Using Multiple Data Sources to Identify Types and Sources of Coverage Error on an American Indian Reservation

Laurie Schwede
U.S. Census Bureau, 4600 Silver Hill Road, Washington DC 20233

Abstract

This paper reports the results of a small-scale study to improve understanding of household coverage issues on the Cheyenne River Indian Reservation in the 2006 Census Test. Multiple matched data sources from 31 reservation households are used to identify and explore types and sources of coverage error to suggest improvements to enumeration methods, not to estimate coverage rates. Matched data sources include 1) the household roster from the actual Update/Enumerate Operation; 2) the household roster from the Census Coverage Measurement (CCM) Operation three months later; 3) CCM interview audiotapes; and, in 14 cases, 4) immediate qualitative debriefings to resolve coverage anomalies. The paper identifies types of coverage errors (omissions, duplications), sources of error (e.g., mobility, living situations, non-city style addresses, cultural differences, question wording, recall issues, and interviewer/respondent familiarity), and types of households and persons possibly miscounted. Benefits of using multiple matched sources are discussed.

Key Words: 2006 Census Test, Coverage Measurement, Coverage Error, Methodology, Cheyenne River Sioux Indian Reservation

1. Introduction

Improving enumeration of American Indian reservations and other hard-to-count areas is a long-standing goal of and challenge to the Census Bureau. In recent censuses, the Census Bureau has conducted independent sample surveys to use dual systems estimation to estimate coverage of demographic groups in the census, including hard-to-count subpopulations such as American Indian and Alaska Native (AIAN) reservations. The Post-Enumeration Survey (PES) estimated a 1990 Census net undercount of AIAN reservations of 12.22 percent (S.E., 5.29 percent): the highest for any group (Fenstermaker and Haines 2002: Table 1). The Accuracy and Coverage Evaluation (ACE) in Census 2000 estimated a slight net reservation overcount of 0.88 percent (S.E. 1.53 percent), not statistically different from 0, and the second lowest race/ethnic match rate of 86.7% (Fenstermaker and Haines 2002, Tables 1, 9; National Research Council 2004, Table 6.7). Also, these reservations had the highest race/ethnic duplicate enumeration rate: 2.74 percent (Mule 2002 F-1).

Ethnographic studies of American Indians done in conjunction with the 1988 Dress Rehearsal, 1990 Census (e.g., Ackerman 1988, de la Puente 1993, Lobo 1990, Jojola 1992) and Census 2000 (Tongue 2006; Schwede, Blumberg, and Chan 2006; Schwede 2003 has lists of 1990 and 2000 Census ethnographic studies) and on counting Indians in general (e.g., Lujan 1990) have identified omissions, erroneous enumerations, and household misses. Sources of errors include irregular housing, complex households, mobility, cultural differences, residence concept mismatches and interviewer error. Large households, renter status and coverage errors are linked.

The purpose of this paper is to match multiple data sources to identify types and sources of coverage error in a small nonrandom sample of 31 Cheyenne River Sioux Indian Reservation households in the 2006 Census Test. Data sources used in this study include: 1) test census data from the Update/Enumerate (U/E) Operation, matched to 2) independent test survey data from the Census Coverage Measurement Personal Interview (CCM PI) Operation done several months later, matched to 3) audiotapes of the 31 CCM PI interviews, matched to 4) qualitative debriefings in 14 of those and to 5) field notes. I start by describing the Cheyenne River Sioux Indian Reservation. I discuss the data sources in depth and describe methods used to conduct observations and qualitative debriefings. I summarize the matched households in terms of cases 1) where census and CCM rosters match and 2) where they are inconsistent. I use case studies to show living situations and summarize types and sources of coverage error and the benefits of using multiple data sources.

2. Description of the Cheyenne River Sioux Indian Reservation

The Cheyenne River Reservation is in the rolling prairies of central South Dakota, with a few very small towns along the main roads in a rural area about the size of Connecticut. This is the reservation of the Lakota Sioux, a High Plains Indian tribe who formerly were nomadic, hunting buffalo on horses until settlers exterminated the buffalo. Today, people live in houses, apartments, trailers, ranches and farms but are still mobile in and out of households, on and off the reservation. When asked for address, most give just a P.O. Box number, for two reasons. First, this rural area has few named streets; houses are located by distance from landmarks and descriptions, like "white house with red door." Even in towns with street names, there is little or no reason to know them; most do not. Nearly everyone picks up mail at the post office as there is almost no home mail delivery. People use local road names, not official map names.

The Sioux have developed a sophisticated culture based on living in harmony with nature; valuing the group's welfare over that of the individual, maintaining relationships, meeting the needs of others and taking them in, integrating spirituality with everyday life, and

being principled in tradition (Olson 2005). In 2000, the reservation population was 8,470–about 74 percent American Indian with the rest whites mostly living on farms and ranches on private land (Olson 2005). In this rural area, many know each other.

