

Estimating the Error in Labor Force Data Using Markov Latent Class Analysis

Paul P. Biemer

Research Triangle Institute, Research Triangle Park, NC 27709

John M. Bushery

U.S. Census Bureau, Washington, D.C. 20233

Key Words: panel surveys; nonsampling error; unemployment; data quality

1. Introduction

Beginning in 1942, the U.S. Census Bureau has conducted the Current Population Survey (CPS) each month to monitor unemployment and labor force participation. Since the early 1950s the Census Bureau has evaluated the CPS with a reinterview program. This reinterview evaluation aims to replicate the CPS interview process to estimate test-retest reliability. Before 1994 the Census Bureau also conducted a reconciled reinterview which attempted to obtain the most accurate response (see Forsman and Schreiner, 1991).

Several researchers have analyzed CPS reinterview data to estimate classification error (Sinclair and Gastwirth, 1996,1998; Biemer and Forsman, 1992; Chua and Fuller, 1987; Poterba and Summers, 1986, 1995). Some of these analyses use the reconciled reinterview responses as a “gold standard” which can be considered the truth. However, research has shown that the reconciled reinterview data are themselves subject to substantial classification errors and cannot be considered the truth (Sinclair and Gastwirth, 1996,1998; Biemer and Forsman, 1992; Forsman and Schreiner, 1991; Schreiner, 1980).

As an alternative to the “gold standard” assumption, Chua and Fuller (1987) and Fuller and Chua (1985) estimate CPS response probabilities by applying a latent structure model to reconciled reinterview data. For model identifiability, they impose complex restrictions on the response probabilities and assume independent classification errors (ICE) between the interview and reinterview and across the months in sample. However, research suggests that the ICE assumption does not hold for the CPS reinterview (O’Muircheartaigh, 1991; Singh and Rao, 1995), so the error probabilities estimated with their approach may be biased.

Sinclair and Gastwirth (1996 and 1998) apply a latent class modeling approach to the CPS interview-reinterview data, using model restrictions originally proposed by Hui and Walter (1980) for medical diagnostic testing. Unfortunately, model parameter estimation consumes all the available degrees of freedom, so no residual degrees of freedom are available to test model fit.

Shockey (1988) applies latent class analysis to examine rotation group bias in the CPS. He used confirmatory factor analytic methods rather than reinterview data to support his claims. Shockey reports much larger error rates than those reported by other authors, suggesting bias in his model. Like Sinclair and Gastwirth, Shockey’s data set is not adequate to fully test his model assumptions.

Markov latent class (MLC) analysis offers a promising method to estimate the classification error in panel survey data. Wiggins (1973) first proposed Markov latent class models and Poulsen (1982) refined the method. MLC analysis exploits the repeating nature of panel surveys to extract information on classification error from the interview data. It uses a combination of two models:

- a latent Markov chain model representing the month-to-month transitions among the true labor force classifications.
- a classification error model representing the deviations between the true and observed labor classifications.

An important advantage of the MLC method over the Census Bureau's reinterview and these other methods of estimating measurement error is that it requires only the survey panel data. It does not need reinterview or other supplementary data. MLC analysis may be the only way to evaluate measurement error in panel surveys where a reinterview program is not possible. MLC analysis also allows bias estimation, which traditional reinterviews do not. Finally, applied to reinterview data combined with panel data MLC can provide even more information about classification error (van de Pol and Langeheine, 1997).

This paper reports our findings regarding the validity of the MLC modeling approach for estimating labor force classification error in the CPS. Several sources provide software for fitting MLC and other latent class models. We used the software, *lEM* (Vermunt, 1997). Section 2 describes the MLC model in the context of the CPS and applies the MLC methodology, fits several models, and estimates the classification error, using the best MLC model. Section 3 examines the validity of the MLC estimates against four criteria. Finally, Section 4, summarizes our findings and recommends appropriate uses of the MLC method for evaluating labor force classification error.

