332

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A TIME SERIES MODEL WITH QUALITATIVE VARIABLES

David M. Grether
California Institute of Technology

G. S. Maddala
University of Florida and
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 332

July 1980

ABSTRACT

This paper considers a distributed lag model in which the dependent variable is observed qualitatively. The relation of our "lagged index" model to other models that have appeared in the literature is discussed and a computationally tractable method of obtaining consistent estimates is presented. The model is applied to data on party identification in the United States. The results obtained indicate that party identification is responsive to changes in individual opinions, especially regarding the performance of an incumbent president.

A TIME SERIES MODEL WITH QUALITATIVE VARIABLES*

David M. Grether
California Institute of Technology

G. S. Maddala
University of Florida

I. INTRODUCTION

The literature on time-series analysis has been largely confined to the analysis of time-series data where:

(i) The number of observations is large, and

(ii) the variables are all observed as continuous variables.

There are a large number of problems of practical interest that depend on time-series analysis of a different type of data sets. These data sets, which are being made available through numerous longitudinal surveys, have the following characteristics:

(1) The number of cross-section units is large and the number of time periods is small. We thus have a large number of short time series.

(2) Very often the dependent variable is either a categorical variable or a censored variable.

In the present paper we consider the estimation of a distributed lag model based on panel data where the time series on the dependent variable is observed as a categorical variable. In Section 2 we provide examples of some situations where this type of model is applicable. In later sections we present the estimation method and some empirical results.

Earlier discussions of similar problems can be found in Chamberlain (1980) and Heckman (1979). The model considered here is different from the models discussed by these authors. Briefly stated, the main differences are as follows: Chamberlain considers the model

$$y^*_{it} = \eta_i + \sum_{j=0}^{k} \beta_j x_{i,t-j} + u_{it} \qquad i = 1,2,\dots,N \quad t = 1,2,\dots,T \tag{1.1}$$

where y^*_{it} is the "index" variable that is observed only as a qualitative variable. A major emphasis in his paper is on how to handle the "incidental" parameters η_i. He suggests some conditional maximum likelihood methods for this model (consider the likelihood function conditional on sufficient statistics for the incidental parameters).

The model that Heckman considers is of the form:

$$y^*_{it} = \alpha y_{i,t-1} + \beta x_{it} + u_{it} \tag{1.2}$$

where

$$y_{it} = 1 \quad \text{if} \quad y^*_{it} > 0 \tag{1.3}$$
$$\quad = 0 \quad \text{otherwise.}$$

Thus, it is the realized values of the "index" variable in the previous periods that affect the current value of the index variable. In the case where y^*_{it} determines the probability of a person finding employment in period t, the model says that this depends on whether

*Financial support from the National Science Foundation is gratefully acknowledged. We would like to thank R. P. H. Fishe for the computations done in Section V; and D. Roderick Kiewiet and Morris P. Fiorina for their many helpful comments. Also we gratefully acknowledge the Sherman Fairchild Distinguished Scholars Program and the Program for the Study of Enterprise and Public Policy.

the person was employed or not in the previous period (what Heckman calls "state dependence"). Heckman's model contains more lags than one and there are other complications in the formulations adopted by Heckman, but the essence of his models is that it is the lagged values of the dichotomous realizations that occur as explanatory variables.

The model we consider can be termed a "lagged index" model, as contrasted to a "lagged dummy" model that Heckman considers. It is

$$y^*_{it} = \alpha y^*_{i,t-1} + \beta x_{it} + u_{it} \qquad (1.4)$$

where we do not observe y^*_{it} but observe the variable y_{it} as defined in (1.3) (or alternatively a polychotomous variable). In the case where y^*_{it} determines the probability of finding employment in period t, model (1.4) says that this depends on the corresponding index in period t - 1. The model (1.2), on the other hand, captures previous employment experience through the variables $y_{i,t-j}$. Thus, in the labor supply case the model (1.2) is more reasonable than (1.4), though one can make a case for (1.4) as well (or perhaps a combination of (1.2) and (1.4)).

