DATA MINING THE CPS REINTERVIEW: DIGGING INTO RESPONSE ERROR

Pamela D. McGovern and John M. Bushery
U.S. Bureau of the Census

1.   Introduction

The Current Population Survey (CPS) is a monthly survey of approximately 60,000 households and is the primary source of information on labor force characteristics for the U.S. Data obtained from the CPS is used to calculate the U.S. unemployment rate.

This paper illustrates how reinterview data can help identify sources of error and focus research to improve data quality. Because it's experimental, we often present point estimates only. However, we make statistical comparisons to confirm characteristics associated with inconsistent reporting.

Response error causes moderate unreliability in the CPS estimates of "Unemployed" (Waite 1990, 1991, 1993, 1994). In categorical data, like labor force status, poor reliability due to measurement error is mathematically associated with bias (U.S. Census Bureau, 1985).

Unfortunately, the traditional reliability measure of the CPS, the index of inconsistency, is a difficult concept to interpret.   Most users can grasp the effect of response error more easily from the Interview-by-Reinterview cross-tabulation (Table 1). The denominators of the marginal percentages in Table 1 are the total numbers of respondents in the reinterview. The percentages in the interior cells use the column totals as the denominators. Table 1 shows that only two-thirds (66.3 percent) of people classified as "Unemployed" in the original interview are classified as "Unemployed" in the reinterview. Considered from another viewpoint, of the 960 people classified as "Unemployed" in one interview or the other, only about half (49.0 percent) were classified as "Unemployed" in both interviews.

The reliability problems in the estimates of "Unemployed" also cause bias. Using the estimates of the CPS response probabilities from a Markov latent class analysis of CPS data (Biemer and Bushery, 1999), we estimate that "Unemployed" is understated by about 13 percent.

Reinterview analyses can identify problematic questions. "Data mining" the CPS reinterview data also can identify demographic groups most associated with inconsistent reporting of unemployment and help determine which questions generate inconsistency in the major labor force classifications.

2.   Methodology

We analyzed CPS reinterview data from January 1995 through December 1997, from both CATI and CAPI. The CPS has used a reinterview program to evaluate data quality since the early 1950s. In the reinterview a subsample of the survey households receives a second interview, covering the same time period as the original survey. The CPS reinterview program has two components: response

error (RE) and interviewer quality control. This paper discusses only the RE component, which measures simple response variance (test-retest reliability). Since 1994 the CPS and its reinterview have been conducted using Computer Assisted Interviewing (CAI). The RE reinterview re-asks all the questions in the CPS from a random sample of about 500 households each month. The RE sample excludes all original noninterviews, cases previously reinterviewed, and cases without telephone numbers.

Senior interviewers conduct the reinterviews, almost exclusively by telephone, between one and ten days after the original interview. From January 1994 through January 1998, the original CPS respondent was reinterviewed. If that respondent could not be reinterviewed, the reinterviewer accepted a noninterview for that household. Only household members who are 15 years and older are included in the CPS and reinterview.

2.1   Analysis Methods

Random errors of measurement in the survey process add variability to the data we collect from respondents. When the errors are not correlated with the answers or with each other, we call this variability "simple response variance". The index of inconsistency is a relative measure of response variance. It is defined as the ratio of simple response variance to the total variance (U.S. Census Bureau, 1985). The index can identify unreliable questions or estimates, but cannot explain why the data are unreliable.

Original-by-reinterview cross-tabulations of labor force classifications, such as Table 1, show changes between the two interviews. They provide more detail about the unreliable data.

The next level of analysis compares the characteristics associated with inconsistent responses with those associated with consistent responses. Differences in these characteristics might provide clues for improving data reliability.

The CPS assigns a Major Labor Force Recode (MLR) to each sample person, based on an algorithm using responses to several key labor force questions. We examined responses to individual questions used to assign the MLR and we examined demographic characteristics associated with inconsistent responses. We concentrated the analysis on inconsistent classifications between "Unemployed Looking for Work" and "Not in the Labor Force (NILF) Other."

