



Measuring education in the Current Population Survey

Robert Kominski and Paul M. Siegel

The 1940 Decennial Census of Population included a question about educational attainment that was a marked departure from the way in which an individual's level of education had been measured in previous censuses. Prior to 1940, the educational level of respondents was determined by asking whether they could read and write. In the 1940 census, however, a question about the highest grade of schooling completed was asked of all persons. This item (later modified to determine "highest grade or year of school" completed) has for nearly 50 years been the fundamental means of assessing educational attainment, both in the census and in a variety of surveys, including the Current Population Survey (CPS).

That the years of schooling completed measure served for so long is a strong testament to its utility and face validity. Nevertheless, in 1990, the U.S. Bureau of the Census implemented a major change in the form of the question and its possible response categories. The decision to do so was difficult, but one that became necessary because of changes in the way persons acquire education and view it as an economic and social resource. Beginning in January 1992, the CPS adopted a question similar to that used in the 1990 census. This report briefly discusses the form of the new question, examines some of the factors motivating the change, and presents data

from a national CPS test conducted in February 1990 that shows the empirical relationship between the old attainment question and the new one.

The old and new items

The educational attainment question used in the CPS has been virtually unchanged since the 1940's. This question appears on the control card of the interview, and consists of two parts (exhibit 1, panel a). The first part asks, "What is the highest grade or year of regular school . . . has ever attended?" This is followed with the question, "Did . . . complete the grade?" Response codes range from 00 to 26, where the series 21 through 26 is used to represent years of college. Persons having attended more than 6 years of college are coded as '26'. The two-part question allows the respondent to indicate a grade that was attended, but not completed. This would include many persons who are currently enrolled in that grade.

The new education-related survey item introduced in January 1992 involves several changes to the older item (exhibit 1, panel b). A single question is now asked: "What is the highest level of school . . . has completed or the highest degree . . . has received?" Response categories range from 31 to 46, an intentional change to prevent field staff from attempting to code the new question using the old years-of-schooling answers. In the new item, response categories for lower levels of schooling have been collapsed into several summary categories. A new category, "12th grade, No Diploma," has been added.

The major change in the item occurs in the categories for high school completion and beyond. Beginning with the response, "High School Graduate—high school diploma or the equivalent (for example, GED)," the categories identify specific degree completion levels, rather

than years of schooling. The focus of the item remains "regular" schooling—that is, schooling that is a part of the collegiate system. No attempt has been made to incorporate postsecondary educational attainment from institutions other than the regular college system. Five different levels of degree attainment are identified—Associate, Bachelors, Masters, Professional, and Doctoral degrees. Associate degrees are further distinguished between those awarded in academic programs and those given for completing an occupational or vocational program. A residual category of "some college but no degree" is used to identify persons who have not completed a degree program.

Reasons for the change

Why was the educational attainment item changed? The answers are several, most suggested by a series of tests of results using the old item and possible new alternatives, conducted during the 1980's by the Education and Social Stratification Branch of the U.S. Census Bureau. These analyses are documented in a series of papers and memoranda,¹ and are available by contacting the branch directly (301-763-1154). The main findings of this research are summarized here.

Under the old procedures, years of schooling completed tended to be misclassified into degree status.

Users of CPS data have become accustomed to equating persons with 4 years of college or more with graduates holding a Bachelors degree. Research based on four independent data sources, however, suggests that this is not always the case. Estimates of the proportion of persons with at least 4 years of college who did not hold at least a Bachelors degree varied, from 13 percent in the 1980 census Content Reinterview Survey and the 1986 National Content Test, to 7 percent in data taken from the Survey of Income

Robert Kominski is Chief of the Education and Social Stratification Branch and Paul M. Siegel is Chief of the Education Analysis Staff in the Population Division of the U.S. Bureau of the Census.

and Program Participation in 1984, and to 6 percent in a February 1990 cps test. Each of these estimates is based on somewhat different methods of survey administration, and this may account for at least some of the variation. Nevertheless, the evidence indicates that a non-trivial amount of misclassification occurs when one infers attainment of a degree from years of schooling.