2. Markov Latent Class Analysis of the CPS

As a condition of identifiability, we need data from people who responded to the CPS in at least three consecutive months. We will fit MLC models to CPS data from January, February, and March of three different years – 1993, 1995, and 1996. The sample sizes for these three data sets are 45,291 (for 1993), 49,347 (for 1995), and 41,751 (for 1996). We excluded from the analysis noninterviews and cases where the whole household changed in one or more of the three months. To facilitate comparisons with reinterview estimates and the cited research on classification error this analysis uses unweighted cell counts. Weighted MLC estimates of classification error showed no important differences from the unweighted estimates.

In 1994, the CPS underwent a major redesign, including restructuring the questions used to determine labor force status (Rothgeb, 1994). We expect these improvements to reduce classification error in the post-1994 CPS, relative to 1993.

2.1 Theoretical Development and Model Specification

Divide the CPS population into L groups (such as age or race groups), with the variable G indicating group membership. For example, $G_i = 1$ if the i th population member is in group 1. Let X_{gi} , Y_{gi} , and Z_{gk} denote the true labor force classifications for the i th person in group $G = g$ (for $g = 1, \dots, L$ and

$i = 1, \dots, n_g$), with X_{gi} defined as:

$$X_{gi} = \begin{cases} 1 & \text{if person } (g,i) \text{ is employed in time period 1} \\ 2 & \text{if person } (g,i) \text{ is unemployed in time period 1} \\ 3 & \text{if person } (g,i) \text{ is not in the labor force in time period 1} \end{cases}$$

and Y_{gi} and Z_{gi} defined similarly for periods 2 and 3. Let $\pi_{x,y,z|g}$ denote $\Pr(X=x, Y=y, Z=z | G=g)$, $\pi_{y|g,x}$ denote $\Pr(Y=y | X=x, G=g)$, and $\pi_{z|g,y,x}$ denote $\Pr(Z=z | Y=y, X=x, G=g)$. Then, the probability that an individual in group g has labor status x , y , and z in periods 1, 2, and 3 is

$$\pi_{x,y,z|g} = \pi_g \pi_{x|g} \pi_{y|g,x} \pi_{z|g,y,x}.$$

Model identifiability requires the Markov assumption (Van de Pol and de Leeuw, 1986): given a person's true status in period 2, his true status in time period 3 does not depend on his period 1 status. This assumption is violated for people like the chronically unemployed. Unfortunately, we cannot explore the validity of this assumption using the observed data in this research. Classification errors distort the observed data to an unknown extent. We assess the overall validity of the MLC estimates using the methods discussed in Section 3 below. We write the Markov assumption

$$\pi_{z|g,x,y} = \pi_{z|g,y}.$$

The outcome variables in our analysis are the monthly CPS labor force classifications. Denote the observed labor force classifications by A_{gi} , B_{gi} , and C_{gi} for periods 1, 2, and 3, respectively. A_{gi} represents the person's *classification* (1, 2, or 3), similar to the definition of X_{gi} for the person's *true labor force status*. Define indicators, B_{gi} , and C_{gi} for periods 2 and 3 analogously.

Denote the response probabilities in each of these classifications as $\pi_{a|g,x} = \Pr(A = a | X = x)$. Define $\pi_{b|g,y}$ and $\pi_{c|g,z}$ similarly.

Finally, assume "local independence," that the classification error is independent across the three months: $\pi_{a,b,c|x,y,z} = \pi_{a|g,x} \pi_{b|g,y} \pi_{c|g,z}$. Meyers (1988) concluded that this assumption "seems a reasonable approximation." Rao and Singh (1995) reached a similar conclusion. In this paper, we assume "local independence" with no further investigation of its validity.

Under these assumptions, the probability for classifying a CPS sample member in cell (g,a,b,c) of the *GABC* table is $\pi_{g,a,b,c} = \sum_{x,y,z} \pi_g \pi_{x|g} \pi_{a|g,x} \pi_{y|x,g} \pi_{b|y,g} \pi_{z|g,y} \pi_{c|g,z}$. (1)

We use maximum likelihood methods to estimate the model parameters from the likelihood function for the *GABC* table, $\Pr(GABC) = k \prod_{g,a,b,c} \pi_{g,a,b,c}^{n_{g,a,b,c}}$ (Van de Pol and de Leeuw 1986).

