a lower rate than included children.

Thus, although missing data may bias the CPS:Oct94 estimate, they are not likely to affect it greatly. The effect of the lower NHES:96 household response rate cannot be estimated.

Impact of Observational Errors

Observational errors can be introduced by data collection procedures, survey instruments, and respondents.

Data Collection

The surveys differ with respect to data collection procedures in at least three ways. First, although both surveys are conducted with computer-assisted interviewing (CAI), CPS interviewers use both personal interviewing (CAPI) and telephone interviewing (CATI), whereas NHES interviews are conducted entirely via telephone interviewing. Whether and how personal interviewing, compared with telephone interviewing, might produce different results with respect to home schooling is unknown.

Second, CPS is a panel survey, whereas NHES is not. The effects of this, aside from potential differences in response rates (which were examined separately), cannot be assessed with available data.

Third, the surveys also differ with respect to timing. In addition to the 15- to 18-month span between surveys' administration, the two surveys differ in the time of year at which they were administered. The NHES is administered from January through April, in contrast to the October administration of the CPS:Oct94. To the degree that parents are more likely to home school their children at some times of the year than at others, the difference in survey timing may contribute to the difference between the estimates.

Instrument Error

This report examined question wording, question sequencing, and the location of home-

schooling items as potential sources of instrument error.

Question Wording and Sequencing

The questions regarding home schooling were worded differently between the two studies and even among interviews in CPS:Oct94. CPS:Oct94 interviews varied depending on the age and enrollment status of the person about whom the interview was being conducted. Regarding enrolled children, the question as to whether the child was "schooled primarily at home" allows for the possibility that children who were schooled partly at home and partly at school were not identified as home schooled. NHES:96 PFI/CI interviewers first asked whether children were enrolled and then asked, regardless of enrollment status, whether children were schooled at home. When respondents indicated that a child was home schooled, the interviewer clarified the response by asking whether the child was schooled at home "instead of at school." It is not clear how parents who schooled their children partly at home and partly at school might have responded to these items.

In addition to the difference in wording discussed above, the number of items and the complexity of their sequence are considerably greater in CPS:Oct94 than in NHES, creating more opportunities for missing or inaccurate responses. Although the greater number and complexity of the items in CPS:Oct94 does not appear to have affected response rates, as they were consistently high, whether it affected the quality of responses cannot be determined with the available data.

Respondent Fatigue

When surveys become too long, respondents often begin to tire or lose interest, a phenomenon known as "respondent fatigue." As a consequence of this fatigue, questions near the end of a long survey often have higher rates of nonresponse and responses to these questions can be less accurate than responses to questions near the beginning of the survey.

The issue of respondent fatigue was addressed because the CPS education supplement questions regarding children's schooling occur near the end of the interview, after the basic labor force and supplement items for adults are asked. In contrast, the NHES items regarding children's schooling occur at the very beginning of the PFI/CI interview.

It appears unlikely that this difference has affected these data. As noted above, the response rates to the supplement items regarding home schooling are high, which indicates that fatigue did not affect response rates greatly. In addition, in CPS:Oct94, household interviews that included supplement interviews for children aged 6 to 17 years old averaged 15 minutes in length. Given this relatively short duration, fatigue is not likely to have been a problem. However, whether fatigue did occur and affected the quality of responses cannot be determined with these data.[1]

Respondent Error

Respondents' knowledge of the survey topic affects their ability to answer questions accurately. Therefore, respondents' relationships to the chil-

[1]Although the NHES:96 PFI/CI interviews were longer—19 minutes in addition to the 6-minute Screener interview—the home-schooling questions were asked at the beginning of the extended interview and are thus relatively safe from the effects of respondent fatigue.

dren about whom the home-schooling questions were asked may affect the accuracy of their answers. In addition, the political and cognitive contexts within which questions are asked and answered may affect respondents' answers.

Respondents' Relationships to Children

The CPS:Oct94 respondents could be different from the respondents to the NHES PFI/CI interviews because the instructions given to interviewers for choosing respondents differed between the two surveys. In CPS:Oct94, any household member 15 years old or older was eligible to respond for all household members, although interviewers were instructed to interview the most knowledgeable adult in the household if possible. In NHES:96 PFI/CI, interviewers asked to speak to the parent or guardian who knew most about the sampled child's education. Respondents were required to be 18 years old or older.

It is not possible to establish empirically whether and how the respondents for the two studies differed. Although data regarding the relationship of the respondent to the child are available for all children in NHES, these data are available only for 15- to 17-year-olds in CPS. The available data indicate that parents were the most frequent respondents in both surveys, and it seems quite likely that if parents were the most common respondents for 15- to 17-year-olds in CPS:Oct94, they would also be so for younger children.

Political/Legal and Cognitive Contexts

The political/legal and cognitive contexts within which surveys are conducted can affect respondents' answers to particular questions. Home-schooling researchers have suggested that home-schooling families may be more reticent than others to participate in government research, particularly research that might address the issue of home schooling, because of the often ambiguous legal status of home schooling (Kaseman and Kaseman 1991; Ray 1997). On the other hand, to the degree that in recent years parents have become more interested in home schooling and in working with schools and districts to facilitate home schooling, there may be less reason for concern in this regard.

The household- and item-level response rates provide relevant but conflicting evidence in this regard. The household response rate for NHES:96 (which respondents were told was sponsored by the U.S. Department of Education and concerned education issues) was lower than the corresponding rate for CPS:Oct94 (which was conducted by the Census Bureau and which respondents were told covered labor force participation issues). This is consistent with the hypothesis that home-schooling parents may be more reluctant to discuss education issues, although the impact of the lower household response rate is somewhat mitigated because 30 percent of households, not 100 percent, are likely to include school-aged children. The high item response rates in both surveys indicate that respondents in participating households were no less likely to discuss home schooling than other issues. Unfortunately, whether the political/legal context of home schooling affected the quality of response cannot be determined with the existing data.

The cognitive context may also have been affected by the different sponsors and purposes of the two surveys. In general, participating respondents want to cooperate with interviewers, and in their attempts to do so, use all available information to determine what the interviewer wants to know so they can provide the best information. Therefore, respondents are likely to have consid-

ered the different sponsors when they responded to questions, although any particular effects of these considerations upon their responses cannot be predicted or measured.

Impact of Data Processing Errors

Whereas the NHES:96 PFI/CI interview included on-line edits and all NHES:96 data were edited after data collection concluded, the CPS:Oct94 supplement did not include on-line edits and the home-schooling items were not edited after data collection. As noted above, without editing some cases could not be included in the CPS:Oct94 analysis due to missing information. Furthermore, not correcting errors that could be identified through consistency and plausibility checks in the CPS data may have contributed additional error to the CPS:Oct94 estimates relative to the NHES:96 estimates. The available data do not permit estimation of the direction or magnitude of this potential error in this instance.

Conclusion

This study examined several differences between the methods used in the CPS:Oct94 and NHES:96 that might contribute to the observed difference in the two surveys' estimates of the number and proportion of home-schooled children. The potential direction and magnitude of estimate differences could not be predicted for most of these methodological differences between the surveys, however.

This study raises a number of research questions regarding survey research and home schooling. First, it would be useful for researchers to address whether and how the political context of home schooling or other factors affect respondents' willingness to participate in the respective surveys and the accuracy of their answers to questions about home schooling. Second, research should explore the variety of schooling arrangements—exclusively at home, exclusively at school, and various combinations thereof—that parents make for their children, the frequency of these arrangements, and the factors that affect the kind of arrangement parents choose. Third, the results of cognitive laboratory research into parents' understanding of the term "home schooling" would aid in interpretation of responses to survey questions. Future research—using the NHES:99 data or cognitive laboratory studies of alternative question wording, for example—may address some of the issues raised in this report.

Foreword

In 1994 and 1996 two national surveys—the October 1994 Current Population Survey, conducted by the U.S. Bureau of the Census, and the 1996 National Household Education Survey, conducted by the National Center for Education Statistics—asked representative, national samples of households a series of questions regarding home schooling in the United States. Although it is believed to be a growing phenomenon, home schooling remains a relatively rare event, and therefore difficult to study with precision. This study was undertaken to compare the methods used in the two surveys and study the possible effects of differences in survey methods on estimates of the number and proportion of U.S. children who are home schooled. The results of this study have already been used to improve data collection on home schooling in the 1999 National Household Education Survey, and will influence future research on the topic with national survey data.

Acknowledgments

A great many people contributed to the completion of this report. At MPR Associates, Tim Urdan contributed to the literature review and early conceptualization of the project. Ellen Liebman wrote the programs that analyzed much of the data, and Susan Choy added helpful suggestions at a point when they were most needed. Andrea Livingston edited the report, Francesca Tussing created the tables and figures, and Barbara Kridl managed its final production.

At the U.S. Census Bureau, Rosalind Bruno, Timothy Marshall, and Ronald Tucker provided helpful information at key points during report production. At NCES, Dan Kaspryzk's knowledge, experience, and commentary improved the report immeasurably. Tom Snyder, Bruce Taylor, and Beth Aronstamm Young reviewed the report near its completion and also provided many useful suggestions.

Several others reviewed the report at various stages. Leslie Scott of the Education Statistical Services Institute provided a thorough technical review, and final reviewers included Jack Klenk of the U.S. Department of Education's Office of Non-Public Education; Patricia Lines, formerly of the Department of Education; Brian Ray of the National Home Education Research Institute, and Scott Somerville of the Home Schooling Legal Defense Association. Their contributions are much appreciated.

Table of Contents

List of Tables

List of Figures

1. Introduction

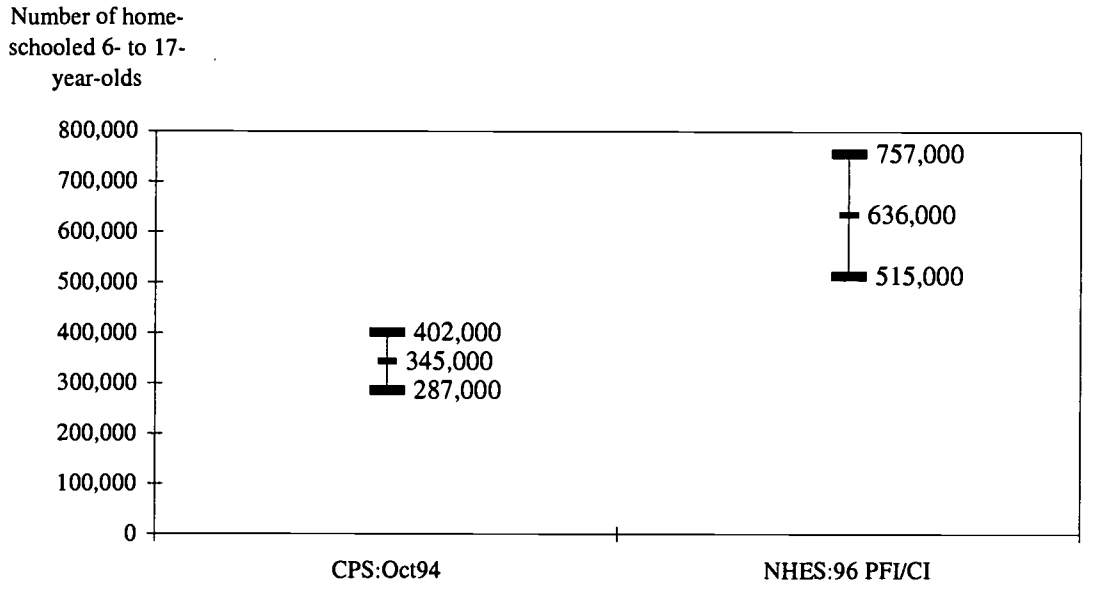
Home schooling in the United States has become a topic of interest to education policymakers, administrators, and the general public. For policymakers and education administrators, home schooling raises issues regarding whether and how to regulate parents' education of their children at home and the types of services school districts can and should provide to home-schooling parents and their children. For example, when school voucher initiatives were proposed in Colorado, legislators discussed whether home-schooled children should receive vouchers as well as their private school educated peers (Walsh 1998). Also, administrators in higher education are determining how to assess home-schooled students' qualifications when making admissions decisions (Belluck 1998; Shea 1996). Many of these issues concern taxpayers and citizens at large, and therefore, the general public also has substantial interest in home-schooling policy discussions. Informed discussions of home-schooling policy are compromised without accurate estimates of how many children are educated at home and whether the proportion of children who are so educated is changing.

However, measuring the prevalence of home schooling in the United States has proven to be a difficult task. Estimates of the number of children who are home schooled vary by hundreds of thousands of children. In the last decade, there have been several attempts to determine the number of children who are home schooled. Some studies have attempted to assess the size of the home-schooled population by identifying and surveying home-schooling families and extrapolating from those surveys estimates of the number of children who are schooled at home. Other researchers have collected data from state administrative records to develop estimates. All of these estimates, over time, have ranged from 200,000 children in 1988 (Kohn 1988) to 1.15 million children in 1995 (Ray 1997). However, most of these researchers recognize that their estimates of the number of home-schooled children include unknown sources of error (Kohn 1988; Lines 1991, 1996, 1998; Ray 1997).

In order to overcome some of the weaknesses of these studies and obtain accurate estimates of the number of children who are home schooled, two surveys with large, representative samples of the U.S. population included questions regarding home schooling: the 1994 October supplement to the 1994 Current Population Survey (CPS:Oct94) and the 1996 National Household Education Survey (NHES:96). However, although these surveys used scientific sampling methods, they have produced quite different results. The point estimates derived from these data sets

were 345,000 6- to 17-year-olds in CPS:Oct94 and 636,000 6- to 17-year-olds in NHES:96. The 95-percent confidence interval around the CPS:Oct94 point estimate ranges from 287,000 to 402,000, and the 95-percent confidence interval from 515,000 to 757,000 around the NHES:96 point estimate (figure 1).

Figure 1—CPS:Oct94 and NHES:96 point estimates and their 95-percent confidence intervals of number of home-schooled 6- to 17-year-olds



SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement/Civic Involvement Component (NHES:96, PFI/CI).

Although the differences between these two surveys' point estimates may reflect true growth in the proportion of students being home schooled between October 1994 and January through April 1996, it seems unlikely that the number of home-schooled children nearly doubled in less than two years (Ray 1999; Lines 1998). However, other differences between the CPS and the NHES, such as variation in question wording, make it impossible to determine whether the incidence of home schooling changed this dramatically between the two data collections. The purpose of this report is to explore differences in the designs and executions of the two surveys that may have contributed to the different estimates of the number of children who were home schooled.

This report is based on the premise that for any given year there is some "true" number of home-schooled children in the population. Point estimates derived from the CPS and NHES de-

part from this true value by some amount of error. Errors in surveys include errors in nonobservation, errors in observation, and data processing errors (Groves 1991).[1] After describing the data sources and presenting more details concerning the estimates, this report examines each type of error. First, the report explores differences between the surveys in the potential for nonobservation errors. These nonobservational errors are classified into two sources: errors in sample coverage and errors due to nonresponse. Next, the report examines potential sources of observational errors within the surveys, including errors associated with the data collection design, instrument error, and respondent error. Finally, the report examines the potential for error in the data processing of each survey.

[1] In this report, "error" is used to denote both random and nonrandom errors. "Bias" is also used to denote nonrandom error, e.g., "nonresponse bias."

2. Data Sources

This report examines data from the 1994 October Supplement to the Current Population Survey (CPS:Oct94) and the 1996 National Household Education Survey (NHES:96), surveys that resemble and differ from each other in a number of ways. Although most of this report discusses these similarities and differences in detail, this section provides a brief overview of the surveys.

2.1 Current Population Survey (CPS:Oct94)

Each month for over 50 years the Census Bureau has conducted the CPS, primarily to collect national data on labor force participation. Education is strongly related to several aspects of labor force participation, and therefore, questions on adults' education attainment are included among the questions that are asked each month, which are referred to as the "basic" survey or items. Supplements to the basic labor force participation survey focus on different topics each month. The October survey's supplement focuses on civilian participation in education programs each year, and includes separate instruments regarding the enrollment of children (ages 3 to 14) and adults (ages 15 and older). In October 1994, the education supplement to the CPS included additional questions on home schooling in both the adult and child instruments.

The CPS sample is selected from addresses from the previous Decennial Census and updates from building permits and other information sources. Sampled housing units are in the sample for 8 months out of 16 consecutive months: in 4 months, out 8 months, and in again during the same 4 months of the following year. The first and fifth month's interviews are conducted in person, and the remainder by telephone if the household agrees to telephone interviews. Most households' interviews are conducted by telephone, either by staff at one of three telephone centers or by the CPS personal interviewer who originally interviewed the household. Any household member 15 years old or older can serve as the respondent, although interviewers attempt to interview the most knowledgeable person in the household. Labor force participation data are collected regarding each household member 16 years old and older. Supplement data are collected about each household member for whom the questions are appropriate. In the case of the education supplement, questions regarding school enrollment are asked about all household members 3 years old and older (U.S. Department of Commerce 1999; U.S. Department of Labor 1997).

In January 1994 CPS began using computer-assisted personal interviewing (CAPI) and computer-assisted telephone interviewing (CATI) to collect data. In CAPI, field staff have laptop computers with them when they conduct the household interview. Similarly, in CATI, interviewers use computers to conduct interviews over the phone. In both CAPI and CATI, computer programs take interviewers through the questionnaire, preventing errors in skip pattern implementation. CAPI/CATI programs also can be programmed for on-line editing in which answers provided in previous interviews or earlier in the current interview can be confirmed or applied in subsequent questions, reducing errors and response burden. Furthermore, conflicts between responses can be resolved with the respondent, rather than during post survey edits.[2] CPS data are collected during the week that includes the 19th of each month, and are asked regarding the previous week, i.e., the week including the 12th.

2.2 National Household Education Survey (NHES:96)

The U.S. Department of Education's National Center for Education Statistics (NCES) began collecting data through the NHES in 1991, and since then has administered the survey in 1993, 1995, 1996, and 1999. NHES was developed to collect information on education-related issues that are not easily studied through institution-based surveys. Such issues include the use of nonparental childcare, participation in early childhood education programs, parents' and families' involvement in children's education, and school readiness. Topics are rotated among survey administrations to allow monitoring of trends over time (see NHES Web site, http://www.nces.ed.gov/nhes, for details). The NHES data used in this report were collected in the Parent and Family Involvement/Civic Involvement (PFI/CI) component of NHES:96.[3]

Households are selected for participation in NHES through random digit dialing of telephone numbers from January through April of the administration year. Households without telephones are excluded from the sample. Using CATI, interviewers in telephone centers conduct the interviews, beginning with Screener interviews that determine whether sampled telephone numbers serve a residence, and if so, whether household members are eligible to be sampled for extended topical interviews.

Screener respondents must be at least 18 years old or the head of the household, and extended-interview respondents are chosen in accordance with the population chosen for each ex-

[2]Section 6 discusses the editing (on-line and post-data-collection) procedures used in CPS:Oct94 and NHES:96.

[3]The NHES:95 Early Childhood Program Participation (ECPP) interview also included questions regarding home schooling. The ECPP interview was conducted regarding a sample of children who were in grades three or below and no older than age 10. In contrast, the NHES:96 PFI/CI interview was conducted on a sample of children enrolled in grades 12 and below and no older than age 20, a population that can be replicated (for the most part) with CPS:Oct94 data. Therefore, this report focuses on the CPS:Oct94 and NHES:96 data.

tended interview. The population of interest for the 1996 PFI/CI interview was children age 3 to 12th grade, with a maximum age of 20. These children were sampled for the interviews based on information gathered from the Screener. When a child in the household was selected for the extended-interview sample, the interviewer asked to speak with the parent or guardian who knew the most about the child's education in order to conduct the PFI/CI interview.

3. Comparison of Home-Schooling Estimates

The population of children chosen for this analysis is 6- to 17-year-olds who 1) lived with a parent, guardian, or nonrelative at least 12 years older than the child; 2) had neither graduated from high school nor earned a GED; and 3) were reported to be in kindergarten through grade 12, the equivalents thereof, or an ungraded class or level of instruction. The age and grade limitations were imposed partly to restrict the study to children who were expected to be enrolled in school across the nation. Therefore, 5-year-olds were excluded because the modal grade for 5-year-olds is kindergarten and enrollment in kindergarten is not required in all states. Similarly, children who had already earned a high school diploma or GED were excluded because they were no longer required to attend school.

