Some Markov-Chain Models For Nonresponse In Categorical Data From Panel Surveys

Elizabeth A. Stasny, Ohio State University

I. Introduction

Many large-scale sample surveys, such as the Current Population Survey and the National Crime Survey, use rotating panel designs under which individuals are interviewed several times before rotating out of the sample. Typically, these large-scale surveys are used to produce point-in-time estimates. The rotating panel structure of the survey is a result of the need to reduce costs by keeping the same interviewers and subjects for more than a single interview. Recently, however, there has been increasing interest in using the longitudinal data bases available from such surveys to estimate gross change over time.

In this paper, I consider the problem of estimating period-to-period gross change over time using categorical data from a panel survey where, as one would expect in a sample survey, there is nonresponse in the data so that some of the surveyed individuals are completely cross-classified while others are partially cross-classified or completely missing. In particular, I consider some Markov-chain models for the gross flow process along with Markov-chain models for the process generating the nonresponse. The models presented here are an improvement over previous models considered for estimating gross flows in the presence of non-random nonresponse (see for example Stasny (1983, 1985, and 1986) and Stasny and Fienberg(1985)) because they allow a person to be missing at both of two interview periods.

Section II of this paper presents a two-stage model for the panel data with nonresponse. Section III describes the Markov-chain models for nonresponse. In Section IV, I fit the models to employment data from the Canadian Labour Force Survey. Section V gives extensions of the models.

II. A Two-Stage Model for Panel Data

One possible approach to the problem of estimating gross change over time using panel data is to use only the information from individuals who are respondents in both of two consecutive interview periods. In order to use this approach, we must assume that individuals who do not respond in one or both periods are a random sample of all individuals (Rubin, 1976). However, in most cases, we do not believe that nonresponse occurs at random. For example, using data from the Canadian Labour Force Survey, Paul and Lawes (1982) and Fienberg and Stasny (1983) give evidence that nonresponse is related to labor force classification. Since there is evidence that nonresponse does not occur at random, we would like to consider some models for estimating gross flows that allow us to treat nonresponse as related to the survey classifications.

Suppose that the result of each interview is the classification of the subject into one of K non-overlapping categories. Consider estimating gross flows among these categories using records of surveyed individuals matched over consecutive interview periods. It will be impossible to obtain matches for individuals who were nonrespondents in one or more of the interview periods or who rotated into or out of the survey during the period being considered. Thus, as a result of matching the survey data, we will have a group of individuals for whom we have survey classifications at each interview period, a group of individuals for whom we have classifications in some but not all periods, and a group of individuals who never responded to the survey.

The survey classification data for individuals who responded at two consecutive interview times, say t-1 and t, can be summarized in a K×K matrix. The available information for individuals who were nonrespondents for the time t-1 interview but who responded to the time t interview may be summarized in a column supplement. The available information for individuals who were nonrespondents for the time t interview but who responded to the time t-1 interview may be summarized in a row supplement. Individuals who were nonrespondents at both times t-1 and t are counted in a single Missing cell. Therefore, the observed time t-1 to time t gross flow data can be displayed as in Table 1.

TABLE 1: OBSERVED GROSS FLOW DATA

		Time t				
		1	2	...	K	*Row Supp.*
	1	x_{11}	x_{12}	...	x_{1K}	R_1
Time t-1	2	x_{21}	x_{22}	...	x_{2K}	R_2
	⋮	⋮	⋮	⋮	⋮	⋮
	K	x_{K1}	x_{K2}	...	x_{KK}	R_K
Column Supp.		C_1	C_2	...	C_K	M

where x_{ij} = number of sampled individuals with classification i at time t-l and j at time t,
R_i = number of individuals who were nonrespondents at time t and had classification i at time t-1,
C_j = number of individuals who were nonrespondents at time t-1 and had classification j at time t, and
M = number of individuals who were nonrespondents at both times t-1 and t.

Extending the ideas of Chen and Fienberg (1974) for maximum likelihood estimation in contingency tables with partially cross-classified data, I take the observed gross flow data to be the end result of a two-stage process where, in the unobserved first stage, individuals are allocated to the nine cells of a K×K matrix according to probabilities from a Markov-chain. Let

π_i = initial probability that an individual is in state i at time t-1, where $\Sigma_i\pi_i = 1$ and
p_{ij} = transition probability from state i to state j, where $\Sigma_j p_{ij} = 1$ for all i.

At the second stage of the process each individual in the (i,j) cell of the gross flow matrix may either a) be a nonrespondent in month t-1 and, hence, lose its row classification, b) be a nonrespondent in month t and lose its column classification, or c) be a nonrespondent in both months and lose both its row and column classification. Let

$\xi(i,j)$ = initial probability that an individual in the (i,j) cell of the matrix responds in month t-1,
$\rho_{RR}(i,j)$ = transition probability from respondent in month t-1 to respondent in month t, and
$\rho_{MM}(i,j)$ = transition probability from nonrespondent in month t-1 to nonrespondent in month t.