The simplest MLC model specifies *homogeneous* response probabilities and transition probabilities – that they are the same for all people in the target population. However, an earlier analysis (Biemer, Bushery, and Flanagan, 1997) indicated that these probabilities are not homogeneous. We explored several covariates to account for this heterogeneity, including: gender, education, interview mode, proxy/self-response, and race. A variable indicating proxy/self response best accounted for population heterogeneity. We define this variable, denoted by P :

$$P = \begin{cases} 1 & \text{all three interviews are self-response (Self)} \\ 2 & \text{two of the three interviews are self-response (Mostly Self)} \\ 3 & \text{two of the three interviews are proxy response (Mostly Proxy)} \\ 4 & \text{all three interviews are proxy response (Proxy)} \end{cases}$$

Based on O’Muirchartaigh’s research (1991), we expect the Self group ($P=1$) to have less classification error than the Proxy group ($P=4$). About one-third of the 1993 sample is in the Self group, one-fourth in the Proxy group, with the remainder distributed roughly equally between the Mostly Self and Mostly Proxy groups. For 1995 and 1996, the Proxy group comprises about one-third, rather than one-fourth of the sample.

To fit an MLC model with the single grouping variable, P , the input data are the $4 \times 3 \times 3 \times 3$ cell counts in the cross-classification table of $P \times A \times B \times C$, where A , B , and C are the labor force classifications for January, February, and March, respectively.

We fit five increasingly complex MLC models for each of the three data sets. The simplest assumes the transition probabilities and response probabilities are homogeneous – they do not differ across the groups P . It also assumes these probabilities are stationary – they remain the same for all three months). This model may be written as:

$$\text{Model 1: } \pi_{p,a,b,c} = \sum_{x,y,z} \pi_p \pi_{x|p} \pi_{a|x}^3 \pi_{y|x}^2,$$

obtained from (1) by imposing a constraint on transition probabilities: $\pi_{z|yp} = \pi_{y|xp} = \pi_{y|x}$ and a constraint on response probabilities $\pi_{a|xp} = \pi_{b|yp} = \pi_{c|zp} = \pi_{a|x}$ for all p .

Model 2 relaxes the homogeneity of transition probabilities constraint, allowing the transition probabilities to vary by Self/Proxy Group, P : $\pi_{z|yp} = \pi_{y|xp}$ for $p = 1, \dots, 4$.

Model 3 keeps Model 1’s homogeneity constraint on transition probabilities, but relaxes the stationarity constraint, written as: $\pi_{y|xp} = \pi_{y|x}$ and $\pi_{z|yp} = \pi_{z|y}$ for all p .

Model 4 relaxes both the homogeneity and the stationarity constraints on the transition probabilities, allowing them to vary by group and by month, $\pi_{y|xp} \neq \pi_{z|yp}$. Note that Models 2, 3, and 4 all retain the second constraint of Model 1, equal response probabilities across groups and months.

Model 5 is the most general model we considered. The second constraint of the earlier models becomes: $\pi_{a|xp} = \pi_{b|yp} = \pi_{c|zp}$, for all p . Under these constraints, (1) can be written as

$$\pi_{p,a,b,c} = \sum_{x,y,z} \pi_p \pi_{x|p} \pi_{y|x,p} \pi_{z|p,y} (\pi_{a|p,x})^3.$$

In Model 5 the month-to-month transition probabilities vary independently across the four proxy/self

groups. The response probabilities but may vary across the four proxy/self groups, but must remain the same for January, February, and March.

To assess model fit we used:

- the likelihood ratio chi-square statistic, L^2
- the dissimilarity index (d).

A p-value on L^2 of 0.05 or greater is the usual criterion for adequate model fit. However, due to the large sample sizes in our analysis, we consider a p-value as small as 0.01 to be acceptable. The dissimilarity index (d) is the proportion of observations that would have to change cells for the model to fit perfectly. As rule of thumb, models with $d \leq 0.05$ (i.e., 5 percent model error) are considered to fit the data well.

