

ISSN 1183-7284
ISBN 0-7714-2238-5

**Theoretical and Analytical Aspects
of Longitudinal Research**

by
Fernando Rajulton
Zenaida R. Ravanera

Discussion Paper no. 00-5

May 2000

On the web in PDF format: <http://www.sscl.uwo.ca/sociology/popstudies/dp/dp00-5.pdf>

**Population Studies Centre
University of Western Ontario
London CANADA N6A 5C2**

Theoretical and Analytical Aspects of Longitudinal Research

Fernando Rajulton and Zenaida R. Ravanera
Population Studies Centre

This paper has a three-fold aim: To outline briefly the historical development of ideas that have given rise to the current interest in longitudinal studies, particularly in the social sciences; to discuss relevant theoretical considerations when analyzing longitudinal data; and to illustrate approaches toward analyzing longitudinal data when measured as binary and categorical variables. The overall aim is to promote a better understanding of the type of information that longitudinal data provide and of the appropriate techniques needed to analyze such data.

A. Longitudinal Research -Development of Ideas

Although the very first longitudinal study is traced back to 1759 when Gueneau de Montbeillard recorded his son's growth to the age of 18 (Buffon, 1837), questions on advantages and disadvantages of longitudinal research were raised as early as 1920s when for the first time criticisms were directed against cross-sectional methods. In essence, most of those arguments were to discard age as a definition of a population and replace it with duration since the latter, it was argued, was the one that could explain growth. The key ideas associated with a longitudinal study were already contained in those arguments, and these ideas gave rise to the quest for developmental sequences and their inter-relations.

More elaborate discussions on advantages and disadvantages of longitudinal research had to wait for four more decades. Two important studies are worth citing here. One is that of an eminent psychologist, Rene Zazzo (1967) of the University of Paris (a paper presented in 1966 to the Symposium on Longitudinal Studies) and the other is the Report of the National Foundation for Educational Research that was commissioned in 1967 by the Social Science Research Council in UK to carry out a review of longitudinal studies and to identify the distinctive contribution that studies of this kind could make to the development of the social sciences (Wall and Williams, 1970). After a careful and scrupulous evaluation of the advantages and disadvantages of longitudinal studies, both of them were finally not very much in favor of pursuing longitudinal studies - for a variety of reasons that are valid even today and will be pointed out in the subsequent pages. As the Commission's Report finally read: "We do not share the pessimism of many, but we are still aware that unbounded optimism as to the outcome of continued study of representative national or regional samples of the same individuals over time is *not*

now and probably *will never be justified*" (Wall and Williams, 1970:70 , italics mine).

Many other studies since then have clarified the basic distinction between cross-sectional and longitudinal data that we readily accept today but not necessarily apply in practice. Cross-sectional information deals with *status*, while longitudinal information is concerned with *progress and change* in status. Sir Cyril Burt, who is well known for his innovative techniques in the 1920s, used the term "*conspective*" for cross-sectional as opposed to "*prospective*" for longitudinal data. Zazzo proposed that a study which reexamines the same population at recurring intervals (that is, longitudinal) was nothing but "*evolutive transverse*" (that is, cross-sectional-developmental).

We now understand the term "longitudinal" as one which is based on repeated measurements of the *same* individuals over a time span *long enough to encompass a detectable change* in developmental status. The span of time over which observations are made depends therefore on the issue of investigation and its rate of change. The major virtue of any longitudinal information is that it is inherently socio-or-psycho-dynamic. Longitudinal studies are meant to uncover that dynamism which can be done only by examining both stability and change, and not one or the other. The general motivation underlying our interest in longitudinal studies is that they can show the nature of growth, trace patterns of change in an individual, and *possibly* give a true picture of cause and effect relationships over time.