The data are observed after this second stage. From the observed data, we want to make inferences about both the probabilities of the Markov-chain generating the labor force data and the probabilities of the Markov-chain generating nonresponse. In the context of this two-stage model, the underlying probabilities for the observed gross flow matrix are as given in Table 2.

III. Markov-Chain Models for Nonresponse

We wish to use maximum likelihood estimation to obtain estimates of the parameters shown in Table 2. The likelihood function for the observed data under the two-stage model described in Section II is as follows:

$$\{\Pi_i\Pi_j[\xi(i,j)\rho_{RR}(i,j)\pi_i p_{ij}]^{x_{ij}}\}$$
$$\times \{\Pi_i[\Sigma_j\xi(i,j)[1-\rho_{RR}(i,j)]\pi_i p_{ij}]^{R_i}\}$$
$$\times \{\Pi_j[\Sigma_i[1-\xi(i,j)][1-\rho_{MM}(i,j)]\pi_i p_{ij}]^{C_j}\}$$
$$\times \{[\Sigma_i\Sigma_j[1-\xi(i,j)]\rho_{MM}(i,j)\pi_i p_{ij}]^{M}\}$$

where i and j take on the values from 1 to K. There are, however, $4K^2+K$ parameters with K+1 constraints on the parameters in the above likelihood and only $(K+1)^2$ cells of observed counts with the single constraint that the observed counts sum to the total sample size. Thus, we must reduce the number of parameters to be estimated. We will do this by considering 4 models for the ξ and ρ parameters, the parameters pertaining to nonresponse. The models are as follows:

Model A: $\xi(i,j) = \xi$, $\rho_{RR}(i,j) = \rho_{RR}$, $\rho_{MM}(i,j) = \rho_{MM}$
Model B: $\xi(i,j) = \xi(i)$, $\rho_{RR}(i,j) = \rho_{RR}$, $\rho_{MM}(i,j) = \rho_{MM}$
Model C: $\xi(i,j) = \xi$, $\rho_{RR}(i,j) = \rho_{RR}(i)$, $\rho_{MM}(i,j) = \rho_{MM}(i)$
Model D: $\xi(i,j) = \xi$, $\rho_{RR}(i,j) = \rho_{RR}(j)$, $\rho_{MM}(i,j) = \rho_{MM}(j)$

Models A and B have 2K-2 and K-1 associated degrees of freedom respectively while there are no degrees of freedom associated with models C and D. The π and p probabilities for the gross flow process are as defined in Section II.

Under Model A, the initial probability that an individual responds at time t-1 is the same for all survey classifications. The transition probabilities from respondent to respondent or from nonrespondent to nonrespondent also do not depend on the survey classification. Under Model B, the initial probability that an individual responds at time t-1 depends on the individual's classification at time t-1 while the transition probabilities from respondent to respondent or from nonrespondent to nonrespondent do not depend on classification. Under Models C and D, the initial probability that an individual responds at time t-1 is the same for all survey classifications. The transition probabilities from respondent to respondent or from nonrespondent to nonrespondent under Model C depend on the survey classification at time t-1 while under Model D they depend on the classification at time t. Note that Model A is a special case of each of the other three models. The methods used to fit each of these models are described below.

Model A

The likelihood function for the observed data under Model A can be written as the product of two factors, one involving only the π and p parameters. Thus, estimates of the π and p parameters can be found separately from the estimates of the ξ and ρ parameters. This means that, under Model A, the Markov-chains for the gross flow process and for nonresponse are separate, independent Markov-chains.

Factor f_{A1} is maximized using Lagrange multipliers to impose the constraints that $\Sigma_i\pi_i=1$ and $\Sigma_j p_{ij}=1$ for all i. In general, an iterative procedure must be used to provide estimates of the π and p parameters. Initial estimates for the π and p parameters are as follows:

$$\pi_i^{(0)} = x_{i\cdot}/x_{\cdot\cdot} \quad\text{and}\quad p_{ij}^{(0)} = x_{ij}/x_{i\cdot}$$

where a '.' in a subscript indicates summation over that subscript. The iterative steps for obtaining the parameter estimates are given in Table 3a.

The estimates of the ξ and ρ parameters have the following closed forms:

$$\xi = [x_{\cdot\cdot} + R_\cdot]/[x_{\cdot\cdot} + R_\cdot + C_\cdot + M],$$
$$\rho_{RR} = x_{\cdot\cdot}/[x_{\cdot\cdot} + R_\cdot], \text{ and}$$
$$\rho_{MM} = M/[C_\cdot + M].$$

Model B

The likelihood function for the observed data under Model B can be written as the product of two factors, one involving only the π, p, and ξ parameters. Thus, estimates of the π, p, and ξ parameters can be found separately from the estimates of the ρ parameters.