Four questions determine the MLR for the "Unemployed Looking for Work" category. We compared responses to these questions for people with consistent and inconsistent MLRs to determine which questions generated inconsistent MLRs.
- *Have you been doing anything to find work during the last four weeks?* Yes/No
- *What are all of the things you have done to find work during the last four weeks?*
- *Last week, could you have started a job if one had been offered?* Yes/No
- *Why is that?* Waiting for new job to begin/Own temp. illness/Going to school/Other

A negative response to the first question results in an MLR of "NILF Other." The second question, "*What are all of the things you have done to find work during the last four weeks?*" elicits

information about active versus passive job searching. People were classified as actively looking for work if they mentioned at least one active job search response. People not actively looking for work were classified as "NILF Other."

Active job search
- Contacted employer directly / interview
- Contacted public employment agency
- Contacted private employment agency
- Contacted friends or relatives
- Contacted school / university employment center
- Sent out resumes / filled out applications
- Checked union / professional registers
- Placed or answered ads
- Other active (specify)

Passive job search
- Looked at ads
- Attended job training programs / courses
- Other passive (specify)
- Nothing

We conducted a demographic analysis to determine if relationships exist between demographic characteristics and inconsistent classification as "Unemployed Looking." We looked at the age, marital status, relationship to the household reference person, gender, race, and education of respondents and subject persons. We performed a similar analysis for the question, *"Have you been doing anything to find work during the last four weeks?"* to find any groups more likely to answer this question inconsistently.

2.2   Limitations

A response variance reinterview relies on two assumptions: independence and replication. The reinterview may not have been independent of the original interview to the extent that respondents remembered and repeated their answers from the original interview. Lack of independence would cause an understatement of the number of inconsistencies.

Operational constraints make it difficult to conduct the reinterview as an exact replication of the original interview. When a reinterview does not replicate the original interview perfectly, the differences in methodology may cause an overestimation or underestimation of the response variance. Aspects of the reinterview which did not replicate the original interview are:
- Only senior interviewers conducted the reinterview.
- Almost all reinterviews were conducted by telephone, even if the original interview was conducted in person.
- The reinterview may not perfectly "anchor" respondents in the original interview's reference period. This lack of replication, discussed more below, may generate spurious inconsistencies.

When the reinterview starts, it "anchors" the reference period to the same week as the original interview. However, it does not reinforce this anchoring for every time-dependent question. In particular, the question, *"Have you been doing anything to find work during the last four weeks?"* uses exactly the same wording in both the reinterview as in the original interview and relies on the initial reference period anchor. Because it doesn't explicitly refer to the same four weeks as the

original interview, the respondent might refer to the four weeks preceding the reinterview.

Some data suggest a reference effect might be possible. For example, fewer inconsistent reports of "Unemployed" come from CATI data than CAPI data. CATI reinterviews are conducted within three days of the original interview, versus up to ten days for CAPI. The shorter lag between interview and reinterview would weaken any reference effect in CATI and reduce inconsistency.

But other factors also can explain the lower inconsistency in CATI data.
- Respondents can remember their earlier responses more easily in CATI because the two interviews are closer together.
- Respondents interviewed by CATI are different from those interviewed by CAPI. Their labor force status may be more clear cut and easier to report consistently.
- CATI enforces better adherence to interview procedures, enhancing consistency between the two interviews.

We believe that the reinterview's failure to anchor explicitly to the original interview week plays only a small part in inconsistent responses to this question because:
- The beginning of the reinterview sets the same reference period as the original interview.
- Even if respondents fail to anchor properly, three weeks of the four are correct.
- A Markov latent class (MLC) analysis using only original CPS responses (Biemer and Bushery, 1999) predicts roughly similar inconsistency for "Unemployed" as the reinterview -- 63.6 percent of original interview reports of "Unemployed" would remain "Unemployed" in the reinterview. Because MLC analysis uses only original interviews, no differential reference effect is operating.
- Proxy reporting is more inconsistent than self reporting. A reference effect would be unlikely to vary by proxy versus self reporting.

The CAI instrument used until 1998 did not collect date and time from the reinterview, but the new CASES instrument implemented in 1999 collects the date and time of the reinterview. We will analyze the effect of reinterview-interview lag on inconsistent reporting in these new data.

3.   Results

Table 1 shows 490 inconsistent responses involving Unemployed. Of these, 338 (69.0 percent) switched between Unemployed and Not in Labor Force (NILF). The more detailed labor force categories in Table 2 show that over half of the inconsistent reports of "Unemployed" (57.6 percent) involve switching to or from "NILF Other." Of the 338 inconsistent "Unemployed" ↔ "NILF" reports, 282 (83.4 percent) switch between "Unemployed Looking for Work" and "NILF Other."