In the following section we will give some examples of cases where the "lagged index" model makes more sense.

II. SOME EXAMPLES OF THE LAGGED INDEX MODEL

Consider the following model

$$y^*_{it} = \alpha y_{i,t-1} + \beta x_{it} + u_{it} \qquad i = 1,\ldots,N \qquad (2.1)$$
$$t = 1,\ldots,T$$

$$y_{it} = 1 \quad \text{if} \quad y^*_{it} > 0$$
$$= 0 \quad y^*_{it} \le 0.$$

Thus, this is a standard distributed lag model with the dependent variable observed qualitatively. For example, y^*_{it} might be the posterior log-odds in favor of some hypothesis, $y^*_{i,t-1}$ the prior log-odds, and x_{it} the log likelihood ratio. Suppose that individuals on the basis of certain observed data state which of a pair of hypotheses they believe is correct, and suppose further that they receive a valuable prize if their guess is correct. Then their choice of hypotheses is equivalent to the statement that y^*_{it} is positive or negative. If individuals' decisions follow Bayes' rule, then $\alpha = \beta = 1$. Alternatively, if people use the representativeness heuristic of Kahneman and Tversky (1972), they tend to ignore their priors, so according to this theory α is less than β. Of course, if instead people use the anchoring and adjustment heuristic (Tversky and Kahneman 1974), one might expect α to be greater than β. The same prediction would follow from the theories of Howell (1967, 1971). Thus, there are a number of alternative theories that bear on this setup. For an example of an experimental design that could be used to generate this type of data see Grether (1979).

The study of party affiliation by political scientists provides another example for the application of this model. There is a substantial empirical literature for both the United States and the United Kingdom on the concept of party identification. Party identification has been defined as "the individual's effective orientation to an important group-object in his environment" (Campbell, Converse, Miller, and Stokes 1960). One feature of party identification that has been subject to considerable debate is its stability. Based upon a survey Campbell, Converse, Miller, and Stokes estimated that only 20 percent of the population had (as of the survey) ever changed party identification. According to the traditional approach, people have a long-term orientation towards a political party which is much less volatile than voting intentions. This orientation is more stable than positions on issues (Converse 1964, Converse and Markus 1979).

Recently some have questioned the importance of the concept and have challenged its stability as well (Brody 1977, Dreyer 1973, Fiorina 1979). In the United States party identification is a measure of degree of affiliation with either the Democratic or Republican party. This sense of identification is according to the traditional interpretation supposed to change slowly in response to the performance of the party, performance of the president, general economic conditions, domestic turmoil, personal experience, and so forth. These considerations suggest that a partial adjustment model may be appropriate. Let $\tilde{y}^*_{it} = \beta' x_{it}$ be the desired or long-term equilibrium party identification of voter i at time t. If there were no social or psychological costs in changing political affiliations, then we would have

$$y^*_{it} = \tilde{y}^*_{it} + u_{it}$$

where y^*_{it} is the actual party identification. Since there are possibly some social and psychological costs, we have

$$y^*_{it} - y^*_{i,t-1} = \alpha(\tilde{y}^*_{it} - y^*_{i,t-1}) + u_{it}$$

so that

$$y^*_{it} = (1 - \alpha) y^*_{i,t-1} + \alpha \tilde{y}^*_{it} + u_{it}$$

$$= (1 - \alpha) y^*_{i,t-1} + \alpha\beta' x_{it} + u_{it}.$$

We further assume that

$$E(u_{it}) = 0$$

$$E(u^2_{it}) = \sigma^2_u .$$

In practice, of course, y^*_{it} is not observed as a continuous variable; party identification being measured as a polychotomy on a five or a seven point scale (e.g. strong Democratic, independent, to strong Republican).