The estimates of the ρ parameters under Model B have the same closed-form solutions as under Model A. Factor f_{B1} is maximized using Lagrange multipliers to impose the constraints that $\Sigma_i\pi_i=1$ and $\Sigma_j p_{ij}=1$ for all i. In general, an iterative procedure must be used to provide estimates of the π, p, and ξ parameters. Initial estimates for the π and p parameters are as given for Model A. Initial estimates for the ξ parameters are as follows:

$$\xi(i)^{(0)} = [x_{\cdot\cdot} + R_\cdot]/[x_{\cdot\cdot} + R_\cdot + C_\cdot + M].$$

The iterative steps for obtaining the parameter estimates are given in Table 3b.

Model C

The likelihood function for the observed data under Model C can be written as the product of two factors, one involving only the π, p, and ρ_{MM} parameters. Thus, estimates of the π, p, and ρ_{MM} parameters can be found separately from the estimates of the ξ and ρ_{RR} parameters.

The estimate of ξ under Model C has the same closed-form solution as under Model A. The closed-form estimator of $\rho_{RR}(i)$ is:

$$\rho_{RR}(i) = x_{i\cdot}/(x_{i\cdot} + R_i).$$

Factor f_{C1} is maximized using Lagrange multipliers to impose the constraints that $\Sigma_i\pi_i=1$ and $\Sigma_j p_{ij}=1$ for all i. In general, an iterative procedure must be used to provide estimates of the π, p, and ρ_{MM} parameters. Initial estimates for the π and p parameters are as given for Model A. Initial estimates for the ρ_{MM} parameters are as follows:

$$\rho_{MM}(i)^{(0)} = M/[C_\cdot + M].$$

The iterative steps for obtaining the parameter estimates are given in Table 3c.

Model D

The likelihood function for the observed data under Model D can be written as the product of two factors, one involving only the π, p, and ρ parameters. Thus, estimates of the π, p, and ρ parameters can be found separately from the estimate of the ξ parameter.

The estimate of the ξ parameter under Model D has the same closed-form solution as under Model A. Factor f_{D1} is maximized using Lagrange multipliers to impose the constraints that $\Sigma_i\pi_i=1$ and $\Sigma_j p_{ij}=1$ for all i. In general, an iterative procedure must be used to provide estimates of the π, p, and ρ parameters. Initial estimates for the π and p parameters are as given for Model A. Initial estimates for the ρ_{RR} and ρ_{MM} parameters are as follows:

TABLE 2: PROBABILITIES FOR OBSERVED GROSS FLOW DATA

	Time t: 1 2 ... K	*Row Supp.*
Time t-1: 1, 2, ..., K	$\{\xi(i,j)\rho_{RR}(i,j)\pi_i p_{ij}\}$	$\{\Sigma_j\xi(i,j)[1-\rho_{RR}(i,j)]\pi_i p_{ij}\}$
Column Supp.	$\{\Sigma_i[1-\xi(i,j)][1-\rho_{MM}(i,j)]\pi_i p_{ij}\}$	$\{\Sigma_i\Sigma_j[1-\xi(i,j)]\rho_{MM}(i,j)\pi_i p_{ij}\}$

$$\rho_{RR}(j)^{(0)} = x_{..}/[x_{..} + R_{.}] \text{ and}$$
$$\rho_{MM}(i)^{(0)} = M/[C_{.} + M].$$

The iterative steps for obtaining the parameter estimates are given in Table 3d.

Fitting the Models

For each of the models, the steps of the iterative procedure given in Table 3 are repeated for $\nu = 0, 1, 2, \ldots$ until the parameter estimates converge to the desired degree of accuracy. The formulas given for $\pi_i^{(0)}$, $p_{ij}^{(0)}$, $\xi(i)^{(0)}$, $\rho_{RR}(i)^{(0)}$, and $\rho_{MM}(j)^{(0)}$ are merely suggested initial estimates. Any values between 0 and 1 satisfying $\Sigma_i\pi_i=1$ and $\Sigma_j p_{ij}=1$ for all i may be used. In the data analysis reported in Section IV, a number of different starting values were used. In each case, the final estimates were the same.

After any one of the above models has been fit to the data, the cell probabilities underlying the observed data may be estimated following the formulas given in Table 2. These estimated probabilities may then be multiplied by the total sample size to obtain the expected observed cell counts. Either the Pearson X^2 or the likelihood ratio statistic, G^2, can be compared to a χ^2 distribution with the appropriate degrees of freedom to help assess the fit of the model.