In addition to including only children expected to be in school, other sample limitations were imposed due to choices made by the data collectors. In the NHES, parents of 6-year-olds who were in preschool were not asked the questions relating to home schooling, and parents of youth who were 18 years old or older were likewise not taken through the items about home schooling. Therefore, these groups were excluded from analyses of the CPS:Oct94 data because they were not captured in the NHES:96 data collection.[4]

Table 1 displays the estimated number of home-schooled children and percentage of all children who were home schooled from CPS:Oct94 and NHES:96.[5] Column 5 of table 1 shows the absolute differences between the two data sources in the number of home-schooled children. Column 6 shows the percentage difference between the estimated number of home-schooled children as derived from the two data sources. For example, the first row shows that the point estimates of the number of 6- to 17-year-olds in kindergarten through grade 12 who were home

[4]Emancipated minors—defined as children not living with someone identified as a parent or guardian or who was at least 12 years older than the child and not the child's spouse—were not included in the population of interest for the PFI/CI interview. Although the NHES definition of emancipated minors could not be replicated with the CPS data, analyses of the CPS data indicated that very few 15- to 17-year-olds were living alone or with a spouse, and none of them was home schooled. See appendix A for further details regarding this issue.

[5]Both surveys allowed distinction between children who were schooled at home for various reasons, including illness or disability. In addition, CPS:Oct94 included a question about whether the home-schooled child's instructor was a member of the school district staff or a parent or person employed by the parent. One might argue that children who were schooled at home because of illness or disability or by school personnel should not be included in estimates regarding home-schooled children. Relatively few children fell into this category, however, and therefore they were included in the estimates presented in this report. See appendix A for further discussion of this issue.

Table 1—CPS:Oct94 and NHES:96 estimates of the number and percentage of 6- to 17-year-olds who were home schooled in grades K–12 and difference between estimates

	CPS (October 1994)		NHES:96 (Spring 1996)		Difference between NHES:96 and CPS:Oct94 estimates	
	Number educated at home	Percent educated at home	Number educated at home	Percent educated at home	Number (NHES:96 minus CPS:Oct94)	Difference as percent of NHES:96 estimate
Total	345,000	0.8	636,000	1.4	292,000	45.9
Sex						
Male	161,000	0.7	268,000	1.1	107,000	39.9
Female	183,000	0.9	368,000	1.7	185,000	50.3
Race/ethnicity						
White, non-Hispanic	316,000	1.1	552,000	1.8	236,000	42.8
Black, non-Hispanic	10,000[1]	0.2	14,000[1]	0.2	4,000	28.6
Hispanic	16,000[1]	0.3	51,000[1]	0.9	35,000	68.6
All others	3,000[1]	0.2	20,000[1]	1.0	17,000	85.0
Age						
14 and younger	290,000	0.9	455,000	1.3	165,000	36.3
6	42,000[1]	1.1	31,000[1]	0.8	-10,000	-32.3
7	40,000[1]	1.1	43,000[1]	1.1	3,000	7.0
8	36,000[1]	1.0	60,000[1]	1.5	25,000	41.7
9	44,000[1]	1.2	59,000[1]	1.6	14,000	23.7
10	29,000[1]	0.8	46,000[1]	1.1	17,000	37.0
11	41,000[1]	1.1	51,000[1]	1.3	10,000	19.6
12	16,000[1]	0.4	53,000[1]	1.4	36,000	67.9
13	20,000[1]	0.5	50,000[1]	1.3	31,000	62.0
14	22,000[1]	0.6	62,000[1]	1.7	40,000	64.5
15 to 17	55,000	0.6	181,000	1.6	127,000	70.2
15	26,000[1]	0.7	69,000[1]	1.8	43,000	62.3
16	4,000[1]	0.1	65,000[1]	1.7	60,000	92.3
17	24,000[1]	0.8	47,000[1]	1.4	23,000	48.9
Family income[2]						
Less than $20,000	87,000	0.8	177,000	1.4	90,000	50.8
$20,000–34,999	80,000	0.8	172,000	1.7	93,000	54.1
$35,000–49,999	90,000	1.1	142,000	1.6	52,000	36.6
$50,000 or more	89,000	0.7	146,000	1.0	57,000	39.0
Father's education attainment						
Unknown	32,000[1]	0.3	158,000	1.0	126,000	79.7
No high school diploma	17,000[1]	0.4	56,000[1]	1.4	39,000	69.6
High school diploma	74,000	0.7	73,000	0.8	-1,000	-1.4
Some postsecondary education	101,000	1.2	189,000	2.5	87,000	46.0
Bachelor's degree	84,000	1.6	89,000	1.7	4,000	4.5
Graduate degree	35,000[1]	1.1	72,000[1]	1.9	37,000	51.4

Table 1—CPS:Oct94 and NHES:96 estimates of the number and percentage of 6- to 17-year-olds who were home schooled in grades K–12 and difference between estimates—Continued

	CPS (October 1994)		NHES:96 (Spring 1996)		Difference between NHES:96 and CPS:Oct94 estimates	
	Number educated at home	Percent educated at home	Number educated at home	Percent educated at home	Number (NHES:96 minus CPS:Oct94)	Difference as percent of NHES:96 estimate
Mother's education attainment						
Unknown	11,000[1]	0.6	49,000[1]	1.2	38,000	77.6
No high school diploma	27,000[1]	0.4	82,000[1]	1.2	55,000	67.1
High school diploma	100,000	0.7	132,000	1.0	32,000	24.2
Some postsecondary education	133,000	1.1	251,000	2.0	117,000	46.6
Bachelor's degree	64,000	1.2	97,000	1.5	33,000	34.0
Graduate degree	10,000[1]	0.5	26,000[1]	1.0	16,000	61.5
Highest parental education attainment						
Unknown	5,000[1]	—	(†)	(†)	(†)	(†)
No high school diploma	22,000[1]	0.4	42,000[1]	1.0	20,000	47.6
High school diploma	59,000	0.5	134,000	1.0	75,000	56.0
Some postsecondary education	130,000	1.0	243,000	1.8	113,000	46.5
Bachelor's degree	87,000	1.3	130,000	1.8	42,000	32.3
Graduate degree	41,000[1]	0.9	87,000	1.3	46,000	52.9
Number of parents						
Mom and dad	(†)	(†)	504,000	1.6	(†)	(†)
Mom only	(†)	(†)	100,000[1]	0.9	(†)	(†)
Dad only	(†)	(†)	11,000[1]	0.8	(†)	(†)
Nonparent guardian	(†)	(†)	21,000[1]	1.3	(†)	(†)
Number of children in family						
One	(†)	(†)	98,000	1.3	(†)	(†)
Two	(†)	(†)	157,000	0.9	(†)	(†)
Three	(†)	(†)	187,000	1.4	(†)	(†)
Four or more	(†)	(†)	194,000	2.7	(†)	(†)
Region						
Northeast	68,000	0.9	81,000[1]	0.9	13,000	16.0
South	125,000	0.8	207,000	1.3	82,000	39.6
Midwest	55,000	0.5	153,000	1.4	98,000	64.1
West	97,000	1.0	196,000	2.0	99,000	50.5

—Denominator sample size too small for a reliable estimate.

†Not available.

[1]Sample size too small for a reliable estimate. Interpret with caution.

[2]The income variable for the NHES was household income rather than family income, as in CPS. In addition, the income categories for the NHES were defined slightly differently, as follows: $20,000 or less; $20,001–35,000; $35,001–50,000; and $50,001 or more.

NOTE: Total numbers are rounded to the nearest thousand. Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement/Civic Involvement Component (NHES:96, PFI/CI).

schooled were 345,000 in CPS:Oct94 and 636,000 in NHES:96. The estimate from the NHES data is thus 292,000 children and 46 percent more than that from CPS.

3.1 Characteristics of Children

The data indicate that there were differences in the home-schooling rate by children's sex, race/ethnicity, and age. However, the findings were not always consistent across the two data sets. Although the CPS:Oct94 data indicate that there was no difference in the rates at which boys and girls were home schooled, the NHES:96 data indicate that girls were more likely than boys to be home schooled.

With respect to race/ethnicity, the CPS data indicate that white, non-Hispanic 6- to 17-year-olds were more likely than others to be home schooled. In addition, the NHES data indicate that white, non-Hispanic 6- to 17-year-olds were more likely to be home schooled than black, non-Hispanic and Hispanic 6- to 17-year-olds.

Finally, with respect to age, the CPS:Oct94 data indicate that children younger than 15 were more likely to be home schooled than were children 15 and older. However, the NHES data indicate that there was no difference in the home-schooling rates of children in these two age groups.

3.2 Characteristics of Families

As with the characteristics of children, whether some characteristics of families are associated with differential rates of home schooling depends on which data set one consults. Both data sets indicate that children in families with higher incomes were more likely to be home schooled. However, the NHES data also indicate that at the highest income level, the rate of home schooling declines. Although this also appears to be the case in the CPS data, the statistical tests do not support this conclusion.

Findings regarding the relationship between home schooling among children and their parents' education vary between the two data sets. The CPS data indicate that the rate of home schooling among 6- to 17-year-old children was related to their parents' (whether fathers', mothers', or the highest of either parent) level of education attainment. With respect to fathers' and the highest of the two parents' education attainment, the relationship was positive in that the more educated the parent the more likely the child was home schooled. With respect to mothers' education attainment, children whose mothers had graduate degrees were home schooled at a lower rate than were children whose mothers did not have a graduate degree. In contrast, how-

ever the NHES data indicate no linear relationship between the rate of home schooling and children's fathers', mothers', or highest parental education attainment.

The NHES:96 data indicate that children who lived with two parents were no more likely than children who lived with only one parent or a nonparent guardian to be home schooled. However, children in families of four or more children were more likely than children in two-child families to be home schooled.

Finally, neither the CPS:Oct94 nor the NHES:96 data indicate that home schooling is more common in some parts of the country than in others. Although in each data set it appears as if children in the West were more likely than children in other regions to be home schooled, these differences were not statistically significant.

Beginning with the next section, the remainder of the report discusses the estimated or potential impact of several types of error on the difference between the two surveys' point estimates. Section 4 addresses nonobservational errors, section 5 observational errors, and section 6 data processing errors.

4. Impact of Nonobservational Errors

Nonobservational errors result from missing portions of the targeted population. These errors can occur for a variety of reasons. In simple terms, if members of the targeted population are excluded from the sample frame, i.e., not covered by the sample frame, the sample cannot represent them, and therefore cannot represent the entire population. Similarly, if some members of the sample are less likely than others to participate in the survey, i.e., if there are low response rates overall or among certain groups, response bias may occur because some sample members' information is less well represented. Thus, issues of sample coverage and response rates are examined in this section.

4.1 Sample Coverage

The validity of estimates derived from a sample survey depends a great deal on the quality of the sampling frame and the sample design. The degree to which the sampling frame, that is, the list of population members from which the sample is drawn, covers all segments of the desired population affects the quality of the sample drawn from that frame. Equally adequate frames, however, can yield samples of greater or lesser quality depending on the design for drawing the sample. This section discusses the sampling frames and sample designs used in these studies.

4.1.1 Sampling Frames

Both the CPS:Oct94 and NHES:96 sampled children by sampling households, although they drew their samples of households from quite different sampling frames. This section discusses the differences between the sampling frames and the potential for, and potential effects of, biases introduced by these differences.

4.1.1.1 CPS:Oct94

The CPS sampling frame is the list of addresses used in the most recent Decennial Census updated using building permits in order to capture housing units created since that Census. Using data from the 1990 Decennial Census, the Census Bureau estimates that the overall undercoverage rate, that is, the proportion of individuals omitted from the frame due to missed housing units

and missed persons within sampled households, is 8 percent (U.S. Department of Commerce 1994, Attachment 18). The undercoverage rate for children under 15 years old is 7 percent, for 15-year-olds it is 10 percent, and for 16- to 19-year-olds it is 13 percent.

Among noninstitutionalized children under 20 years old, Census Bureau research indicates that the rate of undercoverage varies with sex, race/ethnicity, and age. Rates for males are typically higher than those for females, rates for blacks are higher than those for nonblacks, and rates for older children are typically higher than those for younger children. The lowest rate of undercoverage reported is 4 percent among nonblack females under age 15, and the highest rate is 29 percent among black males between 16 and 19 years old, inclusive (U.S. Department of Commerce 1994, Attachment 18).

Sampling weights can be adjusted to compensate for undercoverage. The CPS sampling weights are adjusted to independent population controls, in this case, projections based on 1990 Decennial Census figures; information on births, deaths, and migration from various sources; counts of people in the Armed Forces and other institutions; and estimates of the Decennial Census undercount. In order to take into account differences among subgroups in the rate of undercoverage, the CPS weights are adjusted to the following population controls: the civilian noninstitutionalized population 16 years old and older, by state; the national civilian noninstitutionalized population of 14 Hispanic and 5 non-Hispanic age-sex categories; and the national civilian noninstitutionalized populations of 66 white groups, 42 black groups, and 10 groups of other racial/ethnic backgrounds (U.S. Department of Labor 1997).

However, if undercoverage is not random with respect to a particular characteristic, such as being home schooled, estimates related to that characteristic may be biased. Bias is the product of 1) the difference between the rates of home schooling among omitted and included population members and 2) the rate of undercoverage in the population. Therefore, if the rate of home schooling among undercovered subpopulations were significantly different from that of covered subpopulations, the estimates relating to home schooling could be biased. For example, if boys who are omitted from the sampling frame were more likely to be home schooled than were boys who are included in the sampling frame, estimates of the number and proportion of home-schooled boys will be biased downwards.

The data in tables 2 and 3 present two scenarios, one in which undercoverage leads to overestimation of the number and proportion of children who are home schooled and another in which undercoverage leads to underestimation. For example, if no nonblack, 6- to 14-year-old girl who was omitted from the sample (highlighted row in table 2) were home schooled, the

Table 2—CPS:Oct94 estimates of the number of 6- to 17-year-olds enrolled in school or home schooled in grades K–12, percentage home schooled, undercoverage rate, and percentage and number overestimated, assuming no home schooling among undercovered population, by age, race/ethnicity, and sex

	Total number of children	Estimated percent home schooled	Difference between estimated and minimum rate (0 percent)	Undercoverage rate (percent)	Overestimated	
					Percent	Number
Total	43,163,000	0.8	0.8	(*)	0.05	24,000
6- to 14-year-olds						
Black						
Males	2,583,000	0.2	0.2	15.0	0.03	1,000
Females	2,537,000	0.1	0.1	16.0	0.02	0
Nonblack						
Males	14,389,000	0.9	0.9	8.0	0.07	10,000
Females	13,710,000	1.1	1.1	4.0	0.04	6,000
15-year-olds						
Black						
Males	269,000	0.0	0.0	24.0	0.00	0
Females	278,000	0.0	0.0	18.0	0.00	0
Nonblack						
Males	1,514,000	0.7	0.7	7.0	0.05	1,000
Females	1,425,000	1.1	1.1	10.0	0.11	2,000
16- to 17-year-olds						
Black						
Males	510,000	0.0	0.0	29.0	0.00	0
Females	535,000	0.6	0.6	20.0	0.12	1,000
Nonblack						
Males	2,808,000	0.6	0.6	12.0	0.07	2,000
Females	2,605,000	0.4	0.4	11.0	0.04	1,000

*Obtained from U.S. Department of Commerce 1994, Attachment 18. This source did not include an undercoverage rate for all children or all 6- to 17-year-olds.

NOTE: Numbers are rounded to the nearest thousand. Note that although rates of home schooling were calculated for 6- to 17-year-old children, rates of undercoverage were not published for children in these age groups. The estimates of bias reported here depend on the assumption that 6- to 14-year-olds were undercovered at the same rate as children 0 to 14 years old and that 16- to 17-year-olds were undercovered at the same rate as 16- to 19-year-olds.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table 3—CPS:Oct94 estimates of the number of 6- to 17-year-olds enrolled in school or home schooled in grades K–12, percentage home schooled, undercoverage rate, and percentage and number underestimated, assuming rate of home schooling among undercovered population is three times the observed rate, by age, race/ethnicity, and sex

	Total number of children	Estimated percent home schooled	Three times the estimated rate	Undercoverage rate (percent)	Underestimated	
					Percent	Number
Total	43,163,000	0.8	2.4	(*)	0.17	72,000
6- to 14-year-olds						
Black						
Males	2,583,000	0.2	0.6	15.0	0.09	2,000
Females	2,537,000	0.1	0.3	16.0	0.05	1,000
Nonblack						
Males	14,389,000	0.9	2.7	8.0	0.22	31,000
Females	13,710,000	1.1	3.3	4.0	0.13	18,000
15-year-olds						
Black						
Males	269,000	0.0	0.0	24.0	0.00	0
Females	278,000	0.0	0.0	18.0	0.00	0
Nonblack						
Males	1,514,000	0.7	2.1	7.0	0.15	2,000
Females	1,425,000	1.1	3.3	10.0	0.33	5,000
16- to 17-year-olds						
Black						
Males	510,000	0.0	0.0	29.0	0.00	0
Females	535,000	0.6	1.8	20.0	0.36	2,000
Nonblack						
Males	2,808,000	0.6	1.8	12.0	0.22	6,000
Females	2,605,000	0.5	1.5	11.0	0.17	4,000

*Obtained from U.S. Department of Commerce 1994, Attachment 18. This source did not include an undercoverage rate for all children or all 6- to 17-year-olds.

NOTE: Numbers are rounded to the nearest thousand. Note that although rates of home schooling were calculated for 6- to 17-year-old children, rates of undercoverage were not published for children in these age groups. The estimates of bias reported here depend on the assumption that 6- to 14-year-olds were undercovered at the same rate as children 0 to 14 years old and that 16- to 17-year-olds were undercovered at the same rate as 16- to 19-year-olds.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

estimated number of these girls who were home schooled would be 0.04 percent (1 percent of 4 percent) too high (table 2, row 5, column 6).[6] This overestimate would occur because the weights used in generating these estimates were calculated under the assumption that the 6- to 14-year-old nonblack girls who were omitted from the sampling frame were home schooled at the same rate as their counterparts who were in the sample. In total, the maximum overestimate of home-schooled children due to this potential bias would be 24,000 children.

In contrast, if nonblack, 6- to 14-year-old girls who were omitted from the sampling frame were home schooled at a much higher rate than those in the sample, e.g., three times the rate observed among their counterparts in the sample, the estimate of the number of home-schooled, nonblack, 6- to 14-year-old girls would be 0.13 percent (3 percent of 4 percent) too low (table 3, row 5, column 6). Under this scenario, the total underestimate of home-schooled 6- to 17-year-olds would be 72,000 children.[7] The error due to undercoverage of the sampling frame is probably between these extremes.

4.1.1.2 NHES:96

Rather than sampling households through addresses, the NHES samples households via telephone numbers. Sampling households through telephone numbers introduces one potential source of coverage bias, and the particular method of telephone sampling used in the NHES, random digit dialing (RDD) with the list-assisted method (Brick et al. 1995), introduces another. This section discusses evidence regarding these potential biases in NHES estimates of the number and proportion of children who were home schooled.

4.1.1.2.1 Excluding Nontelephone Households

All surveys that use telephone numbers as frames for sampling households exclude households without telephones, raising concerns regarding potential for coverage bias. The magnitude of this bias with respect to estimates of the number and proportion of children who are home schooled depends on 1) the proportion of children who are excluded from the sample because their households did not have telephone service at the time of the survey; and 2) whether telephone and nontelephone households home school children at different rates.

[6]Coverage rates are provided in CPS documentation (U.S. Department of Commerce 1994, Attachment 18). Note that although rates of home schooling were calculated for 6- to 17-year-old children, rates of undercoverage were not published for children in these age groups. The estimates of bias reported here depend on the assumption that 6- to 14-year-olds were undercovered at the same rate as children 0 to 14 years old and that 16- to 17-year-olds were undercovered at the same rate as 16- to 19-year-olds.

[7]See footnote 6.

CPS:Oct94 data indicate that 94 percent of households had telephone service. Research indicates that nontelephone households differ from telephone households in a number of ways. With respect to studying home schooling, the most relevant of these differences is that nontelephone households are more likely than telephone households to include at least one child (Montaquila, Brick, and Brock 1997).

Sampling weights adjust the NHES:96 estimate of the total number of children to population values derived from the March 1995 CPS, which includes children in nontelephone households, and therefore corrects for at least some of the bias introduced by omitting nontelephone households. In addition to adjusting for telephone undercoverage, the household-level sample weights adjust the estimates to known population totals with respect to state, whether children under 18 were present in the household, whether the home was rented/owned or occupied on some other basis, urbanicity, and race/ethnicity of the oldest member of the household. The person-level sample weights adjusted estimates to totals that included information on children's race/ethnicity, household income, region, urbanicity, home type, and children's grade in school (Collins et al. 1997).