The poverty rate on this reservation is high. In 2000, its two counties were in the lowest one percentile of per capita U.S. income; people earned just 40 percent of the national per capita income (Olson 2005). This area also has high rates of unemployment, public assistance, vacancy, and lack of telephones. These factors associated with census coverage error are included in the Census Bureau's Planning Data Base (PDB) as scale variables to produce relative hard-to-count scores for each county (Robinson et al. 2007); half of these counties' census tracts have disproportionately high PDB hard-to-count scores (Robinson, May 2008 personal communication). About 44 percent of the housing is rental. Some is owned by the government: Housing and Urban Development, the Bureau of Indian Affairs and Indian Health Service (Olson 2005). The tribe owns housing that it rents to its employees (who may wait five years) and to disabled people.

3. Data Sources and Methods

This is a study of 31 reservation households included in the Census Coverage Measurement Person Interview (CCM PI) Operation for which we have four sources of data. The first source is lists of people rostered in these households during the Update/Enumerate (U/E) census collection in this Census Test site in March and April 2006. This involves personal visit interviews to complete paper census forms in every household in the site. The U/E Operation is conducted on Indian reservations and in remote locations where the lack of city-style addresses and other factors preclude use of the mailout form used in most U.S. areas (see Heimel, King, and Sheppard 2006).

The second source is lists of people rostered in a sample of CCM PI households. This is the independent sample survey conducted some months after the census to provide data to estimate census coverage using dual systems estimation. The CCM is the current equivalent of the 1990 PES and the 2000 A.C.E. CCM PI interviewers conducted computer-assisted personal interviews on laptops from June to September 2006. The instrument starts with the roster question: "First, please tell me the names of everyone who lives here now." It then asks roster probes to compile an inclusive roster of the names of 1) "...anyone else who has another place to live but who stays here often"; 2) "...anyone else who is staying here until they find a place to live"; 3) "…any babies, foster children or other children who stay here that you didn't mention yet"; 4) "…any relatives or unrelated people who live here"; and 5) "…anyone else living or staying here during March or April of this year who is no longer here." Each had a followup: "Anyone else?" The CCM also has residence rule questions to decide which people on the roster should and should not be counted in the census. The Person Follow-up (PFU) Operation, done in early 2007 to resolve ambiguous household and person matches and nonmatches, altered some of our residence status codes.

The third data source consists of audiotapes and observations in our 31 households. In 14 of these, we have a fourth source: qualitative unstructured debriefings done right after the CCM PI interview to resolve any ambiguities in who should be counted there. Researchers Tamara Adams and I collected these two data sources while accompanying interviewers as they conducted CCM PI interviews during one week in July 2006. Before the interview, the interviewer introduced the researcher, allowing the researcher to request permission to tape; all respondents agreed. The interviewer then did the interview. The researcher taped, observed and listened intently to the interviewer-respondent interaction to identify any clues of persons missed or incorrectly listed or for any other residence ambiguity. If any clues appeared, the researcher did an informal debriefing with the respondent to resolve any issues on who to include on the final census and/or CCM roster (see Nichols and Childs 2007a and 2009). A case study will be presented later. The fifth source is field notes.

This analysis was done in 2008, when the final matched Census-CCM PI rosters were accessible in the Census Bureau's Per Marcs electronic system. I compared the matched census/CCM data for these households using transcripts of our taped CCM interviews and debriefings to determine which people were on one or both lists, the reasons for discrepancies and whether they may be coverage errors.

4. Limitations

It is important to note that while these households were originally part of the CCM PI scientific sample to estimate census coverage, this set of 31 households is a non-random sample; it is not representative of any larger area or group. The results on inconsistencies and possible coverage errors cannot be generalized to this reservation as a whole in this test. This was an observational study to document what actually happened in the field and to conduct qualitative debriefings when necessary. We researchers had no role in choosing enumeration areas or cases. We either accompanied assigned interviewers or asked interviewers in the office if we could go with them.

5. Results

Our 31 households comprise about 6 percent of the 500 CCM survey households in the 2006 Census Test on this reservation, in 13 (4.9 percent) of the 265 sampling clusters. They were in government-, tribal- and privately-owned houses, trailers, and apartments in towns and rural areas; we do not know how typical they were of the entire reservation. We start with consistency of the census and CCM rosters, leaving aside for a moment whether coverage error occurred. Table 1 shows that 17 (55 percent) of our 31 households were matched in numbers and names of people on both rosters. In the other 14 (45 percent), there was at least one census/CCM inconsistency. In four households (column 3), the Census list had one or more persons the CCM did not. In seven households (column 5), the CCM list had one or more persons the census did not. In the last three households (column 4), we saw both inconsistencies.