We often tend to view, and define, the term "longitudinal" as an antonym of "cross-sectional", and yet do not hesitate to apply to the former the design and analytical procedures devised for the latter. Zazzo's criticisms were precisely based on this anomaly, and he aptly pointed out (even in the 60s) how longitudinal studies invariably go for larger and larger samples when a few subjects would suffice for the discovery of developmental sequences, and how standardized tests (involving rigidity in theory and in instrumentation) are invoked in such studies while their essential aim is freedom in search of the unknown. He dubbed them all as illusions, the most perverse of them being the "imperative that length of observation and of development cannot be coincident and coterminous if the nature and rhythm of the latter are to be revealed - the ancient fallacy that confounds the reality of what is observed with the process of observation" (Wall and Williams, 1970:16).

Though in general researchers accept that longitudinal information is necessary especially for studies on individual behaviour and attitudes, debates on advantages and disadvantages of longitudinal versus cross-sectional studies continue even today. It is expensive to collect longitudinal data, in terms of

money, time and energy; it requires more complex and unfamiliar statistical procedures to analyze the collected data; and, so few computer software are available to do the job properly. What is the point, then, of all the trouble and expense to collect data of such richness when that richness cannot be tapped? Those who are content with cross-sectional data argue that after all cross-sectional data are not uninformative about the processes of change. For example, most censuses ask questions on where respondents lived one year or five years earlier and we are able to make inferences on changes that have taken place during the intermediate period. Demographers in particular have for long been collecting at least a partial demographic history through cross-sectional surveys. Similar tactics can be used to capture many other processes of change.

Debates on advantages and disadvantages of cross-sectional versus longitudinal data can and will continue *ad infinitum* unless and until researchers can show that longitudinal data do make a difference in social research. Three decades have passed since the debates on the issue started and things have definitely changed during these three decades. Although we are still discussing about the costs and benefits of collecting longitudinal data in the same way as it was done three decades ago, we can see a remarkable shift in researchers accepting longitudinal studies as more appropriate and even required for certain kinds of research. These include, but not limited to, such topics as the causes of changes or the role of changes in attitudes. Furthermore, in Canada, as in many other developed nations, prospective surveys are in the process of replacing transversal (and even retrospective) surveys. In such an atmosphere of acceptance and trend toward collecting more and more longitudinal data, it is time to set aside debating on the costs and benefits of such endeavours and instead expend one's energy on tapping the rich information being collected. Longitudinal surveys are here to stay and social scientists are fortunate to have rich data sets at their disposal, thanks to the generous funding gestures of many governments of developed nations and of many research funding organizations.

In spite of the general acceptance of usefulness of longitudinal data, it is not an exaggeration to say that many researchers are not ready to use adequate techniques of analyzing such data. This situation cannot be rectified unless we find a way to disseminate new and correct techniques of analysis to would-be users of longitudinal data. This paper is a modest attempt towards achieving that purpose by presenting some useful paradigms and initial steps at undertaking longitudinal analysis.

Before getting into this task, a few observations to temper the "unbounded optimism" are in order. First, one of the most frequently advanced arguments in favour of longitudinal data is the prevalence of ambiguities in causality with cross-sectional data. This ambiguity can arise in many different ways and it

is still not very clear how longitudinal data will help “solve” some of these ambiguities. For instance, social research is replete with examples of uncertainties about the direction of causality. The uncertainty problem is somewhat serious in attitudinal research that examines the relationship between attitudes and behavior. If data from different waves (on the *same* individuals) are available, the direction of causality has a better chance of being identified. This is especially true in the case of a non-recursive causal relationship (that is, X to Y as well as Y to X) as frequently happens in attitudinal research:- attitudes influence behavior and behavior results in modification of attitudes. Or it can happen even in simpler contexts such as the Malthusian vicious circle:- increased food supply per capita leads to increased fertility and increased fertility leads in turn to decreased food supply per capita. It is obvious that such a non-recursive causal relationship cannot be clearly established with cross-sectional data and we hope that longitudinal data will help solve the problem. Will longitudinal information be able to prove such a relationship? Theoretically yes, but in practice, there are many doubts, because even some simple questions relating to the modes of collecting longitudinal data do not have clear answers; questions like what is the optimal length of time between interviews, what is the number of interviews (or waves) necessary to achieve the research objectives, how long should the data collection be done before an appreciable change is observed or a causal mechanism is identified, and so on.

