Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model with Errors in Classification

Author(s): James M. Poterba and Lawrence H. Summers

Source: *The Review of Economics and Statistics*, May, 1995, Vol. 77, No. 2 (May, 1995), pp. 207-216

Published by: The MIT Press

Stable URL: https://www.jstor.org/stable/2109860

The Review *of* Economics *and* Statistics

VOL. LXXVII | MAY 1995 | NUMBER 2

UNEMPLOYMENT BENEFITS AND LABOR MARKET TRANSITIONS: A MULTINOMIAL LOGIT MODEL WITH ERRORS IN CLASSIFICATION

James M. Poterba and Lawrence H. Summers*

Abstract—This paper utilizes validation data on survey response error in the Current Population Survey to generalize the standard multinomial logit model to allow for spurious events that result from classification error. Our basic approach could be used with other stochastic models of discrete events as well. We illustrate our algorithm by studying the effect of unemployment insurance (UI) on transitions from unemployment to employment and on labor force withdrawal. Our results confirm earlier work suggesting that UI lengthens unemployment spells, and show that correcting for classification error strengthens the apparent effect of UI on spell durations.

RESPONSE errors are a problem in all research utilizing sample surveys. Significant rates of response error have substantial implications for research on labor market dynamics, since they result in spurious transitions between labor market states. A number of previous studies, including Abowd and Zellner (1985), Fuller and Chua (1985), and Poterba and Summers (1986), have analyzed the effect of response error on estimates of the gross flows of individuals between the labor market states of employment, unemployment, and not-in-the-labor force (NILF). This paper addresses a distinct issue: the effect of classification error on microeconometric studies of labor market behavior. We generalize standard discrete-event models of individual labor market transitions to allow for the possibility that some observed transitions are the result of classification error, and apply this model to studying the effect of unemployment insurance on the duration of unemployment.

The paper is divided into four sections. The first, in the spirit of Bound et al. (1990), Duncan and Hill (1985), and Poterba and Summers (1984), documents the incidence of response errors in labor market survey data. We focus on errors in the employment status questions in the Current Population Survey (CPS), and show that misclassification between "unemployed" and "not in the labor force" appears to be a particularly substantial problem. In section II, we develop a probabilistic model for the labor market transitions of unemployed individuals, generalizing the multinomial logit model to allow for the possibility of misclassification. Because of the particular features of the data set we subsequently analyze, our model assumes that an individual's true labor market status is observed in the first of two surveys. The probability that an individual who is known to be unemployed in the first survey is recorded as moving from unemployment to employment is therefore the sum of the probability that he actually finds a job and is correctly classified as unemployed, and the probability that he remains unemployed or leaves the labor force but is misclassified as employed.

Section III presents empirical results illustrating our technique. We estimate our model using data from matched CPS records for May and June 1976 along with information from the May 1976 Job Search Supplement. We focus on the effect of unemployment insurance (UI) benefits on an individual's re-employment probability. We find that UI has a substantial effect in depressing re-employment probabilities and increasing the duration of unemployment spells. There is also a

Received for publication August 2, 1993. Revision accepted for publication August 30, 1994.

* Massachusetts Institute of Technology and National Bureau of Economic Research; and U.S. Department of the Treasury, respectively.

We are grateful to the NSF and NBER for financial support, to Daniel Feenberg, Jonathan Gruber, two anonymous referees, and especially Jerry Hausman for helpful comments, and to Irv Schreiner of the Census Bureau for providing us with unpublished data. Opinions are those of the authors and should not be attributed to any institutions with which they are affiliated.

significant reporting effect. Receipt of unemployment benefits requires that an individual remain in the labor force looking for work; this causes a substantial reduction in the labor force exit rate amongst individuals receiving UI. Allowing for the possibility of classification error does not change the general pattern of the empirical results, but in some cases the magnitude of the estimated effects does change substantially. A brief conclusion interprets our findings and suggests directions for future work.

I. Employment Status Misreporting in the CPS

Reporting errors are a substantial problem in the Current Population Survey. The incidence of errors due to response and coding mistakes is well documented by the Census Bureau's Reinterview Surveys.[1] These surveys reinterview a subsample of the households included in the CPS, typically about a week after the original survey. Reinterview respondents are asked to describe their activities in the preceding week, the same week that the original survey asked about. In the "nonreconciled" component of the Reinterview Survey, there is no attempt to determine which, if either, of the two responses is correct. However, for the "Reconciled" subgroup of the Reinterview Survey, which accounts for about one-third of the reinterviewed households, the second interviewer compares the results on the first survey with the reinterview answers. Then, before leaving the household, he attempts to decide which, if either, of any conflicting responses is correct.[2] The Reinterview responses for those in the reconciled subsample, therefore, are the "truth" as determined by the second interviewer.