IV. Example From the Canadian Labour Force Survey

The Labour Force Survey

The Canadian Labour Force Survey (LFS) is based on monthly interviews with respondents in approximately 56,000 households. Sampled households are retained in the sample for six months before being rotated out of the sample. Under this LFS scheme, the month-to-month overlap of sampled housing locations is 83%. A detailed description of the LFS can be found in *Methodology of the Canadian Labour Force Survey 1976*, Statistics Canada (1977).

Month-to-month gross flows in labor force participation show how persons with each labor force classification in one month are classified in following month providing, for example, estimates of the numbers of persons who were employed in one month and unemployed in the next, unemployed in one month and employed in the next, employed in both months, and so forth.

A single panel of micro-data from the LFS is available for our use. The data set contains responses for a subset of the survey questions for all individuals from the panel that rotated into the sample in August 1979 and remained in the sample through January 1980. Information is available for each individual in that panel who responded at least once during the six month period. Unweighted cell counts for a gross flow matrix can be obtained from this micro-data. The models described in this paper are suitable for unweighted data from a simple random sample. The LFS uses a multi-stage cluster sample. Thus, the models proposed in this paper are not ideally suited to describe the data. We will, however, fit the models to the data for illustrative purposes and as a first attempt at modeling the nonresponse in the data.

Persons interviewed for the LFS in a given month are classified as employed, unemployed, not in the labor force, or outside the population of interest. Naturally, persons outside the population of interest are not intentionally included in the LFS sample. The relatively few persons classified as outside the population of interest who do appear in the sample are included by accident rather than by design. Thus, the out-of-population cells based on the available panel

TABLE 3: EQUATIONS FOR ITERATIVE PROCEDURES FOR MODEL FITTING

3a. Model A

$$\pi_i^{(\nu+1)} = \{x_{i.} + R_i + \Sigma_j[C_j\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}]\} \times \{x_{..} + R_{.} + C_{.}\}^{-1} \text{ and}$$
$$p_{ij}^{(\nu+1)} = \{x_{ij} + [C_j\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}]\} \times \{x_{i.} + \Sigma_j[C_j\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}]\}^{-1}$$

3b. Model B

$$\pi_i^{(\nu+1)} = \{x_{i.} + R_i + \Sigma_j[C_j[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}] + [M[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}]\} \times \{x_{..} + R_{.} + C_{.} + M\}^{-1},$$
$$p_{ij}^{(\nu+1)} = \{x_{ij} + [C_j[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}]\} \times \{x_{i.} + \Sigma_j [C_j[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}]\}^{-1}, \text{ and}$$
$$\xi(i)^{(\nu+1)} = \{x_{i.} + R_i\} \times \{x_{i.} + R_i + \Sigma_j[C_j[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}] + [M[1-\xi(i)^{(\nu)}]\pi_i^{(\nu)}/\Sigma_k[1-\xi(k)^{(\nu)}]\pi_k^{(\nu)}]\}^{-1}.$$

3c. Model C

$$\pi_i^{(\nu+1)} = \{x_{i.} + R_i + \Sigma_j[C_j[1-\rho_{MM}(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\rho_{MM}(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}] + [M\rho_{MM}(i)^{(\nu)}\pi_i^{(\nu)}/\Sigma_k[\rho_{MM}(k)^{(\nu)}\pi_k^{(\nu)}]]\}$$
$$\times \{x_{..} + R_{.} + C_{.} + M\}^{-1},$$
$$p_{ij}^{(\nu+1)} = \{x_{ij} + [C_j[1-\rho_{MM}(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\rho_{MM}(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}]\}$$
$$\times \{x_{i.} + \Sigma_j [C_j[1-\rho_{MM}(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\rho_{MM}(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}]\}^{-1}, \text{ and}$$
$$\rho_{MM}(i)^{(\nu+1)} = \{M\rho_{MM}(i)^{(\nu)}\pi_i^{(\nu)}/\Sigma_k[\rho_{MM}(k)^{(\nu)}\pi_k^{(\nu)}]\}$$
$$\times \{\Sigma_j [C_j[1-\rho_{MM}(i)^{(\nu)}]\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k[1-\rho_{MM}(k)^{(\nu)}]\pi_k^{(\nu)}p_{kj}^{(\nu)}] + M\rho_{MM}(i)^{(\nu)}\pi_i^{(\nu)}/\Sigma_k[\rho_{MM}(k)^{(\nu)}\pi_k^{(\nu)}]\}^{-1}.$$