III. ESTIMATION OF THE LAGGED INDEX MODEL

Rewrite the model (2.1) as

$$y^*_{it} = \beta x^*_{it}(\alpha) + \alpha^t \eta_{i0} + w_{it} \tag{3.1}$$

where

$$x^*_{it}(\alpha) = \sum_{j=0}^{t-1} \alpha^j x_{i,t-j}$$

$$w_{it} = \frac{u_{it}}{1 - \alpha L}$$

$$\eta_{i0} = \beta \sum_{j=0}^{\infty} \alpha^{i} x_{i,-j} = E_{i}(y^{*}_{i0}).$$

If y^{*}_{it} were observed, then one could obtain consistent estimates of the parameters of the model using the method suggested by Klein (1958). For the case at hand the same method works using a logit or probit estimation. The proof of consistency is basically that given by Lee (1980) who considers the Tobit estimator and parallels that given by Amemiya (1973) for the case of serially independent disturbances. For a proof of the strong consistency and a derivation of the limiting distribution of the Tobit estimator with serial correlation see Robinson (1980). In general, w_{it} are going to be autocorrelated so these estimates will not be efficient. Of course, efficient estimates can be obtained by the method of maximum likelihood. While in principle this is the appropriate estimation method, in practice it is not of much use. The reason is that with serial correlation the evaluation of the likelihood function involves the calculation of the T-fold integrals of the multivariate distributions. This leads to excessive computational problems if T is greater than two. The only other problem would be that if N is large relative to T, then there is an incidental parameter problem. If there is only one observation per individual on the dependent variable (i.e. T = 1), then one can treat the initial conditions as being random across individuals which simply adds another component to the disturbance term. The contribution of the component is not identified, but the estimation is otherwise straightforward; (3.1) is estimated using a standard probit program searching over admissible values of α. Those estimates corresponding to the maximum value of the likelihood function

are chosen. As there is only one observation over time, there cannot be any serial correlation present so these estimates are efficient and consistent as N goes to infinity provided the x's satisfy appropriate regularity conditions.

Now consider the more common situation in which one has many observations (over time) on the y_{it}'s and on the x_{it}'s. Notice that one cannot use the preceding procedure with all the data and obtain consistent estimates. The reason is that the disturbances in the model

$$w_{it} = \alpha^{t}\eta_{i0} + \frac{u_{it}}{1 - \alpha L} \tag{3.2}$$

are heteroscedastic due to the variation in α^{t}, and in logit or probit type models heteroscedasticity causes estimates to be inconsistent. One could assume $\eta_{i0} = \eta_{0}$, but for the application at hand this is equivalent to assuming that, prior to the sample, all voters had identical views of business conditions, of their own financial conditions, of the president's performance, etc. which is surely false. This problem does not arise if one were studying the effect of the state of the economy on party identification, as in this case the x_{it}'s would (for each t) be the same for all individuals. In this case one could simply include α^{t} as an explanatory variable in the logit or probit estimation. This is not a feasible alternative when studying party identification as the data sources are large but infrequent surveys. Thus, having party ID a function of aggregate measures only would lead to a substantial degrees-of-freedom problem.

Chamberlain (1980) considers the conditional likelihood approach to a similar problem. The model he deals with may be written as

$$y^*_{it} = \beta x_{it} + \eta_i + u_{it}. \qquad (3.3)$$

Unfortunately the procedure he uses does not work with dynamic models. To see this, consider the model (3.3) for the case where the observed variable, y_{it}, is a dichotomous variable. For T = 2 the conditional likelihood function for individual i for whom the $y^*_{i1} > 0$, $y^*_{i2} \leq 0$ is

$$\frac{\dfrac{e^{\beta x_{1i}+\eta_i}}{1+e^{\beta x_{1i}+\eta_i}}\;\dfrac{1}{1+e^{\beta x_{2i}+\alpha\eta_i}}}{\dfrac{e^{\beta x_{1i}+\eta_i}}{1+e^{\beta x_{1i}+\eta_i}}\;\dfrac{1}{1+e^{\beta x_{2i}+\alpha\eta_i}} + \dfrac{1}{1+e^{\beta x_{2i}+\alpha\eta_i}}\;\dfrac{e^{\beta x_{1i}+\eta_i}}{1+e^{\beta x_{1i}+\eta_i}}}$$

$$= \frac{e^{\beta x_{1i}+\eta_i}}{e^{\beta x_{1i}+\eta_i}\;1+e^{\beta x_{2i}+\alpha\eta_i}} = \frac{1}{1+e^{\beta(x_2-x_1)-(1-\alpha)\eta_i}}.$$

Thus the η_i do not drop out and this form of the conditional likelihood approach does not provide the necessary simplification.