Inconsistencies between the two lists do not necessarily indicate coverage errors. Some inconsistencies will be natural differences resulting from the different time periods for which lists of persons were collected: April 1 for the census, and both April 1 and the Interview Day for the CCM PI. I later distinguish those who may have had coverage error from those who did not. The 2006 Census Test Residence Rules were used to decide who to count (Vitrano 2006). For example, coverage errors include omissions of those who should have been counted in the household in the Census or CCM and those counted in more than one household in either operation.

Table 1: Consistency of Census and CCM Rosters in Persons Listed and Whether Debriefings were Conducted

Debriefing conducted	*Consistency: Census & CCM roster same*	*Inconsistency: Census has persons CCM does not*	*Inconsistency: Both have persons other does not*	*Inconsistency: CCM has persons Census does not*
Yes	5	3	3	3
No	12	1	0	4
Totals	17	4	3	7

5.1 Households with Consistent Rosters

In 15 of the Table 1 consistent households, the matches were complete. In the other two a rostered person would later be classified as out of scope for the census. The debriefings showed that both were due to living situations and mobility. In one, a college student on both had been at school off the reservation but was back in the household in July in the CCM PI; he should not be counted there.

In the second, a classification error occurred due to a conceptual difference between the census rules and the respondent's understanding of residence rights. The CCM PI instrument classified three family members as inmovers between the Census and CCM interview days. The observation and debriefing revealed that this family had not yet moved in by the CCM PI interview day; when we arrived, the house was vacant and the respondent was repairing it. Before the interview began, he said he had started renting this tribal house a month earlier after getting to the top of the waiting list, but he wanted to finish fixing it up before moving in. Despite this statement and the lack of furnishings, the interviewer began the interview, maybe thinking she would learn who had lived there earlier. But when asked for the names of everyone living here now, the respondent gave their names, despite having just said they had not moved in yet. The instrument did establish that this family should be counted at another address on Census Day (I later learned this family was tabulated correctly at that other address). However, the instrument also classified this house as occupied on CCM Interview Day, although the debriefing showed that the family had not yet slept there. This house should have been classified as vacant, not occupied, in the CCM.

This case reveals a mismatch between the Census Bureau's usual residence definition and this man's concept of residence rights. His family had not yet slept or moved here, but had rights to the house, were renting and fixing it and choosing when to move in; it is their house. But the residence rule states that we count people where they live and sleep most of the time—they should not be counted here.

5.2 Households with Inconsistent Rosters I: Census has Persons CCM Does Not

We now focus on households with inconsistent names of one or more persons on matched census/CCM PI rosters. We consider first the inconsistency occurring when persons were on the census, not the CCM PI roster. CCM PI interviews were done three to four months after the census. We aimed to identify outmovers since Census Day with the probe, "Was there anyone else living or staying here during March or April of this year who is no longer living here?" The probe did not work consistently; as seen in Table 1 columns 3 plus 4, seven of the 31 households had one or more persons on the census but not the CCM roster. Four (column 3) had just census people not on the CCM or those who should not be listed. The three in column 4 had both CCM people not on the census and the converse.

Of the four households with just this inconsistency type, three had one person on the census form not on the CCM roster. In the first, a middle-aged woman was listed as a roomer in a married couple's house in the census, but she was not identified in the CCM outmover probe above. In a later followup this woman was checked and coded as a correct enumeration. In the second, a young adult son was on the census, but not CCM roster. I searched his name in all CCM households in the dataset but did not find him. Lacking other data, we assume he was correctly counted there in the census; he may or may not have been in this test site in a non-sample household during the CCM PI. In the third household, it was a brother living with two sisters, included by one of them on the census, but not CCM, roster.

In the fourth household, two men on the census roster did stay at this address around Census Day and should be counted here. They came onto the roster in response to the "anyone staying here until they find a place to live" probe. Before the July CCM interview day these brothers of the respondent had each moved out. The debriefing revealed that one brother had lived here from January to June, but then moved to an off-reservation institution. The second brother was visiting this household during the CCM interview. He had lived here for two years and was counted here in the census, but was now staying with others "nearby" while still looking for his own place, without giving the address. This man's residence in the CCM remained unresolved; he was not followed up by CCM staff by design due to this debriefing. He did not turn up in my full CCM dataset search, but "nearby," implies he is living in the test site. He is a CCM missing data case; tenuously attached men like these can be missed or counted more than once.