3d. Model D

$$\pi_i^{(\nu+1)} = \{x_{i.} + R_i + \Sigma_j[C_j\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}] + \Sigma_j[M\rho_{MM}(j)^{(\nu)}\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\Sigma_h[\rho_{MM}(h)^{(\nu)}]\pi_k^{(\nu)}p_{kh}^{(\nu)}]\} \times \{x_{..} + R_{.} + C_{.} + M\}^{-1},$$
$$p_{ij}^{(\nu+1)} = \{x_{ij} + [R_i[1-\rho_{RR}(j)^{(\nu)}]p_{ij}/\Sigma_k[1-\rho_{RR}(k)^{(\nu)}]p_{ik}] + [C_j\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}]$$
$$+ [M\rho_{MM}(j)^{(\nu)}\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\Sigma_h[\rho_{MM}(h)^{(\nu)}\pi_k^{(\nu)}p_{kh}^{(\nu)}]]\}$$
$$\times \{x_{i.} + R_i + \Sigma_j[C_j\pi(i)^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\pi(k)^{(\nu)}]p_{kj}^{(\nu)}] + [M\Sigma_j\rho_{MM}(j)^{(\nu)}\pi_i^{(\nu)}p_{ij}^{(\nu)}/\Sigma_k\Sigma_h[\rho_{MM}(h)^{(\nu)}]\pi_k^{(\nu)}p_{kh}^{(\nu)}]\}^{-1},$$
$$\rho_{RR}(j)^{(\nu+1)} = x_{.j} \times \{x_{.j} + \Sigma_i[R_i[1-\rho_{RR}(j)^{(\nu)}]p_{ij}^{(\nu)}/\Sigma_k[1-\rho_{RR}(k)^{(\nu)}]p_{ik}^{(\nu)}]\}^{-1}; \text{ and}$$
$$\rho_{MM}(j)^{(\nu+1)} = 1 - \{[C_j\Sigma_k\Sigma_h\rho_{MM}(h)^{(\nu)}\pi_k^{(\nu)}p_{kh}^{(\nu)}]/[M\Sigma_k\pi_k^{(\nu)}p_{kj}^{(\nu)}]\}.$$

of data are mostly empty and I do not include them in the analysis given here. I will consider estimating gross flows among the three labor force classifications using records of individuals matched over two consecutive months.

The observed gross flow data is given in Appendix I.

The Fits of the Models

All four models described in Section III were fit to the five possible observed gross flow matrices constructed from the available panel of data. Since there are K=3 possible survey classifications, Model A has 4 associated degrees of freedom and Model B has 2 degrees of freedom. Models C and D, which have no associated degrees of freedom, will both fit the data exactly although they need not necessarily produce the same parameter estimates.

The criterion for stopping the iterative procedures necessary for obtaining some of the parameter estimates was that the maximum difference between estimates at two consecutive steps was less than .0005. The iterative procedure for fitting Model A converged quickly, requiring only 2 steps to converge for each of the 5 observed gross flow matrices. The iterative procedure for fitting Model B converged relatively slowly requiring between 24 and 58 iterations. The iterative procedures for fitting Models C and D converged in between 16 and 25 iterations, and between 9 and 13 iterations respectively. The parameter estimates, X^2, and G^2 values for all models are given in Appendix II.

The fits of Models A and B to the August to September data are similar. For all other gross flow matrices, Model B provides a better fit to the data. (Note that, given the large cell counts in the observed gross flow matrices, we find the fits of Model B reasonable even though the X^2 and G^2 values are larger than the value of $\chi^2_{.99}(2) = 9.21$.) Recall that under Model A the probabilities of nonresponse are the same for individuals in all employment classifications while under Model B the initial probability of being a nonrespondent in month t-1 depends on the employment classification in that month. Thus, since Model B provides a better fit than Model A, we have some evidence that nonresponse does depend on employment status.

In part B of Appendix II, we see that the estimated initial probabilities of falling in each labor force classification are similar under Models A, C, D, and in August to September under Model B. Under Model B in all other months, however, the estimated initial probability that a person is unemployed is higher than under other models. For example, in October to November, the proportion of persons initially unemployed is estimated to be about 3.8% under Models A, C, and D while it is about 5.1% under Model B.

Part C of Appendix II shows that the estimated transition probabilities among the various employment classifications do not vary greatly from model to model. Note, however, that the estimated transition probabilities do appear to change over time. This change may be due to actual changes in the labor force over time but it may also but due, at least in part, to the effects of rotation group bias (see, for example, Bailar (1975) and (1979)).