The following procedure, on the other hand, should produce consistent estimates of all the parameters (provided $N \to \infty$). First, estimate

$$y^*_{it} = \beta x^*_{it}(\alpha) + w_{it}$$

using a standard logit or probit program using data on the dependent variable for time t only. Then reestimate the model using data for some other time period. These estimations provide estimates of α and of

$$\frac{\beta}{(\alpha^{2t}\sigma^2_\eta + \sigma^2_u)^{1/2}}$$

for two different values of t. This allows one to obtain an estimate of σ^2_η/σ^2_u; call it r. Finally, using all the data we can estimate

$$y_{it} = \beta\tilde{x}^*_{it}(\alpha) + \alpha^t\tilde{\eta}_{it} + u_{it}$$

$$= \frac{\beta x^*_{it}(\alpha)}{(\alpha^{2t}r+1)^{1/2}} + \frac{\alpha^t\bar{\eta}}{(\alpha^{2t}r+1)^{1/2}} + w_{it}$$

where w_{it} is serially correlated but has equal variances for each observation. Obviously there are a variety of ways that one can pool the estimates from separate cross sections to obtain a final set of estimated parameters.

An obvious alternative procedure would be to try to estimate all the parameters by a two step procedure analogous to two stage least squares. First, one would estimate the reduced form equation to obtain $\hat{y}^*_{i,t-1}$ and then estimate the model (1.4) directly by probit or logit using $\hat{y}^*_{i,t-1}$ as an explanatory variable. The trouble with this procedure is that while it produces consistent estimates of the β's (up to a common scale factor) α is not identified. The reason is that the disturbances in the reduced form and structural equations have different (and unidentified) variances.

If N equals one so that the data are for a single time series y^*_t, then things are much simpler. In this case one simply performs a grid search for α using the maximum likelihood probit method applied directly to

$$y_t^* = \beta x_t^*(\alpha) + \eta z_t^*(\alpha) + w_t \qquad (3.1)'$$

where

$$z_t^*(\alpha) = \alpha^t.$$

The parameter estimates chosen are those corresponding to the $\hat{\alpha}$ which produces the maximum value of the likelihood function.

IV. AN EMPIRICAL ILLUSTRATION

The data are from the 1972, 1974, 1976 election panel study administered by the Center for Political Studies at the University of Michigan. The panel consisted of 1320 individuals who were interviewed both before and after the 1972 election and either before or after the 1974 and 1976 elections. The questions used for this example concerned: party identification, civil rights, the performance of the president, the government's economic performance, future expectations concerning the economy, and personal financial conditions. The party ID variable was measured as a seven point scale (strong Republican, weak Republican, Independent-Republican, Independent-Independent, Independent-Democrat, weak Democrat, strong Democrat), and the presidential performance was dichotomous (approve, disapprove). All other variables were either three point scales or were collapsed to three point scales which we coded as binary variables with the center category as the control. For example, regarding civil rights, the respondents were asked whether they thought that civil rights leaders were pushing too fast, about right, or too slowly. For this question two binary variables (one for too fast and one for too slow) were created.

Table 1 shows the results of estimating the equation for party identification for 1976. Note that over half the sample was lost due to missing data and split-form questionnaires--not all respondents were asked every question for each election. Note also that since only one year's data on the dependent variable is used, the serial correlation problem does not matter so that the standard errors, etc. are consistently estimated (conditionally upon the value of α). It is clear that presidential performance is highly significant, while chi-square tests indicate that the other variables are not statistically significant at conventional levels. The evaluation of the government's economic performance and the pace of civil rights actions are the only variables that are close to being significant. These results are generally consistent with those of Fiorina (1979) who found that presidential performance and a Nixon pardon variable were highly significant in 1976, but that other variables were marginal. Table 2 shows the results when the personal financial condition and economic expectations variables are dropped. Both sets of results are quite similar. In both cases the maximum likelihood estimate of α is .4, which supports the more recent or revisionist arguments and runs counter to the traditional view that party identification changes only very slowly. Hypothesis tests that $\alpha = .8$ or $\alpha = .9$ are rejected ($\chi^2(1) = 10\text{-}14$ respectively). From the likelihood ratio statistics one can calculate approximate confidence intervals and estimate the standard deviation of $\hat{\alpha}$. In this case it appears that the standard error of $\hat{\alpha}$ is approximately .14. Notice that in addition to the coefficient estimates being stable the estimated cutoffs are nearly the same, and both sets

