Survey Methodology, December 1998
Vol. 24, No. 2, pp. 157-169
Statistics Canada

Estimates of the Errors in Classification in the Labour Force Survey and Their Effect on the Reported Unemployment Rate

MICHAEL D. SINCLAIR and JOSEPH L. GASTWIRTH[1]

ABSTRACT

This paper studies response errors in the Current Population Survey of the U.S. Bureau of the Census and assesses their impact on the unemployment rates published by the Bureau of Labour Statistics. The measurement of these error rates is obtained from reinterview data, using an extension of the Hui and Walter (1980) procedure for the evaluation of diagnostic tests. Unlike prior studies which assumed that the reconciled reinterview yields the true status, the method estimates the error rates in both interviews. Using these estimated error rates, we show that the misclassification in the original survey creates a cyclical effect on the reported estimated unemployment rates. In particular, the degree of underestimation increases when true unemployment is high. As there was insufficient data to distinguish between a model assuming that the misclassification rates are the same throughout the business cycle, and one that allows the error rates to differ in periods of low, moderate and high unemployment, our findings should be regarded as preliminary. Nonetheless, they indicated that the relationship between the models used to assess the accuracy of diagnostic tests, and those measuring misclassification rates of survey data, deserves further study.

KEY WORDS: Misclassification errors; Unemployment rates; Diagnostic tests; Reconciliation; Reinterview surveys; Response errors.

1. INTRODUCTION

Several articles, Poterba and Summers (1986 and 1995) and Abowd and Zellner (1985) used the data from the U.S. Bureau of the Census' reinterview program to estimate the misclassification rates of the Current Population Survey (CPS) and assessed their impact on estimates of labour market transition rates. The estimated misclassification rates were based on the assumption, that a particular reinterview method, reconciliation, yields the "truth." Biemer and Forsman (1992), Forsman and Schreiner (1991) and unpublished research of the U.S. Bureau of the Census (1963), have questioned this assumption. The purpose of this paper, is to provide estimates of the misclassification rates, from response errors in all interviews and reinterviews and to explore their impact on the reported unemployment rates. In contrast to the earlier papers that were concerned with gross flow, we emphasize the accuracy of the labour force estimates themselves. Our approach is based on extending the Hui and Walter (1980) paradigm, for estimating error rates of medical diagnostic tests to trinomial classifications. An advantage of this method is that, no single interview needs to be considered as perfect.

Under certain assumptions, Hui and Walter (1980) developed a method for estimating the error rates associated with a new diagnostic screening test, using a confirmatory test with an unknown low error rate. By treating the reinterview as the confirmatory test, and the original survey as the screening test, this methodology can be used to estimate the error rates in the original survey, and the reinterview and the prevalence rates of the trait screened for. The Hui and Walter (1980) method requires two subpopulations with different prevalence rates of the characteristic. While the two tests may have different error rates, the error rates for each test are assumed equal in the two subpopulations. Furthermore, the model (described in more detail in the appendix) assumes that the errors from the two tests conditioned on the subject's true status, are independent.

The Hui and Walter method was developed for dichotomous test outcomes, and was adapted by Sinclair and Gastwirth (1996) to study misclassification of labour force participation rates. Here, we extend the approach to account for three classifications: unemployed, employed and not in the labour force (NLF), and assess the effect of the misclassification on the reported unemployment rates. The basic model is presented in section two. The reinterview program data, to which the model will be fitted, are described in section three. The resulting error rates are given in section four, along with the "adjusted" unemployment rates, which account for the estimated classification errors. In addition, a measure of accuracy, the predictive value, used in the medical screening literature, is applied to the unemployment rate in section four. It shows that the probability an individual classified as unemployed in the CPS is actually unemployed, varies with the true level of unemployment.

2. THE DATA AND THE MODEL

Labour force reinterview data consists of trinomial responses from both the original survey and a subsequent reinterview. This data for a given subpopulation and year,

[1] Michael D. Sinclair, Senior Statistician, Mathematical Policy Research, 101 Morgan Lane, Plainsboro, N.J. 08536, U.S.A.; Joseph L. Gastwirth, Professor of Statistics and Economics, George Washington University, 2201 G Street Rm. 315, Washington, D.C. 20052, U.S.A.

is summarized in a 3×3 table, where the observed frequency counts of persons in the table, is denoted by, n_{ygij}. With this notation:

- y denotes the year;
- g denotes subpopulation membership, $g = 1$ or 2;
- i denotes the subject's classification by the original survey, $i = 1$ for unemployed, $i = 2$ for employed and $i = 3$ for NLF; and
- j denotes the same subject's classification by the reinterview, $j = 1, 2$ and 3.

We denote the true prevalence rate for each labour force status, $i = 1, 2$ and 3, by π_{ygi}, for subpopulation g and year y. Throughout this paper, we will use the term prevalence rate, to refer to the proportion of persons in one of the three labour force categories (*e.g.*, π_{yg1}). Note that the fraction, π_{yg3} of the population in the NLF category equals $(1 - \pi_{yg1} - \pi_{yg2})$, and that the true unemployment rate in year y for subpopulation g, is equal to π_{yg1} divided by $(\pi_{yg1} + \pi_{yg2})$.

Each classification rate, β_{ygrij}, is defined as the probability that the r-th data collection process, $r = 1$ for the original survey, and $r = 2$ for the reinterview, will classify a person in year y from subpopulation g, to be in category i, $i = 1$, 2 and 3 when the true status of the individual is category j. For example, β_{11131} denotes the probability that in the first year ($y = 1$), a person from the first subpopulation ($g = 1$), was classified by the original survey ($r = 1$) as NLF ($i = 3$) when the person's true status is unemployed ($j = 1$). The classification rates can be divided into two groups, corresponding to those associated with a correct classification, and those associated with an erroneous classification. For each y, g and r, the probability that survey method r, classifies a truly unemployed person in year y from subpopulation g correctly as unemployed, is equal to $\beta_{ygr11} = (1 - \beta_{ygr21} - \beta_{ygr31})$. The corresponding probabilities for employed and NLF are respectively, $\beta_{ygr22} = (1 - \beta_{ygr12} - \beta_{ygr32})$, and $\beta_{ygr33} = (1 - \beta_{ygr13} - \beta_{ygr23})$. With conditional independence of the original survey and the reinterview classification rates, the expected observed frequencies, as expressed in terms of the given notation, for each of the nine cells associated with a particular year y and subpopulation g are:

$$E(n_{yg11}) = n_{yg..}(\pi_{yg1}(1-\beta_{yg121}-\beta_{yg131})(1-\beta_{yg221}-\beta_{yg231}) + \pi_{yg2}\beta_{yg112}\beta_{yg212} + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg113}\beta_{yg213})$$

$$E(n_{yg12}) = n_{yg..}(\pi_{yg1}(1-\beta_{yg121}-\beta_{yg131})\beta_{yg221} + \pi_{yg2}\beta_{yg112} * (1-\beta_{yg212}-\beta_{yg232}) + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg113}\beta_{yg223})$$

$$E(n_{yg13}) = n_{yg..}(\pi_{yg1}(1-\beta_{yg121}-\beta_{yg131})\beta_{yg231} + \pi_{yg2}\beta_{yg112}\beta_{yg232} + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg113}(1-\beta_{yg213}-\beta_{yg223}))$$

$$E(n_{yg21}) = n_{yg..}(\pi_{yg1}\beta_{yg121}(1-\beta_{yg221}-\beta_{yg231}) + \pi_{yg2}(1-\beta_{yg112}-\beta_{yg132})\beta_{yg212} + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg123}\beta_{yg213})$$

$$E(n_{yg22}) = n_{yg..}(\pi_{yg1}\beta_{yg121}\beta_{yg221} + \pi_{yg2}(1-\beta_{yg112}-\beta_{yg132}) * (1-\beta_{yg212}-\beta_{yg232}) + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg123}\beta_{yg223})$$