Attempts to estimate demographically the population stock of Bachelors degree-holders since 1940 (that is, by accounting for births, deaths, and migration) yield an interesting pattern. Through the 1970 census, efforts to balance the number of Bachelors degrees awarded through time (as recorded in administrative records) with the number of Bachelors degree-holders as inferred from the census showed a fair amount of correspondence—at most a difference of several hundred thousand. By the 1980 census, the difference had jumped to well over a million “excess” Bachelors degrees (as measured by years of schooling in the census), and extending the method to the mid-1980’s using data from the March 1985 cps showed the excess remaining over a million. This discrepancy between administrative records and the “years-of-school-completed” data corresponds well to the documented increase in the length of time students are taking to obtain a Bachelors degree: if respondents report the number of years they have spent in

the postsecondary system rather than their academic standing, this would exacerbate the problem. Apart from this, however, the number of degrees inferred could be inflated simply because more adults are undertaking additional years of schooling without any intention of acquiring a degree. In this respect also, the years-of-schooling measure becomes a misleading indicator of the educational credentials individuals hold.

The old procedures made it impossible to identify specific degrees.

If the ability to infer the number of “traditional” Bachelors degrees deteriorated over time, the ability to measure other kinds of educational credentials was poor from the start. Most notable was the inability to identify holders of Associate (generally 2-year) degrees from years-of-schooling data. During the 1950’s and 1960’s, the early era of growth in the 2-year college system, Associate degrees were often seen as stepping stones to a 4-year credential. However, during the 1970’s and 1980’s, the Associate degree became recognized as a credential in its own right. The development of many occupational and job-specific curriculums in 2-year colleges offered students who did not wish or could not afford a 4-year degree the opportunity to obtain a postsecondary credential useful in obtaining employment. The structure of the old years-of-school-

ing question made it impossible to identify these persons.

Similarly, the years-of-schooling question could not provide direct evidence about the growing pool of persons holding degrees beyond the Bachelors level. The February 1990 cps test provides estimates showing that more than 10 million U.S. adults hold advanced degrees of some sort, based on answers to the new, degree-based question. These individuals represent a critical national resource, the size of which formerly could only be roughly approximated in the cps by the number of persons with 5 or more years of college completed. Under the old procedures, an estimate from the February 1990 data would incorrectly attribute such advanced degrees to nearly 4 million more persons.

Use of the former item led to uncertainty in the classification of high school graduates.

Many people, although they did not complete 12 years of schooling, hold a high school diploma because of programs like the GED, night school, and other equivalency mechanisms. Individuals who hold a high school equivalent were supposed to be coded as high school graduates in the decennial census and other Census Bureau surveys. Analysis of selected data sources indicates that this was not always the case. Be it due to respondent or interviewer error, there are numerous instances (about

Exhibit 1. Educational items as contained in the "old" and the "new" CPS

(a) "Old" CPS		(b) "New" CPS									
<p>What is the highest grade or year of regular school ... has ever attended?</p> <p>00 - Never attended or kindergarten 01-08 - Elementary 09-12 - High school 21-26 - College (Academic)</p> <p style="text-align: center;">23a</p> <table border="1"> <tr> <td>First code</td> <td>Update code</td> </tr> <tr> <td> </td> <td> </td> </tr> </table>		First code	Update code			<p>What is the highest level of school ... has completed or the highest degree ... has received?</p> <p>Update in 5th month AND in Feb., July, and Oct.</p> <p style="text-align: center;">23</p> <table border="1"> <tr> <td>First code</td> <td>Update code</td> </tr> <tr> <td> </td> <td> </td> </tr> </table>		First code	Update code		
First code	Update code										
First code	Update code										
		EDUCATION CODES FOR 23									
		<p>31 Less than 1st grade</p> <p>32 1st, 2nd, 3rd, or 4th grade</p> <p>33 5th or 6th grade</p> <p>34 7th or 8th grade</p> <p>35 9th grade</p> <p>36 10th grade</p> <p>37 11th grade</p> <p>38 12th grade NO DIPLOMA</p> <p>39 HIGH SCHOOL GRADUATE -- high school DIPLOMA, or the equivalent (For example: GED)</p>	<p>40 Some college but no degree</p> <p>41 Associate degree in college -- Occupational/vocational program</p> <p>42 Associate degree in college -- Academic program</p> <p>43 Bachelor's degree (For example: BA, AB, BS)</p> <p>44 Master's degree (For example: MA, MS, MEng, MEd, MSW, MBA)</p> <p>45 Professional School Degree (For example: MD, DDS, DVM, LLB, JD)</p> <p>46 Doctorate degree (For example: PhD, EdD)</p>								

1 million persons in the February 1990 cps test) in which number of school years completed have been recorded instead of the fact that a high school equivalent was earned.