TABLE 1

PROBIT ESTIMATES PARTY ID 1976*

Variable	Coefficient	t-ratio**
Constant	.05	.3
Presidential performance	.97	10.7
Financial condition - good	.03	.4
- poor	.02	.3
Government's economic performance - good	.16	1.4
- poor	-.04	.4
Economic expectation - good	-.06	.7
- poor	-.03	.2
Civil rights - too fast	.03	.4
- too slow	-.19	1.3

α = .4
n = 622
ln L = -1097.95
$\hat{R}^2$ = .23

Estimated Cutoffs	Standard Errors
0.0	n.a.
.8	.06
1.1	.07
1.4	.07
1.8	.08
2.4	.09

*Data for independent variables taken from 1976, 1974, and 1972 surveys.

**Calculated conditional upon α = .4.

TABLE 2

PROBIT ESTIMATES PARTY ID 1976*

Variable	Coefficient	t-ratio**
Constant	.04	.3
Presidential performance	.91	10.8
Government's economic performance - good	.15	1.4
- poor	-.02	.2
Civil rights - too fast	.03	.4
- too slow	-.17	1.2

α = .4
n = 622
ln L = -1098.2
$\hat{R}^2$ = .23

Estimated Cutoffs	Standard Errors
0.0	n.a.
.8	.06
1.1	.07
1.4	.07
1.8	.08
2.4	.09

*Data for independent variables taken from 1976, 1974 and 1972 surveys.

**Calculated conditional upon α = .4.

suggest that the thresholds for the different points on the seven point scale are not evenly spaced (especially towards the ends of the scales). This is of interest as political scientists occasionally code these ordinal variables as interval levels (0, 1, 2, 3, 4, 5, 6) and use them in regressions, which can have unfortunate consequences (Grether 1974, 1976).

The grid search using the 1974 data on party identification (and 1974 and 1972 data on the explanatory variables) did not converge to $\hat{\alpha}$ equal to .4, but produced the boundary solution $\hat{\alpha}$ equal 1.0. As the data for 1976 are richer (three years as opposed to two for the independent variables) we take .4 as the preliminary estimate of α. Table 3 shows the estimates obtained for the equation eliminating the financial and expectational variables. The correction factor for the heteroscedasticity was obtained from the ratio of the 1976 and 1974 coefficients for presidential performance, and the model reestimated using all 1244 observations. The results are shown in Tables 4 and 5. Note that estimates of α using both data sets is .5. The presidential performance variable is the only substantive variable that is highly significant, though the civil rights variable is nearly so. It is unreasonable to assume that the average initial condition (presample) was zero and the shift variable is included for this reason. The overall constants in 1974 and 1976 respectively are $a + \alpha^t \bar{\eta}$ and $a + \alpha^{t+1} \bar{\eta}$. Thus the constant terms in Table 4 and Table 5 are estimates of $a + \alpha^t \bar{\eta}$ and the coefficients of the shift variables are estimates of $(\alpha^{t+1} - \alpha^t)\bar{\eta}$ and are negative and statistically significant. This suggests that prior to sample period individuals on average

TABLE 3

PROBIT ESTIMATES PARTY ID 1974*

Variable		Coefficient	t-ratio**
Constant		.37	2.4
Presidential performance		.55	6.3
Government's economic performance	- good	-.22	1.6
	- poor	-.08	1.0
Civil rights	- too fast	.06	.9
	- too slow	-.29	1.9

α = .4

n = 622

ln L = -1157.64

$\hat{R}^2$ = .10

Estimated Cutoffs	Standard Errors
0.00	n.a.
.57	.05
.89	.06
1.16	.06
1.49	.07
2.06	.08

*Data for independent variables taken from 1974, and 1972 surveys.