Model 5 fits best for all three years. It estimates 83 parameters and leaves 24 degrees of freedom to assess fit. For 1993, 1995, and 1996, L^2 is 23.0 (p-value = 0.50), 25.0 (p-value = 0.41), and 39.3 (p-value = 0.03), respectively. In fact, only Model 5 provides an acceptable fit when the p-value criterion is considered. The other models generally had p-values well below 0.01. Model 5 also has the smallest dissimilarity index, d , $\leq .0003$ which indicates a very good fitting model.

Model 5 is also plausible from a response theory perspective. It postulates that classification error varies by self/proxy group, which is consistent with the survey methods literature. We used Model 5 to generate the estimates of labor force classification error.

2.2 *Estimation of Classification Error*

Table 1 shows estimated response probabilities from Model 5 for the Self group, the Proxy group, and all four groups combined. The probability of a correct response is high for those truly employed or Not in the Labor Force, 97 percent or higher. But the true Unemployed display a much lower probability of a correct response. Our discussion will concentrate on these respondents. The probability of a correct response from the truly Unemployed ranges between 72 percent (Proxy group, 1996) and 86 percent (Self group, 1993). As expected, the probability of a correct response from the Self group is significantly higher than from the Proxy group (in 1993 and 1996). Unemployed people tend to be misclassified as Not in the Labor Force roughly twice as often as they are misclassified as Employed, although both error rates are high.

Surprisingly reporting accuracy for the true Unemployed was significantly higher in 1993, before the CPS redesign than after – 82 percent, compared with about 75 percent in 1995 and 1996. Further investigation is needed to understand the causes underlying this result.

3. **Validity of the MLC Methodology**

The goal of this paper is to evaluate the validity of the MLC approach for estimating labor force classification error in the CPS. Do the model estimates of error probabilities, reflect the actual levels of error in the CPS labor force classifications? Because no “gold standard” exists to assess the accuracy of the CPS (Sinclair and Gastwirth 1996, 1998; Biemer and Forsman, 1992; Schreiner

1980), we cannot evaluate the bias of MLC estimates. Instead, we assess the validity of the MLC estimates of CPS classification error using four criteria:

- 1 *Model diagnostics and goodness of fit across years of the CPS.*
- 2 *Agreement with other estimates in the literature.*
- 3 *Agreement with test-retest estimates of the Index of Inconsistency.*
- 4 *Plausibility of patterns of classification error.*

3.1 *Model Diagnostics and Goodness of Fit across Years of the CPS*

A necessary condition for model validity is that the model fits the data adequately. Our discussion of the chi-square statistic and the dissimilarity index suggest that model fit is reasonable.

A common validation technique assesses fit using data independent of the data used for modeling. In this study, fitting the same model to data for the three years is a form of independent model verification. Despite the redesign, the mechanism governing the CPS response process has remained relatively stable. We would not expect the best model to differ appreciably from one year to another. A single model structure that adequately fits the data for all three years would support model validity.

Model 5 fit the best all three years. The 1995 and 1996 parameter estimates showed little evidence of differences. Only the 1993 estimates of the response probabilities of the truly Unemployed displayed a significant difference from the other two years. Finally, although we tested other grouping variables, proxy/self proved to be the best grouping variable for all three years.

3.2 *Agreement with Estimates in the Literature*

Agreement between the MLC estimates of classification probabilities and similar estimates obtained through other methods in the literature would help validate both the MLC and other methods. Disagreement suggests one or more of methods lack validity.

Table 2 compares the MLC estimates of classification probabilities for 1993 with similar estimates by Chua and Fuller (1982), Poterba and Summers (1991), using reinterview data, and the CPS reconciled reinterview. The relative magnitude of the MLC estimates across the labor force categories generally agrees with these estimates. The greatest differences occur for the true unemployed population. The literature yields estimates of response accuracy for this group that are three to seven percentage points higher than the MLC estimates. Correlations between the errors in interview and reinterview may understate the error and cause an upward bias of the response accuracy rates. Alternatively, failure of the Markov assumption may bias the MLC estimates downward. We suspect that both explanations are true to some extent.