$$E(n_{yg23}) = n_{yg..}(\pi_{yg1}\beta_{yg121}\beta_{yg231} + \pi_{yg2}(1-\beta_{yg112}-\beta_{yg132})\beta_{yg232} + (1-\pi_{yg1}-\pi_{yg2})\beta_{yg123}(1-\beta_{yg213}-\beta_{yg223}))$$

$$E(n_{yg31}) = n_{yg..}(\pi_{yg1}\beta_{yg131}(1-\beta_{yg221}-\beta_{yg231}) + \pi_{yg2}\beta_{yg132}\beta_{yg212} + (1-\pi_{yg1}-\pi_{yg2})(1-\beta_{yg123}-\beta_{yg113})\beta_{yg213})$$

$$E(n_{yg32}) = n_{yg..}(\pi_{yg1}\beta_{yg131}\beta_{yg221} + \pi_{yg2}\beta_{yg132}(1-\beta_{yg212}-\beta_{yg232}) + (1-\pi_{yg1}-\pi_{yg2})(1-\beta_{yg123}-\beta_{yg113})\beta_{yg223})$$

$$E(n_{yg33}) = n_{yg..}(\pi_{yg1}\beta_{yg131}\beta_{yg231} + \pi_{yg2}\beta_{yg132}\beta_{yg232} + (1-\pi_{yg1}-\pi_{yg2})(1-\beta_{yg123}-\beta_{yg113})(1-\beta_{yg213}-\beta_{yg223})),$$

where, the total sample size for year y and subpopulation g is denoted by $n_{yg..}$.

The model has 14 parameters (six error rates for the original survey, $r = 1$, six error rates for the reinterview, $r = 2$, and two unique prevalence rates) for each subpopulation and year. On the other hand, the 3×3 table for a given year and subpopulation has only 8 independent frequencies, or degrees of freedom. As a result, the model is overparameterized and the number of parameters must be reduced for estimation purposes. The Hui and Walter paradigm enables us to accomplish this.

3. APPLICATION OF THE MODEL AND THE CPS REINTERVIEW PROGRAM

The U.S. Bureau of the Census' Current Population Survey Reinterview Program (U.S. Bureau of the Census 1963) is conducted approximately two weeks after the initial survey, to measure response errors, and to evaluate interviewer performance. The sample design for the reinterview, consists of the self-weighting random sample of households (Levy and Lemeshow 1980) among the selected interviewer assignments. The sample size is about 1/18 of the monthly CPS sample of 50,000 to 60,000 household interviews. Two reinterview procedures are conducted. Three-fourths to four-fifths of the sample cases participate in a response-bias study. Here, an initial reinterview is conducted and after this interview is completed, the reinterviewer reconciles disagreements with the respondent, between the original and the initial reinterview responses. Hence, in the response-bias study, up to two reinterview responses may be obtained from each subject; the first unreconciled reinterview response and a reconciled reinterview response. The remaining one-fifth to one-fourth of the sample households receive a reinterview without reconciliation.

In the response bias study, the reinterviewer is instructed not to look at the original survey responses until the initial reinterview is completed. Forsman and Schreiner (1991) and Schreiner (1980) suggested that the reinterviewers may change the initial reinterview responses to match the original response, as they observed that the rate of disagreement between the original responses and the initial reinterview responses were greater in the unreconciled sample. Sinclair (1994) and Sinclair and Gastwirth (1996) showed that these differences were statistically significant. As a result, the reconciliation process creates a correlation between the original and unreconciled reinterview responses, in the reconciled sample. Hence, we decided to limit our analysis to the original and unreconciled reinterview data from the unreconciled study sample. For the purposes of this study, we will assume that in the unreconciled sample, the errors from the original survey and the unreconciled reinterview conditioned on the respondent's true status, are independent.

To apply the Hui and Walter approach, one needs two subpopulations with different prevalence rates. As males and females are known to have different labour force participation rates, we use them. We also need to assume, that the classification error rates are equal in the two subpopulations, males and females, *i.e.*, $\beta_{y1rij} = \beta_{y2rij}$. At this stage, we assume that the classification error rates for the original survey and the unreconciled reinterview, may be different, and that they may differ by year. With this reduction, for the two subpopulations, in a given year, we now have a total of 12 error rate parameters and 4 prevalence rates, yielding 16 parameters. Since two 3×3 tables contain a total of 16 degrees of freedom, estimation is possible. In this paper, we have analyzed the CPS unreconciled reinterview sample data for the period 1981 through 1990. Complete yearly data for 1987 as well as more recent data, were not available from the U.S. Bureau of the Census.

The CPS estimates of the unemployment rate are published regularly by the Bureau of Labour Statistics (BLS) (see Bureau of Labour Statistics 1992). Since the reinterview is a sub-sample of the full CPS sample, the original survey estimates of the unemployment rate from the reinterview sample, will differ from the BLS published results. Data processing procedures are used on the full sample CPS, that are not applied to the reinterview data. For example, the full CPS sample is weighted, based on the sample selection probabilities, and nonresponse adjustment factors are applied to the data. Given these differences, the estimated prevalences from our model, based solely on the reinterview data, are not directly comparable to the BLS reported values. We have used the CPS reinterview data, primarily to estimate the error rates in the original survey. Furthermore, we have treated the unreconciled reinterview data as a simple random sample of the population, for analysis and hypothesis testing purposes, throughout this paper. Using these error rate estimates, we estimate adjusted Bureau of Labour Statistics (BLS) unemployment rates, where the term adjusted, means that the reported values have been modified to account for the misclassification in the survey. The formula for estimating the true unemployment rate as a function of the reported BLS prevalences from the full CPS sample, and the estimated classification error rates as obtained from the unreconciled reinterview data, is given in the appendix.

4. DATA ANALYSIS AND RESULTS

The first step in preparing our final estimates, was to obtain the parameter estimates, for each of nine yearly data tables, using the SAS NLIN procedure with the Gauss-Newton weighted least squares method. As the reinterview procedures remained constant during the period, we decided to test the hypothesis, that each of the error rates remained equal across the years studied, *i.e.*, $\beta_{ygrij} = \beta_{y'grij}$ for all years $y \neq y'$. In conjunction with the basic assumption, that the error rates for males and females are equal, *i.e.*, $\beta_{y1rij} = \beta_{y2rij}$, this implies, $\beta_{ygrij} = \beta_{y'g'rij}$ for all $y \neq y'$ and $g \neq g'$.

From the two sets of results, we conducted a likelihood ratio test under the assumption, that the reinterview sample is a simple random sample of the population, to test the assumption that each of the error rates was the same for all years. The likelihood ratio statistic, $-2 \log \lambda$ with 96 degrees of freedom (144 parameters in the full model less 48 parameters in the reduced model) yielded a value of 84.06 with a p-value of 0.8027. Hence, the data is consistent with the reduced model, enabling us to use the reduced model estimates and to simplify the notation. We will now use β_{rij} to denote β_{ygrij} for all g and y.

The estimated error rates for the original survey and for the unreconciled reinterview, are presented in Tables 1 and 2, respectively, with their estimated standard errors. The estimated reinterview error rates in Table 2, are similar to corresponding error rate estimates for the original survey. This similarity indicates that the U.S. Bureau of the Census unreconciled reinterview serves as an effective replication. The error rate estimates show that the CPS survey procedures are able to classify the employed, and those not in the labour force, quite accurately. On the other hand, these procedures do not perform well for classifying the unemployed, as the proportion of truly unemployed persons who are classified as unemployed, $(1 - \beta_{121} - \beta_{131})$, is only 0.8397.