In the last three households, both types of inconsistencies occurred: persons on the census, not the CCM roster and vice versa. In each, subfamilies were on the census, but not CCM roster. In household five, a sister and four "other relatives" (likely her children) appeared

only on the census roster of a woman with five resident children. We did not find them in any CCM household—we lack data to decide if they were valid census residents or just visitors. In household six, a daughter in her early twenties, a man of the same age and his 1-year old son were listed in the census. A later search of the whole CCM dataset revealed that the man was rostered in both the CCM and census in another household; his duplication was resolved by counting him at the other household, not here. The son was counted here in the census, then moved to his father's house and was counted there in the CCM. The daughter was not located in the CCM.

Debriefings were done in five of these households for other reasons; in just one was a census outmover identified in the interview or the debriefing. Debriefings were triggered by perceived ambiguities during the CCM PI interview. Respondents who do not recall that additional people had stayed there several months ago would not give any sign of anomalies, so a debriefing would not be done. This is a design limitation of just observing in CCM households and not conducting debriefings in all households to probe for outmovers.

A solution to recall error is elusive. We can consider adding a new CCM PI probe to try to further stimulate memories of residents on April 1. We could modify the study design to do a debriefing in every observed household to try to elicit memories of people around April 1. But we should also assess whether most people can still remember who was there on an indefinite date 90 or more days ago.

In household seven, a man, his wife, and two children were on his mother's census roster, but not all were on her CCM roster. The debriefing collected critical information. This case study summarizes the CCM interview and shows what was learned in the debriefing.

5.2.1 A Case Study Illustrating the Survey Interview Interaction and Debriefing

On the third visit to a house, we found a man in the driveway. The interviewer asked, "Do you currently live here?" After hesitating, he said, "No, I don't live here–my mother does." Classifying him as a proxy respondent, she asked for "the names of everyone who lives here now." He gave his mother's name. She asked, "Is there anyone else who has another place but who stays here often?" He hesitated and said tentatively, "I stay here often." Instead of listing him, she asked an unscripted probe as to whether he lived there most of the time. After pondering, he answered no. She did not roster him. Asked if any babies or other children stayed there, he named two of his mother's grandchildren. Asked if any relatives or unrelated persons living here had been missed, he named his brother who comes and goes. She rostered this brother as the householder's son. Because he had not been living here around Census Day and had left, it seemed he should not be counted here in the census (we return to him later). The final roster listed his mother and her two grandchildren.

I did a debriefing to ascertain the Census Day residence status of the grandchildren and to learn why the respondent hesitated about "staying here often." I learned that the grandchildren had been here for ten months—they should therefore be counted here in the census. The respondent's brother, when not here, sometimes stayed with his sister and at another place. He was gone now. Around April 1, he did not stay here most of the time, so he should not be counted here. But then who were the grandchildren's parents?

"My wife and I." Since his children were here, I asked where he and his wife are living. He explained that he and his wife are not "living" in this household, they are just "staying" here with his mother during a time of transition. They have been staying here for the last ten months while preparing to move their trailer here from town and it will take another few months to finish. The debriefing thus revealed how this man and his family came to be "staying" at his mother's house, not "living" here, for what may stretch into a full year. According to the census residence rule, this man and his family have been living or staying here most of the time, including Census Day, so this household is their "usual residence"; they should be counted here. This respondent, however, did not consider himself to "live" here and so did not give his or his wife's name to either the roster question or probes–this explains his hesitations.

This man's decision that he does not "live" here is consistent with Gerber's research (1994) on conceptual systems respondents use to make residence decisions. The *intention* to stay just temporarily with the *expectation* of leaving in the near future links persons loosely to households and may result in them not including themselves as household members on rosters. Gerber's theory has been confirmed in a study of complex households in six race/ethnic groups (Schwede 2006).

The case study reveals one type of coverage error—CCM person omissions—and three sources of error. The first source of error related to survey question wording. The main roster question asked, "Do you currently live here?" Wording asking just for people who "live here" was not consistent with his concept of "staying here." There is thus a mismatch in residence concepts between the survey instrument and the respondent (the second error source). Question rewording to "Do you currently live or stay here?" fixed this. The third was interviewer error–she should not have asked an unscripted probe and should have rostered him at the "stays often" probe; the instrument would have classified him automatically. Had she listed him, he might have named his wife, completing the roster correctly.

In sum, seven households were inconsistent in having one or more persons rostered on the census, but not the CCM PI rosters (or not counted there in the CCM). The case study household had a CCM coverage error—2 person omissions—but this was corrected later in CCM. The household with the middle-aged woman was resolved later as a correct enumeration. In another, I confirmed with the outmover brother that he had been a Census Day resident but was staying elsewhere in the site on our CCM Interview Day. We did not get an adequate address, so his CCM address is unresolved. The household with the daughter, man and son in the census was later partly resolved by tabulating the man (a duplicate), in another CCM household for both the census and CCM, and his son here in the census and there with his father in the CCM. We did not find in the CCM 1) the woman in that house, 2) the son in another, or 3) the woman and four children; we do not know if they should be counted in these census households. I discuss the case study brother below.