The remaining discussion examines characteristics of subject persons, information providers, and the interview associated with consistent reporting of "Unemployed Looking" and inconsistent reporting between "Unemployed Looking" and "NILF Other."

- Do any methodological effects contribute to inconsistent reporting?
- Are any demographic groups over-represented among the inconsistent reports?

3.1 Methodological effects

Proxy reporting is more likely to yield inconsistent classifications between "Unemployed Looking" and "NILF Other." About half (50.4 percent) of the proxy responses were inconsistent, versus only 30.0 percent of self responses (z-statistic = 5.51). Put another way, 69.9 percent of the inconsistent classifications come from proxy response while only 49.4 percent of the consistent classifications come from proxy (z-statistic = 5.51).

The CPS uses four questions to determine whether someone is "Unemployed Looking" (See Section 2.1). Only the first two had enough responses to analyze.

- *Have you been doing anything to find work during the last four weeks? Yes/No.*
- *What are all of the things you have done to find work during the last four weeks?* (active / passive job search)

Inconsistent answers to these two questions account for 84.0 percent of the 282 inconsistent responses between "Unemployed Looking" and "NILF Other."

Of 268 people answering the question, "*Have you been doing anything to find work during the last four weeks?*" in both interviews, 78.4 percent answered inconsistently. This result strongly suggests that uncertainty about whether a person was looking for work generates the great majority of these inconsistent responses.

Because this question is not specifically anchored to the original survey's reference period, it is possible that the reinterview data for this question suffer from a differential reference effect. However, based on the discussion in the Limitations section, we believe that a differential reference effect is not a major factor in inconsistent reporting.

The second question, "*What are all of the things you have done to find work during the last four weeks?*" elicits information about active versus passive job searching. Of the 50 people with inconsistent MLRs asked this question in both interviews, 56.0 percent changed from active to passive, or vice versa.

3.2 Demographic Analysis

We found several demographic groups over-represented in the set of inconsistent responses ("Unemployed Looking" ↔ "NILF Other"), compared with the set of consistent responses ("Unemployed Looking" both times).

- people age 21 and under
- never married people
- females
- Blacks
- children of the reference person
- people with a high school diploma or less education

For example, 45.7 percent of inconsistent responses were for people age 21 and under, while only 27.7 percent of the consistent responses were for people age 21 and under (z-statistic 4.83).

Table 3 compares the proportions of inconsistent and consistent responses for subjects in these demographic groups. Table 3 makes the same comparisons for the information providers (people who talked to the interviewer).

Except for race (Black) and education (high school or lower), the information providers show no over representation among inconsistent responses versus consistent responses. This result suggests that the characteristics of the person who provides the labor force information do not affect the consistency of that information. We hypothesize that the exceptions, race and educational level of the information provider, are correlated with the subject's race and education. This is certainly true for race. We need to conduct more analysis to understand the link between education of the information provider and inconsistent reporting.

We compute the proportions from these demographic groups as:

proportion of inconsistent (or consistent) responses from demographic group = number inconsistent (or consistent) and from demographic group / number inconsistent (or consistent).

Because proxy response is so strongly associated with inconsistent reporting, Table 4 presents the same comparisons by demographic characteristics of the subject as Table 3, but for proxy and self response separately. We see some interaction between response type and the demographic characteristics associated with inconsistent response.

All the demographic characteristics except "high school education" are over represented among inconsistent proxy responses, relative to consistent proxy responses. "Female", "Black", and "high school education" are over represented among inconsistent self responses, relative to consistent self responses.

Females are over represented in the inconsistent responses regardless of type of response. This result suggests that females and their proxy reporters are more likely to provide inconsistent information regarding whether she is looking for work.

While Tables 3 and 4 depict the relationship between demographic characteristics and inconsistent reporting, they don't address more complex questions about the combined relationship between demographics and inconsistent reporting. We performed a logistic regression analysis to model inconsistent reporting as a function of these demographic characteristics.