For comparative purposes we conducted an analysis of the 75% sample reconciled reinterview data, for the same

1981-1990 period, under the assumption that the reconciled responses were error-free. We created a 3 × 3 table for the number of persons classified by the original interview, in each labour force category, by the number of persons classified by the reconciled reinterview, in each labour force category. The data is given in Table 3. The table frequencies report aggregate data, by year and sex, so that the error rates derived from this table, are comparable to our model. Using the column status, as the true status, one computes an estimate of the error rates. For example, the estimate of β_{121}, the probability that an unemployed person will be classified in the original survey as employed, is 332/17,681 = 0.0188. These error rates are presented in Table 1, to illustrate how the estimated error rates from our method, based on the unreconciled data, differ from those relying on the assumption that the reconciled reinterview is perfect.

Table 1 also presents the estimates of the original survey error rates, as obtained by Poterba and Summers (1986), using reinterview data (combined for both sexes) for the first half of 1981. The Poterba and Summers' method uses both the data from the unreconciled and reconciled samples to estimate the error rates. These authors assume that in the reconciled sample, the interviewers use the original survey data provided, to influence the initial reinterview response. As a result, they assume that a reconciled value is only obtained for a portion of persons, that should have had a discrepancy between the original survey and the initial reinterview. When a reconciled value is obtained, Poterba and Summer assume that the reconciled data is error-free. With these assumptions, they use the unreconciled sample to estimate the incidence of the error, and the reconciled data to provide the information on the true labour force status. In summary, both the Poterba and Summers method, and the reconciled reinterview estimates, rely on the reconciled reinterview data being perfect.

Table 4 presents the reported BLS yearly unemployment rates among those in the labour force, for males and females combined, in comparison to the estimated adjusted unemployment rates based on: (1) our error rate estimates, (2) Poterba and Summers (1986) error rates, and (3) error rates assuming the reconciled reinterview is perfect. If the results in Table 4, are sorted by the value of the BLS reported unemployment rate, an apparent trend is observed in the bias in the original CPS estimates. Figure 1 shows that the reported values, tend to overestimate the actual unemployment rate of persons in the labour force in low unemployment years (1989, 1988 and 1990), and to underestimate the unemployment rate in high unemployment years (1982-1983). Furthermore, the bias associated with our method is shifted upward from the two other approaches. All three methods indicate cyclical effect, the smallest of which is obtained when the reconciled reinterview is assumed perfect.

Table 1

Estimated Error Rates in the Original CPS Estimates

Error Rate Parameter	Description		Estimated Value β_{1ij}			Estimated Standard Error
	Classified as	True Status	Our Method	P&S (1986)	Recon. Reint. Perfect	Our Method
β_{121}	Employed	Unemployed	0.0407	0.0378	0.0188	0.01892
β_{131}	NLF	Unemployed	0.1196	0.1146	0.0838	0.01463
β_{112}	Unemployed	Employed	0.0049	0.0054	0.0017	0.00124
β_{132}	NLF	Employed	0.0100	0.0172	0.0098	0.00154
β_{113}	Unemployed	NLF	0.0110	0.0064	0.0034	0.00155
β_{123}	Employed	NLF	0.0205	0.0116	0.0053	0.00247

Table 2

Estimated Error Rates in the Unreconciled Reinterview CPS Estimates

Error Rate Parameter	Description		Estimated Value Our Method β_{2ij}	Estimated Standard Error
	Classified as	True Status		
β_{221}	Employed	Unemployed	0.0333	0.01772
β_{231}	NLF	Unemployed	0.1128	0.01360
β_{212}	Unemployed	Employed	0.0057	0.00135
β_{232}	NLF	Employed	0.0145	0.00160
β_{213}	Unemployed	NLF	0.0157	0.00171
β_{223}	Employed	NLF	0.0248	0.00238

Table 3
Cross-tabulation of the Aggregated 1981-1990 Original/Reconciled Reinterview Responses
75% Reconciled CPS Reinterview Data

Survey Result	Reconciled Reinterview			
Original CPS	Unemployed	Employed	NLF	Total
Unemployed	15,868	372	480	16,720
Employed	332	213,987	744	215,063
NLF	1,481	2,123	138,077	141,681
Total	17,681	215,482	139,301	373,464

Table 4
Implications of the Error Rate Estimates

Year y	BLS Reported Unemployment Rate UE_y^{BLS}	Prob. Unemp. Given Classified Unemp.	Adjusted Estimate of BLS Reported Unemployment Rate AUE_y^{BLS} Our Method	Poterba and Summers (1986)	Reconciled Data (1981-1990) Perfect	Difference in Reported *vs.* Adjusted Our Method	Estimated Standard Error in Difference Our Method
1990	5.44%	.8135	5.27%	5.36%	5.63%	0.17%	.27%
1989	5.20%	.8052	4.99%	5.09%	5.37%	0.21%	.26%
1988	5.43%	.8113	5.25%	5.35%	5.62%	0.18%	.27%
1986	6.89%	.8503	6.97%	7.04%	7.22%	-0.08%	.33%
1985	7.09%	.8531	7.20%	7.27%	7.44%	-0.11%	.34%
1984	7.41%	.8581	7.56%	7.63%	7.79%	-0.15%	.36%
1983	9.47%	.8894	9.99%	10.00%	10.04%	-0.52%	.48%
1982	9.54%	.8902	10.08%	10.09%	10.12%	-0.54%	.49%
1981	7.50%	.8581	7.66%	7.72%	7.88%	-0.16%	.36%

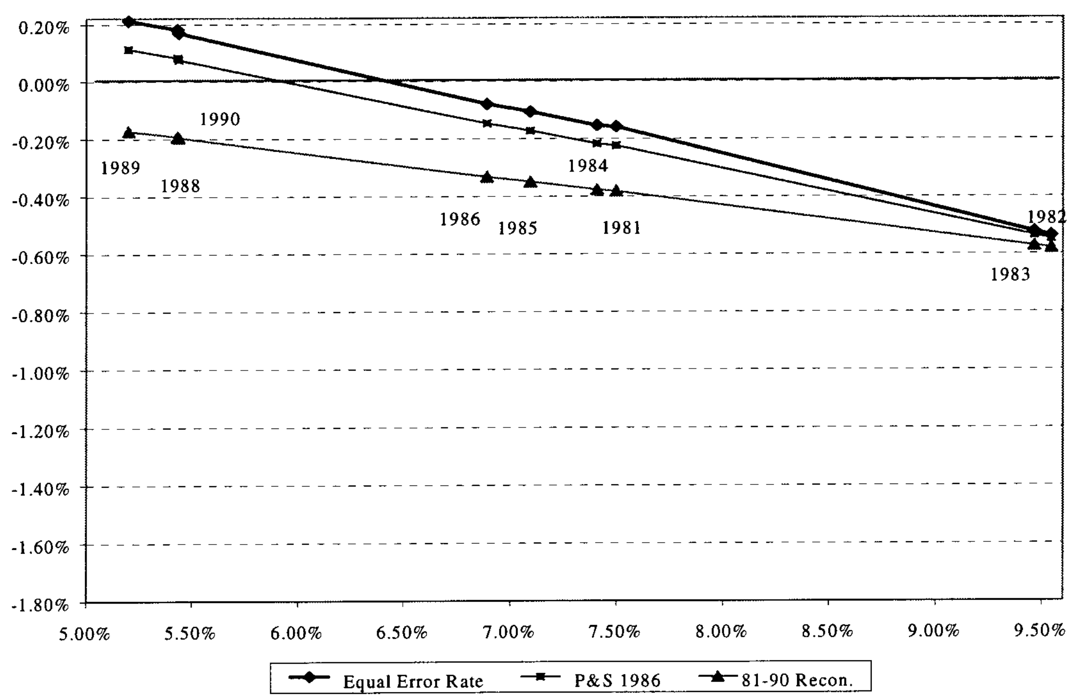

Figure 1. A Comparison of the Bias in the Reported Unemployment Rates as Computed Using Three Methods

In the screening test literature (Gastwirth 1987), the fraction of positive classifications which are correct, called the predictive value of a positive test, is known to vary directly with the prevalence of the characteristic. This is why, quite accurate diagnostic tests can have unacceptably high misclassification rates when populations with a low prevalence of a disease, are screened with them. The analog of this measure in our context, is the proportion of individuals classified as unemployed who are truly unemployed. This proportion is given in the third column of Table 4. Even though the range of reported unemployment rates is fairly narrow, a similar relationship with the unemployment rate can be seen.