What factors explain why there was no mention or clue in the CCM interview that one or more additional people had been counted in five of these inconsistent households some months earlier in the census? We do not have any concrete data; our observations were done during the CCM PI Operation alone and these census people were not identified in the CCM interview. We can only speculate and identify possible error sources. One is that different people were the census and CCM respondents for the same household and may have differed in listing persons. The second is recall decay–the census interviews had taken place several months earlier and respondents may have forgotten that others were in the household at that time. A third error source might be the lack of a specified U/E interview date. Respondents often say they cannot remember what happened on April 1; recalling an indeterminate interview day would be harder. A fourth error source could be the high level of mobility among households for indeterminate time periods both on and off this reservation. We do not know whether these extra people captured in census households were core residents, tenuously attached persons, or just visitors staying for short periods who should not be counted there. A fifth possible source is concealment; I did not see this.

5.3 Households with Inconsistent Rosters II: CCM has Persons Census Does Not

We turn to ten households with inconsistencies in which persons rostered on the CCM PI roster were not listed on the census roster (Table 1 columns 4 and 5). Seven (column 5) had this type of error only, while three (column 4) also had census people not on the CCM roster. This relatively large number of people on the CCM but not census roster is partly due to a design feature and does not necessarily imply coverage error. Designed to measure coverage, the CCM PI instrument casts a wide net with inclusive probes (shown in Section 3) to identify persons and situations known to be hard to count. Once rostered, other residence rule probes determine if each person should be counted as a Census Day resident or not. As noted, the CCM list includes new inmovers since Census Day, identifies their Census Day residence and counts them at the other place; these are not necessarily coverage errors.

These ten households fall into five categories based on living situations of persons identified in the CCM PI who were not on the census rosters: 1) mobile, tenuously attached persons; 2) mobile students going on and off the reservation; 3) reservation students sleeping at the school dorm in town during the school week and at home on the weekends; 4) mobile visitors from outside the reservation; and 5) unknown reason. I describe households in each category, specifying which inconsistencies indicate potential coverage errors.

5.3.1 Mobile, Tenuously Attached Persons

Three households are in the first category in which mobile, tenuously attached persons are identified in the CCM. In one, a young homeless man cycled among reservation homes of friends and relatives and his girlfriend's off-reservation home. He shows up unexpectedly, stays indefinitely, and leaves without notice of departure or destination. The respondent has had a long relationship with this man and has kept his home open to him especially since his mother died. We were lucky to catch him on this household's CCM roster; he had arrived the day before and might soon be gone again. The instrument recorded that around April 1, he had mostly been staying off reservation with his girlfriend; the debriefing confirmed he was out of scope. The respondent did not know the man's age or birthdate, his girlfriend's phone number, or where he was when not here. Without these data, it would be very difficult to match him.

In the second (case study) household, I noted that the respondent's brother had been added to the roster in a probe, had moved among his sister's and other places, and should not be counted at his mother's on Census Day. I am pleased to report that in analyzing the rosters for this paper, I found this man in the census in an unexpected place–his sister's apartment (our third household in Section 5.2 with one brother on the census, but not CCM roster). The mention of his sister's place during the CCM interview at his mother's house validates that he was properly matched to his sister's house in the census. There are two reasons for this miss: 1) his sister did not identify him in her CCM interview and 2) the respondent in the mother's household (his brother) could not give his age and birthdate. These are critical match variables for identifying census duplications. If they are inaccurate or missing, the match probability score may not be high enough to be valid. [Some other Sioux respondents did not recall or give full birthdates and/or ages for relatives other than children; I am exploring whether Sioux and/or Navajo people (during 2008 research) are similar to or different from others in recalling and reporting relatives' exact ages and birthdates).] In a third household, it is not clear if the young man on the CCM roster is a tenuously attached person without his own place or a new unmarried partner staying sometimes, with another place. The respondent listed him in the "staying often" probe as cycling between her house and one off-reservation monthly. His status there seems tenuous.

In sum, just the tenuously attached brother newly matched in this study had a coverage error. This match was not identified in the CCM.

5.3.2 Students Going on and off the Reservation for School

The second category includes two households in which new CCM persons live off-reservation most of the time for school. In one household, they were teenagers visiting their father on vacation, but living off reservation with their mother to attend school. In the other, the college student on the CCM roster had been properly omitted from his mother's census roster as living at school.