However, bias with respect to the number and percentage of children who are home schooled may persist after weighting if the excluded children are home schooled relatively more or less often than children who are represented. The CPS:Oct94 data indicate that the estimated rate of home schooling did not differ between 6- to 17-year-olds in households with telephones and those in households without telephones (table 4). Moreover, even if the apparent difference (0.8 percent among children in telephone households versus 0.5 percent among children in

Table 4—CPS:Oct94 estimates of the number of 6- to 17-year-olds who were enrolled in school or home schooled in grades K–12 and percentage who were home schooled, by whether telephone in household

	CPS (October 1994)	
	Total number	Percent educated at home
Total	43,163,000	0.8
Telephone in household		
Telephone	40,316,000	0.8
Nontelephone	2,848,000	0.5

NOTE: Total numbers are rounded to the nearest thousand. Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

nontelephone households) were real, the NHES estimate of the number of home-schooled children would be only about 10,000 children too high.[8] Thus, any bias in the NHES estimate of the number or proportion of children who were home schooled due to excluding nontelephone households from the sample is likely be quite small.

4.1.1.2.2 Excluding the Zero-Listed Stratum

The list-assisted method of RDD, which is used to reduce survey cost by eliminating many calls to nonresidential telephone numbers (e.g., business telephone numbers), adds another source of coverage bias. In the list-assisted method, all active telephone numbers are divided into groups of 100. Within each group of 100 numbers, numbers are checked against residential white page directories to determine whether they are listed. As soon as one number in a group has been identified as listed, that group is included in the listed stratum. Groups with no listed number are allocated to the zero-listed stratum. The sample is drawn from the listed stratum, which includes both listed and unlisted residential numbers.

The potential for bias is introduced because unlisted residential numbers that were allocated to the zero-listed stratum are excluded from the sample. Thus far, three out of four studies of the list-assisted method have found that the proportion of residential telephone numbers excluded by using this method ranged from 3 to 4 percent (Brick et al. 1995). Moreover, it appears that there are very few differences between households whose telephone numbers are in the zero-listed stratum and those whose numbers are in the listed stratum, and where found, those differences are small. However, these studies examined differences in the demographic characteristics between listed and zero-listed stratum households, not whether households were more or less likely to home school children. Nevertheless, because the proportion of households excluded from the sampling frame is small and the rate at which children are home schooled is generally very low, any coverage bias in estimates of the number and proportion of children who are home schooled due to excluding households in the zero-listed stratum would be quite small even if the rate of home schooling among excluded population members were two or three times that of included population members.

[8]The 10,000-child overestimate was calculated as follows: 7 percent of 6- to 17-year-olds lived in households without telephones and of these children, 0.3 percent were estimated as home schooled when they were not actually home schooled. The latter occurred because the sampling weights assume that children in households without telephones were home schooled at the same rate as children in households with telephones. 45,710,000 * .07 * .003 = 9,599 children estimated to be home schooled who were not.

4.1.2 Sampling Among and Within Households

As discussed above, there is some potential for error in the studies' sampling frames. However, the sampling frame is only the first step in drawing a sample from the universe of households. Once the frame is determined, one must design strategies for drawing a sample of households from the frame and, depending on the study's population of interest, for selecting cases within households. The remainder of this section describes the procedures used to sample cases from each study's respective sampling frame, and indicates that neither of the studies' sample designs appears likely to induce any particular source of error. Both studies sample randomly from households within their frames and oversample some minority groups to collect sufficient data for reliable estimates concerning those groups. In addition, the NHES:96 PFI/CI randomly sampled children within households, depending on the number of children who were eligible to participate within a household.

4.1.2.1 CPS:Oct94

The sample design for the CPS:Oct94 was based on the 1980 census, although households from the 1990 census were phased into the sample design between April 1994 and July 1995. Sampling was conducted in two stages. First, primary sampling units (PSUs)—geographic areas usually corresponding to a county or several contiguous counties within a state—were selected from the 1,973 PSUs in the United States. The universe of PSUs was stratified within each state.

At the second stage of sampling, households were selected within PSUs. A household's probability of selection depended on the size of its PSU and the sampling ratio of its state, which in turn depended on the population of the state. Selection probabilities were calculated to ensure that the state-level unemployment rate estimates were sufficiently reliable. Each month the Census Bureau samples approximately 70,000 housing units, of which approximately 10,000 are ineligible for participation in the CPS because they are vacant, demolished, used for commercial purposes, and so on.[9]

The stratified design of the CPS:Oct94 sample selection ensures that all states, the District of Columbia, and the metropolitan areas of New York City and Los Angeles are represented proportionally and that the sample includes subgroups such as minority group members, renters, and

[9]The November 1994 issue of Employment and Earnings, which reported CPS:Oct94 labor force estimates, reported that 72,000 households were assigned each month, of which 60,000 were occupied and therefore eligible for interview. The Bureau of Labor Statistics *Handbook of Methods* (U.S. Department of Labor 1997) reports that the 1990 sample design includes 66,000 households, but does not indicate how many of those households are typically eligible for interview.

female householders. Random selection within strata and clusters[10] ensures against bias in the sample.

In order to improve the reliability of estimates of change from month to month, the sample is divided into eight panels that are rotated into and out of the sample over a period of 16 months. Each panel is interviewed for 4 consecutive months, removed from the sample for the next 8 months, and then interviewed for the following 4 consecutive months (one year after the first 4 months). Therefore, in any given month, seven-eighths of the interviewed households will have been interviewed previously, and one-eighth will be new to the survey.

4.1.2.2 NHES:96

The NHES sample design differs from that used for CPS in a number of ways. First, households were sampled using telephone numbers as discussed above. In order to ensure that members of minority racial/ethnic groups were adequately represented, telephone numbers in the listed stratum were further stratified into high- and low-minority concentration substrata, and telephone numbers in the high-minority concentration stratum were sampled at twice the rate used in the low-minority stratum. Weighting procedures adjusted for oversampling of minority population members. In addition, the sample design for NHES:96 included stratification by state in order to obtain state-level estimates for certain items in the Screener interview.

Second, eligible members of sampled households were sampled for the extended interviews. Each of the extended interviews in the NHES is targeted toward a specific subpopulation, and therefore in NHES:96 many telephone numbers were screened in order to achieve adequate sample sizes for all of the extended interviews. After interviewers establish that they have reached a residential telephone number, they conduct Screener interviews to determine which household members are eligible for one of the extended interviews. When multiple household members are eligible for one or more extended interviews, household members are randomly sampled to limit the number of interviews per household.

The within-household sampling strategy for the PFI/CI interview was to sample 1) no more than one child who was at least 3 years old but in no grade higher than the fifth and 2) no more than one child who was younger than 21 years old and enrolled in grades 6 through 12.[11] Children who met these criteria were excluded from the population of the study, and therefore the

[10]In most sampling methods, the sampling units are basic elements of the population, e.g., individual children or adults. "In cluster sampling, sampling units are not basic elements but are groups or collections of elements. These groups or collections of elements are termed *clusters*." (Jaeger 1984, 173).

[11]Weights did not rake to totals for children in families of various sizes, and therefore cannot account for the difference in the rates of home schooling between families with four or more children and families with fewer children.

sample, if they were neither enrolled in school nor home schooled or if they were emancipated minors. In the NHES:96 PFI/CI, an emancipated minor was identified when a sampled child was the Screener respondent, identified himself or herself as the person in the household who was most knowledgeable about his or her education, and did not live with anyone who was at least 12 years older and not a spouse or boyfriend/girlfriend.

Third and finally, households participate in NHES only once, in contrast to the CPS panel design. Thus, the sampling frames and sample designs vary considerably between the two surveys.

4.2 Response Rates

Response rates affect the validity of estimates in at least two ways. First, they determine the sample size—that is, the number of sampled cases from which data are actually obtained, which is related to the reliability of estimates. In general, the smaller the sample, the less reliable the estimate. Second, the lower the response rate, the greater the chance of bias in the sample and in the estimates derived from the data.

The structure of these surveys requires the examination of response rates at each of several stages of data collection. First, some households did not participate in the survey for which they were sampled. Second, within participating households, some individuals who were eligible for the CPS Education Supplement or who were sampled for the NHES PFI/CI interview did not do so. Third, some individual respondents did not provide responses to some of the questions they were asked. Therefore, this section discusses the various response rates for each survey, beginning with the CPS:Oct94.

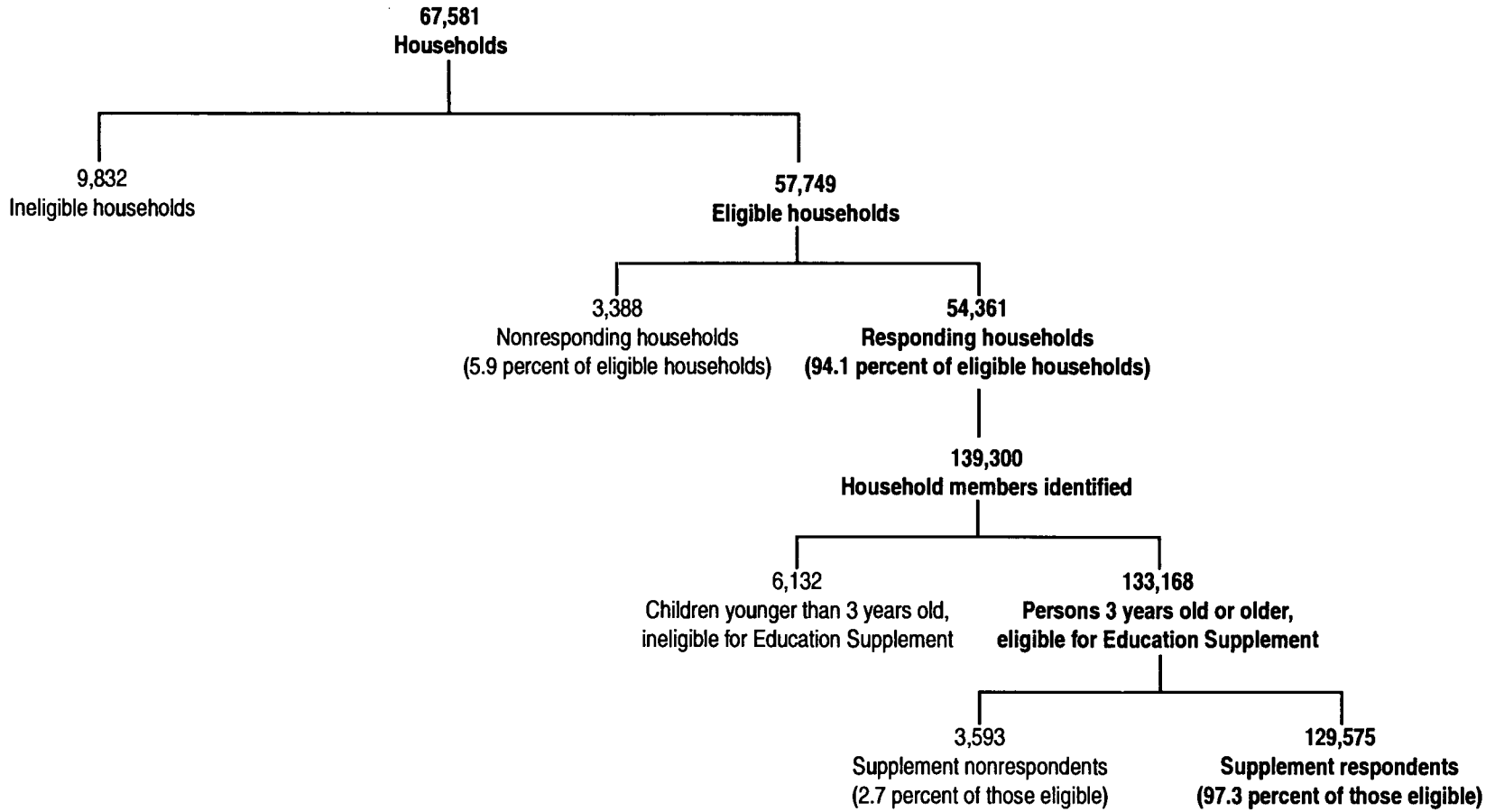
4.2.1 CPS:Oct94

CPS:Oct94 response rates are calculated at three levels. First, the survey response rate measures the percentage of eligible households that completed the survey. Second, the supplement response rate measures the percentage of eligible household members for whom the education supplement was completed. Third, the item response rate measures the percentage of eligible household members who completed each item.

4.2.1.1 Survey Response Rate

Of the 67,581 housing units sampled for the October 1994 CPS, 9,832 were ineligible to participate in the study because they were vacant, had been demolished, were under construction, or were being used as storage space (figure 2). Another 3,388 households were unavailable for

Figure 2—Disposition of sampled households and household members eligible for Education Supplement in CPS:Oct94, including response rates



NOTE: The number of responding households is the number of householders about whom interviews were conducted.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct 94).

interview after many attempts or they refused to participate. Responding households were defined as those in which the basic labor force survey was completed for at least one eligible household member. In October 1994, 54,361 households were classified as responding households for an unweighted survey response rate of 94 percent.

4.2.1.2 Supplement Response Rate

Within responding households, all household members age 3 and older are eligible to participate in the Education Supplement. Within the responding households in October 1994, 139,300 household members were identified, of whom 133,168 were 3 years old or older. Supplement respondents were defined as those household members 3 years old or older for whom the first supplement question, which asked whether the respondent was enrolled in school, was completed. Among the 133,168 respondents eligible to participate in the supplement, the first question was completed for 129,575 respondents, resulting in an unweighted supplement response rate of 97 percent. When weighted, the supplement response rate remained 97 percent.

4.2.1.3 Item Response Rates

Overall, weighted item response rates for the items used in this analysis were high, ranging from 92 percent to 100 percent (table 5). However, in order to compute estimates with the CPS:Oct94 data that would be comparable with the NHES:96 estimates, information from several items (age, enrollment status, grade level, whether high school graduate, whether GED recipient) was used to identify a comparable subsample of children. In addition, the homeschooling items were needed to identify home-schooled children. Although missing data were imputed for some variables, such as age and enrollment status, missing data were not imputed for all necessary variables. The cumulative effect of missing data on the variables for which missing data were not imputed may be larger than that of any single variable.

The CPS:Oct94 data indicate that there were missing data on a least one variable for an estimated 2,345,000, or 5 percent, of the estimated 45,835,000 6- to 17-year-olds.[12] This amounts to a weighted cumulative item response rate of 95 percent.[13]

[12]Data on age and enrollment status were imputed and therefore children who were missing these data were not included among the 2.345 million.

[13]Details regarding these calculations are available in appendix A.

Table 5—Weighted CPS:Oct94 response rates for selected items

Item	Weighted item response rate
Age[1]	99.6
Family income	91.8
Grade[1]	96.7
Grade equivalent	100.0
Graduation status	
Education attainment[1]	99.1
Diploma by GED[2]	95.5
Home-schooling status	
Among not enrolled	92.9
Among enrolled	99.1
Race[1]	99.1
Respondent's relationship to subject of interview/child[2]	97.1
Sex[1]	99.9
Telephone in household[1]	99.7

[1]Missing responses for these items were imputed in the datafiles used in this analysis.
[2]Available for children 15 years old and older only.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Assuming that all of the interviews with missing data (representing an estimated 2,345,000 children) were conducted with children who were eligible for the study (i.e., met the grade-level and graduation status criteria), and that these children were home schooled at the same rate as children who were not missing data, approximately 30,000 of these children would be home schooled. Adding this estimate to the previously identified 345,000 home-schooled children, the total estimate of the number of home-schooled children would be 375,000, and the estimated percent of 6- to 17-year-olds who were home schooled would remain 0.8 percent.

However, it is unlikely both that all of the 2,345,000 children who were missing data were eligible for the analysis and that those who were eligible were home schooled at the same rate. Thus, although missing data could alter the estimate of the number of children who were home

schooled by approximately 9 percent (30,000 / 345,000), the actual error associated with missing data is likely to be considerably lower.

4.2.2 NHES:96

As with the CPS response rates, NHES:96 response rates are calculated at three levels. The first is the survey level, which in the NHES case is the rate at which households participated in the Screener interview used to gather household information and sample individuals for the extended interviews. The second is the extended-interview level, which is the rate at which sampled individuals completed extended interviews, and the third is the item level, which is the rate at which those who participated in extended interviews completed each item.

4.2.2.1 Survey Response Rate[14]

Computing the Screener, or survey-level, response rate for the NHES is complicated by the fact that not all sampled telephone numbers were residential. Sampled telephone numbers that were not residential must be excluded from the base of eligible telephone numbers. The challenge lies in determining the number of sampled telephone numbers that are nonresidential.

When interviewers reach nonresidential numbers, their ineligibility is determined and these numbers are excluded from the calculation of the Screener response rate. However, it is not always possible to determine whether unreached sampled telephone numbers were residential. Therefore, in order to determine the Screener response rate, one must estimate the number of residences among unreached telephone numbers.

The number of unreached sampled telephone numbers that belonged to residences was estimated in three ways. Two of these methods are simply to assume either that none of the unreached telephone numbers was residential (liberal response rate) or that all of them were residential (conservative response rate). A third estimate of the number of unreached telephone numbers eligible to be in the sample is the business office method. This method involves verifying the residential status of a sample of unanswered numbers with the telephone company and then using this information to estimate the proportion of unanswered phone numbers that were

[14]In its discussion of response rates, the NHES documentation distinguishes between completion rates and response rates when discussing Screener and interview rates. Using this terminology, the extended-interview completion rate is the proportion of households who completed the Screener who then completed an extended interview. By contrast, the extended-interview response rate is the extended-interview completion rate multiplied by the Screener response rate. Whereas the completion rate measures response at each level on its own, the response rate takes into account nonresponse to previous levels of the survey. For the sake of clarity, this report discusses only the proportion of responding units (whether households or respondents) at each level, and uses the term "response rate" at each level. Thus, what the NHES documentation refers to as the "interview completion rate" is referred to as the "interview response rate" in this report.

working residential phone numbers. The product of this proportion and the total number of unanswered phone numbers is then added to the number of known residential responses, producing an overall estimate of the number of residential numbers dialed, which is used to estimate the Screener response rate (Collins et al. 1997).[15]

The business office method of estimating the number of unreached telephone numbers that are residential is well accepted (Collins et al. 1997), and therefore estimates derived with this method are used in this report. In NHES:96, 161,446 telephone numbers were sampled, of which 9,452 were not reached (figure 3). Using the business office method, 3,828 of the unreached telephone numbers were estimated to be residential. Combined with the 76,258 sampled numbers that were reached and determined to be residential, the total estimated number of sampled residential telephone numbers was 80,086. From these numbers, 55,838 Screener interviews were completed, leading to an unweighted Screener, or survey-level, response rate of 70 percent. The weighted estimated Screener response rate was also 70 percent (Collins et al. 1997).

However, households with children in grades K–12 do not appear to have refused at a higher rate than households without children. Approximately 30 percent of households in the United States have children in grades K–12, and approximately 30 percent of the households that responded to the NHES:96 Screener interview had children in these grades.