While the results of the likelihood ratio test indicated, that the error rates were constant throughout the period, the referees suggested a further analysis to explore this assumption. We divided each of the nine survey years into three groups, according to the year's reported unemployment rate. Survey years, 1990, 1989 and 1988 were classified as having low unemployment, with reported rates from 5.20% to 5.44%. Similarly, survey years 1982 and 1983 were classified as having high unemployment, with reported rates of 9.54% and 9.47%, respectively. The remaining years with rates ranging from 6.89% to 7.5%, were classified as having moderate unemployment rates. With this three group structure, we developed an alternative model that assumed that the error rates were constant within each of the three rate size groups, but allowed each of these groups to have different error rates. The estimated error rates for the original interview are presented in Table 5. The error rates from Table 1, using the equal error rate model, are presented for comparative purposes.

We conducted a likelihood ratio test, to test the assumption that each of the error rates was the same, within each of these three groups, in comparison to the initial nine year model. The likelihood ratio statistic, $-2\log\lambda$ with 72 degrees of freedom (144 parameters in the full model less 72 parameters in the three-group model), yielded a value of 69.25 with a p-value of 0.5697.

In general, the error rate estimates for the three unemployment rate classes, appear to be similar. Because the standard errors of the estimated error rates are quite large, a formal homogeneity test would have insufficient power to detect any variation in an error rate over the three periods.

To assess the sensitivity of the adjusted unemployment rate estimates in Table 4, we recomputed them using the error rates from the three-group model. The results are given in Table 6, which also provides the standard error of the unemployment rate estimates, ranging from a low of about 1.4% to a high of about 2.6%.

Figure 2 presents a graph of the bias in the unemployment using the three group model, and for comparison, the original equal error rate model. The results in Figure 2 are quite interesting. While the cyclical effect is still apparent, the estimated bias is shifted downward and shows a consistent negative bias throughout the business cycle.

Table 5

Error Rates in the Original CPS Data Estimated for Three Unemployment Rate Classes

Error Rate Parameter	Description		Error Rate Estimates							
			Model in Table 1 Assumes Constant Error Rates Across Years		Estimates Using Three Group Model					
	Classified as	True Status			Low Years 1990,1989, & 1988		Moderate Years 1981, 1984-1986		High Years 1982, 1983	
			Est.	STE	Est.	STE	Est.	STE	Est.	STE
β_{121}	Employed	Unemployed	0.0407	0.0189	0.0635	0.1061	0.1113	0.1258	0.0974	0.0717
β_{131}	NLF	Unemployed	0.1196	0.0146	0.1680	0.0538	0.1000	0.0246	0.1084	0.0221
β_{112}	Unemployed	Employed	0.0049	0.0012	0.0000	0.0047	0.0000	0.0098	0.0000	0.0069
β_{132}	NLF	Employed	0.0100	0.0015	0.0080	0.0038	0.0096	0.0025	0.0096	0.0031
β_{113}	Unemployed	NLF	0.0110	0.0015	0.0096	0.0040	0.0109	0.0024	0.0103	0.0029
β_{123}	Employed	NLF	0.0205	0.0025	0.0187	0.0065	0.0202	0.0034	0.0227	0.0044

Table 6
Implications of the Error Rate Estimates Using Three Group Model

Year *y*	BLS Reported Unemployment Rate	Prob Unemp. Given Classified Unemp. Three Group Model	Adjusted Estimate of BLS Reported Unemployment Rate		Difference in Reported *vs.* Adjusted		
			Original Equal Error Rate Model	Three Group Model	Original Equal Error Rate Model	Three Group Model	Estimate Standard Error of the Difference Three Group Method
1990	5.44%	0.9124	5.27%	6.43%	0.17%	-0.99%	1.40%
1989	5.20%	0.9088	4.99%	6.12%	0.21%	-0.93%	1.35%
1988	5.43%	0.9105	5.25%	6.41%	0.18%	-0.98%	1.41%
1986	6.89%	0.9170	6.97%	8.01%	-0.08%	-1.12%	2.35%
1985	7.09%	0.9178	7.20%	8.25%	-0.11%	-1.16%	2.42%
1984	7.41%	0.9199	7.56%	8.64%	-0.15%	-1.23%	2.53%
1983	9.47%	0.9400	9.99%	11.18%	-0.52%	-1.71%	2.05%
1982	9.54%	0.9404	10.08%	11.27%	-0.54%	-1.73%	2.08%
1981	7.50%	0.9191	7.66%	8.74%	-0.16%	-1.24%	2.56%

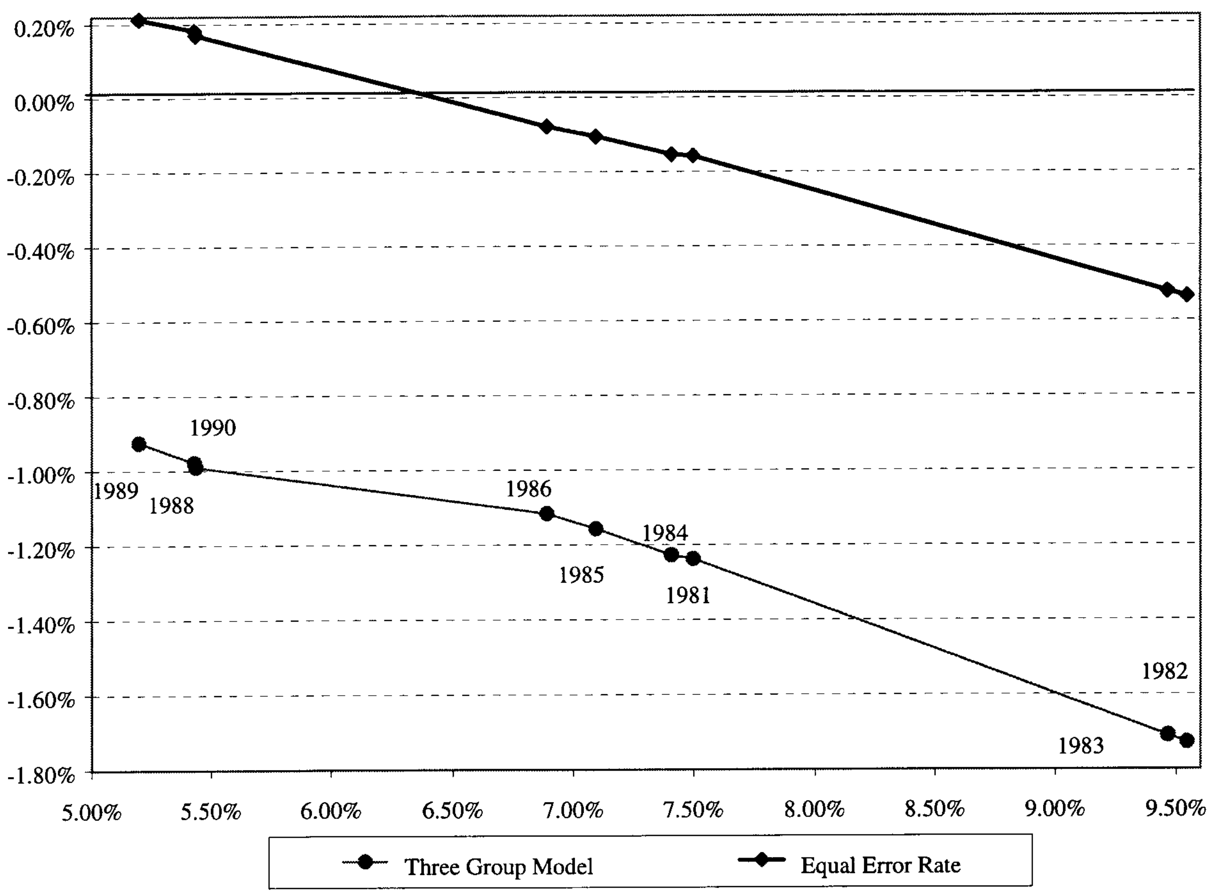

Figure 2. A Comparison of the Bias in the Reported Unemployment Rates as Computed Using the Equal Error Rate Model and the Three Group Model