5.3.3 Reservation Students Cycling between Sleeping at the Dorm during the Week and at Home on Weekends

This category has a household in which three children under 17 living with relatives were on the CCM but not census roster. The CCM revealed that in the March/April period, two of the children were staying overnight in the school dorm in town during the week and returning to this relative's home on weekends; the other was staying here. According to the residence rules (Vitrano 2006), they should have been counted in the census but were not: these are coverage errors. This is important: American Indian children below college age

often go to boarding schools or dorms on- or off-reservation (see Tongue 2006, Lobo 1990). They should be counted in their parents' homes, but may be missed. Children under 17 on Indian reservations were undercounted in the 1990 Census (West and Robinson 1999).

5.3.4 Mobile Visitors from off the Reservation

Three households had people who live off-reservation and come back to visit. They should not be counted on the reservation as of Census Day unless they were there then and had no usual residence elsewhere. In one of these, the respondent listed 6 visitors; none should be counted there. Even this respondent and spouse were later identified as cross-household duplicates and tabulated elsewhere.

5.3.5 Unknown Reason

Two households had CCM residents who respondents claimed were also Census Day residents, but they are not on the households' census rosters. We do not know why. One is a cousin. The second household includes a young adult son and a grandchild with an unresolved flag. This household's debriefing had focused on the respondent (the son's mother) and did not include questions about him because she gave no sign he had been away on Census Day. As in our case study, this respondent insisted she did not "live" here, but on a second visit the interviewer, who knew her, got her to start the interview by agreeing that she had two places and stayed here sometimes (she was verified as living here in the CCM and counted here in the census). Her granddaughter was now staying with her mother "over there" for two months, but she did not give address identifiers, leaving the girl's CCM address unresolved. Tamara Adams ran a customized search of the census dataset to try to find these people in this site. She did not find the cousin, but did find the son tabulated at another household in both the census and CCM. He had been identified by CCM staff as a cross-household CCM duplicate. This shows that decisions about who should be counted in the household during the observations and debriefings in the PI Operation may later change during processing. The cousin is a possible census omission. The son was correctly unduplicated.

In these ten households with persons on the CCM, but not on census rosters, the large number of inconsistencies has been whittled down to potential coverage errors in four households for five persons. The tenuously attached brother in his mother's CCM house newly matched to his sister's census household is a coverage error; this match was not identified in CCM processing due to the absence of critical matching data and his not being rostered in his sister's CCM interview. In a third household, two middle-school students cycling between sleeping at the school dorm in town for five nights and at home on weekends and a younger child were not rostered: 3 census omissions. The cousin expected, but not found, in a fourth household was a possible census omission. The granddaughter's CCM residence is unresolved, but the girl seemed to be nearby (in the site). In a real census, this would be resolved or imputed.

6. Types and Sources of Coverage Error

In the preceding, I wove together information from the four matched data sources for each of the 14 households with inconsistent census and CCM rosters to identify those with potential coverage errors. Five types of coverage error were identified in nine of our 31 households: 1) likely or possible census omission; 2) unidentified cross-household match; 3) CCM omission; 4) duplicated persons across households; and 5) unmatched census persons. Note that some households fall into more than one of these five categories.

Data from our linked observations and debriefings enabled me to identify four households in categories one to three. In the first, likely or possible census omissions of four persons were found in two households: the one with students cycling between sleeping at the dorm and at home and the one with the missing cousin. In the second category, one new cross-household match (the case study brother) was missed by the CCM. In the third, our case study household had two CCM person omissions; these PI errors were corrected later.

Use of the final post-processing Per Marcs matched rosters enabled me to identify five households with category 4 cross-household duplications (one type of erroneous enumeration). The three "inmovers" in the vacant house were correctly tabulated back at the right census residence. The other cross-household duplications in three study households of five persons resulted in census tabulations of the 1) one-year old child in the study household and 2) father, respondent and spouse, and adult son in other households (erroneous enumerations in study households). Further, the Per Marcs system identified three households with seven unmatched, unresolved census people in category 5. As noted, the CCM does not need to capture all persons to work; it just needs to be independent. In a real census, greater efforts would be made to follow up on these census-only people (in this test, we did not calculate coverage rates and we did not send the debriefing households to PFU). In a real census, CCM would impute residence and enumeration probabilities of correct enumerations and errors for these persons before estimating coverage rates. But in a real census, the blanket person search in the whole CCM/Census dataset used in this qualitative study in a small circumscribed test site could not be done at the national level.