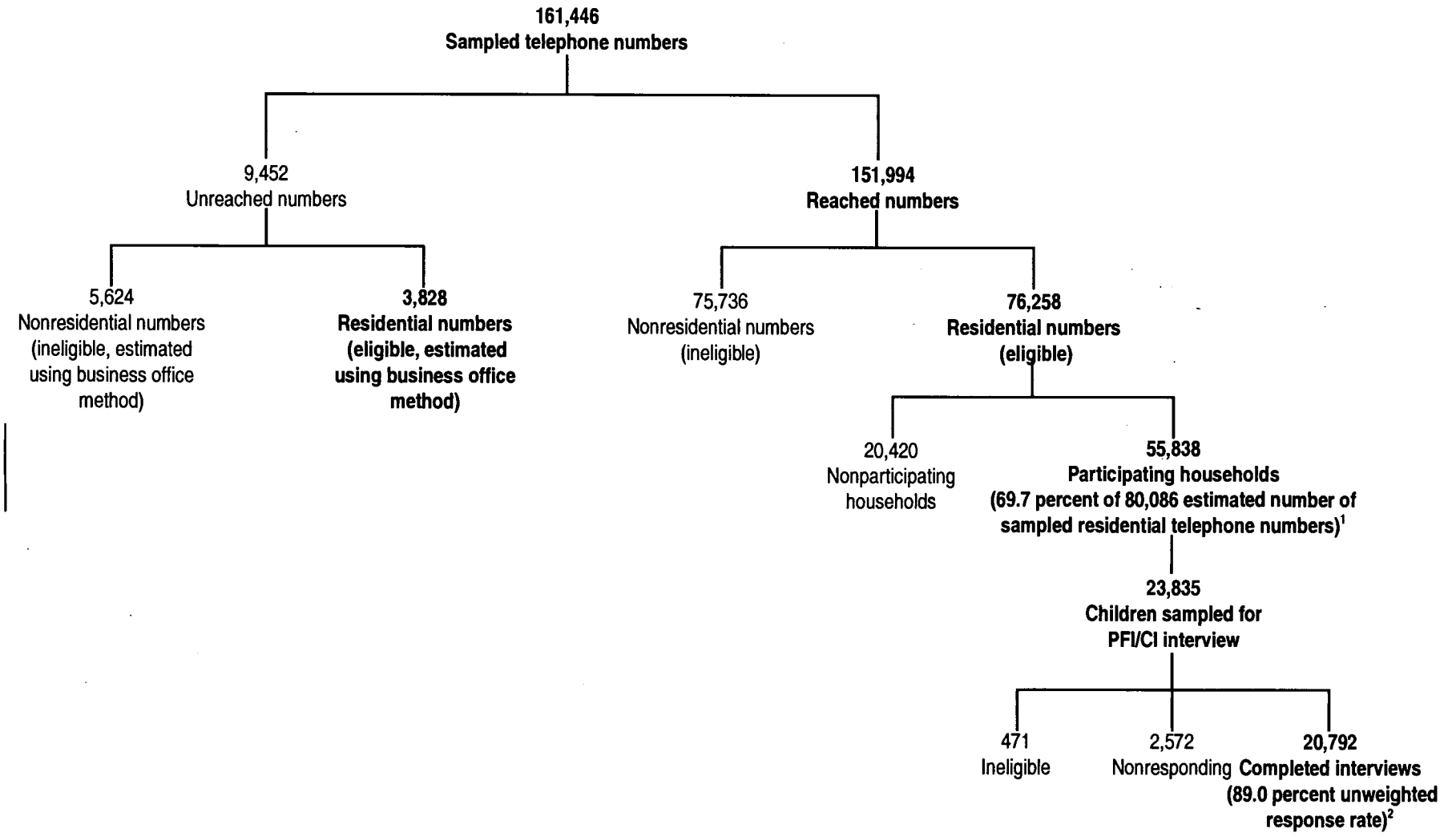
4.2.2.2 Interview Response Rate

From the 55,838 households that participated in the Screener interviews, 23,835 children were sampled for the PFI/CI extended interview. Among these children, 471 were determined ineligible during the extended interview, and therefore were excluded from the interview response rate calculation. Extended interviews were defined as "complete" at varying points, depending on the age/grade level of the child, but in all cases respondents had to complete many questions before an interview was classified as complete.[16] Of the remaining children, interviews

[15]The *Data File User's Manuals* (Collins et al. 1997) reports a fourth estimate of the Screener response rate that is derived using the Council for American Survey Research Organizations (CASRO) rate of estimating the number of residential numbers among unreached sampled telephone numbers. Montaquila and Brick (1997) do not recommend using this rate for estimating response rates with NHES data, because analyses done with data from the 1991 and 1995 NHES indicate that the residence rate for the unreached numbers is lower than that implied by the CASRO rate. Therefore, this report does not present the CASRO rates or response rates derived from it for the PFI/CI data.

[16]Interviews were classified as complete at the following points: for children in kindergarten or younger, respondents had to complete the series of questions on children's health; for children from kindergarten through fifth grade, respondents had to complete the series of questions on children's disabilities; and for children in grades 6–12, respondents had to complete the questions on community and government (Chris Chapman—personal communication, April 1999). Interested readers may refer to the NHES:96 questionnaires to determine more exactly what proportion of the interviews were necessary for classification as "complete."

Figure 3—Disposition of sampled telephone numbers and of children sampled for the PFI/CI interview in NHES:96, including response rates



[1]Denominator includes known residential numbers and 40.5 percent of unreached numbers (3,828 numbers), estimated to be residential using business office method. Weighted screener response rate was 69.9 percent.
[2]Weighted extended interview response rate was 89.4 percent.

SOURCE: Collins et al. (1997) *National Household Education Survey of 1996, Data File User's Manual, Volume I* (NCES 97–425). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

were completed for 20,792, resulting in an unweighted extended interview response rate of 89 percent. The weighted extended interview response rate was also 89 percent.

It is possible that the 11 percent of sampled households that did not participate in the extended interviews included a disproportionate number of home-schooled children. However, even if as many as 5 percent of children in nonresponding households were home schooled, compared with 1.4 percent of children in responding households, the weighted estimate of the proportion of children who were home schooled would be off by 3.6 percent of 11 percent, or 0.4 percent of all children. Thus, the 11 percent extended interview nonresponse rate is likely to have had a small impact on the overall estimates.

4.2.2.3 Item Response Rates

In the NHES, weighted item response rates, including those for home-schooling questions, were consistently high, ranging from 89 percent to 100 percent (table 6). Most of the items examined in this study, including questions about home schooling, were answered by at least 99 percent of the extended-interview respondents. Questions about family income were the only questions used in these analyses that were answered by less than 99 percent of the respondents.

When missing data occurred, that is, when the item response rate was not 100 percent, the missing values for that item were imputed using a hot-deck procedure. Manual imputation was done for a few cases. The analyses reported here used imputed data, and therefore no cases were eliminated from analysis due to missing data. For details concerning imputation procedures, see Collins et al. (1997) and Montaquila and Brick (1997).

4.3 Conclusion on Nonobservational Errors

This section has examined the CPS:Oct94 and NHES:96 PFI/CI with respect to important sources of nonobservational error, including sampling (sampling frames and sample design) and response rates. Although the surveys do differ with respect to these characteristics, it is rarely possible to determine the direction or magnitude of error associated with their different methods. Moreover, it is quite possible that errors may negate each other, e.g., if one source of error leads to an underestimate and another source leads to an equivalent overestimate.

Although the sample designs of both surveys rely on lists of household members that may be incomplete, there is no evidence of bias in either survey's sample design. However, the sampling frames of the CPS:Oct94 and NHES:96 PFI/CI leave open the possibilities for small biases in estimates of the number and proportion of children who are home schooled. Previous studies

Table 6—Weighted NHES:96 PFI/CI response rates for selected items

Variable	Weighted item response rate
Age	100.0
Enrolled in school	100.0
Family income	89.4
Father's education attainment	98.8
Grade	100.0
Grade equivalent	99.7
Home-schooling status	99.9
Mother's education attainment	99.3
Race	99.5
Respondent's relationship to child	99.9
Sex	99.9

NOTE: Missing responses for all items were imputed.

SOURCE: J. Montaquila and J.M. Brick. (1997). *Unit and Item Nonresponse, Weighting, and Imputation Procedures in the 1996 National Household Education Survey* (NCES 97–40). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

of CPS indicate that between 7 and 13 percent of the under-20 population is excluded from the sampling frame due to coverage biases, and this undercoverage could bias the estimates related to home schooling. If such bias exists it is likely to be quite small, however, because of the low rates of home schooling that have been observed. CPS data allow empirical exploration of the potential for telephone undercoverage bias in the NHES data, and provide no evidence of such a bias. Although data to study the potential for bias due to excluding the unlisted stratum in NHES:96 are not available, previous research indicates that if such bias exists it is in all likelihood very small, again because of the low rate of home schooling and because only 3 to 4 percent of residential telephone numbers are excluded by this method.

Response rates and their potential effects on the estimates related to home schooling vary between the two studies. The CPS:Oct94 survey response rate is considerably higher than the analogous NHES:96 Screener response rate. The higher survey-level response rate for the CPS:Oct94 may be at least partly attributable to the fact that it is a panel survey begun with an in-person interview. Respondents may find it more difficult to refuse a person at the door than a

voice on the telephone, and once the interviewer has built rapport with the household, maintaining cooperation throughout the remaining 7 months of participation may be relatively easy in most cases. Regardless of the reason for the higher NHES:96 survey nonresponse, although approximately 30 percent of households did not respond to the NHES:96 Screener, only about 9 percent of households did not respond and could have had children in the desired age/grade range who are home schooled.

Response rates at other levels more closely resemble each other. The CPS:Oct94 supplement response rate (97 percent) is 8 percent higher than the NHES:96 PFI/CI participation rate (89 percent). This is likely to be related to differences between CPS:Oct94 and NHES:96 PFI/CI in the criteria used to define the supplement and extended-interview response rates, respectively. CPS supplement interviews were included in the numerator of the supplement response rate when the first question of the supplement was answered, whereas PCI/CI interviews were included in the numerator of the extended-interview response rate only after many more questions had been answered.

Item response rates were high for nearly all items used in these analyses, regardless of survey. Missing data in CPS:Oct94 that were not imputed may bias the CPS:Oct94 estimate of the number of home-schooled children downward relative to the estimate derived from the fully-imputed NHES:96 data. It is not possible to determine the effects of missing data exactly. Similarly, it is not possible to determine whether the lower NHES:96 household or interview response rates led to greater nonresponse bias than may have occurred in CPS:Oct94 because it is not possible to determine whether home-schooling families were more or less likely than other families to participate. In sum, although the response rates of the two surveys differ considerably, the net effect of these differences on the estimates cannot be determined.

5. Impact of Observational Errors

Errors of observation are, essentially, measurement errors (Groves 1991). Following the Groves (1991) taxonomy, this section discusses the surveys in terms of three types of observational or measurement error. First, the section describes the data collection methods used in the two surveys and their potential to introduce error. Second, the section describes the surveys' potential for instrument error, which arises from the particularities of the questions, their response categories, and the sequence of questions in the survey. Third, the section describes respondent error, which arises from respondents' differing abilities to respond accurately to a given question, differences between the surveys in the context of the questions, and therefore respondents' interpretations of those questions.

5.1 Data Collection Procedures and Potential Associated Error

Once the sample is designed and drawn, data collection begins. This section first describes the monthly CPS data collection procedures and then the procedures used for the NHES:96.

CPS interviews are conducted during the week that includes the 19th day of each month, and questions are asked regarding the previous week, that is, the week including the 12th day of the month. As noted above (section 2), CPS is a panel survey in which households are sampled to be interviewed for 8 months of a 16-month period. Each household's first- and fifth-month interview is conducted in person, and most other interviews are conducted by telephone. Interviewers in telephone centers use computer-assisted telephone interviewing (CATI) to interview 12 to 15 percent of households each month, from Sunday through Wednesday of the interview week. Sampled households whom interviewers in the telephone centers have not reached by Wednesday of the interview week are turned over to the interviewer who conducts the in-person first- and fifth-month interviews. This interviewer has until the following Tuesday to complete the interview or declare the household a noninterview for that month. If the household does not have a telephone or access to one, is composed of people with limited English-speaking skills, or does not agree to telephone interviews, the interviewer who conducts the first- and fifth-month inter-

views visits the home each month (U.S. Department of Commerce 1999). In all, approximately 60 percent of CPS interviews are conducted by phone and the remainder in personal interviews.[17]

Since January 1994 interviewers have administered the CPS using computer-assisted interviewing (CAI), either by telephone or in person, rather than having interviewers complete a paper-pencil questionnaire. CAI saves time because information on each sampled household is preloaded, which allows interviewers to confirm information gathered from previously interviewed households rather than repeat each question each month. CAI also reduces errors by routing the interviewer through the instrument based on the responses he or she keys in, ensuring that skip patterns are followed accurately (U.S. Department of Labor 1997).

In contrast to CPS, NHES:96 was conducted from January through April 1996 by telephone center interviewers. The CATI system assigned telephone numbers to interviewers. Many numbers were finalized within one or two calls, including 51 percent of numbers that resulted in completed Screener interviews, 87 percent of numbers that were identified as nonworking, and 68 percent of numbers that were identified as nonresidential. Numbers for which interviewers had neither completed an interview nor received a refusal were tried a maximum of seven times for at least two weeks before they were put aside and refielded (that is, tried again at a later date). Unless both the interviewer and supervisor identified a refusal as hostile, cases that refused were refielded up to two times before the conclusion of data collection (Vaden-Kiernan et al. 1997).

Thus, the surveys differ with respect to data collection procedures in at least three ways. First, although both surveys are conducted with CAI, CPS interviewers use both personal and telephone interviewing, whereas NHES interviews are conducted entirely via telephone. Previous research indicates that personal interviewing yields more accurate responses to questions regarding sensitive topics such as alcohol and illicit drug use (Clark and Schober 1992; Fowler 1993). Whether and how personal interviewing, compared with telephone interviewing, might produce different results with respect to home schooling is unknown.

Second, CPS is a panel survey, whereas NHES is not. The effects of this, aside from the potential differences in response rates that were discussed above, cannot be assessed with available data.

[17]The Bureau of Labor Statistics *Handbook of Methods* (U.S. Department of Labor 1997) reports that about 60 percent of interviews are conducted by telephone, including 12 percent by interviewers in telephone centers. The Census Bureau's Web site reports that, overall, 90 percent of the households in their second through fourth and sixth through eighth months in sample are interviewed by phone. Assuming that first and fifth month-in-sample households comprise one-quarter of the sample, 90 percent of the remaining three-quarters yields 68 percent of households interviewed by phone. The Web site indicates that 15 percent of households are interviewed by telephone center personnel each month. For purposes of this analysis, the figures cited in the *Handbook of Methods* were used.

Third, the surveys also differ with respect to timing. In addition to the 15- to 18-month span between the surveys, they differ in the time of year at which they were administered. The NHES is administered from January through April, in contrast to the October administration of the CPS:Oct94. To the degree that parents are more likely to home school their children at some times of the year than at others, the difference in survey timing may contribute to the difference between the estimates.

5.2 Instrument Error

This section explores three potential sources of instrument error. These sources include the wording of the home-schooling questions in each survey, the sequence of these questions within each survey, and the home-schooling questions' susceptibility to respondent fatigue due to the location of these questions within the overall survey.

5.2.1 Question Wording

Responses to survey questions depend in large part on how questions are worded (Tanur, 1992; Sudman, Bradburn, and Schwarz 1996). Changing the wording of an item can have a dramatic impact on survey responses. For example, in order to respond appropriately, respondents must share the meaning of an item wording with the survey researchers. Furthermore, an "appropriate" response to an item requires the respondent not only to understand the literal meaning of the question but also to infer the questioner's intention in order to determine the *practical* meaning of an item (Schwarz 1999).

In the present situation, the meaning of "home schooling" as used by education researchers may not be familiar to many respondents—it is a fairly rare event and many respondents may not have heard of home schooling. The wording of the CPS:Oct94 and NHES:96 items on home schooling are discussed in the sections below.

5.2.1.1 CPS:Oct94

In CPS:Oct94 the household informant was asked whether the child was "attending or enrolled in regular school" (figure 4). If the child was enrolled, the respondent was asked the child's grade and whether the child was "attending a regular day school, boarding school, schooled primarily at home by someone paid by the school, schooled primarily at home by a parent or other person paid or chosen by the parent, or someplace else." The use of the word "primarily" in the question used to identify children who were being home schooled may affect the

Figure 4—Items related to home schooling in CPS:October 1994 interview, by subject age and enrollment status

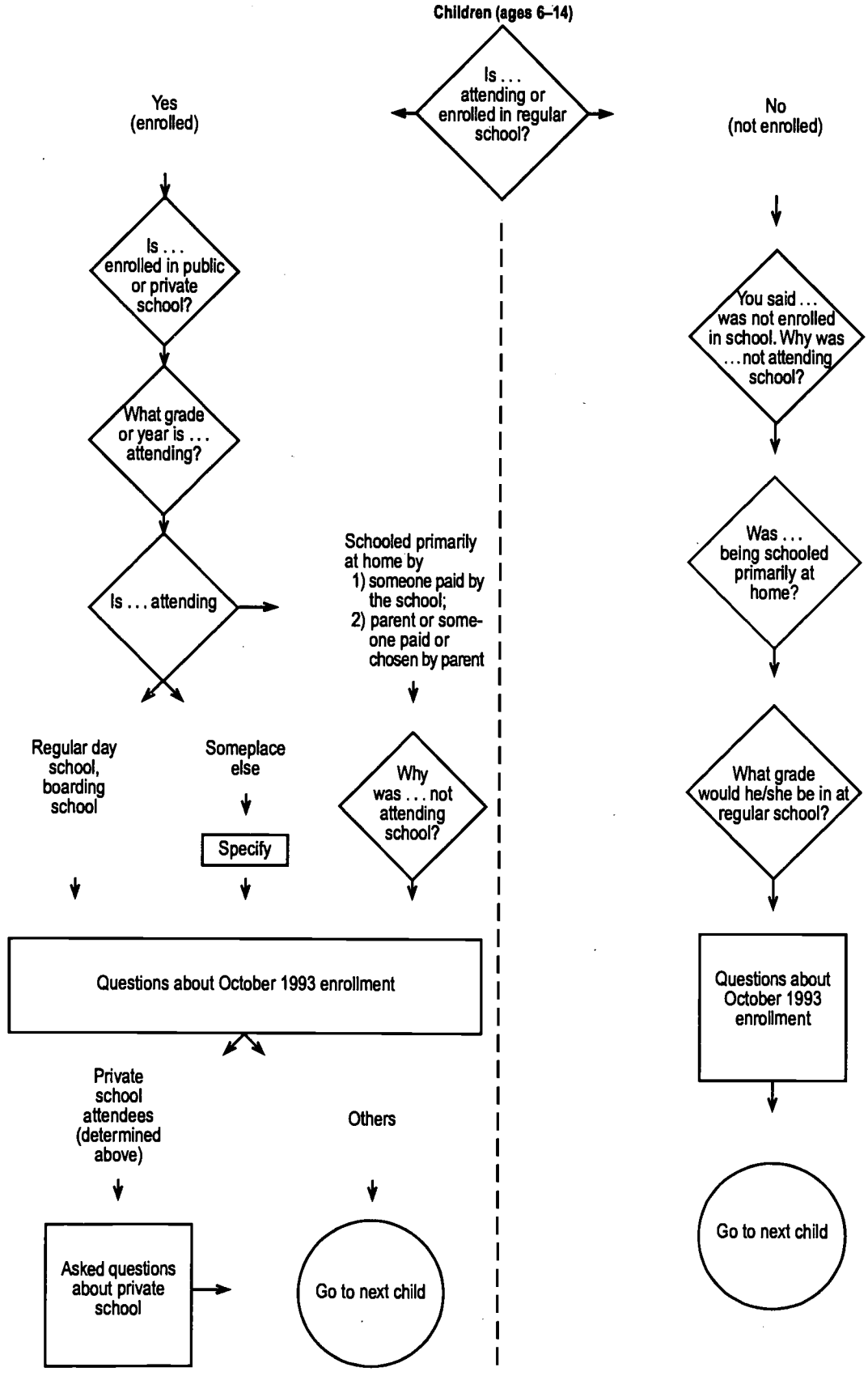
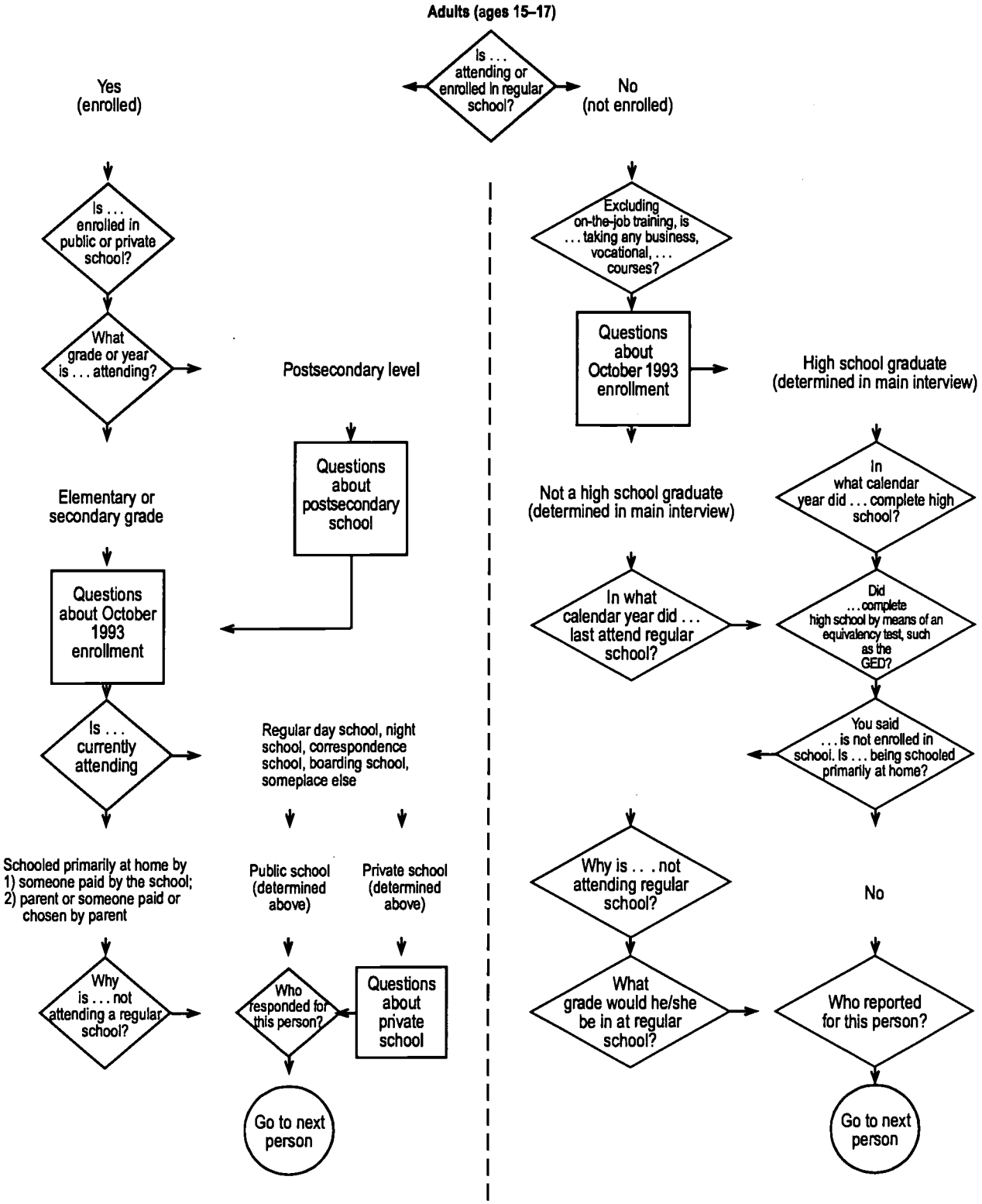