An associated problem is the growing number of persons who have completed 12 years of school, but who do not have a high school diploma. This problem is especially common in areas that require a final graduation test or certification. Individuals who fail this test but have completed their 12 years are free to leave school, but are not awarded a high school diploma. Instead they often are given a "certificate of completion" or other document to indicate that they have "served their time." The February 1990 test showed that nearly 4 million people fell into this category.

The old questionnaire design did not meet agency and programmatic data needs.

As part of the content development for the 1990 Decennial Census, interagency working groups were formed for each substantive topic area. These groups met to discuss specific agency data needs, reflecting programmatic and legislative data requirements, as well as a full array of research and analytic needs. The Education working group identified the need for degree-specific data as the highest priority item, and suggested potential item response categories. This process also revealed that detailed attainment information was not legislatively required (or generally needed) below the fifth grade. The group suggested that the collapsing of lower levels of schooling into broader categories would not be a serious loss. Faced with the addition of new categories beyond the high school level, and a serious space constraint in the decennial census survey instrument, the decision was made to use collapsed categories of first through fourth and fifth through eighth grades; discussions with the Bureau of Labor Statistics in the development of the attainment item for the cps led to establishment of separate categories for fifth through sixth grade and seventh through eighth grade.

Potential problems with change

The switch to the new attainment item is somewhat controversial. One major ob-

jection has been raised regarding the loss in continuity with time series that date back as far as 50 years. On one level this argument must be accepted; the new item, after all, is different in terms of both structure and wording and the range of possible response categories. One must remember, however, that the meaning of the old question itself did not remain constant over the past 50 years—that is, in fact, a primary reason why it *needed* to be changed. "Years of schooling completed" does not connote the same thing now that it did in 1940. In that year, one could say with a high degree of certainty that a person with 4 years of college completed had a Bachelors degree, but that same assertion could not be made in 1990. To use the old item today and attribute the 1940 meaning to the results would be misleading, and the problem would become more serious with each passing year.

The new item, on the other hand, does allow comparable construction of many of the educational attainment concepts and benchmarks that are frequently used in research, analysis, and devising legislation. Persons with less than a high school degree, for example, can easily be identified, including the 12th grade "completers" who are not graduates. Persons with college yielding less (or more) than a Bachelors degree also can be determined—as, obviously, can those who have attained a Bachelors as their highest degree. In this respect, the new item provides better year-to-year comparability, because it relies on known degree levels rather than inference of degree status from years of schooling.

A second source of concern is cited by some researchers who feel that one of their independent variables of choice for regression models has been lost. While it is true that the apparent continuous form of the years-of-schooling item made the educational data a handy (and empirically useful) addition to many regression models of subjects such as earnings and employment, most analysts would recognize that the estimates yielded by the old item do not reflect the reality of credentialing and the rewards based thereon. In fact, a perusal of the literature shows that in most empirical analyses years of schooling is summarized into a small set of convenient educational benchmarks, generally represented by

binary (dummy) variables. (For example, a code of '0' could indicate that an individual did not have a college degree, while a code of '1' would mean that a degree had been earned.) Analyses such as these are commonplace, particularly in the context of the growing use of categorical data methods during the past decade.

A related concern for some analysts is the elimination of summary measures, such as median or average years of schooling completed. While these measures will no longer be derivable in the same metric, this may not be a major analytical loss. During the entire decade of the 1980's, median years of schooling for persons aged 25 and older changed by 0.2—from 12.5 to 12.7 years—pointing up the fact that the median is not a very good summary measure of the pace of change in educational attainment. The tabulation below shows four different summary measures, including the median. Note that the three proportionate measures all tell a much different story than does the median about how levels of educational attainment changed over the 16-year period 1975–91:

<i>Measure</i>	<i>1975</i>	<i>1991</i>
Median years of schooling	12.3	12.7
Percent high school graduates	62.5	78.4
Percent with 1 year of college or more	26.3	39.8
Percent with 4 years of college or more	13.9	21.4

While the change in the attainment question will cause the loss of the time series of median years of schooling, one would be hard pressed to explain the utility of keeping a measure that almost totally misses the dynamic in the Nation's changing educational structure.

February 1990 cps results

As indicated earlier, a test of the new educational attainment item was conducted in the February 1990 cps. The conventional years-of-schooling question, which was still on the control card—the portion of the survey instrument that remained with the household throughout its participation in the survey—was updated at the beginning of the interview for all persons. After completion of the regular cps-1 labor

not to report this experience under the old procedure, and reported their highest credential (the high school diploma) instead.