Table 3 compares the MLC response probabilities from the 1996 CPS with those obtained from unreconciled reinterview data using the Hui-Walter (H-W) method (Sinclair and Gastwirth 1996, 1998). Although the MLC and H-W models are both forms of latent class models, the model specifications are very different. For example, the H-W method does not require the Markov assumption for model identifiability.

The H-W method requires grouping the population to produce a group×interview×reinterview table, GAA' . Using the notation developed in section 2.1, $\pi_{gaa'}$ is the probability of classifying an individual from group g into cell (a, a') of the table. The interview and reinterview response probabilities, $\pi_{a|x}$ and $\pi_{a'|x}$, do not depend on the group, G . The H-W model also relaxes the interview-reinterview assumption of parallel measures, estimating the response probabilities for the two interviews separately. However, the prevalence rates of employed, unemployed, and NLF, $\pi_{x|g}$, still depend on G and the ICE assumption is still needed to achieve identifiability. The model is written:

$$\pi_{gaa'} = \sum_x \pi_g \pi_{x|g} \pi_{a|x} \pi_{a'|x}.$$

We used the H-W method to estimate CPS response probabilities for the time periods pre-1994, 1995, and 1996 and compared these with the MLC estimates. We found reasonably good agreement for 1995 and 1996, the two years for which the time periods for the reinterview data and the CPS data were closely matched (see **Table 3**). For 1993, we found significant but small differences between MLC estimates and the corresponding H-W estimates. The difference in time periods might explain the differences between the 1993 MLC estimates and the H-W estimates, because the most recent pre-1994 reinterview data available for H-W were collected from 1985 through 1988. All the H-W estimates were computed from a complete year of reinterview data, because the January-March estimates were very unstable.

3.3 *Agreement with Test-Retest Estimates of the Index of Inconsistency*

Our evaluation of the MLC analysis approach includes comparisons between MLC estimates of classification error and estimates derived from interview-reinterview data. The MLC analysis of the CPS longitudinal data provides maximum likelihood estimates of the response probabilities and the true prevalence probabilities. We compute an expected interview×reinterview tabulation from these probabilities. Then we compute the traditional measure of test-retest reliability for CPS labor force data, the index of inconsistency, $I = 1 - R$, (U.S. Census Bureau, 1985).

Table 4 compares the reinterview and MLC estimates of the index for all three time periods. The Unemployed category is of particular interest because of its large error rate. Standard errors are not available for the MLC estimates of the index so formal hypothesis tests are not possible. However, standard errors for the traditional estimates are provided which can be used as rough approximations of the standard errors for the H-W estimates.

The patterns and magnitudes of the MLC and test-retest estimates generally agree well for all three years. However, for the NLF category in 1995 and 1996, the test-retest estimates of I are somewhat larger than the MLC model estimates. Further analysis shows that the traditional estimate's parallel measures assumption fails to hold, biasing the reinterview's estimates of I , which might explain this difference. We were able to test the assumption of parallel measures by comparing the fit of a H-W type model with and without the restriction $\pi_{a|x} = \pi_{a'|x}$. This test rejected the additional for 1995 and 1996 at the 10 percent level of significance. This result suggests that the difference in the NLF estimates for 1995 and 1996 may be partly due to bias in the test-retest estimates of I . The time period difference between the MLC and test-retest estimates of I confound that comparison.

3.4 *Plausibility of Classification Error Patterns*

Finally, the plausibility (or face validity) of the response probability estimates can also support the MLC method's validity. For example, it seems implausible that proxy responses should be more accurate than self-responses. The plausibility of the model estimates supports their face validity. The MLC estimates of misclassification probabilities appear plausible. The estimates across proxy/self groups were consistent with prior expectations of lower error rates for self respondents than for proxy respondents.

4. Summary and Conclusions

The primary goal of this research was to investigate the validity of MLC estimates of CPS labor force classification error. We evaluated the MLC approach against the four validity criteria described in Section 3 and found no evidence from these analyses to question the validity of the MLC approach. We recommend that the MLC method be considered an alternative method to evaluate the accuracy of the CPS labor force estimates. The strong agreement between the MLC and H-W estimates supports the validity of the H-W method as well. We recommend considering both methodologies in future studies of CPS data quality.