**Calculated conditional upon α = .4.

TABLE 4

PROBIT ESTIMATES 1974 AND 1976 POOLED*

Variable	Coefficient	t-ratio**
Constant	.43	3.5
Presidential performance	.82	11.8
Financial condition - good	.01	.2
- poor	.03	.3
Government's economic		
performance - good	.02	.2
- poor	-.03	.5
Economic expectation - good	-.04	.5
- poor	-.13	1.1
Civil rights - too fast	.04	.8
- too slow	-.21	1.9
Dummy (1976 = 1)	-.42	4.9

α = .5

n = 1244

ln L = -2265.88

$\hat{R}^2$ = .16

Estimated Cutoffs	Standard Errors
0.0	n.a.
.7	.04
1.0	.04
1.3	.05
1.6	.05
2.2	.06

*Data for independent variables taken from 1976, 1974, and 1972 surveys.

*Calculated conditional upon α = .5.

TABLE 5

PROBIT ESTIMATES 1974 AND 1976 POOLED*

Variable	Coefficient	t-ratio**
Constant	.34	4.8
Presidential performance	.84	13.2
Civil rights - too fast	.04	.7
- too slow	-.22	2.0
Dummy (1976 = 1)	-.43	6.0

α = .5

n = 622

ln L = -2266.83

$\hat{R}^2$ = .15

Estimated Cutoffs	Standard Errors
0.0	n.a.
.7	.04
1.0	.04
1.3	.05
1.6	.05
2.2	.06

*Data for independent variables taken from 1976, 1974, and 1972 surveys.

**Calculated conditional upon α = .5.

were inclined toward the Republican end of the scale. As before $\hat{\alpha}$ is significantly different from the extreme values, e.g. .8 or .2 ($\chi^2(1)$ = 4.18 and 5.44 respectively). Thus, to the extent that the traditional view can be fairly represented as arguing that α is close to one, then that position is not supported.

Note that though ignoring serial correlation in the residuals and using the usual probit ML method gives us consistent estimates of the regression parameters, there still remains the problem of getting consistent estimates of the standard errors.

Denote by L^* the (pseudo) likelihood function, i.e. the likelihood function one would have by assuming serially independent residuals. Let θ be the set of parameters to be estimated. Then the appropriate covariance matrix for $\hat{\theta}$ obtained by maximizing the pseudo likelihood function is given by:

$$\text{Plim} \left(- \frac{\partial^2 \text{ Log } L^*}{\partial\theta\partial\theta'}\right)^{-1} \left(\frac{\partial \text{ Log } L^*}{\partial\theta} \frac{\partial \text{ Log } L^*}{\partial\theta'}\right) \left(- \frac{\partial^2 \text{ Log } L^*}{\partial\theta\partial\theta'}\right)^{-1}.$$

Though this expression can, in principle, be computed, it is very cumbersome to do so. In the computation of the standard errors we have reported, we have just used the expression

$$\left(- \frac{\partial^2 \text{ Log } L^*}{\partial\theta\partial\theta'}\right)^{-1}.$$

As noted earlier, this expression is correct in the case of a single cross-section and thus the standard errors in Table 1 are consistently estimated. Since the main qualitative conclusions following from Tables 1 and 4 are the same, it would seem that not much would be

gained from the extra computation of the correct expression in the case of the results of the pooled sample presented in Table 4.

An alternative model for political affiliation would be that some people hold strong political opinions and others do not. These transitions are always small for one group while the other will occasionally jump from one extreme to the other. The data in Table 6 suggest that the model is not adequate to describe our data. Note that nearly all transitions are to neighboring cells and there are almost no transitions from one extreme to the other.