The reconciled Reinterview Surveys permit analysis of employment status coding errors. The first panel of table 1, which is labelled Error Probabilities A, shows the fraction of individuals in each employment category in the reconciled subgroup by their category on the first survey. While most of the employed CPS respondents are correctly classified, a substantial fraction of the unemployed from the Reinterview Survey were reported in other categories on the original survey. Ten percent of the truly unemployed were classified as not in the labor force (NILF), and a further 3.6% as employed, on the first survey. The accuracy of responses by those truly out of the labor force was also quite high, with 99.2% correctly classified.

The finding that many unemployed individuals are misclassified is important for studies of unemployment dynamics, because many observed transitions between labor force states may be spurious. For example, if 15% of unemployed individuals are incorrectly classified in a given month, if individuals were never misclassified in two consecutive months, and if there were *no* true transitions, then the expected duration of measured unemployment spells would be 26.7 weeks (1/.15

[1] Graham (1979), Woltman and Schreiner (1979), Census Bureau Technical Report No. 19 (1969), and Biemer and Forsman (1992) provide further details on the Reinterview Surveys.

[2] This fails to detect those individuals who report consistent, but incorrect, responses in both months. Bound and Krueger (1991) present evidence of positive persistence of measurement errors in CPS earnings data.

TABLE 1.—MISCLASSIFICATION PROBABILITIES

True State	Recorded State: Employed	Unemployed	NILF
Error Probabilities A (Calculated Reinterview Error Probabilities)			
Employed	.9905	.0016	.0079
Unemployed	.0356	.8602	.1041
NILF	.0053	.0025	.9923
		$N = 7079$	
Error Probabilities B			
Employed	.950	.040	.010
Unemployed	.070	.720	.210
NILF	.020	.180	.800

Source: Error Probabilities A were computed from a table of "General Labor Force Status in the CPS Reinterview by Labor Force Status in the Original Interview, Both Sexes, Total, After Reconciliation," May 1976, provided from unpublished records at the Census Department. Error Probabilities B were constructed by the authors to illustrate a plausible scenario for measurement errors in unemployment transition studies.

months). In this example, however, the *true* mean spell duration for unemployed workers would be infinite, and all recorded labor market flows would be the result of classification errors.

Error Probabilities A describes the measurement error problem in the Current Population Survey as a whole. It does not reflect the errors associated with individuals who are unemployed in a particular month and then experience transitions. Clark and Summers (1979) and Poterba and Summers (1984) report that many individuals, when monitored for three consecutive months in the CPS, report themselves as experiencing unemployment–labor force withdrawal–unemployment. The U-N-U transitors also tend to report long spell durations at their third interview, suggesting that they perceive themselves as having experienced an ongoing spell of unemployment. While the Reinterview Survey reveals that only one-quarter of one percent of NILF-reported individuals are actually unemployed, this is because many individuals are genuinely not in the labor force and are rather unlikely to be experiencing an unemployment spell.[3] However, conditional upon having been unemployed the month before, the measurement error rates for the NILF category may be large.

The lower panel of table 1, which we label Error Probabilities B, presents our conjecture of plausible measurement error rates conditional upon unemployment in the previous quarter. We double the probabilities of classification error for individuals who are unemployed, and we introduce substantial error probabilities for those reported as NILF. In our estimation results below, we use both error matrices to determine the effect of large error rates on our estimated coefficients.

The substantial error rates in the CPS suggest that empirical investigations based on these and similar data should allow for response errors. This is especially true in studies that involve discrete choices, or that rely on survey questions that ask respondents to describe their activities at some previous time. The problem of response error may become acute when studies are focused on the *difference* in discrete variables reported in two surveys. Some allowance for spurious transitions must also be made in applying duration models to panel data on unemployment, since when some of the hazarded events occur because of measurement errors, the resulting hazard function parameter estimates will be inconsistent.

In a recent study that focuses on issues similar to those in the current paper, but proposes an alternative solution, Hausman and Scott-Morton (1994) show that classification errors will lead to inconsistent estimates of the behavioral parameters in non-linear models of labor market transitions. Even in linear probability models, when classification errors are uncorrelated with individual characteristics they will result in inconsistent estimates of the constant as well as the slope coefficients, although in the linear case the *ratios* of the slope coefficients are estimated correctly. If classification error rates are correlated with individual characteristics, but are not modelled as such, then even algorithms such as those developed in this paper may yield inconsistent parameter estimates.

The assumption that classification error rates are independent of individual characteristics, which is maintained in our algorithm, is unlikely to hold in practice. Yet data limitations make this assumption difficult to test. We were able to obtain separate CPS Reinterview Survey tabulations for men and women, and in table 2 we report separate error probabilities for those groups. Women appear more likely to be categorized as NILF when they are unemployed, and employed women are also more likely than employed men to list themselves as out of the labor force. More men than women who are out of the labor force report themselves as unemployed or employed.