Figure 4—Items related to home schooling in CPS:October 1994 interview, by subject age and enrollment status —Continued



SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994: School Enrollment, Technical Documentation, Attachments 9 and 10B.

estimate of the number of children who were home schooled. Children who were schooled at home but not "primarily" at home—that is, partially at home and partially at school—may not have been identified as home schooled because some respondents in this situation may have indicated that the child was attending a regular day school or some alternative.[18]

5.2.1.2 NHES:96

Respondents in NHES were asked whether the sampled child was enrolled in school (figure 5). Regardless of the child's enrollment status, interviewers followed up with "Some parents decide to educate their children at home rather than sending them to school. Is (CHILD) being schooled at home?" Positive responses to this question were clarified with the follow-up question "So your child is being schooled at home *instead* of at school?" If the respondent answered "yes" to this question, the response to the first question remained unchanged, but if the respondent answered "no" to this question, the response to the first question was recoded as negative, in other words, that the child was not home schooled.[19] The follow-up question was added because researchers found that parents who supplemented their children's education at school with instruction at home reported their children were being schooled at home when asked only the first question.[20]

5.2.2 Question Sequencing

The number of items needed to estimate the number of home-schooled children and the complexity of those items' sequence is considerably greater in the CPS:Oct94 interview than in the NHES interviews. As illustrated in figure 4, the items related to home schooling in the

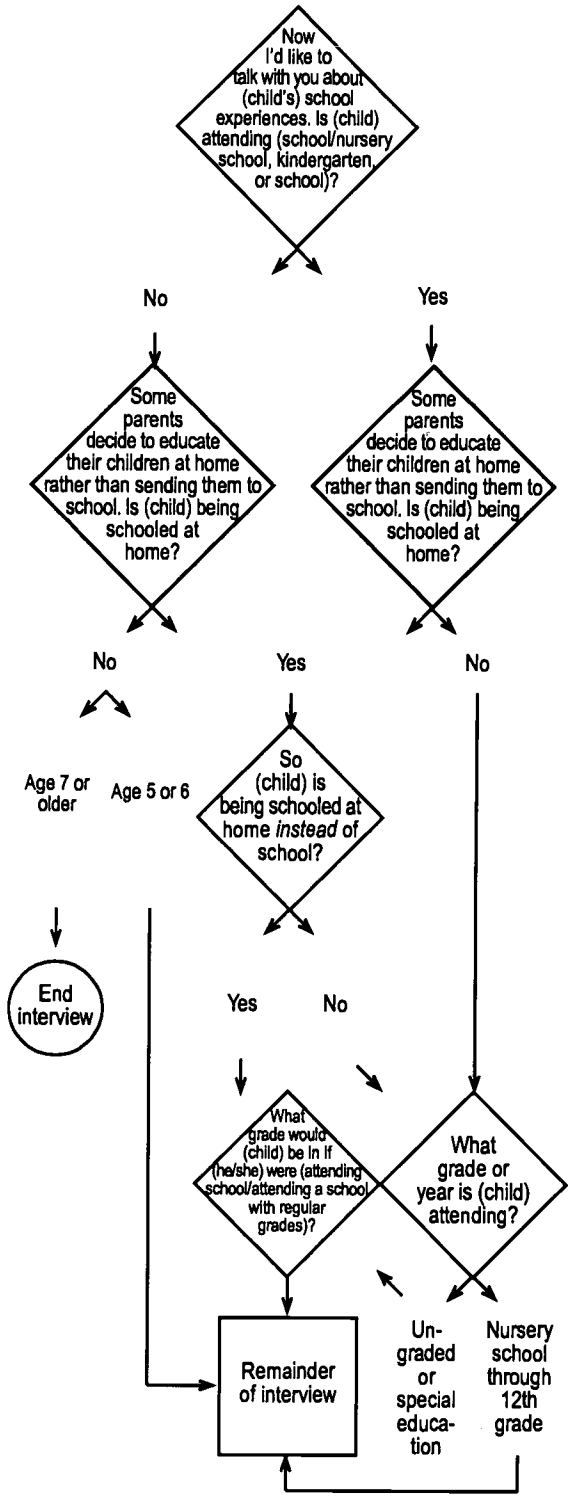
[18]Respondents were also asked whether nonenrolled children were schooled "primarily" at home. However, it seems unlikely that the word "primarily" would affect the estimates of nonenrolled children who were schooled at home as it is unclear what alternatives could have been made available.

[19]This recoding was done on-line, rather than after data collection, which prevents examination of whether the 1996 responses would have more closely resembled the 1995 responses if the follow-up question had not been asked.

[20]This follow-up question was not asked in the NHES:95. The addition of the follow-up question "So your child is being schooled at home *instead* of at school?" in the 1996 interview appears to have decreased the estimate of home-schooled children from that generated from the 1995 data. According to these estimates, a higher percentage of 6- to 8-year-olds were home schooled in 1995 (2 percent) than in 1996 (1 percent). This apparent change was observed at a time when home-schooling advocates claim that home schooling was rapidly becoming more prevalent. However, it is not possible to determine the degree to which the apparent difference between the estimates results from the additional question in 1996 rather than changes in population values over time.

In addition, the possibility that this question addition could affect the estimates raises the issue of whether home schooling is an all-or-nothing phenomenon or a continuum between full-time, full-week schooling at school and full-time, full-week schooling at home. Children might, for example, receive instruction at school on some days per week or in some subjects while being schooled the remainder of the time at home. Alternatively, home-schooled children may receive some instruction from school personnel and the rest from their parents. In either instance, it is unlikely that parents of all children in these situations would have answered the home-schooling questions in the same way.

Figure 5—Items related to home schooling in NHES:96 interview



SOURCE: Collins et al. (1997) *National Household Education Survey of 1996, Data File User's Manual, Volume I* (NCES 97–425). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

CPS:Oct94 are many and follow a complex skip pattern. Within the adult and child interviews, the question "Is . . . attending or enrolled in regular school?" determined which of two paths the interview then took, as both interviews approached home schooling differently for enrolled versus non-enrolled subjects. Children who were both enrolled and home schooled were identified through the question regarding the type of school they attended. Children who were not enrolled were identified by responses to the question "Is/Was . . . being schooled primarily at home?" Other questions regarding grade level, enrollment status in October 1993, and reasons why children were not enrolled in regular school were inserted in various places among the four sequences. In contrast, identifying home-schooled children in the NHES interview relied on no more than three items that were asked in the same order regarding all children (figure 5).

In general, the more questions involved in determining a subject's status, the more opportunities there are for missing data or measurement error to reduce the accuracy of estimates. Although the high CPS response rates indicate relatively few missing data, measurement error remains an issue.[21]

5.2.3 Location of Home-Schooling Items and Respondent Fatigue

Survey designers must consider potential respondent fatigue when developing a survey. In general, items at the end of a long survey become vulnerable to respondent fatigue: the longer a survey, the more likely it is that respondents will become tired or impatient and refuse to answer questions that occur later in a survey or answer them inaccurately (Vaden-Kiernan et al. 1997).

In the CPS interviews, the interviewer first asks the basic monthly labor force questions regarding all adult household members before going on to the supplement questions. Supplement questions regarding children are asked last. Therefore, in households with many adults, respondents could become fatigued or impatient when answering the supplement questions near the end of the household's interview. If the average interview took a long time to complete, or if the length of the interview was particularly high for households with 6- to 17-year-olds, locating the home-schooling questions near the end of the interview could negatively affect data quality.

However, the mean duration of interviews in CPS:Oct94 was 13 minutes for households in which at least one supplement interview was conducted and 15 minutes for that subset of households in which at least one supplement interview was conducted regarding someone 6 to 17 years old. Interviews in children's households may have been longer than those in all households

[21]However, these complex skip patterns in the CPS:Oct94 Education Supplement items were necessary. In order to estimate school enrollment and high school completion and dropout rates consistently over time, two of the primary goals of the Education Supplement, the questions must accommodate both adults and children and both those who are enrolled and those not enrolled in school. They must also be consistently worded in each Education Supplement.

where a supplement interview was completed because more supplement interviews were necessary in households with a 6- to 17-year-old. In households where at least one supplement interview was completed for a 6- to 17-year-old, interviews were conducted on an average of four people. In households where the supplement was completed regarding at least one person, regardless of age, interviews were conducted regarding two to three people, on average.

In contrast, the enrollment and home-schooling items were asked early in the NHES interviews. These items were asked only in the Screener interview if the Screener respondent was the extended interview respondent, and were verified at the beginning of the extended interview if the respondent to the extended interview was different from the Screener respondent. The average length of the PFI/CI interview was 19 minutes, and the average length of a Screener interview where only the parent was sampled was 6 minutes (Vaden-Kiernan et al. 1997). Thus, if two children in the household were sampled for the PFI/CI interview, the second PFI/CI interview would have begun, on average, 25 minutes into the total interview time.

The CPS:Oct94 and NHES:96 interview administration times reported here are not directly comparable at least in part because the CPS:Oct94 administration time data were not edited, whereas the NHES:96 administration time estimates were generated from data that had been edited to reduce the effects of outliers on the estimate of the mean interview administration time (Vaden-Kiernan et al. 1997). Outliers can occur, for example, when interviews are interrupted so a respondent can take care of something in the household (e.g., a pot boiling over or a crying child). Although the time that elapsed while the respondent was away from the phone is counted in the CATI's automatic timing of the interview, this time is not actually part of the interview and therefore leads to measurement error. Such measurement errors no doubt occur during CPS interviews as well, but the CPS data were not edited to take them into account. Nevertheless, the CPS interviews were considerably shorter than the PFI/CI interviews. Therefore, the supplement's position at the end of the household's interview does not appear likely to have fatigued respondents more than they would have been at the end of the NHES:96 home-schooling questions.

5.3 Respondent Error

Respondent error occurs when, for any reason, respondents provide inaccurate responses to a question. Respondents may respond inaccurately for a number of reasons: for example, they may misunderstand the intent of the question, they may not know the correct answer, or they may be unwilling to answer correctly. This section examines various potential sources of respondent error in the CPS:Oct94 and NHES:96.

5.3.1 Respondents' Relationships to Children

The respondent's relationship to the child about whom an interview is conducted is likely to be associated with the quality of the information he or she provides about the child. Primary caregivers, usually parents, are more likely than other household adults to know and be able to report accurately the details of children's educational experiences.

The CPS:Oct94 respondents may be somewhat different from the respondents to the NHES:96 PFI/CI extended interviews because the instructions given to interviewers for choosing respondents differed somewhat. CPS interviewers were instructed to interview only household members who were more than 14 years old and to "try to interview the most knowledgeable member of the household" (Reed 1997). In practice, this respondent is usually the reference person, who is the person who owns or rents the housing unit, or his/her spouse.[22] NHES:96 Screener interviews were conducted with a household member 18 years old or older. If another person in the household was most knowledgeable about the care and education of the child who was sampled for the PFI/CI extended interview, that most knowledgeable person became the respondent for that extended interview.

To the degree that this issue can be examined empirically, the differences in the selection of respondents are reflected in the data. The NHES:96 data indicate that mothers were the most common respondents to the interview: 77 percent of respondents were the children's mothers, and another 18 percent were the children's fathers (table 7). The respondents for home-schooled children were not different from those for children who were not home schooled.

Data regarding the relationship of the respondent to the children in the CPS:Oct94 questions are less available and are not directly comparable to the NHES data. The Education Supplement included two different sets of questions: one for household members aged 15 and older (referred to as the "adult" interview), and one for those age 14 and younger (the "child" interview). The adult questions concluded with "Who responded for this person?" but no such question was asked regarding children. The responses to this question indicate that, as with the NHES data, parents were the most common respondents for teenagers 15 to 17 years old (table 8). However, 9 percent of 15- to 17-year-olds responded to the CPS:Oct94 interview for themselves.

[22]This description of the respondent is synopsized from the CPS Interviewer's Manual. The Bureau of Labor Statistics *Handbook of Methods* (U.S. Department of Labor 1997, p. 10) describes a somewhat different procedure: "At each monthly visit, each household member 16 years of age and over is asked a series of standard questions on economic activity during the preceding week." However, later in the Handbook (p. 12) another statement about the respondents is more consistent with the Interviewer's Manual description: "Generally the persons who provide information for the monthly CPS questions also answer the supplemental questions. Occasionally, the kind of information sought in the special survey requires the respondent to be the person about whom the questions are asked."

Table 7—NHES:96 estimates of the percentage distribution of 6- to 17-year-olds according to relationship to interview respondent, by home-schooling status

	Mother	Father	Grandparent	Other
Total	76.5	18.2	2.6	2.6
Home-schooling status				
Educated at home	79.6	15.6	2.2	2.5
Educated at school	76.4	18.3	2.6	2.6

NOTE: Percentages may not sum to 100 due to rounding.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement in Education and Civic Involvement Component (NHES:96, PFI/CI).

Table 8—CPS:Oct94 estimates of the percentage distribution of 15- to 17-year-olds according to relationship to interview respondent, by home-schooling status

	Unknown	Self	Parent	Spouse	Other relative	Nonrelative
Total	0.1	8.8	79.6	1.5	8.5	1.5
Home-schooling status						
Educated at home	0.0	21.5	64.1	0.0	14.4	0.0
Educated at school	0.1	8.7	79.7	1.5	8.5	1.5

NOTE: Percentages may not sum to 100 due to rounding.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Although it appears that home-schooled 15- to 17-year-olds were more likely to report for themselves than their peers in school, this difference is not statistically significant.

5.3.2 Political and Legal Contexts

Items on home schooling may be affected by the political context in which the questions were asked. It has been suggested that many parents who home school may be reluctant to answer questions about home schooling because they resent government intrusion or fear some sort of stigma or sanction (Kaseman and Kaseman 1991; Ray 1997). Respondents who are generally apprehensive about government intrusion into their lives may have declined to participate in the survey more often than others or may have answered questions less accurately. On the other hand, to the degree that home schooling has become more common and parents more active with respect to working with local education agencies (LEAs) and other government agencies to fa-

cilitate home schooling, parents who home school may not be reticent to participate in interviews or answer questions regarding home schooling.

One component of the political context of home schooling is the legal context, that is, how home schooling is defined or treated legally with respect to compulsory education laws. Legally, "home school" is defined differently among the 50 states (Lines in press). In some states, for example, home schools are clearly defined as private schools, and authorities recognize parents' rights to teach children in private schools in their homes. In other states, home schools are treated as if they were private schools, but some authorities dispute that right. In yet a third category of states, home schools operate under specific legal provisions that distinguish them from private schools. In many states the legal status of home schooling fits in none of these categories (Scott Somerville—personal communication, May 12, 2000). Given the differential legal status of home schooling among states, both the rate of home schooling and the rate at which home-schooling parents report that they home school their children are likely to vary among states. Without additional data, it would not be possible to determine whether variation between the two data sets in state-level rates of home schooling, if such variation exists, is attributable to differences in question wording and their relationships to the legal status of home schooling in different states.

As noted above, the NHES:96 Screener response rate was considerably lower than the CPS:Oct94 household participation rate, although the effects of that rate are somewhat mitigated with respect to home-schooling estimates because many of the nonresponding households would not have had school-aged children. In addition, questions about home schooling were answered by most NHES PFI/CI respondents. Whether the political, including the legal, context regarding education or home schooling in particular affects the rates of home schooling or leads respondents to provide less than accurate responses cannot be studied further with these data.

5.3.3 Cognitive Context

Survey researchers assume that respondents rely on the rules of normal conversation when they attempt to make sense of a survey item or set of items (Levinson 1983). Accordingly, respondents tend to try to provide information that the questioner is interested in. This "cooperativeness principle" implies that respondents will search for cues to the intent of the questioner and use these cues to tailor their response to any particular item. Question wording plays a significant role in respondents' understanding of what information the interviewer is seeking, but other information also provides cognitive context that informs respondents' understandings of what they are being asked.

Although both the CPS:Oct94 Education Supplement and the NHES interview include questions regarding home schooling, they do so within the context of surveys with quite different purposes. In the introductions to the CPS interview, respondents were told that the CPS is conducted by the Census Bureau "to collect information on the number and types of jobs Americans need" (Reed 1997). Respondents were also informed that their participation was voluntary and that all information they provided was confidential. By the time they reach the supplement questions, all respondents would have found that CPS is primarily a labor force survey of the entire population, and therefore largely oriented toward persons 15 years old and older.

NHES:96 PFI/CI respondents were told that NHES is sponsored by the Department of Education, which "is conducting a voluntary and confidential study on the educational experiences of children and how they learn about their communities and government" (Collins et al. 1997). Their experience of the interview began with the Screener interview, which identified children in the household, and then proceeded with questions about the sampled children's educational experiences.

Thus the cognitive contexts within which respondents interpreted questions and provided answers differed between the two surveys. However, whether and how this difference affected respondents' interpretations of the questions, and therefore their answers, cannot be determined from the available evidence.

5.4 Conclusion on Observational Errors

This section examined three types of observational error: error due to data collection methods, instrument error, and respondent error. With respect to data collection, three differences were discussed: the use of personal and telephone interviewing in CPS whereas NHES used telephone interviewing exclusively; panel (CPS) versus cross-sectional (NHES) design; and timing (October 1994 versus January to April 1996). Although each of these could contribute to differences between the surveys' estimates, the direction and magnitude of these differences cannot be estimated with available data.

The surveys' instruments varied as well in that CPS:Oct94 and NHES:96 used different question wording and question sequencing in probing respondents about children's home schooling. Although these differences may have contributed to the observed differences in estimates of the incidence of home schooling, this hypothesis cannot be tested with available data. However, if the "instead of at school" wording of the NHES:96 estimate had the effect of excluding children who were schooled in part at home and in part at school, whereas the CPS:Oct94 wording of "primarily" did not, one would expect the NHES:96 estimate to be lower

than the CPS:Oct94 estimate, and it is not. Confirmation of these effects would have to come from other sources—perhaps experiments in a cognitive laboratory or analysis of additional data.[23] In addition, although the CPS:Oct94 items on home schooling were asked toward the end of the interview, on average CPS:Oct94 interviews were relatively short and therefore respondent fatigue is not likely to have affected respondents' answers.