5. IMPLICATIONS OF THE ADJUSTED ESTIMATES

The results in Figure 1 and 2 show that, all methods for adjusting the unemployment rate for misclassification error, indicate that the degree of bias in the reported rate varies over the business cycle. Given the differences in the estimated bias yielded by the two approaches, it is difficult to determine the magnitude of the bias. Unfortunately, the estimates are sensitive to the model specification, due to the small unreconciled reinterview sample size. This is reflected in the large standard errors of the estimated error rates, and consequently, the estimated bias.

Our approach using the assumption that the error rates remained constant throughout, suggests that bias in the survey estimates is small in years when the unemployment rate is between 5.5% and 7.5%. With this model, the reported unemployment rate appears to be unbiased when the true unemployment rate is around 6.3%, and yields an underestimate when the true rate is above this level, and an overestimate when the true rate is below it. The underestimation bias becomes quite noticeable when unemployment reaches 9%, while the overestimation bias could be meaningful when unemployment is less than 5%.

Using the three-group model results, implies that the reported unemployment rates are underestimates. If the finding is accurate, these results show that the bias in low unemployment years is still about -0.7%, but can be as high as -1.7% in high unemployment years. This contrasts the results obtained from the equal error rate model.

The fact that both the magnitude and direction of the bias in the reported unemployment rate change over the business cycle, may affect the use of that rate in studies of the "natural rate" of unemployment, and the trade-off between inflation and unemployment. Specifically, our results indicate that the range of the true unemployment rate over the business cycle, is larger than the range of the reported rate (see Table 4). Hughes and Perlman (1984) survey the literature on the "natural rate" of unemployment, and the trade-off between inflation and unemployment, as well as the role of search theory in explaining why unemployment is not that low at "full" employment. McKenna (1985) provides a more advanced treatment of job search theory, and its relationship to the duration of unemployment, and the degree to which unemployment is voluntary. Resolving the issue of which model underlies the misclassification error rates in the CPS survey, has important economic implications. If the equal error rate model were correct, in periods of low unemployment, the reported rate would be a slight overestimate. Hence, there would be less true unemployment to explain, by job search and related theories. On the other hand, if the three group model is the correct one, then even at low levels of reported unemployment, there are more persons really unemployed.

6. DISCUSSION

In this paper, we have presented an alternative method for estimating the error rates in the CPS survey. Our study differs from prior work, as we follow the Hui and Walter (1980) approach to estimate the error rates, by assuming that males and females will have the same error rates, and that the errors in the original survey are independent of those in the unreconciled reinterview. While the errors could be slightly correlated, the assumption of independence is standard in data analysis of this type, (see Bailar 1968, Chua and Fuller 1987, and Singh and Rao 1995). A discussion of the bias in the H&W method with dependent errors is given in Vacek (1985). As for the equal error rate assumption, several of the authors cited in this paper (*e.g.*, Poterba and Summers 1986), have noted minor to moderate differences in the error rates between males and females, under the assumption that the reconciled reinterview is perfect. However, this assumption has been questioned. For example, consider the estimate of β_{121}, the probability that an unemployed person, will be classified in the original survey as employed. From Table 3, we estimate this value under the assumption that the reconciled reinterview is unbiased, by dividing n_{21}, divided by $n_{.1}$ (332/17,681 = 0.0188), where n_{ij} is defined previously, with j now corresponding to the classification status in the reconciled reinterview. Using the expected value of these two frequencies from section 2, we can write an expression for the expectation of the estimate in large samples as follows:

$$E(n_{21}/n_{\cdot 1})$$

$$= \frac{\pi_1\beta_{121}(1-\beta_{221}-\beta_{231})+\pi_2(1-\beta_{112}-\beta_{132})\beta_{212}+(1-\pi_1-\pi_2)\beta_{123}\beta_{213}}{\pi_1(1-\beta_{221}-\beta_{231})+\pi_2\beta_{212}+(1-\pi_1-\pi_2)\beta_{213}}$$

$$= \beta_{121} + \beta_{121}\left[\frac{\pi_1(1-\beta_{221}-\beta_{231})}{\pi_1(1-\beta_{221}-\beta_{231}) + \pi_2\beta_{212} + (1-\pi_1-\pi_2)\beta_{213}} - 1\right]$$

$$+ \left[\frac{\pi_2(1-\beta_{112}-\beta_{132})\beta_{212}+(1-\pi_1-\pi_2)\beta_{123}\beta_{213}}{\pi_1(1-\beta_{221}-\beta_{231})+\pi_2\beta_{212}+(1-\pi_1-\pi_2)\beta_{213}}\right]. \quad (1)$$

From (1) it follows that, if the reconciled reinterview error rates, β_{2ij} are equal to zero, that this estimator is unbiased. However, if the reconciled reinterview is not perfect, then the bias in the estimator depends on the prevalence rates in the population studied. As a result, if the actual original survey error rates are in fact equal in the two subpopulations studied, and the reconciled survey classifications are not perfect, the estimated original survey error rates for the two populations will differ. Therefore, one cannot use the similarities or differences in the estimated error rates for males and females from earlier papers, to justify or to contradict the assumptions used here.

We have also conducted a sensitivity analysis of the Hui and Walter (1980) method for dichotomous responses (Sinclair 1994), that indicates that the procedure is sensitive to a violation in the equal error rate assumption, in some circumstances, but the procedure is quite robust in others. Further research is needed to develop reinterview procedures and analytical techniques, to relax the restrictive assumptions currently required in the analysis of the reinterview data.

It should be noted that Chua and Fuller (1987) also obtained estimates of the 3-outcome classification errors in the 1977-1980 CPS 25% sample reinterview data. Analogous to our results, their study found that the largest error rates were associated with classifying the truly unemployed. Poterba and Summers (1995) and Singh and Rao (1995) also found this group to be the hardest to classify. Because all models examined, indicated that the overall misclassification rate of an unemployed individual is around 20%, future reinterviews might focus on understanding why these rates are so high. Hopefully, this will lead to an improved survey.

A potential use of the "adjusted" estimates in Table 4, is in a sensitivity analysis of the literature (*e.g.*, Abowd and Zellner 1985; Poterba and Summers 1995) on gross flows, and labour market dynamics, which assumed that the reconciled interview was perfect. This is equivalent to their adoption of the estimates in the next to the last column of Table 3. Similarly, estimates of the classification errors may be incorporated in procedures, for estimating probit and logit models with misclassified response variables (Hausman and Morton 1994), and in the development of formal statistical procedures for survey data (Rao and Thomas 1991). It should be emphasized, that all the estimates adjusting for misclassification, are still in the research phase, and that the error rates are not yet estimated with sufficient accuracy, to adjust the regular survey data, especially as a new questionnaire and new interviewing procedures were introduced as of January 1994 (Bureau of Labour Statistics 1993).

ACKNOWLEDGEMENTS

The authors wish to thank Irv Schreiner of the U.S. Bureau of the Census for the data used in this research, and helpful discussion. We also wish to thank John Thompson, Henry Woltman and Jon Clark of the U.S. Bureau of the Census, and the referees, and the Associate Editor, for their many helpful comments. This research was supported in part by a grant from the National Science Foundation.