Parts D, E, F, and G of Appendix II give the estimates of the ξ, ρ_{RR}, and ρ_{MM} parameters under Models A, B, C, and D respectively. As noted in Section III, the estimates of ξ, the initial probability of being a respondent, are identical under Models A, C, and D. Under Model B, however, the initial probability of being a respondent depends on the labor force classification. Note that in all months, these initial probabilities of being a respondent are quite similar for persons who are employed or not in the labor force. The estimated value of ξ for unemployed persons, however, is lower than for other persons. This difference is not very large for the August to September data as would be expected since the fits of Models A and B are similar for that data. The difference is fairly large in all other months with the estimates of $\xi(U)$ being about .2 to .3 lower than the estimates of $\xi(E)$ or $\xi(N)$. This again illustrates that response rates appear to differ by labor force classification.

The estimates of ρ_{RR} and ρ_{MM}, the probabilities of transitions from respondent to respondent and nonrespondent to nonrespondent respectively, are identical under Models A and B but depend on employment classification under Models C and D. Note that the estimates of the $\rho_{RR}(E)$, $\rho_{RR}(N)$, $\rho_{MM}(E)$, and $\rho_{MM}(N)$ do not differ much from Model C to Model D although, except in one case, the $\rho_{RR}(E)$ and $\rho_{RR}(N)$ are larger under Model D than under Model C and the estimates of the $\rho_{MM}(E)$ and $\rho_{MM}(N)$ are always smaller under Model D than under Model C. The differences range from about .001 to about .007 between probabilities that are estimated to be from about .74 to about .97. The differences between estimates of the $\rho_{RR}(U)$ and $\rho_{MM}(U)$ under Models C and D are somewhat larger, ranging from about .01 to about .1 on estimates that range from about .58 to about .95. The estimated probabilities of an unemployed person remaining a respondent, $\hat{\rho}_{RR}(U)$, are always larger under Model C than under Model D while the estimated probabilities of an unemployed person remaining a nonrespondent, $\hat{\rho}_{MM}(U)$, are always smaller under Model C than under Model D.

Thus we see that while both Models C and D provide exact fits to the data, they do not result in the same parameter estimates and, hence, they do not give the same estimated expected cell counts after the first stage. This can be seen in Part A of Appendix II. Also note, however, that the estimated expected cell counts after the first stage do not differ by much under any of the Models A, C, and D or in August to September under Model B. In other months, the estimates under Model B do differ somewhat from the estimates under the other models. In particular, note that the expected cell counts in the row corresponding to persons who were unemployed in the first of the two months are larger under Model B.

From this analysis, it is not clear which model we would prefer for modeling nonresponse in this labor force data. It is does seem, however, that unemployeds have response patterns that are different from the response patterns for persons who are employed or not in the labor force. Since response rates appear to be fairly similar for persons who are employed or not in the labor force, it would be worthwhile to consider variations of the models fit above where the probabilities associated with those classifications are the same.

V. Extensions of the Models

An advantage of the Markov-chain models for nonresponse proposed here is that there is a natural way to think about extending the models to allow us to use more than two periods of data in estimating gross flows. This is not the case for the discrete-time models described by Stasny and Fienberg (1985) and by Stasny (1983, 1985, and 1986). Therefore, an important generalization of this work will be to extend the models to handle gross flows over more than two time periods.

After the π_i and p_{ij} have been estimated under one of the above models using one of the procedures described in the previous sections, the estimates may then be used to estimate the intensity matrix for a continuous-time Markov-chain for the gross flow process. Descriptions of estimating continuous-time Markov-chains from data collected at discrete intervals are given by Singer and Spilerman (1976) and Stasny (1983). Thus, another extension of the model would be to allow the Markov-chain model for the gross flow process to be a continuous-time Markov-chain.

APPENDIX I: Observed LFS Data

		September 1979			
		E	U	N	Row Supp.
August 1979	E	9223	128	662	473
	U	221	322	151	59
	N	256	164	5941	292
Col. Supp.		997	69	677	5016

		October 1979			
		E	U	N	Row Supp.
September 1979	E	9698	169	355	475
	U	177	317	143	46
	N	326	159	6522	423
Col. Supp.		554	59	339	4701

		November 1979			
		E	U	N	Row Supp.
October 1979	E	9779	178	392	406
	U	159	362	130	53
	N	212	145	6728	274
Col. Supp.		419	54	269	4920

		December 1979			
		E	U	N	Row Supp.
November 1979	E	9685	202	304	379
	U	129	405	157	48
	N	204	155	6928	232
Col. Supp.		291	60	219	5051

		January 1980			
		E	U	N	Row Supp.
December 1979	E	9413	191	366	339
	U	162	450	168	42
	N	187	180	7004	237
Col. Supp.		252	50	186	5204