Second, closely connected to the above discussion of causality is the requirement that there should be no other plausible explanations for the statistical association. It is practically impossible, even with longitudinal data collected over many waves, to satisfy this requirement. A simplistic and textbook approach to causal analysis may be content with examining a few control variables included in the analysis, but it is quite obvious to any serious researcher that social processes are too “noisy” to yield to any strong evidence of causal relationships. Not only it is impossible to identify all the potential sources of an observed relationship but also many of these potential sources are effectively unmeasurable. Traditional analytical approaches have not been of great help in these circumstances, and newer (more sophisticated) approaches (like unobserved heterogeneity) have neither succeeded in disentangling the real causal relationships.

Third, from a practical point of view, we hope that longitudinal analysis can shed some light on the “true” importance of certain explanatory variables that we have always taken for granted as important in explaining social behaviour and on which we usually frame our social policies. It is well known that cross-sectional analyses consistently over-estimate the importance of explanatory variables, because they do not account for the inherent, and usually strong, behavioural inertia built into a social system. Our analytical procedures need to consider both stability and change in a system before we can boldly

assert how important certain variables are and what behavioural changes can be induced by policies based on the knowledge of impact of such variables. Longitudinal data can provide pictures of both stability and change, and there is a lot of potential here if, of course, proper techniques are used to tap that information.

Fourth, with panel surveys, we need to examine seriously the so-called “panel conditioning” or Hawthorne effect whereby the very act of being interviewed and reinterviewed changes people’s attitudes and behaviour, if not a simple reporting of attitudes and behaviour. If there is a possibility that individuals can modify their behaviour because of the very fact of being included in the study, then the sample may become less and less “randomized” over time. It is also known that panel conditioning may affect the quality of the data reported by the participants in a study:- the longer their participation in the study, the less likely they are to report certain socially unacceptable situations such as unemployment or mental health. Unfortunately, the literature on the impact of conditioning is sparse and there is definitely an urgent need to examine this in depth with the already existing multi-wave surveys. To do it properly, however, we need a control sample alongside the panel, which implies increase in the costs of conducting a study. A rotating panel design may possibly lessen the problem of conditioning since it will replace the panel members regularly. Readers can refer to Waterton and Lievesley (1988) for a detailed discussion on panel conditioning and for some findings from the Social Attitude Panel Study in Germany.

B. Some Useful Paradigms for Longitudinal Studies

1) Planned Behaviour

Since the main purpose of collecting longitudinal data is to follow the socio-psycho-economic development or behavior of individuals over time, the paradigm of planned behaviour (generally used by psychologists) may help explain behavioural changes and adaptation. Demographers have also used this paradigm in many contexts although with different terminologies, starting from the explanations offered for the demographic transition.

A central idea in this paradigm is the individual’s *intention* to perform a given behaviour. Intentions are assumed to capture the *motivations* underlying a behaviour. Psychologists postulate three conceptually independent determinants of intention: a) *Attitude toward the behaviour* which refers to the degree to which the person has a favourable or unfavourable evaluation of the behaviour; b) a social factor called *subjective norm* which refers to the perceived social pressure to perform or not to

perform the implied actions; and, c) the degree of *perceived behavioural control*, which refers to the perceived ease or difficulty of performing the behaviour depending on the past experiences of the individual as well as anticipated impediments, obstacles, costs and benefits.

These three determinants usually have a specific ordering:- the more favourable the attitude, the greater the perceived behaviour control and the stronger will be an individual's intention to perform. Similarly, the greater is the social pressure, the greater is the perceived control and the stronger the intention. And, the stronger the intention is, the more predictable is a given behaviour. These ideas are not be foreign to demographers who are familiar with the three "determinants" of contraceptive practice: ready, willing and able (Coale, 1973; Lesthaeghe and Vanderhoeft, 1998).