Data in the portion of the table showing attainment at the college level or above tend to lend validity to the observations that have been made about the ability of the new item to better classify individuals with respect to the specific degrees they hold. For example, the new item identifies 8.5 million Associate degree-holders, who formerly were distributed primarily among the categories of 1 to 6 years of college. The table also shows that just 81 percent of persons who hold a Bachelors as their highest degree have completed exactly 4 years of college. Overall, the new question counts more than 75 million persons with some college experience, as opposed to just under 70 million counted as such using the old item. In general, while responses to the old and new items demonstrate a high level of consistency, it is the inconsistent responses, particularly for high school completion and beyond, that give the new item a decided advantage in terms of accuracy and usefulness.

CHANGING PATTERNS of enrollment and new perceptions of education as human capital required a serious reexamination of the conventional CPS educational attainment question. The growing importance of postsecondary education to the economy and the increasing credentialing of individuals dictate that more direct measures be used. The new educational attainment item is a major departure from past procedures, but it provides more relevant and useful data for current and future analyses. □

Footnote

¹ Robert Kominski, "Evaluation of the 1980 Decennial Census Education Questions," Preliminary Evaluation Results Memorandum, 104 (Bureau of the Census, 1985); Paul M. Siegel and Robert Kominski, "The Quality of Census Data on Educational Attainment," Paper presented before the Census Advisory Committee on Population Statistics (Bureau of the Census, 1986); Robert Kominski and Paul M. Siegel, "Measuring Educational Attainment in the 1990 Census," Paper presented at the 1987 Annual Meeting of the American Sociological Association, Chicago, IL; and Robert Kominski, "Education and Earnings: Empirical Findings from Alternative Operational-

izations," *Proceedings of the Social Statistics Section* (Washington, American Statistical Association, 1988), pp. 82-87.

Undercoverage of Hispanics in household surveys

Ruth B. McKay

The fact that individuals in some racial and ethnic groups have a high rate of being missed in surveys has become an important subject of discussion and debate by analysts, policymakers, social scientists, and other concerned persons. The focus of this report is on factors contributing to the undercount of the Hispanic population in the Current Population Survey (CPS).¹ The CPS measures labor force activity in a monthly sample of approximately 60,000 households selected to represent the U.S. population.²

Undercounts can occur when the survey enumerator misses an entire household (whole household undercoverage) or when the person reporting for the household does not list all of the occupants of the household (within household undercoverage). The major causes of whole household undercoverage have been identified as missed addresses from the decennial census and deliberately omitted households, for example, a dwelling subdivided into several unreported housing units.³ Within household undercoverage is believed to stem from differences in definitions of household membership used by the survey and respondents in ethnic communities.⁴

Anthropological research methods were used to investigate factors associated with undercount of Hispanics in the CPS and the decennial census. The research consisted of a small-scale field study of a Salvadoran community and an analysis of data collected from eight,

Ruth B. McKay is a behavioral scientist in the Statistical Methods Division, Bureau of Labor Statistics.

predominantly Hispanic, sites studied by the Bureau of the Census in its 1990 Alternative Enumeration (discussed later).

The research methods

The anthropological research methods most often used in undercoverage research are from two areas: *ethnography* and *cognitive anthropology*. *Ethnography* describes the culture, or way of life, of a particular society, from the point of view of members of that society. Ethnographic methods include *participant observation*, which requires a period of residence within the community so that the ethnographer can become immersed in the everyday realities of that society, and the *ethnographic interview*, a particular kind of unstructured interview that attempts to elicit the cultural patterning and context of the subject or topic under discussion. *Cognitive anthropologists* study the conceptual categories which members of a culture use to organize their experience of the world. A good illustration of the application of cognitive anthropological methods to survey research is Elizabeth Gerber's investigation of the concepts of "living with" and "staying at" used by inner city blacks in the District of Columbia. She reported that a man who had not been in his mother's house for 5 months would, under certain circumstances, be classified as "living with" his mother, and that his relationship to a household in which he had eaten and slept every day for the past 5 months would be termed "staying at" that household. The basis for such classification involved, among other factors, intentionality about relationships, and permanent versus temporary affiliation with household units, that were based in the everyday, commonplace experience of domestic and economic relationships within inner city black communities.⁵

The field study

In 1991, an ethnographic field study was conducted in the Washington, DC, area to investigate Hispanic respondents' definitions of household membership and residency. The largest Hispanic group within the area is Salvadoran, and this was the group chosen for study.