II. A Multinomial Logit Model with Classification Errors

In most panel data sources, observed transitions between consecutive interview dates may arise from four sources. First, the individual may have reported correctly in both surveys and actu-

[3] Flinn and Heckman (1983) argue that the states of unemployment and NILF are well-defined and distinct. They draw evidence from the clear differences in the models explaining the probability of unemployed and NILF individuals becoming employed. However, this evidence is not relevant to understanding whether a large fraction of those who are unemployed drift in and out of the NILF category with little or no change in behavior. Many of the individuals classified as NILF are not causal entrants to the labor force; they may be disabled, retired, or otherwise unable to work. They are conceptually distinct from the unemployed, who are searching for work. A small fraction of NILF respondents, but a *substantial* fraction of NILF respondents who were reported as unemployed last month, are in fact searching for work. These are the miscategorized workers on whom we focus.

TABLE 2.—MISCLASSIFICATION PROBABILITIES: MEN AND WOMEN

	Recorded State		
True State	Employed	Unemployed	NILF
	Calculated Reinterview Error Probabilities (Men)		
Employed	.9922	.0013	.0065
Unemployed	.0474	.8720	.0806
NILF	.0062	.0048	.9890
		$N = 3329$	
	Calculated Reinterview Error Probabilities (Women)		
Employed	.9892	.0019	.0089
Unemployed	.0194	.8442	.1363
NILF	.0049	.0015	.9936
		$N = 3750$	

Source: Computed from a table of "General Labor Force Status in the CPS Reinterview by Labor Force Status in the Original Interview, Both Sexes, Total, After Reconcilation," May 1976, provided from unpublished records at the Census Department.

ally experienced a transition. Second, there may be spurious transitions by individuals who were misclassified on the first survey and correctly classified on the second. Third, individuals may be correctly classified on the first survey, experience no change in actual status, but be misreported on the second survey and therefore be recorded as transiting. Finally, some observed transitions may be due to individuals who were misclassified on *both* surveys, but in different ways. All of these possibilities are recognized in the algorithms that Abowd and Zellner (1985), Fuller and Chua (1985), and Poterba and Summers (1986) propose for adjusting the gross labor market flows.

The special nature of the data set we analyze in the next section enables us to assume that we observe an individual's true labor market status on one of the surveys. If the respondent's first survey status is certain, then all of the observed transitions are either true transitions or the result of survey response error in the second period. We obtained a data set in which all of the individuals are known to have been unemployed in the first survey month, but whose subsequent labor market experience might have been recorded incorrectly. We develop the likelihood function for observed transitions conditional on the assumption of correct first-period observations; extending this approach to the case when neither observation is certain is straightforward.

To construct the likelihood function for the observed labor market transitions of a group of unemployed individuals, we assume that the probability of actual, and possibly unobserved, transitions to employment or out of the labor force are described by a multinomial logit model. For each individual, the probability of each type of transition depends upon a vector of individual characteristics, X_i.

$$P_{UE|X_i} = \text{Prob}(\text{Unemployed at } t, \text{Employed at } t+1|X_i) = \frac{e^{-X_i\beta_1}}{1 + e^{-X_i\beta_1} + e^{-X_i\beta_2}} \tag{1}$$

and

$$P_{UN|X_i} = \text{Prob}(\text{Unemployed at } t, \text{NILF at } t+1|X_i) = \frac{e^{-X_i\beta_2}}{1 + e^{-X_i\beta_1} + e^{X_i\beta_2}}. \tag{2}$$

In addition, we assume that the probability of reporting errors depends upon an individual's true labor market state, but is otherwise independent of his characteristics. As we noted above, this assumption is the result of limited data on CPS response errors for individuals with different characteristics, and if the required data were available, it would be appropriate to model error rates as functions of these characteristics. We denote the error probabilities by

$$q_{ij} = \text{Prob}(\text{Recorded State is } j|\text{True State is } i). \tag{3}$$

There are nine misclassification probabilities in the three state model. However, only six are independent since the probabilities satisfy an adding up condition:

$$q_{ii} = 1 - \sum_{j \neq i} q_{ij} \qquad i = E, U, N. \tag{4}$$

The q_{ii} term is the probability of correct classification for an individual in state i.

The likelihood function for observed outcomes involves products of the true transition probabilities and measurement error rates. For example, the probability that an individual with characteristics X_i is observed transiting from unemployment to employment is

$$\tilde{P}_{UE|X_i} = q_{UE}P_{UU|X_i} + q_{EE}P_{UE|X_i} + q_{NE}P_{UN|X_i}. \tag{5}$$

The notation $\tilde{P}_{ij}$ refers to the probability of *observing* a transition, while P_{ij} is the probability of an *actual* transition. The first term in (5) is the probability that the individual actually becomes employed and is correctly classified in the second survey. The next term is the probability that the individual remained unemployed in the second month but was misclassified as having become employed. The final term is the probability that there was a transition out of the labor force that was actually recorded as a transition to employment. Thus, it involves both a transition and a reporting error.