Reinecke and Schmidt (1996) apply this paradigm of planned behavior to examining use of condoms in new sexual contacts in a study of AIDS-risk function, but in a relevant context of missing value treatment. In the experience of many researchers, missing data is one of the practical problems encountered in working with this paradigm (or with the "ready-willing-able" paradigm of demographers- see for example the MA thesis of Mannan (1999)). In addition, longitudinal data bring in the additional dimension of stability and change; intentions, norms and control can all vary over time. Such circumstances can enrich our understanding of causal mechanisms as discussed in the previous section, particularly the mutual influence of each "determinant" on the other two.

2) Life Course Paradigms

Longitudinal data typically provide information on individuals' life courses encompassing various domains of interest. It is no wonder then that theoretical and methodological developments in life course studies have been greatly enhanced with the availability of longitudinal data. Now we are able to examine how individual lives are shaped by personal characteristics as well as by social environment - a clear departure from the emphasis on the former for so long. Featherman and Lerner (1985) envisage a "*developmental contextualism*" that would enable a study of "person-population" processes in the near future. Giele and Elder (1998) have addressed ways of using life events and their timings in a systematic way both theoretically and methodologically.

Cain (1964) used the term "*life course*" to encompass anthropological, sociological, and psychological concepts of "aging" or maturing through an expected sequence of social roles. According to him, the life course patterns in all domains of society (such as the family, the polity, economy and religion) are characterized by "age status synchronization" whereby certain people follow different trajectories from

others. The term (life course) is therefore used in reference to a sequence of socially defined events and roles that an individual experiences over time, in distinction from the term *life cycle* that allows for events and roles that do not necessarily proceed in a socially defined sequence (Elder, 1975).

The concept of the life course helps in dealing simultaneously and adequately with the once problematic age-period-cohort effects because life course links them all together in one perspective: the distinctive historical and cultural events (= period) experienced by persons who not only develop according to a biological clock (= age) but also socially share these experiences with their peers (= cohort). Giele and Elder (1998) suggest a framework consisting of four links that connect all these major elements of life course studies: 1) ***Location in time and place*** or cultural background (period aspect); 2) ***Linked lives*** or social integration (cohort aspect); 3) ***Human agency*** or individual goal orientation (age aspect); and, 4) ***Timing of life events*** or strategic adaptation (longitudinal aspect).

In their view, the fourth dimension, namely timing, integrates the previous three, namely historical, social and individual activities. Obviously, then, timing of life events becomes the corner stone of all life course studies, and longitudinal designs are the primary methods for facilitating such studies. Therefore, either through retrospective or prospective designs, data should be collected, whenever possible, on historical contexts, family, work and social settings, health, well-being, goals and life satisfaction, along with major timings of events.

The life course paradigm is also a very useful tool in examining stability and change not only in individual behaviour but also in structural behaviour. Both stability and change are integral parts of the life course, because as change occurs in a systematic way, it will make for an either stable or changing pattern. In post-war developed societies, for example, increased prosperity, changing value systems and ever-burgeoning individualization are indicators of on-going change, and yet much structural continuity still remains at the same time; for instance the rather stable structure of social inequality.

Change is the result of a tension between the programming by society and the individual's own choice (Kohli, 1986). Thus, the connection between micro-level aging over the life course and macro-level institutional change can be observed through the changing life course patterns of successive cohorts. Some cohorts introduce an "***innovation***" in the life course that usually departs from tradition and prefigures an institutionalization of that way of life so that it becomes the standard (Kohli, 1986). Such innovations, though first viewed as deviance, create new milestones for individuals living in changing times, and at the same time demand the adjustment or creation of social policies that will accommodate and institutionalize the new life patterns. Thus, changes in individual lives can, and often do, result in

structural changes as well. We can think of changes in the roles of women, retirement age, transition to adulthood, family types, etc. as typical examples.