Although the MLC approach performed well in our tests, we recommend caution in applying the methodology in other settings. Reinterview data usually are not available for validation. Analysts may be able to apply only criteria (1), (2), and (5) to check model validity. Some panel data may violate the Markov assumption, a key assumption in the MLC approach. Fortunately, failure of Markov assumption appears not to be an important factor in the validity of the CPS estimates of classification error.

REFERENCES

- Biemer, P., Bushery, J., and Flanagan, P. (1997). "An Application of Latent Markov Models to the CPS," Internal U.S. Census Bureau Technical Report.
- Biemer, P. and Forsman, G. (1992). "On the Quality of Reinterview Data with Applications to the Current Population Survey," *Journal of the American Statistical Association*, Vol. 87, No. 420, 915-923.
- Bushery, J. and Kindelberger, K. (1999). "Simulation Examples for MLC Analysis," internal U.S. Census Bureau Memorandum, January 6, 1999, Washington, DC.
- U.S. Census Bureau (1985). *Evaluating Censuses of Population and Housing*, STD-ISP-TR-5, Washington, D.C.: U.S. Government Printing Office.
- Chua, T.C. and Fuller, W.A. (1987). "A Model for Multinomial Response Error Applied to Labor Flows," Vol. 82, No. 397, pp. 46-51.
- Forsman, G. and Schreiner, I. (1991). "The Design and Analysis of Reinterview: An Overview," in P.P. Biemer, et al. (eds.), *Measurement Errors in Surveys*, John Wiley & Sons, NY.
- Fuller, W. and Chua, T.C. (1985). "Gross Change Estimation in the Presence of Response Error," in Proceedings of the Conference on Gross Flows in Labor Force Statistics, Washington, D.C., U.S. Census Bureau and U.S. Bureau of Labor Statistics, pp. 65-77.
- Hui, S.L. and S.D. Walter (1980). "Estimating the Error Rates of Diagnostic Tests," *Biometrics*, 36, 167-171.
- Meyers, B. D. (1988). "Classification-Error Models and Labor-Market Dynamics," *Journal of Business & Economic Statistics*, Vol. 6, No. 3, pp. 385-390.
- O'Muircheartaigh, C. (1991). "Simple Response Variance: Estimation and Determinants," in P. Biemer, et al (eds.), *Measurement Errors in Surveys*, John Wiley & Sons, 551-574.
- Poterba, J. and Summers, L. (1986). "Reporting Errors and Labor Market Dynamics," *Econometrics*, Vol. 54, No. 6, 1319-1338.
- Poterba, J. and Summers, L. (1995). "Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model with Errors in Classification," *The Review of Economics and Statistics*, Vol. 77, pp. 207-216.

- Poulsen, C.S. (1982). *Latent Structure Analysis with Choice Modeling Applications*, doctoral dissertation, Wharton School, University of Pennsylvania.
- Rothgeb, J. (1994). Revisions to the CPS Questionnaire: Effects on Data Quality, U.S. Census Bureau, CPS Overlap Analysis Team Technical Report 2, April 6.
- Schreiner, I. (1980). Reinterview Results from the CPS Independent Reconciliation Experiment (Second Quarter 1978 through Third Quarter 1979), Internal U.S. Census Bureau Report.
- Shockey, J. (1988). "Adjusting for Response Error in Panel Surveys, A Latent Class Approach," *Sociological Methods and Research*, Vol. 17, No. 1, August, pp. 65-92.
- Singh, A.C. and Rao, J.N.K. (1995). "On the Adjustment of Gross Flow Estimates for Classification Error with Application to Data from the Canadian Labour Force Survey," *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 478-488.
- Sinclair, M. and Gastwirth, J. (1996). "On Procedures for Evaluating the Effectiveness of Reinterview Survey Methods: Application to Labor Force Data," *Journal of the American Statistical Association*, Vol. 91, pp. 961-969.
- Sinclair, M. and Gastwirth, J. (1998). "Estimates of the Errors in Classification in the Labour Force Survey and Their Effects on the Reported Unemployment Rate," *Survey Methodology*, Vol. 24, No. 2, pp. 157-169.
- Van de Pol, F. and de Leeuw, J. (1986). "A Latent Markov Model to Correct for Measurement Error," *Sociological Methods & Research*, Vol. 15, Nos. 1-2, pp 118-141.
- Van de Pol, F. and Langeheine, R. (1997). "Separating Change and Measurement Error in Panel Surveys with an Application to Labor Market Data," in L. Lyberg, et al (eds.) *Survey Measurement and Process Quality*, John Wiley & Sons, NY.
- Vermunt, J. (1997). *LEM: A General Program for the Analysis of Categorical Data*, Tilburg, University.
- Wiggins, L.M. (1973). *Panel Analysis, Latent Probability Models for Attitude and Behavior Processing*, Elsevier S.P.C., Amsterdam.