TECHNICAL APPENDIX A

A Review of the Hui and Walter Method

The Hui and Walter method was developed for the evaluation of diagnostic tests. The advantage of the technique is that, it allows the researcher to measure the error rate in a given test, without requiring the comparison test to be error-free. To accomplish this task, the procedure uses two populations (or subpopulations) with different prevalences, to estimate the parameters. The data from such a study, can be summarized in a 2 × 2 table as given in Figure A below. This Table for a specific subpopulation, is indexed by the letter g. We will denote the frequency of cases from subpopulation g, that have a classification from the first test, of status i ($i = 1$ for those having the trait, and $i = 2$ for those not having the trait), and from the second test of status j ($j = 1$ or 2), by n_{ij}. Let π denote the true unknown prevalence rate of the trait, and let α_r and β_r denote the unknown false positive and false negative rates. These error rates are indexed by the letter r, where $r = 1$ corresponds to the outcome from the first test, and $r = 2$ for the second test, (which, in our context, $r = 1$ corresponds to the original interview, and $r = 2$ to a reinterview). The false positive rate, α_r refers to the probability, that the evaluation from the r-th test, will classify the person as positive when in truth the person should have been classified as negative. Similarly, the false negative rate, β_r, is the probability that evaluation from the r-th test will classify the case as negative, when the case has the trait. One (1) minus each of these parameters, reflects to the specificity and sensitivity of the test (or survey) classification procedures, respectively.

Test 1 Outcome (Original Survey)	Test 2 Outcome (Reinterview)		
	Positive	Negative	Total
Positive	Cell 1	Cell 3	$n_{1.}$
Negative	Cell 2	Cell 4	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{..}$

Figure A. Cross-classification of Test 1 and Test 2 Outcomes

Assuming the errors of the first and second tests are independent of each other (given the true state), the expected probabilities, denoted by P_{ij} associated with the cell frequencies given in Figure A, for a given subpopulation g are as follows:

For

$$\text{Cell 1} \quad P_{g11} = \pi_g(1-\beta_{1,g})(1-\beta_{2,g}) + (1-\pi_g)(\alpha_{1,g}\alpha_{2,g})$$

$$\text{Cell 2} \quad P_{g21} = \pi_g(\beta_{1,g})(1-\beta_{2,g}) + (1-\pi_g)(1-\alpha_{1,g})(\alpha_{2,g})$$

$$\text{Cell 3} \quad P_{g12} = \pi_g(1-\beta_{1,g})\,\beta_{2,g} + (1-\pi_g)(\alpha_{1,g})(1-\alpha_{2,g})$$

$$\text{Cell 4} \quad P_{g22} = \pi_g(\beta_{1,g}\,\beta_{2,g}) + (1-\pi_g)(1-\alpha_{1,g})(1-\alpha_{2,g}). \tag{A.1}$$

From (A.1), we observe that we have a total of five parameters, but only three independent cell entries (or degrees of freedom), from which to estimate them. Therefore, the number of parameters must be reduced.

To reduce the parameters, Hui and Walter first, assume that, the proportion of cases with the trait, differs by subpopulation, which implies that, $\pi_1 \neq \pi_2$. Secondly, they require that two subpopulations can be found, such that the error rates for each test are the same for both subpopulations. The error rates associated with the two tests are allowed to differ. For two subpopulations, this implies that in (A.1), $\beta_r = \beta_{r,1} = \beta_{r,2}$, and $\alpha_r = \alpha_{r,1} = \alpha_{r,2}$, with $\beta_1 \neq \beta_2$, and $\alpha_1 \neq \alpha_2$. Under these conditions, the number of parameters reduces to six, (two prevalence rates, one for each subpopulation, and two error rates each for test 1 and test 2). Given that the two 2×2 tables contain six degrees of freedom, estimation is possible. Notice that if $\pi_1 = \pi_2$, and the error rates were the same in both subpopulations, then the probabilities in (A.1) would be the same for both subpopulations, so we would really have one table, and estimation would not be possible. Weighted nonlinear least squares estimates under the Hui and Walter model, can be computed using the Gauss Newton algorithm from the SAS Nonlinear Regression (NLIN) procedure. With this approach, one can express the observed frequencies, n_{ij}, in terms of the total sample size, $n_{..}$, multiplied by the probabilities in expression (A.1). Hui and Walter also present the closed formed estimators given in (A.2), expressed in terms of the observed cell probabilities denoted by p_{gij}.

$$\hat{\alpha}_r = \frac{(p_{r1\cdot}p_{\bar{r}\cdot 1} - p_{r\cdot 1}p_{\bar{r}1\cdot} + p_{211} - p_{111} + D)}{2E_r}$$

$$\hat{\beta}_r = \frac{(p_{r\cdot 2}p_{\bar{r}2\cdot} - p_{r2\cdot}p_{\bar{r}\cdot 2} + p_{122} - p_{222} + D)}{2E_r} \qquad \text{(A.2)}$$

where,

$$\bar{r} = 2 \text{ if } r = 1,\ \bar{r} = 1 \text{ if } r = 2$$

$$p_{g\cdot j} = \sum_{i=1}^{2} p_{gij},\ p_{gi\cdot} = \sum_{j=1}^{2} p_{gij};$$

$$\hat{\pi}_g = \frac{1}{2} + \frac{[p_{g1\cdot}(p_{1\cdot 1} - p_{2\cdot 1}) + p_{g\cdot 1}(p_{1\cdot 1} - p_{21\cdot}) + p_{211} - p_{111}]}{2D}$$

where,

$$D = \pm\ [(p_{11\cdot}p_{21\cdot} - p_{1\cdot 1}p_{111} + p_{111} - p_{211})^2 - 4(p_{11\cdot} - p_{21\cdot})(p_{111}p_{2\cdot 1} - p_{211}p_{1\cdot 1})]^{\frac{1}{2}}$$

with,

$$E_1 = p_{2\cdot 1} - p_{1\cdot 1},\ E_2 = p_{21\cdot} - p_{11\cdot}.$$

Note that two distinct points exist in the solution set, for either a positive or a negative value of D; however, only one of the values will yield reasonable estimates. Variances for the estimators, derived from the estimated asymptotic information matrix, are given in Hui and Walter's (1980) paper.

TECHNICAL APPENDIX B

Adjusting the Reported Unemployment Rates

To evaluate the implications of the estimated error rates, we needed an expression for estimating the actual prevalence rates (the four π parameters), in terms of the estimated error rates and the observed prevalence rates (or sample frequencies), from a given survey. In this section, we present the formula for these computations. With this expression, we can use the BLS reported unemployed and employed prevalence rates, as the observed values to estimate the adjusted BLS prevalence rates. Such an expression is given in (B.1).

Note that in expression (B.1), we have deleted the g-th subscript from the π parameters, so that the expression represents the prevalence rates among the general population, males and females combined. Note that, in this study, we have assumed that the estimated error rates are equal for males and females.

$$\begin{bmatrix} \hat{\pi}_{y1} \\ \hat{\pi}_{y2} \end{bmatrix} = \begin{bmatrix} 1 - \hat{\beta}_{121} - \hat{\beta}_{131} - \hat{\beta}_{113} & \hat{\beta}_{112} - \hat{\beta}_{113} \\ \hat{\beta}_{121} - \hat{\beta}_{123} & 1 - \hat{\beta}_{112} - \hat{\beta}_{132} - \hat{\beta}_{123} \end{bmatrix}^{-1} \begin{bmatrix} \frac{n_{y1.}}{n_{y..}} - \hat{\beta}_{113} \\ \frac{n_{y2.}}{n_{y..}} - \hat{\beta}_{123} \end{bmatrix}. \qquad \text{(B.1)}$$

In this paper, we have three sets of observed values. We have two observed prevalence rates from the reinterview sample (which is a sub-sample of the full CPS sample), including the unreconciled reinterview sample data, and the reconciled reinterview data, from the response-bias study sample, and BLS reported prevalence rates, as observed from the full CPS original survey. We will concentrate our efforts on the first and last of these three sets of statistics, the unreconciled reinterview sample data, and the published BLS estimates. To keep these two sets separate, we will define,

$$U_y^R = \frac{n_{y1.}}{n_{y..}}$$

$$E_y^R = \frac{n_{y2.}}{n_{y..}} \qquad \text{(B.2)}$$

as the observed unemployed and employed prevalence rates, obtained from the CPS unreconciled reinterview sample data. The corresponding BLS reported prevalence rates based on the full CPS original survey weighted data, are defined by U_y^{BLS} and E_y^{BLS}.