C. Illustrations of Possible Types of Longitudinal Analysis

Longitudinal data consist of time-sequences of measurements, counts or categorical responses from the same (experimental) units. Thus, they are closely related to *time-series data*. Techniques used for time series data, however, may not adequately capture the rich information contained in longitudinal data. [Readers interested in time-series approach to analysing longitudinal data can refer to Jones (1993).] In addition to time series characteristics, longitudinal data have some important aspects which call for different techniques of analysis. It is usually assumed that each unit in the sample comes from an underlying population, information on socioeconomic covariates is usually available and that the response patterns vary with time, treatment or covariates. Unlike time series data, longitudinal data have inherent *state dependencies* that can be examined more adequately only with stochastic frameworks. In particular, the importance of *initial conditions* can never be overstated in analysing longitudinal data. The initial conditions clearly vary with the research designs and can have profound consequences on inferences made from such data. For instance, the initial conditions depend very much on the starting and ending dates of a study (that is, in which status we start observing individuals or in which status we end our observation) which in turn are usually beyond a researcher's control and mostly dependent on the availability of funds! An example below will try to illustrate the importance of initial status even in multivariate contexts. This paper does not illustrate in greater detail the types of life course analysis that can be done with stochastic frameworks (since many readers of this paper may already be familiar with our earlier works), but we suggest to interested readers a visit to our web-site at: <http://www.sscl.uwo.ca/longitudinal/index.html>.

Three salient features are hidden in any type of longitudinal data: subject behaviour, time behaviour and covariate behaviour. Longitudinal studies worth the name should necessarily unpack these various effects either through a pure stochastic model or a model that incorporates structural parameters. Different disciplines approach the problem in different ways, and a challenge facing researchers is to develop a fully unified approach. This section is an attempt to achieve this purpose.

1) Exploratory Techniques: The first step in longitudinal analysis is exploration as in basic statistics. It would be a big mistake to rush on to multivariate analysis without having adequately examined the nature of data at hand. To bring out the distinct information buried in longitudinal data, plots are very

useful, particularly those that can display all the data points over time; for example, Parallel Plots and *Trellis Graphs* (Cleveland, 1993). Not all statistical packages may be able to produce these plots, one may have to use packages such as S-Plus for a good explorative study. Trellis graphs have certain technical features such as “main effects ordering” that help us find hidden relationships in multivariate data sets and how two variables change with one or more “conditioning” variables.

2) Stability and Change: As discussed in the previous sections, these are the two essential characteristics of longitudinal studies in that they encapsulate the distinct advantage that longitudinal data have over other types. The important point is that these two dimensions go together. As far as change occurs in a systematic way, there will always be a stable pattern underneath. Change occurs not only at the individual level but also at the system level. All the changes that we have observed at the individual level since World War II have not taken place in a vacuum. Increased opportunities led to social mobility, which in turn led to changes in value systems and then to changes in individual behaviour. In spite of all these changes, one has to admit that some structural continuities remain. As Engel and Reinecke (1996) point out, a distinct case in point is social inequality in spite of all the progress made in almost all fronts, be they social or economic or cultural.

How, then, to understand the simultaneous occurrence of stability and change? As Coleman (1990) argues, explanations of system behaviour should be based on knowledge about its component parts *below* the system level. How much a system changes depends on how change takes place at a lower level (i.e.) at the individual level. There is no system change if there is no change at the individual level or if there are opposing individual changes that cancel each other out. And, any change at the individual level should necessarily be “diffused” to other individuals such that the system recognizes that a change has taken place and tries to modify itself. Thus, the concepts of stability and change can be viewed as mutually influencing forces of systems on individuals and individuals on systems. Hence, the importance of life course and/or genuine two-way multilevel analysis while working with longitudinal data.

Some societies are characterized more by change while others more by stability. More developed societies are good examples of the former while less developed ones are good examples of the latter. If a society is characterized more by change, then there is a greater need for longitudinal research to examine the processes of both change and stability.

In studying processes of change and stability, researchers will have to deal with complexities arising from independent and/or *dependent samples*, missing information, measurement error, observed and

unobserved heterogeneity, and so on. All these topics cannot be dealt with adequately in a paper like this, but some points will be addressed in specific contexts that give rise to these complexities..