Using our assumption that actual transition probabilities can be modelled in the standard multinomial logit fashion, the probability of the *observed* transition from unemployment to employment may be written

$$\tilde{P}_{UE|X_i} = \frac{q_{UE}\cdot 1}{1+e^{-X_i\beta_1}+e^{-X_i\beta_2}} + \frac{q_{EE}\cdot e^{-X_i\beta_1}}{1+e^{-X_i\beta_1}+e^{X_i\beta_2}} + \frac{q_{NE}\cdot e^{-X_i\beta_2}}{1+e^{-X_i\beta_1}+e^{-X_i\beta_2}}$$
$$= \frac{q_{UE}+(1-q_{EU}-q_{EN})e^{-X_i\beta_1}+q_{NE}e^{-X_i\beta_2}}{\left(1+e^{-X_i\beta_1}+e^{-X_i\beta_2}\right)}. \tag{6}$$

This and analogous expressions for the other observed probabilities form the basis of our likelihood function.

All of the individuals in our sample were employed in the first sample month. The data sample is ordered so that individuals $1,\dots,N_1$ are observed as unemployed in the second month, $N_1+1,\dots,N_2$ are classified as employed, and $N_2+1,\dots,N$ are out of the labor force. The likelihood function is therefore

$$L(q,\beta_1,\beta_2) = \prod_{i=1}^{N_1}\tilde{P}_{UU|X_i}\prod_{i=N_1+1}^{N_2}\tilde{P}_{UE|X_i}\prod_{i=N_2+1}^{N}\tilde{P}_{UN|X_i}. \tag{7}$$

By substituting the expressions corresponding to (6) for both $\tilde{P}_{UE}$ and $\tilde{P}_{UN}$ into (7), we find

$$L(q,\beta_1,\beta_2) = \prod_{i=1}^{N_1}\left(\frac{(1-q_{UE}-q_{UN})+q_{EU}e^{-X_i\beta_1}+q_{NU}e^{-X_i\beta_2}}{1+e^{-X_i\beta_1}+e^{-X_i\beta_2}}\right)$$
$$\times\prod_{i=N_1+1}^{N_2}\left(\frac{q_{UE}+(1-q_{EU}-q_{EN})e^{-X_i\beta_1}+q_{NE}e^{-X_i\beta_2}}{1+e^{-X_i\beta_1}+e^{-X_i\beta_2}}\right)$$
$$\times\prod_{i=N_2+1}^{N}\left(\frac{q_{UN}+q_{EN}e^{-X_i\beta_1}+(1-q_{NE}-q_{NU})e^{-X_i\beta_2}}{1+e^{-X_i\beta_1}+e^{-X_i\beta_2}}\right). \tag{8}$$

We maximize the logarithm of this likelihood function.

We consider the error rates $[q_{ij}]$ as fixed parameters, which we estimate from the Reinterview Survey data, and we do not estimate them simultaneously with the vector of behavioral parameters (β_1,β_2). We maximize the conditional likelihood function of (β_1,β_2) given the values of $[q_{ij}]$. This procedure yields consistent estimates of (β_1,β_2) under the assumption that the error rates are independent of individual characteristics. However, Hausman and Scott-Morton (1994) note that it yields inconsistent estimates of the standard errors for (β_1,β_2) because this procedure ignores the sampling variability of the estimated error probabilities and the relevant information matrix is not block-diagonal.

III. Data and Estimation

To apply our model to labor market transitions in the presence of measurement errors, we utilize data from a May 1976 study of the job search methods used by unemployed workers. A total of 4,668 persons in the May 1976 CPS were classified as unemployed and asked to complete a special supplementary questionnaire concerning previous work experience and earnings, current job-seeking methods, and employment aspirations. Rosenfeld (1977) describes this supplemental survey in detail. In many households, the form was left to be filled in later or was mailed to the unemployed person after a telephone interview. The nonresponse rate was 31% and resulted in a total sample of 3,238 completed questionnaires.[4]

We assume that all of the Job Search Questionnaire respondents were in fact unemployed, so that there were no reporting errors in the May data. This assumption is supported by the fact that all of the surveyed individuals had been recorded as unemployed in the May 1976 Current Population Survey, and that they had persevered and answered a six page survey about their job seeking activities. Since those who felt the survey did not apply to them were allowed to return it unanswered, those who responded seem very likely to be truly unemployed.