APPENDIX II: Parameter Estimates

A. Estimates of the Expected Cell Counts After the First Stage

		Model A			Model B			Model C			Model D		
		E	U	N	E	U	N	E	U	N	E	U	N
8/79-9/79	E	13315	186	954	13315	184	956	13307	185	955	13296	199	952
	U	330	485	225	349	508	239	330	482	226	319	501	217
	N	369	238	8548	366	234	8500	369	236	8561	369	255	8542
9/79-10/79	E	13214	238	481	13202	230	483	13198	230	483	13183	240	490
	U	247	457	198	326	585	263	247	442	199	241	450	197
	N	450	227	8950	436	213	8725	450	219	8994	443	225	8994
10/79-11/79	E	13224	249	528	13166	240	528	13220	241	530	13204	258	529
	U	223	527	182	306	699	250	224	509	183	215	526	175
	N	287	203	9058	278	190	8823	286	196	9091	286	210	9077
11/79-12/79	E	13015	284	409	12716	266	399	13044	272	409	13032	287	407
	U	178	587	217	288	907	350	179	563	218	174	576	210
	N	273	217	9268	267	203	9054	273	208	9283	274	220	9269
12/79-1/80	E	12676	266	492	12464	253	485	12689	258	493	12682	266	492
	U	222	638	230	334	930	346	223	619	231	218	627	226
	N	251	250	9405	244	235	9140	252	242	9425	252	251	9416

B. Estimates of the Initial Probabilities of Being in Each Employment Classification

		Model A	Model B	Model C	Model D
8/79-9/79	$\hat{\pi}_E$	.5864	.5864	.5861	.5861
	$\hat{\pi}_U$	.0422	.0445	.0421	.0421
	$\hat{\pi}_N$	.3714	.3692	.3718	.3718
9/79-10/79	$\hat{\pi}_E$	.5695	.5688	.5687	.5687
	$\hat{\pi}_U$	.0369	.0480	.0363	.0363
	$\hat{\pi}_N$	.3936	.3832	.3950	.3950
10/79-11/79	$\hat{\pi}_E$	.5719	.5692	.5715	.5715
	$\hat{\pi}_U$	.0381	.0513	.0374	.0374
	$\hat{\pi}_N$	.3900	.3796	.3911	.3911
11/79-12/79	$\hat{\pi}_E$	.5607	.5473	.5614	.5614
	$\hat{\pi}_U$	.0402	.0632	.0393	.0392
	$\hat{\pi}_N$	.3991	.3895	.3993	.3994
12/79-1/80	$\hat{\pi}_E$	.5499	.5404	.5501	.5501
	$\hat{\pi}_U$	.0446	.0659	.0439	.0439
	$\hat{\pi}_N$	.4055	.3937	.4060	.4060

C. Estimates of the Transition Probabilities, p_{ij}

		Model A E	Model A U	Model A N	Model B E	Model B U	Model B N	Model C E	Model C U	Model C N	Model D E	Model D U	Model D N
8/79-	E	.9212	.0129	.0660	.9211	.0128	.0661	.9211	.0128	.0661	.9203	.0138	.0659
9/79	U	.3175	.4660	.2165	.3188	.4633	.2178	.3184	.4641	.2175	.3073	.4833	.2094
	N	.0403	.0260	.9337	.0402	.0257	.9340	.0402	.0258	.9340	.0403	.0278	.9319
9/79-	E	.9484	.0171	.0345	.9487	.0166	.0347	.9487	.0165	.0347	.9476	.0172	.0352
10/79	U	.2735	.5069	.2196	.2776	.4986	.2238	.2778	.4978	.2244	.2711	.5067	.2222
	N	.0468	.0236	.9296	.0466	.0227	.9307	.0465	.0227	.9308	.0459	.0233	.9308
10/79-	E	.9445	.0178	.0377	.9449	.0172	.0379	.9449	.0172	.0379	.9438	.0184	.0378
11/79	U	.2397	.5652	.1951	.2439	.5569	.1992	.2442	.5562	.1996	.2345	.5739	.1916
	N	.0300	.0212	.9487	.0299	.0205	.9496	.0299	.0205	.9496	.0299	.0220	.9481
11/79-	E	.9494	.0207	.0298	.9503	.0198	.0298	.9503	.0198	.0298	.9495	.0209	.0296
12/79	U	.1815	.5975	.2210	.1862	.5873	.2265	.1866	.5862	.2272	.1810	.6000	.2190
	N	.0279	.0222	.9498	.0280	.0213	.9507	.0280	.0213	.9507	.0281	.0225	.9494
12/79-	E	.9436	.0198	.0366	.9441	.0192	.0367	.9441	.0192	.0367	.9436	.0198	.0366
1/80	U	.2037	.5853	.2110	.2072	.5779	.2148	.2076	.5770	.2153	.2038	.5853	.2109
	N	.0254	.0252	.9494	.0254	.0244	.9502	.0254	.0244	.9502	.0254	.0253	.9493