The logistic regression attempted to use all the dichotomous variables in Table 3 as explanatory variables for the dependent variable, inconsistent response. Using the SAS forward selection option, only Proxy, Female, Age $\leq$ 21, and Black entered the model. The conditional odds ratios for all four variables were significantly greater than 1.0, suggesting they all contribute to inconsistent reporting. The characteristics Never Married, Child of Reference Person, and High School Education or Lower

did not enter the model. All three correlate highly with Age $\leq$ 21, so they add little additional explanation.

3.3   Demographic Analysis of the "Looking Question"

We have seen that the "Looking Question," *Have you been doing anything to find work during the last four weeks?* accounts for the great majority of inconsistent reporting between "Unemployed Looking" and "NILF Other." Table 5 provides a comparison like those in Tables 3 and 4, for inconsistent versus consistent responses to the "Looking Question". Table 5 displays only the demographic groups that the logistic regression analysis found had a significant relationship to inconsistent reporting.

Subjects 21 years old or younger are over represented among inconsistent responses, compared with consistent responses, whether the information is obtained from the subjects themselves or from proxy respondents.

Females and Blacks seem to show an interaction with type of interview. We see no evidence that females contribute more than their share of inconsistent responses when the information is collected by proxy. Conversely, females contribute much more than their share of inconsistent responses, compared with their contribution to consistent responses, when they respond for themselves. Blacks show just the opposite pattern. They are over represented among inconsistent responses for proxy data, but show no significant over representation for data collected using self response.

4.   Conclusions

This research is experimental in nature, but it identifies characteristics associated with inconsistent reporting of "Unemployed" status. This knowledge might guide improvement of the questions and survey procedures to obtain more reliable estimates of unemployment.

Over half the inconsistent "Unemployed" reports shift between "Unemployed Looking" and "NILF Other."
- Proxy reporting is strongly associated with inconsistent labor force classifications.
- Young people, females, and Blacks are more likely to yield inconsistent reports of "Unemployed Looking."

These results lead us to hypothesize
- Subjects with these demographic characteristics are only loosely connected to the labor force. Whether or not they "did anything to find work" might be as much a matter of perception or interpretation as a hard fact.
- Proxy respondents for these subjects are not particularly knowledgeable about the subjects' job search activity. Perhaps they obtain new information from the subjects between the two interviews.

4.1   Future Work - More Analysis

We plan to perform more detailed analyses of existing reinterview data. For example:
*   Expand the logistic modeling to analyze the interactions among the characteristics associated with inconsistent reporting of "Unemployed."
*   Examine how interviews conducted by CATI and CAPI are associated with inconsistent reporting.
*   Examine the other groups of inconsistent reports of "Unemployed" -- the cases which switch between "Unemployed Looking" and "Employed Working." Responses switching between these two MLRs account for about one-sixth of the inconsistent responses.
*   Extend this analysis to the 1998 CPS data. Beginning in 1998 the reinterview can have different respondents from the original interview. This change generates more inconsistency. What is the effect of different proxies and of proxy response in one interview and self response in the other?
*   Examine the relationship between interview-reinterview lag and inconsistent reporting. This analysis could shed some light on the possibility of a differential reference effect in the reinterview. This work will need to wait until enough sample is available from the 1999 data, when the reinterview began to collect date stamps.

4.2   Future Work - Field Research

The analytic work described above can provide preliminary ideas of factors related to inconsistent reporting, but field research is needed to get a good understanding the process of inconsistent reporting. The analytic work sort of "narrows the search area" so the field work can focus on the demographic groups most likely to report "Unemployed" inconsistently.

This research could attempt to determine how the respondents and the subjects perceive the subjects as looking for work. The Bureau of Labor Statistics and the Census Bureau are considering adding probes to the response error reinterview. These probes could increase understanding about why inconsistent reports happen and guide survey improvements.

For example, we might find that proxy respondents discuss the original interview with the subjects and change their reinterview answers based on the new information. If this is happening, perhaps we should make more effort to obtain self response from subjects in the affected demographic groups. Of course such a change would increase costs.

References

Waite, P.J. (1990). "1988 Reinterview Response Error Results for the CPS," unpublished Census Bureau report, July 25, 1990.

Waite, P.J. (1991). "1989 and 1990 Current Population Survey CPS/CATI Reinterview Response Variance Report," unpublished Census Bureau report, August 21, 1991.

Waite, P.J. (1993). "1991 and 1992 CPS CATI Reinterview Response Variance Report," unpublished Census Bureau report, October 1, 1993.