Three potential sources of respondent error—respondents' relationships to the children, the political context of the surveys, and their cognitive context—were also discussed. Relatively more of the CPS:Oct94 data may have been gathered from sources such as other relatives (not parents or spouses) and nonrelatives, who may be less knowledgeable than parents. However, questions about whether children are home schooled are straightforward and are probably less vulnerable to misinformation than are other questions, such as questions regarding the reasons for which children are home schooled. Therefore, if the surveys differ in the relationships of the interview respondents to the children and if these differences negatively affect the accuracy of the CPS:Oct94 data with respect to identifying home-schooled children, the effect is likely to be minimal. The political and cognitive contexts of these surveys may have affected participation rates and the accuracy of responses, but data to test for or estimate the size or direction of such effects are not available.

[23]The NHES:99 interview included similar questions about home schooling but added questions that will allow researchers to distinguish between complete and partial home schooling. Such discriminations may shed light on the findings presented here.

6. Data Processing Errors

Finally, errors in estimates may occur during the data processing conducted during and after data collection. Examples of these errors include data entry errors and programming errors made when data are manipulated and estimates are computed.

One of the advantages of computer-assisted interviewing is the ability to build error checks into the questionnaire. CATI/CAPI programs often include on-line editing routines, which alert interviewers when respondents provide, or interviewers mistakenly enter, responses that are implausible or inconsistent with previous responses. The interviewer can then verify the data with the respondent and correct errors. In addition to on-line editing, survey researchers routinely, although not always, edit data after they have been collected to correct identifiable errors and impute missing responses.

Although both CPS:Oct94 and NHES:96 used computer-assisted interviewing, their editing procedures varied. On-line edits have been programmed for some of the CPS labor force items, but the education supplement did not include on-line edits. In addition, although many of the education supplement items were edited after data collection was completed, the home-schooling items were not edited during or after the interview (U.S. Department of Commerce, Bureau of the Census, Rosalind Bruno, personal communication, December 1998 and February 2000).

The NHES:96 PFI/CI included on-line edits for items related to children's age, grade or grade equivalent, and relationship to the respondent (Collins et al. 1997), and as these items were used to select the sample of children for the analysis, their quality affects the quality of the estimates presented. Missing data on all NHES items were imputed using a hot-deck procedure.

The lack of editing in CPS:Oct94 leads to the problem of missing data, which could reduce the estimate of the number of home-schooled children relative to an edited CPS:Oct94 estimate. As noted in section 5, however, missing data are not likely to reduce the estimate of the CPS:Oct94 estimate very much. Whether editing the CPS:Oct94 data also affects the estimates because errors were not corrected cannot be determined with available data.

7. Conclusion

The point estimates of the number of 6- to 17-year-olds in kindergarten through grade 12 who were home schooled range from 345,000 to 636,000 between the CPS:Oct94 and the NHES:96. The corresponding point estimates of the proportion of children in this age/grade-level group who were home schooled range from 0.8 percent to 1.4 percent, both rounding to about 1 percent of the age 6 through grade 12 population. Taking sampling variance into account, the 95-percent confidence interval around the CPS:Oct94 estimate ranges from 287,000 to 402,000 children, and the 95-percent confidence interval around the NHES:96 estimate ranges from 515,000 to 757,000 children.

Although it is possible that the number of home-schooled children increased in the 15 to 18 months between the two data collections, such a large increase appears unlikely (Ray 1997; Lines 1998). Therefore, this report examined the data sources with respect to three categories of errors in surveys—nonobservation errors, observation errors, and data processing errors—that may have contributed to the observed difference between the surveys' estimates (table 9). With respect to nonobservation errors, which are errors of population coverage, the report discussed differences between the CPS:Oct94 and NHES:96 in sample coverage and response rates. Examination of the sampling frames indicated that each survey undercovers some portion of the population. However, to the extent that weighting does not offset this undercoverage, analyses indicate that the resulting biases in estimates of the number and percentage of home-schooled children are likely to be small. In addition, although the sample designs of the two studies differed, there is no reason to believe that both of the resulting samples are not representative of the population.

Response rates were computed at three levels for each survey. Although the rates at all three levels of the CPS:Oct94 were at least 90 percent and the NHES:96 extended interview and item response rates were 89 percent or higher, the NHES:96 Screener rate was 70 percent. It has been suggested that home-schooling households may be less inclined to participate in government research, particularly research that might identify them as home-schooling families (Kaseman and Kaseman 1991; Ray 1997). The NHES:96 Screener response rate of 70 percent is not inconsistent with this hypothesis. However, approximately 30 percent of households in the United States include children in grades K–12. Therefore, any children who were home schooled

Table 9—Differences between data sources, by type of error

Type of error	Difference between data sources
Nonobservational errors	
Sample coverage	Sampling frames for both CPS and NHES exclude institutionalized children and some other children (e.g., children in households without telephones are excluded from NHES) but any resulting biases are likely to be small. Sample designs differ but are unlikely to affect estimates.
Response rates	CPS survey-level response rate is higher than NHES rate (94 percent versus 70 percent respectively).
	CPS supplement response rate is 97 percent compared with NHES PFI/CI response rate of 89 percent.
	CPS and NHES both had high item response rates (90 percent or more).
Observational errors	
Data collection	CPS uses both personal and telephone interviewing, NHES uses telephone only. CPS is panel survey, NHES is not. CPS conducted in October, NHES in January through April. If children are more or less likely to be home schooled at one time of the year than another, survey timing will affect estimates. However, whether children's likelihood of being home schooled varies within the calendar year cannot be determined from these data.
Instrument error	
Question wording	For nonenrolled subjects in CPS, interviewers asked whether primarily schooled at home. Interviews for enrolled subjects asked type of school attended. In NHES:96, when respondents indicated that children were being schooled at home, interviewers followed up by asking whether they were being schooled at home instead of at school.
Question sequencing	More items and a complex skip pattern used to identify home-schooled children in CPS:Oct94 than in NHES:96. Although item response rates were high, more items and skips allow more opportunities for error.
Respondent fatigue	There is no reason to suspect fatigue in either survey.
Respondent error	
Relationships to children	Given interviewers' instructions for choosing respondents, respondents may vary between surveys. Degree to which respondents differed with respect to their relationships to the children cannot be estimated with available data.
Political context	Politics of home schooling could affect estimates. There is no evidence available regarding this.
Cognitive context	Different survey sponsors and purposes may affect respondents' interpretations of questions and expectations regarding information desired. Effects of this cannot be estimated with available data.
Data processing	Missing data in CPS biases estimate of number of children downward to unknown extent. Bias with respect to number of home-schooled children may be as high as 9 percent but is likely to be lower.

in the approximately 9 percent of nonresponding households (30 percent of the 30 percent) who had children in grades K–12 were excluded from home-schooling estimates due to household nonresponse. In addition, the high home-schooling item response rates in both surveys raise questions regarding whether respondents perceived home schooling to be a sensitive issue. In general, survey respondents are less likely to answer questions regarding issues, such as income, that they perceive to be sensitive or private.

In addition to examining errors of nonobservation, the report also examined the surveys' potential for observational errors: errors in data collection, instrument error, and respondent error. Data collection procedures varied between the two surveys in several ways. Although at least one-quarter of the CPS:Oct94 interviews were conducted in person, all NHES:96 interviews were conducted by telephone. CPS is a panel survey, and NHES is not. The two surveys were conducted 15 to 18 months apart, one in October and the other from January though April. However, the effects of these differences on the estimates cannot be estimated.

Given the very different purposes of the two surveys, it was not surprising to find that they differed with respect to the potential for instrument error. Questions regarding home schooling were worded differently and occurred in different sequences and different locations in the two interviews. Whether question wording and sequencing might have affected the estimates could not be determined. Whereas NHES:96 located the home-schooling items at the very beginning of the interview, CPS:Oct94 located the home-schooling items near the end of the interview. However, the overall short duration of the average CPS:Oct94 interview indicates that, despite the location difference, respondent fatigue is not likely to have affected CPS:Oct94 respondents' answers more than those of NHES:96 respondents.

The report examined three potential sources of respondent error. First, the respondents' relationships to the sampled children may have differed between the two data sets. If so, this difference is more likely to have affected the accuracy of the home-schooling responses in CPS:Oct94 than that of the NHES:96 home-schooling responses. Without information on the relative accuracy of various types of respondents' answers to questions about home schooling as well as data on the relationships between respondents for 6- to 14-year-olds in CPS:Oct94, this issue could not be pursued more thoroughly.

To the extent that respondents perceived home schooling as a political issue, the political context may have affected the estimates, although whether the political context would have affected one set of estimates differently from the other is unknown. Similarly, the different sponsors and purposes of the two surveys may have provided respondents with distinct cues as to the kinds of information the interviewers were asking for, and these differences may have contrib-

uted to the difference between the surveys' estimates. These issues could not be pursued empirically with the available data.

Finally, with respect to data processing errors, the surveys differed with respect to their use of on-line and post-data collection editing. Whereas missing data on some of the items used to compute the home-schooling estimates were imputed in CPS:Oct94, missing data on all such items were imputed in NHES:96. It is not possible to determine with much certainty how many home-schooled children were not counted in the CPS:Oct94 estimate due to missing data. However, as noted in section 4, it is likely to be less than 30,000 children.

Thus, although the methodologies of these surveys are sufficiently rigorous to produce relatively accurate national estimates regarding home schooling and a variety of other issues, their methods differ in a variety of ways that may be associated with differences between the estimates. The overall direction and magnitude of the expected difference between the surveys' estimates cannot be estimated with the currently available data.

However, the questions used to study home schooling in these surveys raise a number of issues regarding students' schooling that could affect estimates of the number and proportion of children who are home schooled. Students could be schooled partially at home or partially at school, schooled completely at home but only for a limited time due to temporary illness or a school disciplinary action, or schooled at home by school personnel due to disability or illness. Which of these children to define as "home schooled" depends on the researcher or policymaker's perspective. Therefore, in order to obtain the most reliable estimate of the number of home-schooled children, whatever one's definition, survey items must allow respondents to describe their children's schooling arrangements in some detail. Without such detail, it is difficult to know what estimates of the number and proportion of "home-schooled" children mean. Items posed in the NHES:99 allowed respondents to describe their home-schooling arrangements in greater detail than in the CPS:Oct94 or the NHES:96 interviews. Future research with these data, in addition to cognitive laboratory work, may improve subsequent estimates of the number and percentage of home-schooled children.

8. References

Belluck, P. (1998, November 1). Life After Home Schooling. *New York Times,* Education section.

Brick, J.M., Waksberg, J., Kulp, D., and Starer, A. (1995). Bias in List-Assisted Telephone Samples. *Public Opinion Quarterly, 59*: 218–235.

Clark, H.H. and Schober, M.F. (1992). Asking Questions and Influencing Answers. In J.M. Tanur (Ed.), *Questions About Questions* (pp. 15–48). New York: Russell Sage Foundation.

Collins, M.A., Brick, J.M., Nolin, M.J., Vaden-Kiernan, N., Gilmore, S., Chandler, K., and Chapman, C. (1997). *National Household Education Survey of 1996, Data File User's Manual, Volume I* (NCES 97–425). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Fowler, F.J. (1993). *Survey Research Methods*. Newbury Park, CA: Sage.

Groves, R.M. (1991). Measurement Errors Across the Disciplines. In P.B. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (Eds.), *Measurement Errors in Surveys* (pp. 1–25). New York: John Wiley and Sons.

Jaeger, R.M. (1984). *Sampling in Education and the Social Sciences*. New York: Longman.

Kaseman, S. and Kaseman, L. (1991). Does Homeschooling Research Help Homeschooling? *Home Education Magazine, 8*(1): 26–27, 46–49.

Kohn, A. (1988). Home Schooling: Parents Who Educate Their Children at Home Face Little Official Opposition—But a Very Demanding Responsibility. *Atlantic Monthly, 261*(4): 20–25.

Levinson, S.C. (1983). *Pragmatics*. Cambridge, England: Cambridge University Press.

Lines, P.M. (1991). *Estimating the Home-Schooled Population* (Working Paper OR 91–537). Washington, DC: U.S. Department of Education.

Lines, P.M. (1996). Home Schooling Comes of Age. *Educational Leadership, 54*(2): 63–67.

Lines, P.M. (1998). *Homeschoolers: Estimating Numbers and Growth.* U.S. Department of Education. Washington, DC: Office of Educational Research and Improvement.

Lines, P.M. (in press). The Home Education Movement in Context, Practice and Theory. *Peabody Journal of Education, 75*(1, 2).

Montaquila, J., Brick, J.M., and Brock, S.P. (1997). *Undercoverage Bias in Estimates of Characteristics of Households and Adults in the 1996 National Household Education Survey* (NCES 97–39). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Montaquila, J. and Brick, J.M. (1997). *Unit and Item Nonresponse, Weighting, and Imputation Procedures in the 1996 National Household Education Survey* (NCES 97–40). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Ray, B.D. (1997). *Strengths of Their Own.* Salem, OR: National Home Educational Research Institute Publications.

Ray, B.D. (1999). *Facts on Home Schooling.* Salem, OR: National Home Educational Research Institute Publications. Available: http://www.nheri.org/98/research/index.html

Reed, M. (1997). *CPS Basic Monthly Survey Interviewer's Manual.* Washington, DC: U.S. Department of Labor. Available: http://www.bls.census.gov/intmana2.htm

Schwarz, N. (1999). Self-Reports: How the Questions Shape the Answers. *American Psychologist, 53*: 93–105.

Shea, C. (1996). From Home to College: Admissions Officers Grapple with How to Evaluate Applicants Educated Outside of School. *Chronicle of Higher Education, 42*(21): A31–A32.

Snedecor, G.W. and Cochran, W.G. (1967). *Statistical Methods* (6th ed.). Ames, IA: The Iowa State University Press.

Sudman, S., Bradburn, N.M., and Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology.* San Francisco, CA: Jossey-Bass.

Tanur, J.M. (Ed.). (1992). *Questions About Questions.* New York: Russell Sage Foundation.

U.S. Department of Commerce, Bureau of the Census. (1999). *Current Population Survey, Methodology.* Available: http://www.bls.census.gov/cps

U.S. Department of Commerce, Bureau of the Census, Data User Services Division. (1994). *Current Population Survey, October 1994: School Enrollment. Technical Documentation CPS–94*. Washington, DC: Author.

U.S. Department of Education, National Center for Education Statistics. (1997). *1991, 1993, 1995, and 1996 National Household Education Surveys* [CD-ROM] (NCES 97–426). Washington, DC: Author.

U.S. Department of Labor, Bureau of Labor Statistics. (1997). *Handbook of Methods*. Washington, DC: Author.

Vaden-Kiernan, N., Nicchitta, P.G., Montaquila, J., and Collins, M.A. (1997). *Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey* (NCES 97–35). U.S. Department of Education. Washington, DC: National Center for Education Statistics Working Paper.

Walsh, M. (1998). Proposed Tax Credit Drawing Support in Colo. *Education Week, 18*(8): 1, 20–21.

Appendix A—Technical Notes

Given the subject of this report, information on the surveys, sample selection, and component response rates is provided in the text. This section provides additional information regarding overall response rates, the data files used in the analyses, calculations estimating the effects of particular sources of error, variable definitions, software used to compute estimates and variances, and statistical procedures.

Overall Response Rates

Section 4.2 discusses the household- or survey-level, supplement- or interview-level, and item-level response rates in both the CPS:Oct94 and NHES:96 data. Combining these two levels yields an overall response rate. For CPS:Oct94, the overall response rate is the product of the household- and supplement-level rates (94 percent * 97 percent), or 91 percent. For NHES:96 PFI/CI, the overall response rate is the product of the survey- and extended-interview-level rates (70 percent * 89 percent), or 63 percent (Collins et al. 1997).[24]

Although this overall response rate falls below the NCES minimum standard of 70 percent, a response-bias analysis of survey nonresponse by age of household members indicated households with children 0–2 years old and households with persons 60 years old or older were slightly underrepresented relative to CPS estimates (Montaquila and Brick 1997). Therefore, although the overall NHES response rate was lower than desirable, the estimates of households with school-aged children appear not to have been compromised.

Data Files

The CPS:Oct94 data used for this report were obtained from two data files. The first was a Census Bureau file provided to MPR Associates by NCES and is referred to as the "NCES file" in this report. This file did not include some variables that were needed for this analysis, however, including whether the household had a telephone. These additional variables were obtained from a file produced by Unicon Research Corporation in Santa Monica, California, as part of its

[24]Due to rounding, the overall response rate reported in Collins et al. (1997) differs slightly from the product of the component response rates as reported in this sentence.

CPS Utilities series.[25] This file is referred to as the "commercial file" in this report. The NHES data files were extracted using the software provided on the public release compact disk available from NCES (U.S. Department of Education 1997).

Impact of Emancipated Minors

The NHES:96 PFI/CI sample excluded "emancipated minors," defined as youth who were not living with a parent or guardian or with someone at least 12 years older than they who was not a spouse. This sample restriction could not be replicated in the CPS:Oct94 sample because the necessary data were not available. Therefore, it was important to assess whether including emancipated minors in the CPS:Oct94 sample but not in the NHES:96 sample affected the CPS:Oct94 estimates of the number and proportion of home-schooled children relative to the NHES:96 estimates.

Given the data available from CPS, this issue was addressed by dividing 15- to 17-year-olds into one of three categories: not emancipated minors, emancipated minors, and potential emancipated minors. The vast majority of youth—97 percent of the 15- to 17-year-olds—were identified as not emancipated minors because they lived with a parent, foster parent, or grandparent (table A1). Less than 1 percent of 15- to 17-year-olds were identified as emancipated minors because they lived alone or with only a spouse. Potential emancipated minors—those who could not be identified as emancipated or not—made up 3 percent of 15- to 17-year-olds. Of those who were or could be emancipated minors, 9,000 (3 percent) were home schooled.

If the 338,000 emancipated minors and potential emancipated minors were excluded from the analysis, an estimated 336,000 6- to 17-year-olds would be home schooled (345,000 minus 9,000) from a population of 42,825,000 children. The resulting estimate of the percentage of home-schooled children, 0.8 percent, does not differ from that observed when emancipated and potential emancipated minors are included in the analysis, which is presented in table 1. Therefore, the presence of emancipated minors in the CPS:Oct94 but not the NHES:96 sample does not appreciably affect the difference between the two data sets' estimates.

Impact of Reasons for Home Schooling

In this report, children were defined as home schooled regardless of who taught them and why they were taught at home. However, both the CPS:Oct94 and NHES:96 interviews included items that allow identification of children who were taught at home but might not be considered

[25]Further information on the Unicon Corporation can be obtained from the Web site: http://www.unicon.com.

Table A1—CPS:Oct94 estimates of the number and percentage distribution of 15- to 17-year-olds according to whether emancipated minor and number and percentage distribution of emancipated minors according to home-schooling status

	15- to 17-year-olds[1]	
	Number	Percent
Total	9,944,000	100.0
Whether emancipated minor		
Not emancipated minor	9,606,000	96.6
Emancipated minor[2]	9,000	0.0[3]
Potential emancipated minor[4]	329,000	3.3
	Emancipated minors and potential emancipated minors	
Total	338,000	100.0
Home-schooling status		
Educated at home	9,000	2.7
Educated in school	329,000	97.4

[1]Defined as children 15 to 17 years old who had not graduated from high school or earned a GED, and who were enrolled in school or home schooled in grades K–12. Excludes children who did not meet these criteria or whose status regarding these criteria could not be determined due to missing data.

[2]Defined as children living alone or living with only a spouse.

[3]Less than 0.05 percent.

[4]Defined as children who lived with more than one other person but not with a parent, foster parent, or grandparent; or who lived with only one other person who was not a parent, or foster parent. An unknown proportion of these children are emancipated minors.

NOTE: Numbers are rounded to the nearest thousand. Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

"home schooled" by all researchers or policymakers (figure A1). Such children include those who were taught at home by school personnel or who were taught at home because of temporary illness or disability. The CPS:Oct94 data indicate that the vast majority of home-schooled children, 97 percent among 6- to 17-year-olds, were schooled by their parents or someone employed by their parents (table A2). These data also indicate that 94 percent of home-schooled children were educated at home for a reason other than illness or disability.[26]

The NHES:96 data do not indicate whether home-schooled children were served by school personnel. However, they do indicate that 88 percent of home-schooled 6- to 17-year-olds were

[26]These two situations are not mutually exclusive, and therefore there is some overlap. Combining the two variables, i.e., who did the home schooling and whether it was done for illness or disability, 93 percent of all home-schooled children were schooled at home by their parents or people paid or chosen by their parents for a reason other than illness or disability (not shown in table).