D. ξ, ρ_{RR}, and ρ_{MM} Parameter Estimates Under Model A

	ξ	ρ_{RR}	ρ_{MM}	X^2	G^2
8/79-9/79	.7258	.9539	.7421	19	16
9/79-10/79	.7689	.9498	.8316	38	36
10/79-11/79	.7687	.9610	.8690	49	40
11/79-12/79	.7701	.9650	.8986	74	58
12/79-1/80	.7670	.9670	.9143	43	34

Note: $\chi^2_{.99}(4) = 13.28$

F. ξ, ρ_{RR}, and ρ_{MM} Parameter Estimates Under Model C

ξ	$\rho_{RR}(E)$	$\rho_{RR}(U)$	$\rho_{RR}(N)$	$\rho_{MM}(E)$	$\rho_{MM}(U)$	$\rho_{MM}(N)$
.7258	.9549	.9216	.9561	.7411	.7025	.7483
.7689	.9556	.9327	.9431	.8310	.5842	.8552
.7687	.9623	.9247	.9628	.8714	.6474	.8866
.7701	.9641	.9350	.9691	.9101	.6138	.9104
.7670	.9671	.9489	.9688	.9212	.7156	.9264

E. ξ, ρ_{RR}, and ρ_{MM} Parameter Estimates Under Model B

	$\xi(E)$	$\xi(U)$	$\xi(N)$	ρ_{RR}	ρ_{MM}	X^2	G^2
8/79-9/79	.7254	.6870	.7311	.9539	.7421	19	16
9/79-10/79	.7687	.5822	.7926	.9498	.8316	19	19
10/79-11/79	.7719	.5611	.7920	.9610	.8690	26	21
11/79-12/79	.7900	.4784	.7895	.9650	.8986	24	20
12/79-1/80	.7809	.5106	.7909	.9670	.9143	9	8

Note: $\chi^2_{.99}(2) = 9.21$

G. ξ, ρ_{RR}, and ρ_{MM} Parameter Estimates Under Model D

ξ	$\rho_{RR}(E)$	$\rho_{RR}(U)$	$\rho_{RR}(N)$	$\rho_{MM}(E)$	$\rho_{MM}(U)$	$\rho_{MM}(N)$
.7258	.9557	.8852	.9582	.7400	.7364	.7458
.7689	.9567	.9165	.9430	.8271	.7208	.8485
.7687	.9634	.8965	.9643	.8678	.7649	.8811
.7701	.9650	.9138	.9706	.9061	.7588	.9037
.7670	.9677	.9357	.9697	.9178	.8123	.9212

REFERENCES

Bailar, B. A. (1975). "The Effects of Rotation Group Bias on Estimates for Panel Surveys". *Journal of the American Statistical Association*, **70**, 23-30.

__________(1979). "Rotation Sample Biases and Their Effects on Estimates of Change". *Proceedings of the Conference on Recent Developments in Statistical Methods and Applications*, 13-29.

Chen, T. and Fienberg, S. E. (1974). "Two-Dimensional Contingency Tables With Both Completely and Partially Cross-Classified Data." *Biometrics*, **30**, 629-642.

Fienberg, S. E. and Stasny, E. A. (1983). "Estimating Monthly Gross Flows in Labour Force Participation". *Survey Methodology*, **9**, 77-102.

Paul, E. C. and Lawes, M. (1982). "Characteristics of Respondent and nonrespondent Households in the Canadian Labour Force Survey". *Survey Methodology*, **8**, 48-85.

Rubin, D. B. (1976). "Inference and Missing Data". *Biometrika*, **63**, 581-592.

Singer, B. and Spilerman, S. (1976). "Some Methodological Issues in the Analysis of Longitudinal Surveys", *Annals of Economic and Social Measurement*, **5**, 447-474.

Stasny, E. A. (1983). "Estimating Gross Flows in Labor Force Participation Using Data From the Canadian Labour Force Survey". Unpublished Ph.D. dissertation, Department of Statistics, Carnegie-Mellon University.

____________(1985). "Modeling Non-ignorable Non-response in Categorical Panel Data: An Example in Estimating Gross Labor-Force Flows". Technical Report No. 330, Department of Statistics, Ohio State University.

____________(1986). "Estimating Gross Flows Using Panel Data With Nonresponse: An Example From the Canadian Labour Force Survey", *JASA*, March 1986.

Stasny, E. A. and Fienberg, S. E. (1985). "Some Stochastic Models for Estimating Gross Flows in the Presence of Nonrandom Nonresponse". *Proceedings of the Conference on Gross Flows in Labor Force Statistics*, U.S. Department of Commerce and U.S. Department of Labor, 25-39.

Statistics Canada (1977). *Methodology of the Canadian Labour Force Survey 1976*, Ottawa, Canada.