To analyze labor market transitions we combine the Job Search questionnaire with subsequent CPS interview data that document labor market experience. Of the 3,238 Job Search Questionnaire respondents, 1,304 appeared on the CPS match tape which contained the regular CPS questionnaire for both May and June. This reduces our sample size considerably. In addition, some individuals who answered the Job Search Survey did not provide information about their reservation wage or their previous wage. After excluding all individuals with missing data, our final sample contains 908 unemployed men and women.[5]

Two variables are particularly important in our modeling of the transition probability out of unemployment. The first is the unemployment benefit replacement rate. A vast previous literature, surveyed in Atkinson and Micklewright (1991), has explored the effect of unemployment insurance on labor market behavior. In our survey, respondents were asked whether they had received any unemployment insurance benefits during their current spell of unemployment; if they had, they were asked what their weekly benefit was. We use the ratio of this reported UI benefit to previous earnings as our measure of the replacement rate.[6]

The second variable we focus on is the ratio of an individual's reservation wage to his wage at last job, since search theory models of the type surveyed in Lippman and McCall (1976) emphasize this variable as a key determinent of search duration. Our measure of the reservation wage is based on the following pair of questions: (1) "What kind of work were you looking for (in the period April 18 through May 15)?" and (2) "What is the lowest wage or salary you would accept (before deductions) for this type of work?" Individuals who indicated that they were looking for more than one kind of work were asked to specify their reservation wage for the type of job that they preferred. We compute the ratio of this reservation wage to the wage that the individual described as "the usual earnings . . . before deductions" on the "last job at which you worked for two consecutive weeks or more."[7] Since the reservation wage is *chosen by the individual*, it is potentially endogenous. This makes it difficult to

[4] These data were used in Feldstein and Poterba's (1984) study of reservation wages and unemployment insurance, and in Baron and Mellow's (1981) study of unemployment and labor market transitions.

[5] Since one of our explanatory variables is the ratio of the individual's reservation wage to his or her previous wage, we eliminate from the sample all individuals who are classified as new entrants, and therefore have no previous wage, or reentrants, whose previous wage may refer to a much earlier period.

[6] Our UI variable refers to the amount of benefits actually received during the unemployment spell, and not to the benefits to which the individual was entitled under the law. An individual may not receive UI benefits because (1) he is not eligible for benefits, having exhausted benefits or had insufficient previous work experience, or because (2) he has not yet applied for benefits, or because (3) he has applied but has not yet received benefits because of administrative delays in the payment of UI.

[7] Individuals may indicate their usual earnings as a rate per hour, per week, per month or per year. As long as the unit is the same for the reservation wage and the previous wage, the specific choice of unit is irrelevant. When the units are not the same, we convert by assuming 40 hours per week and 4.3 weeks per month. In addition, we define the after tax earnings of the individuals as $(1 - T)^{*}Earnings$, where $T = 0.25$ for everyone in the sample. This has the effect of understating the UI replacement ratio for high income individuals, and overstating it for the low income (and tax rate) individuals.

assign any structural interpretation to the resulting coefficients.

Since data are not available on the amount of supplementary unemployment benefits, welfare, and other forms of nonwage income received by the unemployed, it is not possible to measure their specific effects on transition probabilities. Information is available, however, on whether or not the individual received welfare payments. We included a binary variable, which takes the value of 1 if welfare is received and zero otherwise, in the equations, and regard its coefficients as a weak indication of whether welfare income affects the probability of leaving unemployment.

Several other variables are also included in the logit specification. Indicator variables for the cause of unemployment, whether the individual was a job loser or a job leaver, and for central city residence are also among the explanatory variables. *SMSACEN*, the central city variable, is included to capture the lower search costs and higher probability of finding re-employment attendant with city residence. We also include demographic variables, such as race, marital status, and age–sex subcategory indicators, in all of our models.

Table 3 shows estimates of the basic multinomial transition model for the full sample. There are estimates of the probability of transiting from unemployment to employment, as well as the probability of labor force withdrawal. Corresponding to each transition, three equations are reported. The first, or "No Error" model, assumes that there are no classification errors. This is the standard multinomial logit model. The second model, with "Error Probabilities A," uses Reinterview Survey error probability estimates from the upper panel of table 1 in conjunction with our classification-error augmented likelihood function. The third set of estimates uses the classification error probabilities from the lower panel of table 1; these are labelled "Error Probabilities B."

The coefficient estimates suggest that the receipt of unemployment benefits has an important effect on transition probabilities, and that the correction for measurement error can have substantial effects on the estimated coefficients. We