As a simple illustration for examining stability and change, we present here the so-called *turnover tables* used for analysing categorical data (McCutcheon, 1996). Turnover tables are categorical tables in which response items are classified by earlier and later time points (Time 1 = Row, Time 2 = Column). An example (that was submitted as an assignment by one of my students) is given in Table 2.1 which cross-classifies the observed frequencies of employment status of women living as single-parent or two-parent families in London Ontario. The data are taken from a longitudinal study conducted by Avison (1996); in the first wave conducted in 1991, 1020 interviews were done and in the second wave in 1993, 91% of the respondents were successfully reinterviewed. Even though a two year gap is too short to observe large number of changes, this can still serve as an illustration for examining stability and change with such simple cross-tabulations. This type of analysis can lead to greater understanding of stability and change especially when data from additional waves are also included.

Table 2.1: Observed frequencies of Employment status in 1991 and 1993

		Year 1993					Total	
		Full-time	Part-time	Unemployed	Homemaker	Others		
Year 1991	Full-time	372	31	22	8	9	442	
	Part-time	31	120	15	19	10	195	
	Unemployed		19	17	10	8	4	58
	Homemaker		9	31	16	107	22	185
	Other		9	9	8	15	27	68
Total		440	208	71	157	72	948	

As seen from the table, the majority (67%) of women had the same employment status in both the years, while 33% changed their employment status. Of those who were working full-time in 1991, 70 changed their status, moving primarily to part-time (31) and to unemployed (22) category. The marginal distributions suggest that women who were homemakers in 1991 experienced the largest net decline during the interval whereas the categories Part-time and Other had the greatest net increase.

Categorical data of the above sort can be analysed in different steps for examining stability and change implied in the data:

Step 1: As a first step, we can test the assumption of *independence* using the χ^2 statistic in the traditional manner. The independence hypothesis assumes that employment status in 1991 is wholly unrelated to that in 1993, in other words, there is *no systematic relationship* between statuses at time 1 and time 2. Although this hypothesis can be examined using the traditional packages like SPSS, we shall use the package *IEM* (log-linear and event history analysis with missing data) written by Vermunt (1997) for the sake of using the same package for other types of analyses in the following steps (SPSS and other packages cannot be used for the following types; *IEM* can be downloaded from the Tilburg University's webpage: http://cwis.kub.nl/~fsw_1/mto/mto_snw.htm#software). *IEM* uses the log-linear format and produces additional parameters that can be interpreted in a meaningful way. The results from *IEM* for test of independence are shown in Table 2.2.

Table 2.2: Results of test of independence assumption
 $\chi^2 = 836.26$ $df = 16$

	Status	$\hat{\alpha} = \log\text{-odds}$	$\exp(\hat{\alpha}) = \text{odds}$
1991:	Full-time	1.8718*	6.5
	Part-time	1.0535*	2.9
	Unemployed	-0.1591*	0.8
	Homemaker	1.0008*	2.7
	Other	0.	1.0
1993:	Full-time	1.8101*	6.1
	Part-time	1.0609*	2.9
	Unemployed	-0.0140*	1.0
	Homemaker	0.7796	2.2
	Other	0.	1.0

* = significant at 1% level.

A χ^2 value of 836 for 16 degrees of freedom rejects the hypothesis of independence and confirms that there is a systematic relationship between statuses at the two time points. For each status and for each time point, *IEM* produces parameters that denote the log-odds of *being* in a status; one can also interpret them in terms of odds, if one likes. Since we have rejected the independence model, the interpretation of these parameters is not very useful, but for the sake of illustration we shall do so here. Thus, for example, for the year 1991, the odds of being in a full-time job is 6.5 times greater than the odds of being in the last reference category "other". By 1993, however, these odds declined slightly to 6.1. Similar interpretations can be given for other parameters. Overall, we find that in relation to the same reference group, the most pronounced change in employment status between 1991 and 1993 occurred among homemakers; that is, in 1991 the odds of being a homemaker was 1.25 times greater than in 1993 (2.72/2.18).