Waite, P.J. (1994). "The 1994 Current Population Survey Response Variance Report," unpublished Census Bureau report, December 5, 1994.

U.S. Census Bureau (1985). *Evaluating Censuses of Population and Housing*, STD-ISP-TR-5, Washington, D.C.: U.S. Government Printing Office, p. 231.

Biemer, P. and Bushery, J (1999). "A Markov Latent Class Analysis of Classification Error in the CPS," Internal U.S. Census Bureau Technical Report.

Table 1.   CPS Interview-by-Reinterview Cross-tabulation (1995-1997)

Reinterview Response	Original Response			
	Total	Employed	Unemployed	NILF
Total	22,429	13,712 *(**61.1**)*	709 *(**3.2**)*	8,008 *(**35.7**)*
Employed	13,524 *(**60.3**)*	13,178 *(96.1)*	80 *(11.3)*	266 *(3.3)*
Unemployed	721 *(**3.2**)*	72 *(0.5)*	470 *(66.3)*	179 *(2.2)*
NILF	8,184 *(**36.5**)*	462 *(3.4)*	159 *(22.4)*	7,563 *(94.4)*

Table 2.   CPS Interview-by-Reinterview Cross-tabulation with Seven Labor Recodes (1995-1997)

Reinterview Response	Original Response							
	Total	Emp At Work	Emp Absent	Unemp Layoff	Unemp Look	NILF Retrd	NILF Disab	NILF Other
Total	22,429	13,013	699	110	599	4,141	887	2,980
Employed At Work	12,887	12,467	166	16	46	53	17	122
Employed - Absent	637	109	436	10	8	9	9	56
Unemp/Layoff	97	17	6	57	9	2	0	6
Unemp/Looking	624	39	10	11	393	17	5	149
NILF/Retired	4,152	77	14	3	8	3,917	44	89
NILF/Disabled	871	14	8	0	2	31	755	60
NILF/Other	3,161	290	59	13	133	111	57	2,498

Table 3. Proportions of Inconsistent (UE Looking ↔ NILF Other) and Consistent (UE Looking ↔ UE Looking) Responses by Demographic Characteristics of Subjects and Information Providers

	Subjects			Information Providers		
Demographic Group	Inconsistent	Consistent	Z-stat	Inconsistent	Consistent	Z-stat
Number responses	282	393	--	282	393	--
Age ≤ 21	0.457	0.277	4.83	0.117	0.097	0.82
Never Married	0.638	0.483	4.06	0.252	0.265	-0.38
Female	0.567	0.463	2.68	0.674	0.659	0.41
Black	0.238	0.150	2.83	0.238	0.153	2.73
Child of Ref. Person	0.479	0.300	4.75	0.113	0.081	1.37
H.S. Grad or lower	0.727	0.608	3.29	0.638	0.532	2.78

Table 4. Proportions of Inconsistent (UE Looking ↔ NILF Other) and Consistent (UE Looking ↔ UE Looking) Responses for Proxy and Self Response by Demographic Characteristics of Subjects

	Proxy Responses			Self Responses		
Demographic Group	Inconsistent	Consistent	Z-stat	Inconsistent	Consistent	Z-stat
Number responses	197	194	--	85	199	--
Age ≤ 21	0.574	0.412	3.25	0.188	0.146	0.85
Never Married	0.731	0.603	2.71	0.424	0.367	0.90
Female	0.472	0.371	2.03	0.788	0.553	4.15
Black	0.223	0.160	1.59	0.271	0.141	2.40
Child of Ref. Person	0.614	0.495	2.38	0.153	0.111	0.93
H.S. Grad or lower	0.761	0.727	0.77	0.647	0.492	2.47

Table 5. Proportions of Inconsistent and Inconsistent Responses to "Looking Question" for Proxy and Self Response by Demographic Characteristics of Subjects

	Proxy Responses			Self Responses		
Demographic Group	Inconsistent	Consistent	Z-stat	Inconsistent	Consistent	Z-stat
Number responses	156	235	--	68	216	--
Age ≤ 21	0.609	0.417	3.79	0.235	0.134	1.79
Female	0.436	0.413	0.45	0.824	0.560	4.61
Black	0.244	0.157	2.08	0.294	0.208	1.39