TABLE 3.—LOGIT TRANSITION MODEL ESTIMATES

	Employment Transition			Labor Market Withdrawal		
	No Error	Error Probabilities		No Error	Error Probabilities	
Variable	Model	A	B	Model	A	B
Constant	−0.678	−0.762	−0.481	+0.075	+0.122	+0.657
	(0.729)	(0.675)	(1.126)	(0.567)	(0.736)	(1.633)
Job Loser	−0.379	−0.413	−0.459	−0.551	−1.483	−4.131
	(0.547)	(0.208)	(0.242)	(0.852)	(0.527)	(2.000)
Job Leaver	−0.189	−0.134	−0.132	−0.603	−0.768	−15.51
	(0.732)	(0.260)	(0.303)	(0.945)	(0.541)	(6.263)
UI Ratio	−1.146	−1.189	−1.428	−1.915	−1.859	−2.761
	(1.268)	(0.513)	(0.606)	(1.304)	(1.213)	(3.350)
RW Ratio	+0.107	+0.147	+0.159	−0.072	−0.017	−0.399
	(0.377)	(0.146)	(0.173)	(0.615)	(0.258)	(0.561)
Welfare	−0.278	−0.435	−0.622	+1.381	+2.741	+1.950
(1 = Recipient)	(0.917)	(0.353)	(0.440)	(0.958)	(0.739)	(15.940)
SMSACEN	−0.254	−0.335	−0.425	+0.139	+0.178	+1.451
	(0.398)	(0.187)	(0.212)	(0.730)	(0.392)	(1.124)
Race (1 = Non-	−0.019	+0.034	+0.046	−0.588	−1.895	−6.577
white)	(0.959)	(0.247)	(0.291)	(0.989)	(0.952)	(2.705)
Single Man	+0.609	+0.594	+0.460	−1.333	−14.419	−62.302
	(0.655)	(0.661)	(1.121)	(0.930)	(39.18)	(31.818)
Married Man	+0.471	+0.425	+0.176	−2.360	−12.722	−40.941
	(0.616)	(0.649)	(1.104)	(0.944)	(39.281)	(18.568)
Married	−0.103	−0.259	−0.435	+0.969	+1.391	26.903
Woman	(0.753)	(0.303)	(0.356)	(0.780)	(0.563)	(14.891)
Log Likelihood	−758.78	−758.47	−787.13	−758.78	−758.47	−787.13

Notes: Standard errors are shown in parentheses. Sample size is $N = 908$, of which 234 display transitions to employment, and 113 transit to NILF.

begin by describing the estimates for the model without classification error, and then discuss the changes that occur when we allow for such errors.

The estimates in the first and fourth columns of table 3 correspond to the standard multinomial logit model. The U.I. variable has the predicted negative sign in the employment transition model, and it has an even larger, negative effect on the probability of leaving the labor force. This is presumably due to the "reporting effect," the requirement of on-going search as a precondition for UI receipt. The welfare variable takes its predicted sign in each equation, though it is statistically significant in only about one-half of the estimated equations. The demographic variables also have their predicted signs: married women are less likely to find jobs, and more likely to leave the labor force, than are married men. Job losers and job leavers both have lower probabilities of becoming reemployed than do other groups.

The reservation wage variable does not exhibit any statistically significant effects on labor market transition probabilities. We expected that a higher reservation wage-to-last-wage ratio would be linked to fewer acceptable job offers and therefore to longer expected spell durations. While the reservation wage coefficients in the employment transition equation are always positive, they are statistically insignificant. One argument often made in defense of search models is that they do not apply to individuals on temporary layoff, and therefore may appear inconsistent with the data findings. We tested this view by deleting the 76 temporary-layoff individuals in our sample and re-estimating the model, with results very similar to those for the full sample.

Comparing the parameter estimates in the various columns demonstrates the potential importance of allowing for classification errors in modelling labor market transitions. Many of the coefficients change by more than 30% when the model is estimated using the error probabilities estimated from the Reinterview Survey. When estimated with higher error probabilities, the findings show even more substantial changes.[8] For example, the UI coefficient in the P_{UE} equation rises by 25% between the no-error and the Error Probabilities B models. The coefficient in the P_{UN} equation changes from -1.44 to -2.08, a move of 45%.

The parameters of the estimated logit models may be used to determine the substantive importance of the changes in transition probabilities caused by unemployment insurance. In the standard logit model without classification error, where probabilities take the form

$$P_{UE} = \frac{e^{-X_i\beta_1}}{1 + e^{-X_i\beta_1} + e^{-X_i\beta_2}}, \tag{9}$$

the derivative of the transition probability with respect to X_j is

$$\frac{\partial P_{UE}}{\partial X_j} = -\beta_{1j} P_{UE}\left[1 - P_{UE} - \frac{\beta_{2j}}{\beta_{1j}} P_{UN}\right]. \tag{10}$$

In the case of classification error, expression (10) would still apply, although the relevant parameter estimates would presumably be different. This would be the appropriate derivative to consider if one were interested in evaluating the effect of an explanatory variable on the *true* transition probability. If one is interested instead in how the probability of *observing* a particular outcome, such as *observing* a decline in the measured unemployment rate, responds to an explanatory variable, the corresponding probability derivative is