Table 1. Estimated Labor Force Classification Probabilities (and Standard Errors)

Observed	Group	True Employed (se = 0.1)			True Unemployed (se = 0.1)			True NLF (se = 0.2)		
		1993	1995	1996	1993	1995	1996	1993	1995	1996
Employed	Self	98.9	99.9	99.0	4.2	5.7	5.8	1.0	0.7	0.9
	Proxy	98.7	98.6	98.7	6.6	7.8	8.3	1.6	1.5	1.3
	Total	98.8	98.7	98.8	7.1	7.9	8.6	1.4	1.1	1.1
Unemployed	Self	0.3	0.4	0.4	85.9	80.5	79.7	0.5	0.4	0.4
	Proxy	0.3	0.6	0.4	80.1	77.1	71.6	1.0	0.94	1.2
	Total	0.3	0.5	0.4	81.8	76.1	74.4	0.8	0.7	0.9
NLF	Self	0.8	0.6	0.6	9.8	13.8	14.6	98.5	99.0	98.7
	Proxy	0.1	0.8	1.0	13.3	15.1	20.17	97.5	97.6	97.5
	Total	0.9	0.8	0.8	11.1	16.0	17.0	97.8	98.2	98.0

Table 2. MLC versus Published Estimates

Classification		MLC (1993)	Chua & Fuller (1982, month 1)	Poterba & Summers (1981)	CPS Reconciled RI (1977-1982)
True	Observed				
Employed	Employed	98.8	98.7	97.7	98.8
	Unemp	0.3	0.3	0.5	0.2
	NLF	0.9	1.0	1.7	1.0
Unemp	Employed	7.1	3.5	3.8	1.9
	Unemp	81.8	88.3	84.8	88.6
	NLF	11.1	8.2	11.5	9.5
NLF	Employed	1.4	1.6	1.2	0.5
	Unemp	0.8	1.2	0.6	0.3
	NLF	97.8	97.2	98.2	99.2

Table 3. MLC versus H-W Estimates of CPS Response Probabilities for 1996 (and Standard Errors)

Observed Classification	True Employed		True Unemployed		True NLF	
	MLC	H-W	MLC	H-W	MLC	H-W
Employed	98.8 (0.1)	99.6 (0.1)	8.6 (1.0)	4.6 (15.2)	1.1 (0.1)	2.6 (1.5)
Unemployed	0.4 (0.1)	0.4 (0.1)	74.4 (1.4)	67.6 (11.1)	0.9 (0.1)	0.0 (n/a)
NLF	0.8 (0.1)	0.0 (n/a)	17.0 (1.2)	27.9 (5.3)	98.0 (0.1)	97.4 (1.1)

Table 4. Reinterview versus MLC Estimates of the Index of Inconsistency (and Standard Errors)

Year	Method of Estimation	Employed	Unemployed	Not in Labor Force
1993	MLC	6.4	28.0	7.6
	Reinterview	8.2 (0.2)	33.5 (1.2)	10.0 (0.3)
1995	MLC	6.1	36.2	7.2
	Reinterview	6.7 (0.4)	36.3 (2.8)	10.8 (0.6)
1996	MLC	6.0	37.4	7.8
	Reinterview	5.9 (0.4)	36.0 (2.7)	12.0 (0.6)