Figure A1—CPS:Oct94 items regarding reason for not attending regular school and NHES:96 items regarding reasons for home schooling

CPS: Oct 94

Items		Response Codes
Why is . . . not attending regular school? (15- to 17-year-olds) Why was . . . not attending school? (6- to 14-year-olds, enrolled) You said . . . was not enrolled in school. Why was . . . not attending school? (6- to 14-year-olds, not enrolled)	Verbatim responses hand-coded into one of response categories	Home schooling preferred (religious reasons) Negative feelings about public school Too young Will attend in near future (moved, time in country, lacking proper papers) Sick or disabled Other nonmandating (don't want to be held back, GED classes) Special school/vocational school Attending regular school Don't know Refused

NHES:96

What are the main reasons you decided to school (CHILD) at home?	[CODE ALL THAT APPLY.]	Religious reasons Can give child a better education at home Object to what school teaches Poor learning environment at school School does/did not challenge child Want private school but cannot afford it Could not get into a desired school Child has temporary illness Child has special needs/disability Parent's career Child not old enough for grade/to enter school Student behavioral problems To develop character/morality Problem with available public/private schools Family reason Transportation/distance/convenience Other _____

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994: School Enrollment, Technical Documentation, Attachments 9 and 10B; Collins et al. (1997). *National Household Education Survey of 1996, Data File User's Manual, Volume I* (NCES 97–425). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Table A2—CPS:Oct94 and NHES:96 estimates of the number and percentage of home-schooled 6- to 17-year-olds, by persons providing education and whether schooled at home because of illness or disability

	CPS (October 1994)			NHES (Spring 1996)	
	Total home-schooled children	Schooled by parents or persons paid by parents	Not because of illness or disability	Total home-schooled children	Not because of temporary illness or disability
Number of home-schooled 6- to 17-year-olds in grades K–12	345,000	333,000	322,000	636,000	559,000
Percent of all home-schooled 6- to 17-year-olds in grades K–12	100.0	96.6	93.6	100.0	87.8

NOTE: Numbers are rounded to the nearest thousand.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement in Education and Civic Involvement Component (NHES:96, PFI/CI).

schooled at home for reasons other than temporary illness or disability. This number is not strictly comparable to the statistic derived from the CPS interview because the NHES question allowed respondents to choose more than one reason for home schooling, whereas the CPS question allowed only one response. Thus, the apparently slightly higher rate of children schooled at home because of illness or disability in the NHES data compared with the CPS data is partly an artifact of the question wording. Nevertheless, between the two data sources, it appears that a large majority of home-schooled children were schooled at home by their parents, voluntarily, and not on a temporary basis.

Impact of Missing Data in CPS:Oct94

As discussed above, the NHES:96 sample includes children 6 to 17 years old who were either enrolled in school or home schooled, in kindergarten through the 12th grade or the equivalents thereof, and not emancipated minors. In order to estimate the number and proportion of home-schooled children with the CPS:Oct94 data using a population definition comparable to that of the NHES:96 data, the appropriate subsample must be selected from the full CPS:Oct94 sample.

Selecting CPS:Oct94 cases by age, an estimated 45,835,000 children were 6 to 17 years old (table A3).[27] Data regarding most of the remaining criteria (enrollment in school, home-

[27]Numbers and percentages reported are weighted and numbers are rounded to the nearest thousand.

Table A3—Estimating impact of missing data on CPS:Oct 94 estimates

	Total	Not enrolled	Enrolled
Total 6- to 17-year-olds	45,835,000	745,000	45,091,000
Whether included in analysis			
Excluded because high school graduate, GED recipient, or in pre-kindergarten	327,000	55,000	272,000
Excluded from analysis because missing data	2,345,000	528,000	1,817,000
Not home schooled	443,000	443,000	0
Home-schooling status unknown	1,902,000	85,000	1,817,000
Included in analysis	43,163,000	162,000	43,002,000
Home schooled	345,000	162,000	183,000
Educated at school	42,819,000	0	42,819,000
Percent home schooled	0.8	27.0	0.4
Estimated additional number home schooled if those whose status is unknown were home schooled at same rate as those whose status was known	30,000	23,000	7,000

NOTE: Numbers are rounded to the nearest thousand. Details may not sum to totals due to rounding.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

schooling status, grade or grade equivalent) were not always available due to missing data that were not imputed. In the CPS:Oct94 analyses presented in the body of this report, cases were excluded if they either did not meet the criteria for inclusion in the analysis or if, due to missing data, their status with respect to these criteria could not be determined. This section discusses the potential impact of excluding cases with missing data on the CPS:Oct94 estimates.

Identifying children's home-schooling status differed depending on whether children were enrolled or not enrolled in school. Of the 45,835,000 6- to 17-year-olds, an estimated 745,000 were not enrolled in school and the remaining 45,091,000 were enrolled.[28] Of the 745,000 who were not enrolled, 55,000 were excluded from the sample because they were enrolled in preschool or known to have completed high school or earned a GED. Another 162,000 met sample criteria and were identified as home schooled.

The remaining nonenrolled 6- to 17-year-olds were missing data on one or more of the variables needed to determine sample eligibility and therefore were excluded from the sample. Approximately 443,000 were reported to be neither home schooled nor graduated from high school but were missing grade-level data. An additional 85,000 were missing data on graduation

[28]Breakdowns may not sum to totals due to rounding.

status, grade level, or home-schooling status, and therefore some of these children may have been home schooled. Assuming that (1) all who were missing data met the criteria for inclusion in the population selected for these analyses, and (2) the 85,000 whose home-schooling status was unknown were home schooled at the same rate as the other nonenrolled children, 27 percent, an additional 23,000 6- to 17-year-olds would be home schooled.[29]

It is quite likely that fewer children with missing data were home schooled because nonenrolled children with missing data were disproportionately older compared with nonenrolled children who were eligible for the analysis. Youth 15 to 17 years old made up 13 percent of the nonenrolled children who were selected for the home-schooling analyses (all of whom were home schooled in grades K–12 or their equivalents) (table A4). In contrast, 15- to 17-year-olds made up 73 percent of the 85,000 whose home-schooling status was unknown and 74 percent of the 443,000 who were not home schooled but missing other data. It is quite likely, therefore, that many of the nonenrolled children who were excluded because of missing data were in fact dropouts. If so, the proportion of these children who were home schooled may be less than the 27 percent upon which the 23,000 estimate is based.

Missing data among enrolled children may also affect the estimates. Of the 45,091,000 enrolled 6- to 17-year-olds, 1,817,000 were reported to be neither a high school graduate nor GED

Table A4—CPS:Oct94 estimates of the number of nonenrolled children and percentage distribution according to age, by eligibility for home-schooling analysis

		Percentage distribution by age	
	Number	Percent 6- to 14-year-olds	Percent 15- to 17-year-olds
	Eligible for home-schooling analysis		
Total	162,000	86.8	13.2
	Not eligible for home-schooling analysis because of missing data		
Home-schooling status unknown	85,000	26.9	73.1
Not home schooled	443,000	26.2	73.8

NOTE: Children were eligible for the home-schooling analyses if they were 6 to 17 years old, enrolled in school or home schooled, and in grades 1 through 12 or their equivalents. Children who did not meet these criteria, or who were missing data needed to determine whether they met the criteria, were deemed ineligible for the analyses. Numbers are rounded to the nearest thousand.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

[29]The 27 percent rate is calculated as follows: 162,000 home-schooled / (162,000 home-schooled + 443,000 not home-schooled).

recipient and enrolled in grades K–12, but their home-schooling status was unknown (table A3). If one assumes that these children were home schooled at the same rate as other enrolled children, 0.4 percent, an additional 7,000 children would be home schooled.[30]

In sum, if an additional 30,000 home-schooled children (23,000 + 7,000) were added to the previously identified 345,000, the total estimate of the number of home-schooled children would be 375,000. The resulting estimated percent of 6- to 17-year-olds who were home schooled would remain 0.8 percent. Thus missing data do not affect the CPS:Oct94 estimates substantially.

CPS:Oct94—Definitions of Variables Used in the Report

CPS–Age

AGE: Age reported in years as of the end of the survey week is available for all cases. Missing data were imputed. The variable AAGE provided on the commercial file indicates whether the value of AGE was imputed and, if so, the type of imputation procedure used. The item-level response rate provided in this report is the proportion of cases for which values were not imputed.

CPS–Enrollment

ENROLL: This variable combines results from child and adult enrollment questions. The child interview included two enrollment questions, one for children aged 3 to 5 and another for children 6 to 14. The question for 3- to 5-year-olds (given variable name CHBEG on the commercial file) reads as follows: " Is . . . attending or enrolled in nursery school, kindergarten or elementary school?" The question for 6- to 14-year-olds (CHATT on the commercial file) reads as follows: "Is . . . attending or enrolled in regular school? (Regular school includes nursery school, kindergarten, elementary school and schooling which leads to a high school diploma.)" The enrollment question for the adult interview (SCHATT on the commercial file) reads as follows: "Is . . . attending or enrolled in regular school? (Regular school includes elementary school and schooling which leads to a high school diploma or college, university, or professional degree.)" Each of these components, and the composite, was coded "1" for "yes" and "2" for "no." Missing data were imputed for both of the component variables. The imputation flag variables ACHBEG, ACHATT, and ASCHATT, available on the commercial file, indicated only whether the value was imputed, i.e., they did not indicate the type of imputation procedure. These vari-

[30]The 0.4 percent rate is calculated as follows: 183,000 home-schooled / (183,000 home-schooled + 42,819,000 educated in school).

ables were combined to create ALLATT, which was used to determine the supplement response rate.

CPS–Family Income

INCOME: This variable was derived from a single question asked of the household respondent in the basic labor force survey. It is asked regarding each family within a sampled household. Income includes money income from all sources including jobs, business, interest, rent, social security payments, and so forth. The income of nonrelatives living in the household is excluded, but the income of all family members 14 years old and over, including those temporarily living away, is included. Family income refers to receipts over a 12-month period. Income for families from which no income information was obtained (about 5 percent of families) was imputed using a sequential hot deck procedure. For the purposes of this analysis, the 14 categories of INCOME were collapsed into four categories: less than $20,000, $20,000–34,999, $35,000–49,999, and $50,000 or more. These categories were chosen to match as closely as possible the categories available from the NHES data and to divide children into fairly even quartiles.

CPS–Father's Education Attainment

DADED: The records of fathers who lived in the same household as the child were identified. F_EDAT is the father's value on EDAT (education attainment). To create DADED, F_EDAT was collapsed into six categories as follows:

Unknown	F_EDAT = 99 (Missing)
No high school diploma	F_EDAT < 39 (Less than high school graduate or equivalent)
High school diploma	F_EDAT = 39 (High school graduate or equivalent)
Some postsecondary education	F_EDAT > 39 and F_EDAT < 43 (More than high school graduate or equivalent but less than bachelor's degree)
Bachelor's degree	F_EDAT = 43 (Bachelor's degree)
Graduate degree	F_EDAT > 43 and F_EDAT < 99 (More than bachelor's degree)

CPS–Grade Level

NEWGRADE: Created by combining results from child and adult questions on grade level or grade-level equivalent. See figure 4 for questions. On the NCES file, the two grade-level variables, one for adults and one for children, had been combined into one variable (GRATTEND) as had the two grade-equivalent variables (WHTGRADE). These composites were replicated in the commercial file: CHGRD and GRDATT were combined to create GRADE, and CHREG and

GRDREG were combined to create GRADEREG. In both files, the grade and grade-equivalent variables were combined to create NEWGRADE. The response rates were calculated using the imputation flags ACHGRD and AGRDATT obtained from the commercial file.

CPS–Graduation Status

GRADUATE: Computed from the variables indicating the highest level of school completed or degree received (EDAT on the NCES file and GRDATN on the commercial file) and whether the subject completed high school by passing a graduate equivalency exam such as the GED (HSBYGED on the NCES file and HSGED on the commercial file). Subjects whose reported education attainment was a high school diploma or higher were classified as high school graduates. Those whose reported education attainment was 12th grade, no diploma, or lower were classified as graduates if they were reported to have completed high school by means of an equivalency test, such as the GED. Subjects whose reported education attainment was 12th grade and no diploma were classified as nongraduates if they had not completed high school via an equivalency test, and as "graduation status unknown" if they were missing data on the GED variables. Children younger than age 15 whose education attainment was less than a high school diploma were classified as legitimate skips on GRADUATE because the GED question was not asked in the child interview.

Missing values on EDAT and GRDATN were imputed. The variable AGRDATN identifies cases for which values were imputed and the type of imputation used.

CPS–Highest Parental Education Attainment

PARGRADE: This variable was created to determine the highest level of education among parents who resided with the child. It combined results from DADED (see "Father's education attainment") and MOMED (see "Mother's education attainment") as follows:

Graduate degree	DADED or MOMED = 6
Bachelor's degree	Of all other records, DADED or MOMED = 5
Some postsecondary education	Of all other records, DADED or MOMED = 4
High school diploma	Of all other records, DADED or MOMED = 3
No high school diploma	Of all other records, DADED or MOMED = 2
Unknown	Of all other records, DADED or MOMED = 1

CPS–Home-Schooling Variables

NEWHOME: This variable identifies children who were home schooled. Details regarding definitions for each data set follow. In summary, children were defined as educated at home if they were reported to be enrolled and being schooled at home by parents, people chosen or paid by parents, or school personnel, or if not enrolled and being schooled primarily at home. Children were defined as educated at school if they were enrolled and attending any other type of school. Children who were reported as both not enrolled and not home schooled were defined as educated neither at school nor at home. Cases that were asked neither the home-schooling question (for nonenrolled) or the school type question (for enrolled), that were not asked the enrollment question, or that were older than age 19 were classified as legitimate skip for the purposes of this analysis. Children who were missing on the enrollment, school type, or home-schooling (for nonenrolled children) variables were defined as missing.

NCES file definition:

Children educated at home	ENROLL = 2 and HOMESCH = 1, or ENROLL = 1 and SCHTYPE2 = 5 or 6
Children educated at school	ENROLL = 1 and SCHTYPE2 = 1, 2, 3, 4, or 7
Neither	ENROLL = 2 and HOMESCH = 2
Legitimate skip	ENROLL = 2 and HOMESCH = SYSMIS or ENROLL = 1 and SCHTYPE2 = SYSMIS or ENROLL = -1 or AGE > 19
Missing	ENROLL = 2 and HOMESCH = 9 or ENROLL = 1 and SCHTYPE2 = 9

Commercial file definition:

Children educated at home	ENROLL = 2 and (CHHOME = 1 or SCHHOME = 1) or ENROLL = 1 and (CHTYP = 3 or 4) or ENROLL = 1 and (ATTCUR = 5 or 6)
Children educated at school	ENROLL = 1 and (CHTYP = 1, 2, or 5) or ENROLL = 1 and (ATTCUR = 1, 2, 3, 4, or 7)
Neither	ENROLL = 2 and (CHHOME = 2 or SCHHOME = 2)
Legitimate skip	ENROLL = 2 and CHHOME = -1 and SCHHOME = -1 or ENROLL = 1 and CHTYP = -1 and ATTCUR = -1 or ENROLL = -1 or AGE > 19
Missing	ENROLL = 2 and (CHHOME < -1 or SCHHOME < -1) or ENROLL = 1 and (CHTYP < -1 or ATTCUR < -1)

NEWHOME4: This variable distinguishes between children who were schooled at home by school personnel and those who were schooled at home by parents or people chosen or paid by parents. Children who were classified as educated at school, educated neither at home nor at school, legitimate skip, or missing on NEWHOME were so classified on NEWHOME4. Among those schooled at home, SCHTYPE2 on the NCES file and CHTYP and ATTCUR on the commercial file were used to distinguish between those schooled at home by parents or people chosen by parents and those schooled at home by school personnel.

NEWHOME5: This variable distinguishes between children who were schooled at home because of illness or disability and those schooled at home for other reasons. Children who were classified as educated at school, educated neither at home nor at school, legitimate skip, or missing on NEWHOME were so classified on NEWHOME4. Among those schooled at home, WHYHOME (for enrolled subjects) and WHYHOME2 (for nonenrolled subjects) were used to distinguish among the reasons why children did not attend regular school on the NCES file. CHNOTATT, CHRHOME, RNOTATT, and RNOTREG were used for this purpose on the commercial file. Children for whom there were no data regarding the reason for not attending regular school were classified as missing on this variable.

CPS–Mother's Education Attainment

MOMED: The records of mothers who lived in the same household as the child were identified. M_EDAT is the mother's value on EDAT (education attainment). To create MOMED, M_EDAT was collapsed into six categories as follows:

Unknown	M_EDAT = 99 (Missing)
No high school diploma	M_EDAT < 39 (Less than high school graduate or equivalent)
High school diploma	M_EDAT = 39 (High school graduate or equivalent)
Some postsecondary education	M_EDAT > 39 and M_EDAT < 43 (More than high school graduate or equivalent but less than bachelor's degree)
Bachelor's degree	M_EDAT = 43 (Bachelor's degree)
Graduate degree	M_EDAT > 43 and M_EDAT < 99 (More than bachelor's degree)

CPS–Race/Ethnicity

RACETH: This variable was computed from RACE and HISP to match the race/ethnicity variable available in the NHES data sets. Value definitions were as follows:

- If RACE = White and HISP = non-Hispanic, RACETH = White, non-Hispanic

- If RACE = Black and HISP = non-Hispanic, RACETH = Black, non-Hispanic
- If HISP = Hispanic, RACETH = Hispanic
- If RACE = American Indian or Aleut Eskimo, Asian or Pacific Islander, or other and HISP = non-Hispanic, RACETH = All others.

Missing data for these variables were imputed. The imputation flag ARACE indicated whether the value for RACE had been imputed and, if so, the type of imputation. HISP was derived from ORIGIN, which was not included on the commercial file. Imputation flags were not included for HISP or ORIGIN on the commercial file.

CPS–Region

CENREG: This variable collapses the 50 states and the District of Columbia into four regions as follows:

Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
Midwest	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Virginia, West Virginia
West	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Texas, Utah, Washington, Wyoming

CPS–Respondent's Relationship to Child

WHORPT2: WHORPT was provided on the NCES file and was recoded as follows to include children under 15:

Under age 15, unknown	WHORPT = system missing
Otherwise unknown	WHORPT = 9
Self	WHORPT = 1
Parent	WHORPT = 2
Spouse	WHORPT = 3
Other relative	WHORPT = 4
Nonrelative	WHORPT = 5

CPS–Sex

SEX: Variable found on both files. Missing data were imputed, and cases that were missing data for this variable were identified using the commercial file variable ASEX, which indicated whether, and if so how, a value was imputed.

CPS–Telephone in Household

TELHHD: Household respondents were asked the following question: "Since households included in this survey are interviewed (again/again during the next 3 months), we attempt to conduct the follow-up interviews by telephone. Is there a telephone in this house/apartment?" Missing data were imputed. TELHHD and the variable ATELHHD, which identified cases where values had been imputed and the type of imputation used, were obtained from the commercial file.

NHES:96—Definitions of Variables Used in the Report

Most variables used in this study were provided on the public release file (U.S. Department of Education 1997) and are described in Collins et al. (1997). Exceptions have been noted.

NHES–Age

AGE95: Age reported in years as of December 31, 1995.

NHES–Enrollment

ENROLL: Responses to the enrollment question are provided in figure A1.

NHES–Family Composition

HHPARN1: This variable was constructed from HHMOM and HHDAD, derived variables that indicate which parents live in the household. "Mom" and "dad" in values refer to birth, adoptive, step, or foster parents.

NHES–Family Income

FAMINC: This variable was derived from HINCOME, which was respondents' answers to the following question: "In studies like this, households are sometimes grouped according to income. What was the total income of all persons in your household over the past year, including

salaries or other earnings, interest, retirement, and so on for all household members." Respondents first indicated whether the household income fell above or below $25,000, and then were asked to indicate into which of five or six income ranges the household income fell. HINCOME was recoded into FAMINC as follows:

$20,000 or less	HINCOME < 5
$20,001–35,000	4 < HINCOME < 8
$35,001–50,000	7 < HINCOME < 9
$50,001 or more	8 < HINCOME

NHES–Father's Education Attainment

DADED: DADED recodes DADGRADE—respondents' answers to "What is the highest grade or year of school that [you/(CHILD's) (father/stepfather/foster father) completed?"—into six categories as follows:

Unknown	DADGRADE = -1
No high school diploma	0 < DADGRADE < 4
High school diploma	DADGRADE = 4
Some postsecondary education	4 < DADGRADE < 9
Bachelor's degree	DADGRADE = 9 or 10
Graduate degree	DADGRADE > 10

NHES–Grade Level

NEWGRADE: NEWGRADE is a recode of ALLGRADE, a composite variable made according to NCES specifications to combine into one variable the data on grade (for children in graded schools or classrooms) and grade equivalent (for children in ungraded schools or classrooms or home-schooled children). NEWGRADE further combines kindergarten, transitional kindergarten, and prefirst grade into one category labeled "Kindergarten." All other values were taken as in ALLGRADE.