TABLE 6

PARTY ID TRANSITION 1972-1974 AND 1974-1976

After Transition

Before Transition: Party ID	0	1	2	3	4	5	6	Total
0	147	51	9	5	3	3	0	218
1	57	149	33	11	14	6	0	270
2	23	32	54	20	9	6	0	144
3	6	8	23	55	19	9	2	122
4	3	12	10	21	64	37	13	160
5	6	6	3	8	31	85	26	165
6	2	1	2	3	9	43	105	165
Total	244	259	134	123	149	189	146	1244

V. SOME MONTE CARLO EVIDENCE

To check the practical usefulness of the estimation procedure we used, we conducted two series of sampling experiments. In each case we generated 100 samples of size 100 each.

In experiment 1, we considered the model:

$$y_t^* = \frac{\beta(1-\alpha)}{1-\alpha L} x_t + u_t \tag{5.1}$$

where L is the lag operator defined as $Lx_t = x_{t-i}$. We set $\beta = 1$ and generated x_t as IN(0,1). u_t were $IN(0,\sigma_u^2)$. The variance of u_t was changed for different values of α so that the variance of the systematic part in (5.1) which in this case is $(\frac{1-\alpha}{1+\alpha})$ is four times σ_u^2. The implied theoretical R^2 for equation (5.1) is thus 0.8 in all cases.

The model given in (5.1) was estimated for two cases:

(i) y_t^* observed as a continuous variable

and (ii) y_t^* observed as a dichotomous variable defined as:

$$y_t = 1 \quad \text{if } y_t^* > 0$$
$$= 0 \quad \text{otherwise.}$$

For estimation purposes we write (5.1) as:

$$y_t^* = \beta x_t^*(\alpha) + \alpha^t \eta + u_t \tag{5.2}$$

where $x_t^*(\alpha) = \left[\sum_{i=0}^{t-1} \alpha^i x_{t-i}\right](1-\alpha)$ and $\eta = E(y_0^*)$ is another parameter to be estimated.

In the case where y_t^* is observed as a continuous variable, we estimate (5.2) by searching over α, i.e. estimating (5.2) by OLS for each value of α and choosing the value of α for which the residual sum of squares is minimum. In the case y_t^* is observed as a dichotomous variable, we use the same procedure except that equation (5.2) is estimated by the probit ML method. We choose the value of α for which the likelihood is maximum. Note that β is estimable only up to a scale factor. The search was conducted at intervals of .025 over different values of α.

The summary results are presented in Table 7. We used three values of α in the experiments.

In experiment 2, we considered the model:

$$y_t^* = \alpha y_{t-1}^* + \beta x_t + u_t. \tag{5.3}$$

Again we used the same parameter values as in experiment 1 for β and variance of x_t. The only difference is that the variance of u_t was not varied with α. If we rewrite the model in (5.3) as

$$y_t^* = (\frac{\beta}{1-\alpha L})x_t + (\frac{1}{1-\alpha L})u_t \tag{5.4}$$

the ratio of the variances of the systematic part and the error, with the specifications we made, is constant for different values of α.

For the case where y_t^* is observed as a continuous variable, we estimate equation (5.3) by OLS, since u_t are serially independent. In the case where y_t^* is observed as a dichotomous variable, we estimate equation (5.4) by the probit ML for different values of α, as in the case of experiment 1 but ignoring the serial correlation in the residuals.

The results of these experiments are also presented in Table 7.

The results of experiment 1 shed light on how much information is lost in the fact that y_t^* is observed only as a dichotomous variable rather than a continuous variable. The bias terms are of comparable magnitude and the variances of α when y_t^* is observed as a dichotomous variable are about 2.5-3.0 times the corresponding variances where y_t^* is observed as a continuous variable.

The results of experiment 2 shed light on the consequences of two factors:

(i) y_t^* is observed as a dichotomous variable rather than as a continuous variable.

(ii) The serial correlation in the residuals in (5.4) is ignored in the probit ML estimation of (5.4).

Again, the bias terms are not large. The variances in the case y_t^* is observed as a dichotomous variable are about 2.5-4 times the corresponding variances when y_{it}^* is continuous.