Characteristics of households and persons with coverage error in our study mirror prior census research results showing that the variables most closely associated with census coverage error are tenure (renter), large household size, and complex households for households; relationship (nonrelatives and other than spouse); race, age (young children and adults 18-29) for individuals; and Indian reservation and personal visit enumeration (Ellis 1994, West and Robinson 1998, de la Puente 1993). Respondents in all nine households with possible coverage errors are American Indians; eight are renters. Their average household size is larger—4.55 (CCM) and 4 (census)—than the averages for: 1) this sample of 31 households (3.87), 2) the 15 consistent households (2.93) and 3) the overall U.S. average household size of 2.59. In the CCM, all were complex households, that is, persons other than, or in addition to, married couples with their joint biological kin (Schwede 2003); in the census, six were complex. Of the 23 people with possible coverage errors,

all are American Indian. Most are young, with ten under 17, ten in the 18-29 range and three older. Seven of the ten under age 17 were females, while seven of the ten from 18-29 were males. By relationship, two were respondents and two spouses. Six were biological children, three were adult siblings, eight were nieces/nephews/cousins, one an unmarried partner and 2 were "other nonrelatives."

What were the sources of these errors? Clearly, two major linked factors were the high levels of mobility in and out of households and other places on and off this reservation, and household fluidity. Fourteen (45 percent) of the 31 households had some type of change in people rostered in the three to four month period between the census enumeration and the CCM survey in the 2006 Census Test. This is a lot of change in a short time period. Mobility may be due, in part, to poverty and the lack of reservation jobs. It is also rooted in Sioux nomadic culture with strong cultural obligations to care for kin, according to a Sioux crew leader. As we have seen, Sioux people take in individuals and families for indeterminate periods. Some of our 31 are "anchor households," described by Ackerman (1988), in which one or more core people, often women, remain in the household, allowing others to come and go. Two young CCM interviewers were also cycling among places without their own homes. By the time I returned six months later, each had left the reservation to join the military or go to college. High mobility and fluid households are factors affecting enumeration on this reservation.

Five specific types of living situations associated with mobility and attachments to places were identified as potential error sources. Likely and possible coverage errors were found in two of these, one of which is new and rather particular to Indian reservations which have school dorms and boarding schools: below college-age students moving back and forth on the reservation, sleeping at home on the weekends and the school dorm for the week, who are missed in the census. According to the residence rule, these children are to be counted at their homes. However, if parents/guardians do not list them on household rosters they are missed in the census; we do not enumerate in boarding schools. The category of mobile, tenuously attached persons also included a man with a possible coverage error.

Recall decay is an error source that may have affected coverage. Five CCM PI persons in July did not mention (and likely did not recall) people staying with them several months earlier. Question wording, mismatches of respondents' residence concepts and census terms, and enumerator error were also error sources. Two respondents interpreted the initial roster question asking for names of everyone living here now as not applying to them and did not think they should be listed. One omitted himself and his wife from the roster. The other was contacted twice. Already acquainted with this woman, the interviewer changed the first question so the woman would begin the interview, only to show she did live here. This suggests that pre-existing interviewer-respondent relationships may affect respondent participation in surveys. In this small, close-knit community, interviewers often knew respondents. This may be a new source of error.

A final error source is inadequate or missing data on critical match variables, especially addresses, unit descriptions, birthdates and ages. Due to this source of error, a cross-household match was missed and several persons had unresolved CCM addresses. We might want to emphasize, in our new advertising campaign, links between providing complete address, age, and birth date data and complete counts.

7. Summary and Recommendations on Using Multiple Data Sources

In this paper, I used five data sources to identify types and sources of coverage errors in a small non-random household sample on the Cheyenne River Sioux Indian Reservation in the 2006 Census Test. Observation of CCM interviews with optional immediate qualitative debriefings enabled us to observe interviewer-respondent interactions in the natural setting of respondents' homes and to follow up to resolve coverage issues. This combination of field methods was invaluable in identifying types and sources of coverage error, as shown in the case study; we experienced first-hand the challenges of enumeration in such complex and fluid reservation households. Also, the triangulation of four direct data sources for a small number of households worked well, providing more insights into types and sources of coverage error than any one method could provide alone. We identified possible improvements to overcome limitations of our method: doing debriefings after each interview to try to identify census omissions and observing and debriefing in census operations.

Acknowledgments

Disclaimer: This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on methodological, technical or operational issues are those of the author and not necessarily those of the U.S. Census Bureau. This is a condensed version of a longer paper that will be posted at http://www.census.gov/srd/papers.

The author thanks Elizabeth Nichols for her very helpful reviews. She thanks Tamara Adams for reviews, insights, her interviews on the reservation, her help in using and interpreting matched Per Marcs data, and for customized dataset searches. I also thank Tom Mule for observing and taping four interviews. This study was also done in the Austin, Texas site (Nichols et al. 2007, Adams and Nichols 2007).

References

Ackerman, L. 1988. Residential Mobility among the Colville Indians. Ethnographic Coverage Evaluation of the 1988 Dress Rehearsal Census Report #2. Unpublished report for the Statistical Research Division, U.S. Census Bureau.