Step 2: Once the hypothesis of independence is rejected, we can go for *quasi-independence* hypothesis that focuses *only on those who change their status* between the two time points; in other words, the quasi-independence model is a variant of the independence model, restricting the analysis only to the frequencies that are off of the main diagonal (as if structural zeros are placed on the main

diagonal). Thus we test the hypothesis that there is *no systematic change* (as distinct from *relationship* in the independence model) from time 1 to time 2.

As seen in Table 2.3, the quasi-independence hypothesis must also be rejected ($\chi^2 = 39.47$). The substantial decline in the values of the two χ^2 statistics when the five cells of the main diagonal were set to zero indicates that the quasi-independence model offers a much better fit to the data than the independence model. The parameter estimates given in the table should now be interpreted in terms of “changing” from and to, and not merely in terms of “being” as in the independence model. The parameters associated with 1991 indicate the odds of changing *from* the specific job status and those of 1993 indicate the odds of changing *into* the specific status. For example, *among the changers*, the odds of moving from the “homemaker” category (1.98) is almost twice the odds of moving from the “other” category”. The 1991 parameter estimates indicate that the changes are most likely to originate from the “part-time” category and least likely to originate from the “other” category. Similarly, the 1993 estimates indicate the changes are most likely to end in the “part-time” category, and so on.

Table 2.3: Results of quasi-independence assumption
 $\chi^2 = 39.47$ $df = 11$

	Status	$\hat{a} = \log\text{-odds}$	$\exp(\hat{a}) = \text{odds}$
1991:	Full-time	0.6480*	1.9
	Part-time	0.8294*	2.3
	Unemployed	0.2188*	1.2
	Homemaker	0.6855*	2.0
	Other	0.	1.0
1993:	Full-time	0.5440*	1.7
	Part-time	0.8616*	2.4
	Unemployed	0.3377*	1.4
	Homemaker	0.2475	1.3
	Other	0.	1.0

* = significant at 1% level.

If the quasi-independence hypothesis is rejected, then we can further turn to hypotheses related to systematic change. Here comes the *symmetry* or *quasi-symmetry* models. The former hypothesizes that all changes between 1991 and 1993 are completely balanced. Under this hypothesis, the expected values in the main diagonal will be the same as the observed values (that is, $E_{ii} = f_{ii}$), while the expected values of the off-diagonal elements will be the average of the observed values in the two cells symmetric to the diagonal (that is, $E_{ij} = 0.5 * (f_{ij} + f_{ji})$). In addition, the symmetry model also assumes *marginal homogeneity* which implies that the marginal distributions do not change because changes from one job status to another will be offset by equal changes from the latter to the former. In other words, the symmetry model implies that no status experiences a net gain or loss between the two time periods.

Table 2.4 presents the log-linear results of the symmetry model. The model fits the data quite well ($p=0.47$) with a significant improvement over the quasi-independence model. The symmetry hypothesis cannot be rejected and therefore we can conclude that i) none of the job categories experience a gain or loss from 1991 to 1993 (marginal homogeneity); that is, the *proportion* of women in each of the five job statuses remained constant from 1991 to 1993; and, ii) changes from one status to another was balanced by the changes in the other direction. The parameter estimates given on the main diagonal of Table 2.4 indicate the odds of retaining the same job status, and those in the off-diagonal indicate the odds of changing across

the main diagonal. For example, the odds of a woman having a full-time job in 1991 and 1993 is about 3 times higher than the odds of having a part-time job at both time points. Similarly, the odds of a woman having a part-time job in 1991 and being a homemaker in 1993 is 0.42 (which is the same as the odds of being a homemaker in 1991 and being a part-timer in 1993 - symmetry assumption) which in turn is 1.5 times more likely than a part-timer in 1991 becoming unemployed in 1993.