NHES–Highest Parental Education Attainment

PARGRADE: This variable was derived from MOMGRADE, MOMDIPL, DADGRADE, and DADDIPL by NCES.

NHES–Home-Schooling Variables

HOME2: This variable was derived from HOMESCHL and ENROLL. See figure 5 for question text. Values were derived as follows:

Home-schooled	HOMESCHL = 1
School-schooled	HOMESCHL = 2 and ENROLL = 1

HOME3: This variable distinguishes between children who were schooled at home because of illness or disability and those schooled at home for other reasons.

Home-schooled, not ill or disabled	HOMESCHL = 1 and HSILL ≠ 1 and HSDISABL ≠ 1
Home-schooled, ill or disabled	HOMESCHL = 1 and (HSILL = 1 or HSDISABL = 1)
School-schooled	HOMESCHL = 2 and ENROLL = 1

NHES–Mother's Education Attainment

MOMED: MOMED recodes MOMGRADE—respondents' answers to "What is the highest grade or year of school that [you/(CHILD's) (mother/stepmother/foster mother) completed?"—into six categories as follows:

Unknown	MOMGRADE = -1
No high school diploma	0 < MOMGRADE < 4
High school diploma	MOMGRADE = 4
Some postsecondary education	4 < MOMGRADE < 9
Bachelor's degree	MOMGRADE = 9 or 10
Graduate degree	MOMGRADE > 10

NHES–Number of Children in Family

SIBS: SIBS recodes NUMSIBS into four categories and adds the sampled child to represent the number of children in the family as follows:

One	NUMSIBS = 0
Two	NUMSIBS = 1
Three	NUMSIBS = 2
Four	NUMSIBS > 2

NHES–Race/Ethnicity

RACEETHN: Variable derived by NCES from CRACE and CHISPANI.

NHES–Region

CENREG: This variable collapses the 50 states and the District of Columbia into four regions as follows:

Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
Midwest	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Virginia, West Virginia
West	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Texas, Utah, Washington, Wyoming

NHES–Respondent Relationship to Child

RESP: Recode of ERESRELN, computed as follows:

Mother	ERESRELN = 1
Father	ERESRELN = 2
Grandparent	ERESRELN = 4
Other	ERESRELN = 3, 5, or 6

NHES–Sex

SEX: Used as on file, without modification.

Accuracy of Estimates

This report has focused on three sources of error that arise in survey research: errors of nonobservation, errors of observation, and data processing errors. Another form of nonobservation error, sampling error, was not discussed above and is discussed in this section. Sampling errors occur because observations are made on samples rather than on entire populations, and are perhaps best explained by contrasting it with a hypothetical survey of a population. A survey of a population universe (that is, a census or survey of all members of a population) results in an estimate of the population value that is compromised only by observation and data processing errors. If the measures and execution were perfect, the survey would perfectly describe the population in terms of the measured characteristics.

In contrast, estimates based on a sample will differ somewhat from those that have been obtained by a complete census of the relevant population using the same measures and procedures. The degree to which the sample estimate differs from the population value is highly dependent on the size of the sample: the larger the sample, the fewer population members that were excluded from it and the more accurate the sample estimate. To assess the accuracy of an estimate, researchers compute the standard error of the estimate, a measure of its variation due to sampling and, therefore, its precision. Measures of precision, standard errors, are essential to determining whether two estimates differ from each other: if two estimates appear to be different from each other but are very imprecise, they may not be different after all.

The CPS:Oct94 and NHES:96 data were collected from samples of households that were drawn using complex sampling designs involving stratification, clustering, unequal selection probabilities, and multistage sampling. These features of the sampling designs result in estimated statistics that are usually more variable (that is, have larger standard errors) than they would have been if they had been based on data from a simple random sample of the same size. Therefore, calculation of standard errors requires procedures that are markedly different from the ones used when the data are from a simple random sample. Although the CPS:Oct94 estimates (percentages and numbers of children) were computed using a standard statistical software package, SPSS for Windows, the standard errors for those estimates were computed using the following generalized variance estimation formulae:

Formula for percentages:

$$\text{s.e.} = \sqrt{(b/N)(p)(100-p)}$$

where p = the percentage ($0 < p < 100$),
N = the population on which the percentage is based, and
b = the parameter associated with the characteristic;
b is equal to 2,944 for the total or white population; 2,978 for the black population; and 3,367 for the Hispanic population ages 13 and younger.[31]

Formula for numbers of persons:

$$\text{s.e.} = \sqrt{(bx)(1-x/T)}$$

[31]These parameters are larger than those provided for the total population and therefore result in slightly larger estimates of the standard errors. These parameters were chosen because children 13 years old and younger make up the majority of the sample. Parameters are provided in Attachment 18 of the CPS data file documentation (U.S. Department of Commerce 1994).

where x = the number of persons in a subcategory (i.e., 6- to 17-year-olds in grades K through 12 who were home schooled),
T = population in the larger category (i.e., 6- to 17-year-olds in grades K through 12), and
b = as above.

Estimates and standard errors for the estimates from the NHES included in this report were produced using REPTAB, a SAS procedure that computes variance with the bootstrap variance estimator method using replicate weights provided on the data file. Estimated standard errors are presented in appendix B.

Statistical Procedures

Statistical tests use estimates and standard errors to take the precision of the estimates into account when determining whether apparent differences in the sample are likely to represent population differences. The primary statistical procedure used in this report is based on the Student's *t* statistic, which is the ratio of the difference between the estimates to the precision of the estimates. A student's *t* value is computed with the following formula:

$$t = \frac{E_1 - E_2}{\sqrt{se_1^2 + se_2^2}}$$

where E_1 and E_2 are the estimates to be compared and se_1 and se_2 are their corresponding standard errors.

A difference between two estimates is considered statistically significant, that is, sufficiently likely to be a true population difference rather than an artifact of sampling error, when the Student's *t* value for the difference is greater than a critical value. The critical value is determined by selecting an *alpha* level, which is the probability of inferring that a difference exists when, in fact, it does not. An alpha level of 0.05 was used in the analyses reported here.

In order to interpret these statistics appropriately, three points must be kept in mind. First, comparisons resulting in large *t* statistics may appear to merit special attention. However, this is not always the case because the size of the *t* statistic is related not only to the observed differences in the estimates being compared but also to the standard error of the estimates. A small difference between two groups with a much smaller standard error could result in a large *t* statistic, but this small difference is not necessarily noteworthy.

Second, the formula for the Student's *t* statistic given above is valid only for independent estimates. When the estimates are not independent (for example, when comparing a total per-

centage with that for a subgroup included in the total), a covariance term must be added to the denominator of the formula. The actual covariance terms are not known. Therefore, to be conservative, it is assumed that the estimates are perfectly negatively correlated. Consequently, $2(se_1 * se_2)$ is added to the denominator of the *t*-test formula, as shown below.

$$t = \frac{E_1 - E_2}{\sqrt{(se_1)^2 + (se_2)^2 + 2(r)se_1 se_2}}$$

Third, when multiple comparisons are made within categories of a variable (for example, when comparing the proportion of children who are home schooled among various racial/ethnic groups), it becomes increasingly likely that an indication of a population difference is erroneous. Even when there is no difference in the population, at an *alpha* level of 0.05 there is still a 5 percent chance of concluding that an observed difference, or comparison, between estimates is large enough to be statistically significant. As the number of comparisons increases, the risk of making such an error in inference also increases.

Therefore, to guard against errors of inference derived from multiple comparisons, the Bonferroni procedure was used whenever multiple comparisons were made in this report. Generally, this method adjusts the *alpha* level for the total number of comparisons made within a particular variable to reduce the overall probability of determining that a difference is likely to be real when it is not. For each variable, there are (K*(K–l)/2) possible comparisons (or nonredundant pairwise combinations), where K is the number of categories in the variable. For example, in this report children's race/ethnicity was defined in terms of four categories (white, non-Hispanic; black, non-Hispanic; Hispanic; and all others. Therefore, K=4 and there are 6, or (4*3)/2, possible comparisons among the categories. The Bonferroni procedure divides the *alpha* level for a single *t* test (in this case, 0.05) by the number of possible pairwise comparisons in order to produce a smaller *alpha*, and therefore a higher critical value, for each comparison.

Finally, when home schooling was studied in terms of ordered independent variables, the Student's *t* test was sometimes applied to a measure of linear trend among proportions rather than to the differences between discrete categories. This modification allows researchers to examine whether, for example, the proportion of children who were home schooled significantly increased (or decreased) with their parents' education attainment—in other words, whether there was a linear relationship between the two variables. Based on a simple regression with, in this case, parents' education attainment as the independent variable and the proportion of children who were home schooled as the dependent variable, the test involves computing the regression coefficient (*b*) and its corresponding standard error (se). As described above, the ratio of these two (*b*/se) is the test statistic *t*. If *t* is greater than 1.96, the critical value for one comparison at

the 0.05 alpha level, the hypothesis that there is a linear relationship between parents' education attainment and home schooling is not rejected.[32]

[32]For more information about this modification of the Student's *t* test, see Snedecor and Cochran (1967), pp. 246–247.

Appendix B—Standard Error Tables

Table B1—Standard errors for table 1: CPS:Oct94 and NHES:96 estimates of the number and percentage of 6- to 17-year-olds who were home schooled in grades K–12

	CPS (October 1994)		NHES:96 (Spring 1996)	
	Number educated at home	Percent educated at home	Number educated at home	Percent educated at home
Total	31,700	0.07	118,200	0.14
Sex				
Male	21,800	0.10	29,300	0.12
Female	23,200	0.11	49,800	0.22
Race/ethnicity				
White, non-Hispanic	30,400	0.10	59,400	0.19
Black, non-Hispanic	5,500[1]	0.08	7,600[1]	0.11
Hispanic	6,800[1]	0.13	14,700[1]	0.21
All others	2,900[1]	0.18	8,800[1]	0.45
Age				
14 and younger	86,000	0.09	129,600	0.14
6	11,000[1]	0.30	13,900[1]	0.40
7	10,800[1]	0.29	13,000[1]	0.33
8	10,200[1]	0.28	19,200[1]	0.49
9	11,400[1]	0.30	13,300[1]	0.41
10	9,200[1]	0.25	12,300[1]	0.31
11	10,900[1]	0.30	14,300[1]	0.37
12	6,900[1]	0.19	13,600[1]	0.36
13	7,600[1]	0.21	14,100[1]	0.41
14	8,000[1]	0.22	15,900[1]	0.40
15 to 17	20,700	0.13	53,100	0.28
15	8,700[1]	0.25	17,200[1]	0.41
16	3,600[1]	0.11	16,400[1]	0.44
17	8,400[1]	0.27	19,500[1]	0.61
Family income[2]				
Less than $20,000	16,000	0.14	38,200	0.30
$20,000–34,999	15,300	0.16	27,800	0.27
$35,000–49,999	16,200	0.19	24,300	0.28
$50,000 or more	16,100	0.12	28,300	0.20
Father's education attainment				
Unknown	9,700[1]	0.09	28,100	0.17
No high school diploma	7,100[1]	0.15	18,500[1]	0.44
High school diploma	14,800	0.14	14,600	0.17
Some postsecondary education	17,300	0.20	37,900	0.49
Bachelor's degree	15,700	0.30	18,200	0.34
Graduate degree	10,200[1]	0.31	19,000[1]	0.50

Table B1—Standard errors for table 1: CPS:Oct94 and NHES:96 estimates of the number and percentage of 6- to 17-year-olds who were home schooled in grades K–12 and difference between estimates—Continued

	CPS (October 1994)		NHES:96 (Spring 1996)	
	Number educated at home	Percent educated at home	Number educated at home	Percent educated at home
Mother's education attainment				
Unknown	5,600[1]	0.29	14,900[1]	0.41
No high school diploma	8,900[1]	0.13	22,700[1]	0.34
High school diploma	17,100	0.12	23,600	0.18
Some postsecondary education	19,800	0.16	42,200	0.34
Bachelor's degree	13,700	0.25	20,100	0.31
Graduate degree	5,400[1]	0.26	13,200[1]	0.44
Highest parental education attainment				
Unknown	3,900[1]	—	(†)	(†)
No high school diploma	8,100[1]	0.14	16,900[1]	0.39
High school diploma	13,100	0.10	25,300	0.18
Some postsecondary education	19,500	0.15	42,400	0.30
Bachelor's degree	16,000	0.23	21,500	0.30
Graduate degree	11,000[1]	0.25	19,900	0.30
Number of parents				
Mom and dad	(†)	(†)	52,600	0.17
Mom only	(†)	(†)	24,000[1]	0.22
Dad only	(†)	(†)	6,500[1]	0.46
Nonparent guardian	(†)	(†)	8,500[1]	0.55
Number of children in family				
One	(†)	(†)	18,500	0.20
Two	(†)	(†)	22,100	0.12
Three	(†)	(†)	34,200	0.26
Four or more	(†)	(†)	43,000	0.59
Region				
Northeast	14,100	0.18	22,600[1]	0.25
South	19,100	0.12	36,000	0.22
Midwest	12,700	0.12	27,800	0.21
West	16,900	0.18	34,700	0.40

—Denominator sample size too small for a reliable estimate.

†Not available.

[1]Sample size too small for a reliable estimate. Interpret with caution.

[2]The income variable for the NHES was household income rather than family income, as in CPS. In addition, the income categories for the NHES were defined slightly differently, as follows: $20,000 or less; $20,001–35,000; $35,001–50,000; and $50,001 or more.

NOTE: Total numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement/Civic Involvement Component (NHES:96, PFI/CI).

Table B2—Standard errors for tables 2 and 3: CPS:Oct94 estimates of the number of 6- to 17-year-olds enrolled in school or home schooled in grades K–12 and percent who were home schooled, by age, race/ethnicity, and sex

	Total number of children	Estimated percent home schooled
Total	0*	0.07
6- to 14-year-olds		
Black		
Males	83,700	0.04
Females	83,100	0.03
Nonblack		
Males	155,000	0.09
Females	154,000	0.10
15-year-olds		
Black		
Males	27,000	0.00
Females	27,400	0.00
Nonblack		
Males	50,200	0.24
Females	49,800	0.30
16- to 17-year-olds		
Black		
Males	37,200	0.00
Females	38,000	0.16
Nonblack		
Males	68,400	0.16
Females	67,600	0.13

*Total number is weighted to known population values and therefore has no sampling variance.

NOTE: Numbers are rounded to the nearest hundred. Note that although rates of home schooling were calculated for 6- to 17-year-old children, rates of undercoverage were not published for children in these age groups. The estimates of bias reported here depend on the assumption that 6- to 14-year-olds were undercovered at the same rate as children 0 to 14 years old and that 16- to 17-year-olds were undercovered at the same rate as 16- to 19-year-olds.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table B3—Standard errors for table 4: CPS:Oct94 estimates of the number of 6- to 17-year-olds who were enrolled in school or home schooled in grades K–12 and percentage who were home schooled, by whether telephone in household

	CPS (October 1994)	
	Total number	Percent educated at home
Total	0*	0.07
Telephone in household		
Telephone	88,500	0.08
Nontelephone	88,500	0.22

*Total number is weighted to known population values and therefore has no sampling variance.

NOTE: Numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table B4—Standard errors for table 7: NHES:96 estimates of the percentage distribution of 6- to 17-year-olds according to relationship to interview respondent, by home-schooling status

	Mother	Father	Grandparent	Other
Total	0.50	0.43	0.21	0.18
Home-schooling status				
Educated at home	3.35	3.04	1.21	1.20
Educated at school	0.50	0.43	0.21	0.18

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement in Education and Civic Involvement Component (NHES:96, PFI/CI).

Table B5—Standard errors for table 8: CPS:Oct94 estimates of the percentage distribution of 15- to 17-year-olds according to relationship to interview respondent, by home-schooling status

	Unknown	Self	Parent	Spouse	Other relative	Nonrelative
Total	0.05	0.45	0.64	0.19	0.45	0.19
Home-schooling status						
Educated at home	0.00	8.83	10.32	0.00	7.56	0.00
Educated at school	0.05	0.45	0.64	0.20	0.45	0.19

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table B6—Standard errors for table A1: CPS:Oct94 estimates of the number and percentage distribution of 15- to 17-year-olds according to whether emancipated minor and number and percentage distribution of emancipated minors according to home-schooling status

	15- to 17-year-olds[1]	
	Number	Percent
Total	0[2]	0.00
Whether emancipated minor		
Not emancipated minor	31,000	0.30
Emancipated minor[3]	5,000	0.00[4]
Potential emancipated minor[5]	30,600	0.29
	Emancipated minors and potential emancipated minors	
Total	31,000	0.00
Home-schooling status		
Educated at home	5,000	1.27
Educated in school	30,600	1.25

[1]Defined as children 15 to 17 years old who had not graduated from high school or earned a GED, and who were enrolled in school or home-schooled in grades K–12. Excludes children who did not meet these criteria or whose status regarding these criteria could not be determined due to missing data.

[2]Total number is weighted to known population values and therefore has no sampling variance.

[3]Defined as children living alone or living with only a spouse.

[4]Less than 0.05 percent.

[5]Defined as children who lived with more than one other person but not with a parent, foster parent, or grandparent; or who lived with only one other person who was not a parent, or foster parent. An unknown proportion of these children are emancipated minors.

NOTE: Numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table B7—Standard errors for table A2: CPS:Oct94 and NHES:96 estimates of the number and percentage of home-schooled 6- to 17-year-olds, by persons providing education and whether schooled at home because of illness or disability

	CPS (October 1994)			NHES (Spring 1996)	
	Total home-schooled children	Schooled by parents or persons paid by parents	Not because of illness or disability	Total home-schooled children	Not because of temporary illness or disability
Number of home-schooled 6- to 17-year-olds in grades K–12	31,700	31,200	30,700	61,600	59,700
Percent of all home-schooled 6- to 17-year-olds in grades K–12	0.00	1.67	2.26	0.00	2.45

NOTE: Numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94); U.S. Department of Education, National Center for Education Statistics, National Household Education Survey of 1996, Parent and Family Involvement in Education and Civic Involvement Component (NHES:96, PFI/CI).

Table B8—Standard errors for table A3: Estimating impact of missing data on CPS:Oct 94 estimates

	Total	Not enrolled	Enrolled
Total 6- to 17-year-olds	0*	46,400	46,400
Whether included in analysis			
Excluded because high school graduate, GED recipient, or in pre-kindergarten	30,900	12,700	28,200
Excluded because missing data	80,900	39,200	71,700
Not home schooled	36,000	36,000	0
Home-schooling status unknown	73,300	15,800	71,700
Included in analysis	86,100	21,800	88,500
Home schooled	31,700	21,800	23,100
Educated at school	91,100	0	91,100
Percent home schooled	0.07	5.99	0.05

*Total number is weighted to known population values and therefore has no sampling variance.

NOTE: Numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

Table B9—Standard errors for table A4: CPS:Oct94 estimates of the number of nonenrolled children and percentage distribution according to age, by eligibility for home-schooling analysis

		Percentage distribution by age	
	Number	Percent 6- to 14-year-olds	Percent 15- to 17-year-olds
		Eligible for home-schooling analysis	
Total	21,800	4.56	4.56
		Not eligible for home-schooling analysis because of missing data	
Home-schooling status unknown	14,500	8.28	8.28
Not home schooled	14,500	3.58	3.58

NOTE: Children were eligible for the home-schooling analyses if they were 6 to 17 years old, enrolled in school or home schooled, and in grades 1 through 12 or their equivalents. Children who did not meet these criteria, or who were missing the data needed to determine whether they met the criteria, were deemed ineligible for the analyses. Numbers are rounded to the nearest hundred.

SOURCE: U.S. Department of Commerce, Bureau of the Census, Current Population Survey, October 1994 (CPS:Oct94).

United States
Department of Education
ED Pubs
8242–B Sandy Court
Jessup, MD 20794–1398

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Standard Mail (B)

NCES 2000–311

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

NOTICE

Reproduction Basis

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").