Adams, T. and E. Nichols. 2007. Census Day Residence Evaluation. 2006 Census Test Evaluation 5. US Census Bureau.

de la Puente, M. 1993. Why are People Missed or Erroneously Enumerated in the Census: A Summary of Findings from Ethnographic Research, *Proceedings of the 1993 Research Conference on Undercounted Ethnic Populations*, pp. 29-66.

Ellis, Y. 1994. Categorical Data Analysis of Census Omissions. DSSD 1990 REX Memorandum Series #PP-10. July 26.

Fenstermaker, D. and D. Haines. 2002. A.C.E. Revision II: Summary of Estimated Net Coverage. DSSD A.C.E. Revision II Memorandum Series #PP-54. December 31, 2002.

Gerber, E. 1994. Respondents' Understanding of Residence Terminology in Cognitive Research and in the Living Situation Survey. *Bureau of the Census 1994 Annual Research Conference Proceedings*, pp. 434-449.

Heimel, S., R. King, and D. Sheppard. 2007. 2006 Census Test Evaluation #1: Residence Rules and Instructions. 2006 Census Test Evaluation Series. July 26, 2007.

Jojola, T. 1992. Census Ethnographic Research Project: Isleta Pueblo, New Mexico. Ethnographic Evaluation of the 1990 Decennial Census Report # 23. Bureau of the Census, Washington, D.C.

Lobo, S. 1990. Oakland's American Indian Community History, Social Organization, and Factors that Contribute to Census Undercount. Ethnographic Exploratory Research Report #2. Available at: http://www.census.gov/srd/papers/pdf/ex90-12.pdf

Lujan, C. 1990. As Simple as One, Two, Three: Census Underenumeration among the American Indians and Alaska Natives. Undercount Behavioral Research Group Staff Working Paper # 2. Available at: http://www.census.gov/srd/papers/pdf/ev90-19.pdf

Mule, T. 2002. A.C.E. Revision II Results: Further Study of Person Duplication. *DSSD Revised A.C.E. Estimates Memorandum Series PP-51*. Washington, D.C: U.S. Census Bureau. December 31, 2002.

National Research Council. 2004. *The 2000 Census: Counting under Adversity. Panel to Review the Census*. Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, eds. Committee on National Statistics, Division of Behavioral and Social Science and Education. Washington D.C: The National Academies Press.

Nichols, E. and J. Childs. 2009. Respondent Debriefings Conducted by Experts: A Technique for Questionnaire Evaluation, *Field Methods.*

Nichols, E. and J.H. Childs. 2007. Respondent Debriefings Conducted by Experts: A New Qualitative Methodology for Questionnaire Evaluation. *SRD Report Series: Survey Methodology 2007/39* at www.census.gov/srd/papers/pdf/rsm2007-39.pdf

Nichols, E., J. H. Childs, L. Schwede, J. Bibb, V. Smith, S. Norton, J. Burnham, T. Adams, E. Martin. 2007. 2006 Census Coverage Measurement Person Interview Trip Reports: July, 2006. DSSD 2010 Census Coverage Measurement Series #2006-D7-11.

Olson, T. 2005 Orientation Guide for Working on the Cheyenne River Reservation. Decennial Management Training, Oversight, and Recruiting Branch, Field Division. July, 2005.

Robinson, J. G., C. Johanson, and A. Bruce. 2007. The Planning Data Base: Decennial Census Data for Historical, Real-Time, and Prospective Analysis. Presented at the 2007 Joint Statistical Meetings in Salt Lake City, July 23, 2007.

Schwede, L. 2006. 'Who Lives Here?:' Complex Ethnic Households in America, in *Complex Ethnic Households in America*, ed. by L. Schwede, R. L. Blumberg, and A. Y. Chan. Lanham, MD: Rowman & Littlefield Publishers.

Schwede, L. 2003. Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups. *Census 2000 Testing, Experimentation, and Evaluation Program Ethnographic Evaluation*. U.S. Census Bureau, August 7, 2003. http://www.census.gov/pred/www/rpts/Complex%20Households%20Final%20Report.pdf

Tongue, N. 2006. I Live Here and I Stay There: Navajo Perceptions of Households on the Reservation, in *Complex Ethnic Households in America*, L. Schwede, R. Blumberg, and A. Chan, eds. Lanham, MD: Rowman & Littlefield.

Vitrano, F. 2006. 2006 Census Test Residence Rule and Residence Situations. *2010 Census Test Memoranda Series, Chapter: 2006 Census Test, No. 31*. January 6, 2006.

West, K. and G. Robinson. 1999. What Do We Know About the Undercount of Children? Population Division Working Paper 39.