Similarly, the observed unemployment rate among those in the labour force, from the unreconciled reinterview sample data, is denoted by UE_y^R, equal to U_y^R divided by $(U_y^R + E_y^R)$, and the observed BLS reported unemployment rate, is defined as UE_y^{BLS}.

Simplifying expression (B.1) in terms of the observed reinterview prevalence rates, U_y^R and E_y^R we find:

$$\hat{\pi}_{y1} = \frac{\left\{U_y^R - \hat{\beta}_{113} - \hat{\beta}_{112}U_y^R + \hat{\beta}_{113}\hat{\beta}_{112} - \hat{\beta}_{113}\hat{\beta}_{132} - \hat{\beta}_{123}U_y^R - \hat{\beta}_{112}E_y^R + \hat{\beta}_{123}\hat{\beta}_{112} + \hat{\beta}_{113}E_y^R\right\}}{\left\{1 - \hat{\beta}_{112} - \hat{\beta}_{132} - \hat{\beta}_{123} - \hat{\beta}_{121}(1 + \hat{\beta}_{132} + \hat{\beta}_{123} + \hat{\beta}_{113}) - \hat{\beta}_{131}(\hat{\beta}_{112} + \hat{\beta}_{132} - 1) - \hat{\beta}_{113}(\hat{\beta}_{112} + \hat{\beta}_{132} - 1) + \hat{\beta}_{123}\hat{\beta}_{112}\right\}}$$

$$\hat{\pi}_{y2} = \frac{\left\{-\hat{\beta}_{121}U_y^R + \hat{\beta}_{121}\hat{\beta}_{113} + \hat{\beta}_{123}U_y^R + E_y^R - \hat{\beta}_{123} - \hat{\beta}_{121}E_y^R + \hat{\beta}_{122}\hat{\beta}_{123} - \hat{\beta}_{131}E_y^R + \hat{\beta}_{131}\hat{\beta}_{123} - \hat{\beta}_{123}E_y^R\right\}}{\left\{1 - \hat{\beta}_{112} - \hat{\beta}_{132} - \hat{\beta}_{123} - \hat{\beta}_{121}(1 + \hat{\beta}_{132} + \hat{\beta}_{123} + \hat{\beta}_{113}) - \hat{\beta}_{131}(\hat{\beta}_{112} + \hat{\beta}_{132} - 1) - \hat{\beta}_{113}(\hat{\beta}_{112} + \hat{\beta}_{132} - 1) + \hat{\beta}_{123}\hat{\beta}_{112}\right\}}. \quad (B.3)$$

Using expression (B.3), we can compute estimates of the adjusted unemployment rate among those in the labour force from the reinterview survey, denoted by AUE_y^R equal to $\hat{\pi}_{yg1}$ divided by $(\hat{\pi}_{yg1} + \hat{\pi}_{yg2})$. Note the AUE_y^R can be expressed as follows:

$$A\hat{U}E_y^R = \frac{\left\{-U_y^R + E_y^R + \hat{\beta}_{112}(U_y^R - \hat{\beta}_{113} + E_y^R) + \hat{\beta}_{132}(U_y^R - \hat{\beta}_{113}) + \hat{\beta}_{123}(U_y^R - \hat{\beta}_{112}) - \hat{\beta}_{113}E_y^R\right\}}{\left\{U_y^R + \hat{\beta}_{113}(1 + \hat{\beta}_{112} - \hat{\beta}_{121} - \hat{\beta}_{123}) + \hat{\beta}_{112}(U_y^R + E_y^R - \hat{\beta}_{113}) + \hat{\beta}_{121}(U_y^R + E_y^R - \hat{\beta}_{123}) - E_y^R + \hat{\beta}_{123} + \hat{\beta}_{131}(U_y^R - \hat{\beta}_{123})\right\}}. \quad (B.4)$$

Finally, to obtain the adjusted estimate of the BLS unemployment rate, denoted by, AUE_y^{BLS}, we substitute the values of U_y^{BLS} for U_y^R and E_y^{BLS} for E_y^R, into expression (B.4). Note that the estimated standard errors of the estimates for AUE_y^{BLS}, presented in section four, were computed using a Taylor series approximation method, (Wolter 1985). As a first step in this process, we assumed the variance in the published estimates of U_y^{BLS} and E_y^{BLS} were negligible. While this is not true, this assumption greatly simplifies the computation of the variances, and captures the majority of the total variation. This assumption is supported by the fact, that the size of the variance of these estimates, given the large full CPS yearly sample sizes is negligible in comparison to the sampling error associated with error rate estimates, which are based on the small unreconciled reinterview sample sizes. In summary, once the substitution of U_y^{BLS} for U_y^R, and E_y^{BLS} for E_y^R into expression (B.4) is completed, we assume that U_y^{BLS} and E_y^{BLS} are fixed known values in this equation. Finally, the sampling variance associated with the difference between the adjusted value and the published value, which defines the bias in the original estimate, is computed from the sum of the variances. Hence, by assuming the published value is sampling variance-free, the sampling variability associated with the difference or bias, is simply equal to the sampling variability associated with the adjusted value.

TECHNICAL APPENDIX C

Estimating Standard Errors of the Adjusted Unemployment Rates

For a complex function of several estimated parameters, the estimates of the variances associated with this function, can be computed using a Taylor series approximation as discussed by Wolter (1985). Suppose that the population parameter of interest is $Y = G(\Theta)$. Where Θ represents a n dimensional vector of population parameters, $\Theta = \{\theta_1, ..., \theta_n\}$. If G possesses continuous second derivatives, in an admissible range for Θ and Θ-hat, then Wolter (1985) presents the relationship:

$$\hat{Y} - Y = A + R(\hat{\Theta}, \Theta)$$

where,

$$A = \sum_{k=1}^{n} \frac{\partial G(\Theta)}{\partial \theta_k}(\hat{\theta}_k - \theta_k)$$

$$R(\hat{\Theta}, \Theta) = \sum_{k=1}^{n}\sum_{i=1}^{n} (1/2!)\frac{\partial^2 G(\Lambda)}{\partial \theta_k \partial \theta_i}(\hat{\theta}_k - \theta_k)(\hat{\theta}_i - \theta_i)$$

$$\hat{\Theta} \le \Lambda \le \Theta. \quad (C.1)$$

The remainder term is often regarded of little consequence, and is eliminated from the relationship. Given the first order approximation, Wolter (1985) presents,

$$\begin{aligned} \mathrm{MSE}(\hat{Y}) &= E[G(\hat{\Theta}) - G(\Theta)]^2 \\ &= \mathrm{Var}(A) \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} \frac{\partial G(\Theta)}{\partial \theta_k}\frac{\partial G(\Theta)}{\partial \theta_i}\mathrm{Cov}(\hat{\theta}_k, \hat{\theta}_i) \\ &= d\, \Sigma_{\hat{\Theta}}\, d^T \end{aligned} \quad (C.2)$$

where d is a row vector of dimension n with the elements,

$$d_k = \left[\frac{\partial G(\ominus)}{\partial \theta_k}\right]. \tag{C.3}$$

Wolter calls this estimator, the first order approximation to the mean square error (equal to the sampling variance + the bias of the estimator squared). Higher order approximations can be developed, by retaining additional terms in the expansion. For purposes of variance estimation, we substitute the estimated covariance matrix for $\sum_\theta$, and evaluate d at the estimated values of θ. Specifically, in our problem, we wish to estimate the variance associated with the function of the estimates in expression (C.4), given below.