$$\begin{aligned}\frac{\partial \tilde{P}_{UE}}{\partial X_j} &= \frac{-(1 - q_{EU} - q_{EN})\beta_{1j}e^{-X_i\beta_1} - q_{NE}\beta_{2j}e^{-X_i\beta_2}}{1 + e^{-X_i\beta_1} + e^{-X_i\beta_2}} \\ &\quad + \frac{\left(\beta_{1j}e^{-X_i\beta_1} + \beta_{2j}e^{-X_i\beta_2}\right)}{\left(1 + e^{-X_i\beta_1} + e^{-X_i\beta_2}\right)} \cdot \frac{\left(q_{UE} + (1 - q_{EU} - q_{EN})e^{-X_i\beta_1} + q_{NE}e^{-X_i\beta_2}\right)}{\left(1 + e^{-X_i\beta_1} + e^{X_i\beta_2}\right)} \\ &= -\beta_{1j}P_{UE} \cdot \left[(1 - q_{EU} - q_{EN}) + \frac{\beta_{2j}}{\beta_{1j}} q_{NE} \frac{P_{UN}}{P_{UE}} - \tilde{P}_{UE}\left(1 - \frac{\beta_{2j}}{\beta_{1j}} \cdot \frac{P_{UN}}{P_{UE}}\right)\right].\end{aligned} \tag{11}$$

[8] In some cases the coefficients seem to change by large and implausible amounts. However, these parameter estimates are usually accompanied by large standard errors.

TABLE 4.—UNEMPLOYMENT SPELL DURATIONS AND UI

Concept	Without Reporting Error Correction	With Reporting Error Correction
Probability of Becoming Employed		
Actual (P_{UE})	0.257	0.233
Observed ($\tilde{P}_{UE}$)	0.257	0.257
Probability of Labor Force Withdrawal		
Actual (P_{UN})	0.124	0.048
Observed ($\tilde{P}_{UN}$)	0.124	0.124
Expected Actual Spell Duration ($1/(P_{UN} + P_{UE})$)	2.62 months	3.55 months
Expected Duration of Observed Spells ($1/(\tilde{P}_{UN} + \tilde{P}_{UE})$)	2.62 months	2.62 months
"Indomitable Worker" Expected Spell Duration ($1/P_{UE}$)	3.89 months	4.29 months
$\partial P_{UE}/\partial$ (UI replacement ratio)	−0.157	−0.191
$\partial \tilde{P}_{UE}/\partial$ (UI replacement ratio)	−0.157	−0.215
$\partial P_{UN}/\partial$ (UI replacement rate)	−0.171	−0.052
$\partial \tilde{P}_{UN}/\partial$ (UI replacement rate)	−0.171	−0.109
∂ (Expected Duration) $/\partial$ (UI Replacement ratio)	2.25	3.06
Expected Duration (UI = 0)	2.06 months	2.78 months
Expected Duration (UI = 0.5)	3.18 months	4.32 months

Source: Authors' calculations based on the "No Error" and "Error Probabilities A" equations in table 3.

The derivatives of probabilities with respect to changes in the level of UI benefits are shown in table 4. The calculations correspond to an "average" individual, someone with the sample average transition probabilities of $P_{UE} = 0.233$ and $P_{UN} = 0.048$.[9] A change of 0.50 in the unemployment insurance replacement ratio, from 0 to 0.50, results in a 0.10 decline in the unemployed worker's probability of becoming employed in a given period according to the estimates from the logit model with "Error Probabilities A." The expected duration of *actual* unemployment spells if there were no unemployment benefits, $1/(P_{UE} + P_{UN})$, is estimated to be 2.78 months with "Error Probabilities A," compared to 2.06 months without classification error correction.

The introduction of UI which provides benefits equal to 50% of the worker's post tax wage increases the expected unemployment spell duration from 2.78 to 4.32 months when the experiment is evaluated using the "Error Probabilities A" estimates, compared with a change from 2.06 to 3.18 months when the logit models are not corrected for classification error. The sensitivity of spell duration to UI benefits is substantially greater when the transition probabilities are corrected for classification error, in part because this correction reduces the level of the estimated unemployment escape rate. This finding parallels the literature on gross labor market flows in suggesting that classification errors exaggerate the apparent dynamism of the labor market.

IV. Conclusions

This paper presents an algorithm for analyzing discrete transitions in panel data when these transitions may be the result of either actual transitions or classification errors. We rely on the availability of information from a validation survey, the Current Population Survey Reinterview Survey, to estimate the incidence of re-

[9] The true transition probabilities for this "average" individual are calculated by solving the following system of linear equations for P_{UE}, P_{UU}, and P_{UN}:

$$\tilde{P}_{UE} = .257 = q_{EE}P_{UE} + q_{UE}P_{UU} + q_{NE}P_{UN}$$
$$\tilde{P}_{UU} = .619 = q_{EU}P_{UE} + q_{UU}P_{UU} + q_{NU}P_{UN}$$
$$\tilde{P}_{UN} = .124 = q_{EN}P_{UE} + q_{UN}P_{UU} + q_{NN}P_{UN}.$$

We use the values of the $[q_{ij}]$ from Error Probabilities A for this calculation. The values of observed transition rates shown above are sample average values.

sponse error. We then modify the standard multinomial logit model for discrete transitions to allow for the possibility of classification errors. Our procedure is of more general applicability, however, and could be used in conjunction with other models of discrete events.