These results suggest that the estimation procedure we used in our empirical work is expected to perform well for the sample sizes we had.

TABLE 7

RESULTS OF SAMPLING EXPERIMENTS

Experiment 1:

Model: $y_t^* = \frac{\beta(1-\alpha)}{1-\alpha L} x_t + u_t$

α	y_t^* Continuous			y_t^* Observed as Discrete		
	Mean	Bias	Variance	Mean	Bias	Variance
.5	.498	-.002	.001526	.502	+.002	.004153
.6	.598	-.002	.001409	.598	-.002	.003727
.7	.698	-.002	.001205	.699	-.001	.003508

Experiment 2:

Model: $y_t^* = \alpha y_{t-1}^* + \beta x_t + u_t$

α	y_t^* Continuous			y_t^* Observed as Discrete		
	Mean	Bias	Variance	Mean	Bias	Variance
.5	.499	-.001	.00146	.502	+.002	.004805
.6	.601	+.001	.00121	.603	+.003	.004458
.7	.699	-.001	.00099	.704	+.004	.003392

VI. SUMMARY AND CONCLUSIONS

In this paper we have discussed a number of dynamic models with qualitative variables. For one of these models, the lagged index model, we have proposed a method of obtaining consistent estimates of all the parameters. The method was applied to some United States panel data relevant to the issue of the stability of preferences for political parties. The evidence supports the current view that party identification is subject to short-term fluctuations. Monte Carlo calculations suggest that the method should work reasonably well in practice.

REFERENCES

Amemiya, Takeshi. 1973. "Regression Analysis When the Dependent Variable Is Truncated Normal." Econometrica 41 (November): 997-1016.

Brody, Richard. 1977. "Stability and Change in Party Identification: Presidential to Off-Years." Paper presented at 1977 annual meeting of the American Political Science Association. Mimeographed.

Campbell, A.; Converse, P.; Miller, W.; and Stokes, D. 1960. The American Voter. New York, Wiley, 1960

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." Review of Economic Studies 47:225-238.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In Ideology and Discontent, edited by D. E. Apter. Glenco, Illinois: Free Press.

Converse, Philip E. and Markus, Gregory B. 1979. "Plus ça change . . .: The New CPS Election Study Panel." American Political Science Review 73:32-49.

Dreyer, Edward. 1973. "Change and Stability in Party Identification." Journal of Politics 25:712-722.

Fiorina, Morris P. 1979. "Retrospective Voting in American National Elections." Mimeographed. Pasadena: California Institute of Technology.

Grether, David M. 1974. "Correlations with Ordinal Data." Journal of Econometrics 2:241-246.

__________. 1976. "On the Use of Ordinal Data in Correlation Analysis." American Sociological Review 41:908-912.

__________. 1979. "Bayes Rule as a Descriptive Model: The Representativeness Heuristic." California Institute of Technology Social Science Working Paper No. 245. Quarterly Journal of Economics, forthcoming.

Heckman, James. 1979. "Statistical Models for Discrete Panel Data." Center for Mathematical Studies in Business and Economics Report No. 7902, University of Chicago.

Howell, William C. 1967. "An Evaluation of Subjective Probability in a Visual Discrimination Task." Journal of Experimental Psychology 75:979-986.

__________. 1971. "Uncertainty from Internal and External Sources: A Clear Case of Overconfidence." Journal of Experimental Psychology 89:240-243.

Kahneman, D. and Tversky, A. 1972. "Subjective Probability: A Judgment of Representativeness." Cognitive Psychology 3:430-454.

Klein, L. R. 1958. "The Estimation of Distributed Lags." Econometrica 26:553-565

Lee, L.-F. 1980. "Strong Consistency of the Tobit Estimator in the Presence of Serially Correlated Disturbance." Mimeographed. University of Minnesota.

Robinson, P. M. 1980. "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables." Mimeographed. Vancouver: University of British Columbia.

Tversky, A. and Kahneman, D. 1974. "Judgment under Uncertainty: Heuristics and Biases." Science 185:1124-1131.