$$G(\ominus)=G(\hat\beta_{121},\hat\beta_{131},\hat\beta_{112},\hat\beta_{132},\hat\beta_{113},\hat\beta_{123},U_y^{BLS},E_y^{BLS})=$$

$$\frac{\left\{-U_y^{BLS}+E_y^{BLS}+\hat\beta_{112}(U_y^{BLS}-\hat\beta_{113}+E_y^{BLS}) +\hat\beta_{132}(U_y^{BLS}-\hat\beta_{113})+\hat\beta_{123}(U_y^{BLS}-\hat\beta_{112})-\hat\beta_{113}E_y^{BLS}\right\}}{\left\{U_y^{BLS}+\hat\beta_{113}(1+\hat\beta_{112}-\hat\beta_{121}-\hat\beta_{123})+\hat\beta_{112}(U_y^{BLS}+E_y^{BLS}-\hat\beta_{113}) +\hat\beta_{121}(U_y^{BLS}+E_y^{BLS}-\hat\beta_{123})-E_y^{BLS}+\hat\beta_{123}+\hat\beta_{131}(U_y^{BLS}-\hat\beta_{123})\right\}} \tag{C.4}$$

To create the estimates, we have assumed that the values of U_y^{BLS} and E_y^{BLS} are fixed (*i.e.*, have a negligible sampling variance). Taking the partial derivatives of equation (C.4) with respect to the six error rates, and evaluating these expressions at the estimated values of the error rates, yield a vector d which depends on the values of the error rate estimates and the published BLS unemployed and employed proportions for each year of the study. With our original model, that assumes the error rates are fixed across each year, this d vector for the period of study, only varies from year-to-year for the published values. For illustrative purposes the estimated vector d for 1989 using the BLS published unemployed and employed prevalence rates of .0347 and .6329 is equal to:

$$\hat{d} = \begin{bmatrix} \hat\beta_{121} & .07851 \\ \hat\beta_{131} & .07558 \\ \hat\beta_{112} & -1.2918 \\ \hat\beta_{132} & -.04813 \\ \hat\beta_{113} & -.64214 \\ \hat\beta_{123} & .03884 \end{bmatrix}.$$

The estimated covariance matrix from our SAS NLIN analysis, which, based on the original model that assumes the error rates are fixed by year, and as such, is the same for all years under study, is given below.

$\sum$	β121	β131	β112	β132	β113	β123
β_{121}	0.000358	-4.7E-05	-3.5E-07	-2.6E-08	-3.9E-07	2.9E-07
β_{131}	-4.7E-05	0.000214	-1.7E-07	-5.2E-07	-1.4E-06	-2.8E-07
β_{112}	-3.5E-07	-1.7E-07	1.54E-06	2.14E-07	-2.3E-08	9.9E-10
β_{132}	-2.6E-08	-5.2E-07	2.14E-07	2.37E-06	-1.5E-08	-6.1E-08
β_{113}	-3.9E-07	-1.4E-06	-2.3E-08	-1.5E-08	2.4E-06	-8E-08
β_{123}	2.9E-07	-2.8E-07	9.9E-10	-6.1E-08	-8.0E-08	6.1E-06

Pre and post multiplying the vector d, by the estimated covariance matrix, yields an estimated variance for AUE^{BLS} for 1989 of 6.72 E-6 and a standard error of the estimate equal to .0026 (.26%) as given in Table 4.

REFERENCES

ABOWD, J., and ZELLNER, A. (1985). Estimated gross labor force flows. *Journal of Economic and Business Statistics*, 3, 253-283.

BAILAR, B.A. (1968). Recent research in reinterview procedures. *Journal of the American Statistical Association*, 63, 41-63.

BIEMER, P.P., and FORSMAN, G. (1992). On the quality of reinterview data with application to the current population survey. *Journal of the American Statistical Association*, 87, 915-923.

BUREAU OF LABOR STATISTICS (1993). Overhauling the population survey. *Monthly Labor Review*, 116, 9. Washington DC: U.S. Government Printing Office.

BUREAU OF LABOR STATISTICS (1992). *Employment and Earnings*. 38, 8. Washington DC: U.S. Government Printing Office.

CHUA, T.C., and FULLER, W.A. (1987). A model for multinomial response error applied to labor flows. *Journal of the American Statistical Association*, 82, 46-51.

FORSMAN, G., and SCHREINER, I. (1991). The design and analysis of reinterview: an overview. In *Measurement Errors in Surveys*, (Eds. Paul Biemer *et al.*). New York: John Wiley and Sons.

GASTWIRTH, J.L. (1987). The statistical precision of medical screening procedures: Application to polygraph and AIDs antibodies test data. *Statistical Science*, 2, 213-238.

HAUSMAN, J.A., and MORTON, S. (1994). Misclassification of a Dependent Variables in a Discrete Response Setting. Working paper, Department of Economics, MIT, Cambridge.

HUI, S.L., and WALTER, S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics*, 36, 167-171.

HUGHES, J.J., and PERLMAN, R. (1984). *The Economics of Unemployment*. New York: Cambridge University Press.

LEVY, P., and LEMESHOW, S. (1980). *Sampling for Health Professionals*. California: Lifetime Learning Publications.

McKENNA, C.J. (1985). *Uncertainty and the Labor Market: Recent Developments in Job Search Theory*. New York: St. Martins Press.

POTERBA, J.M., and SUMMERS, L.H. (1986). Reporting errors and labor market dynamics. *Econometrica*, 54, 1319-1338.

POTERBA, J.M., and SUMMERS, L.H. (1995). Unemployment benefits and labor market transitions: a multinominal logit model with errors in classification. *The Review of Economics and Statistics*, 77, 207-216.

RAO, J.N.K., and THOMAS, D.R. (1991). Chi-squared tests with complex survey data subject to misclassification error. In *Measurement Errors in Surveys*, (Eds. Paul Biemer *et al.*). New York: Wiley.

SCHREINER, I., (1980). Reinterview Results From the CPS Independent Reconciliation Experiment (second quarter 1978 through third quarter 1979). Unpublished U.S. Bureau of the Census memorandum, May 7, 1980.

SINCLAIR, M.D. (1994). Evaluating Reinterview Survey Methods for Measuring Response Errors. Phd. Dissertation, George Washington University, September.

SINCLAIR M.D., and GASTWIRTH, J.L. (1996). On procedures for evaluating the effectiveness of reinterview survey methods: application to labor force data. *Journal of the American Statistical Association*, 91, (435) 961-969.

SINGH, A.C., and RAO, J.N.K (1995). On the adjustment of gross flow estimate for classification error with application to data from the Canadian Labour Force Survey. *Journal of the American Statistical Association*, 90, 478-488.

U.S. BUREAU OF THE CENSUS (1963). *The Current Population Survey Reinterview Program: Some Notes and Discussion.* Technical Paper No. 6. Washington, D.C.: U.S. Government Printing Office.

U.S. BUREAU OF THE CENSUS (1985). *Evaluating Censuses of Population and Housing.* Statistical Training Document #ISP-TR-5. Washington, D.C.: U.S. Government Printing Office, 1985.

VACEK, P.M. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics*, 41, 959-968.

WOLTER, K.M. (1985). *Introduction to Variance Estimation.* New York: Springer-Verlag.