Our method could also be applied to other surveys for which information on response error rates is available. Data on survey error rates in the decennial U.S. Census are collected in the Post-Enumeration Survey, and some validation survey data are also available for the Survey of Income and Program Participation. The evidence reported here and elsewhere on the incidence of survey response errors underscores the value of collecting such information for surveys where it is not usually available; Duncan and Hill (1985) provide an example of such a project. It might also be possible, under strong assumptions about the nature of response errors in different surveys, to use information on classification error rates from surveys with validation procedures to adjust estimated behavioral effects in surveys without such procedures.

The algorithm developed here is only one of several that have been proposed to account for measurement error in survey data (Fuller (1987) provides a survey). For example, Krueger and Summers (1988) describe methods for treating measurement error in indicator variable data, again in the context of analyzing labor market data. The algorithm that is most closely related to the one developed here is Hausman and Scott-Morton's (1994) non-parametric method for estimating the incidence of classification errors in probit models. Their approach does not require information on survey responses from a validation survey, so it could be applied in some cases where the present algorithm would be infeasible. It is an important complement to the direct collection of information on classification error rates in validation surveys. Future work should concentrate on applying these algorithms to various data sets and economic problems, to provide better information on the substantive importance of classification error problems.

REFERENCES

Abowd, John, and Arnold Zellner, "Estimating Gross Labor Force Flows," *Journal of Economic and Business Statistics* 3 (1985), 253–283.

Atkinson, A., and J. Micklewright, "Unemployment Compensation and Labor Market Transitions: A Critical Review," *Journal of Economic Literature* 29 (Dec. 1991), 1679–1727.

Baron, J. M., and W. Mellow, "Changes in Labor Force Status among the Unemployed," *Journal of Human Resources* 16 (1981), 427–442.

Biemer, Paul P., and Gosta Forsman, "On the Quality of Reinterview Data with Application to the Current Population Survey," *Journal of the American Statistical Association* 87 (1992), 915–923.

Bound, John, Charles Brown, Greg Duncan, and Willard Rodgers, "Measurement Error in Cross-Sectional and Longitudinal Labor Market Surveys: Validation Study Evidence," in J. Hartog, G. Ridder, and J. Theeuwes (eds), *Panel Data and Labor Market Studies* (Amsterdam: Elsevier Science Publishers, 1990), 1–19.

Bound, John, and Alan B. Krueger, "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," *Journal of Labor Economics* 9 (Jan. 1991), 1–24.

Bureau of the Census, *The Current Population Survey Reinterview Program: January 1961 through December 1966*, Technical Paper No. 19 (Washington, D.C.: Government Printing Office, 1969).

Clark, Kim B., and Lawrence H. Summers, "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity* 1 (1979), 13–60.

Duncan, Greg, and Daniel Hill, "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data," *Journal of Labor Economics* 3 (1985), 508–532.

Feldstein, Martin S., and James M. Poterba, "Unemployment Insurance and Reservation Wages," *Journal of Public Economics* 23 (1984), 141–169.

Flinn, C. J., and J. J. Heckman, "Are Unemployment and Out of the Labor Force Behaviorally Distinct States," *Journal of Labor Economics* 1 (Jan. 1983), 28–42.

Fuller, Wayne, *Measurement Error Models* (New York: John Wiley & Sons, 1987).

Fuller, Wayne, and T. C. Chua, "Gross Change Estimation in the Presence of Response Error," in U.S. Bureau of Labor Statistics, *Proceedings of the Conference on Gross Flows in Labor Force Statistics* (Washington, D.C.: Government Printing Office, 1985).

Graham, Dorcas W., "Estimation, Interpretation, and Use of Response Error Measurements," unpublished paper, U.S. Commerce Department, Washington, D.C., 1974.

Hausman, Jerry A., and Fiona M. Scott-Morton, "Misclassification of a Dependent Variable in a Discrete Response Setting," MIT Department of Economics Working Paper 94-19, 1994.

Krueger, Alan B., and Lawrence H. Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica* 56 (Mar. 1988), 259–294.

Lippman, S., and J. McCall, "The Economics of Job Search," *Economic Inquiry* 14 (1976), 155–189.

Poterba, James M., and Lawrence H. Summers, "Response Variation in the CPS: Caveats for Unemployment Analysts," *Monthly Labor Review* (Mar. 1984), 37–43.

______, "Reporting Errors and Labor Market Dynamics," *Econometrica* 54 (Nov. 1986), 1319–1338.

Rosenfeld, Carl, "Job Search of the Unemployed, May 1976," *Monthly Labor Review* 100 (Nov. 1977), 39–47.

Woltman, Henry, and Irv Shreiner, "Possible Effects of Response Variance on the Gross Changes Data," unpublished Bureau of Census memo, May 11, 1979.