If the symmetry were rejected, then we could go for quasi-symmetry model that drops the marginal homogeneity assumption but retains the cancel effect. In the present case, it is not necessary to do so. Interested readers can find a treatment of all these models in some familiar books for analysing categorical data such as Bishop et al. (1975), Andersen (1990), McCutcheon (1996) and Vermunt (1996, 1997).

3) Binary Response Models : With longitudinal data, we frequently encounter binary sequences or recurrent events. As with a simple binary dependent variable and a number of observed covariates, we can fit a standard logistic model (known as fixed effects model) with any standard packages, but then we would have to ignore the information on repeated measures. Ignoring repeated measures not only would defeat the purpose of collecting longitudinal data but also would ignore a host of other problems

Table 2.4: Results of symmetry assumption
 $\chi^2 = 8.66$ $df = 9$

Log-odds :		1993				
		F	P	U	H	O
1991	F	1.83	-0.65	-1.06	-1.94	-1.89
	P		0.70	-1.31	-0.86	-1.83
	U			-1.78	-1.60	-2.29
	H				0.59	-1.17
	O					-0.78
Odds :		1993				
		F	P	U	H	O
1991	F	6.26	0.52	0.35	0.14	0.15
	P		2.02	0.27	0.42	0.16
	U			0.17	0.20	0.10
	H				1.80	0.31
	O					0.45

all estimates significant at 1% level.

Table 3.1: Insert from the data file wemp.dat from the Sabre package

1.0	1.0	0.0	9.0	1.0	1.0	-12.0
1.0	0.0	0.0	10.0	0.0	1.0	-11.0
1.0	0.0	0.0	11.0	0.0	1.0	-10.0
1.0	0.0	0.0	12.0	0.0	1.0	-9.0
6.0	1.0	0.0	-1.0	0.0	0.0	7.0
6.0	1.0	0.0	0.0	0.0	0.0	8.0
6.0	0.0	0.0	1.0	0.0	0.0	9.0
6.0	0.0	0.0	2.0	0.0	0.0	10.0
6.0	1.0	0.0	3.0	0.0	0.0	11.0
6.0	1.0	0.0	4.0	0.0	0.0	12.0
6.0	1.0	0.0	5.0	0.0	0.0	13.0
6.0	1.0	0.0	6.0	0.0	0.0	14.0
6.0	1.0	0.0	7.0	0.0	0.0	15.0
6.0	1.0	0.0	8.0	0.0	0.0	16.0
6.0	1.0	0.0	9.0	0.0	0.0	17.0
6.0	1.0	0.0	10.0	0.0	0.0	18.0
6.0	1.0	0.0	11.0	0.0	0.0	19.0
6.0	1.0	0.0	12.0	0.0	0.0	20.0
20.0	1.0	0.0	5.0	0.0	0.0	-11.0
20.0	1.0	0.0	6.0	0.0	0.0	-10.0
20.0	1.0	0.0	7.0	0.0	0.0	-9.0
20.0	0.0	0.0	8.0	1.0	1.0	-8.0
20.0	1.0	0.0	9.0	0.0	1.0	-7.0
20.0	1.0	0.0	10.0	0.0	1.0	-6.0
20.0	1.0	0.0	11.0	0.0	1.0	-5.0
20.0	1.0	0.0	12.0	1.0	1.0	-4.0

that naturally arise in longitudinal analysis. This particular illustration has the purpose of highlighting the problem of *unobserved heterogeneity* (or mixture models containing both fixed and random effects) in the case of binary sequences. This specific problem has been over-emphasized in recent literature in spite of the obvious conclusion that there is no adequate and satisfactory solution to this specific problem except a sound theoretical and analytical framework at the very start. This specific illustration shows how even a simple but sound framework makes unnecessary the convoluted procedures suggested in the literature regarding the problem of unobserved heterogeneity.

The package SABRE (Software for the Analysis of Binary Recurrent Events), written by Dave Stott and freely available on the web (at

<http://www.cas.lancs.ac.uk/software/sabre3.1/sabre.html>), is used for this illustration. This package has been chosen for its simplicity as well as its power of analysis. SABRE can be used to examine the effect of a simple heterogeneity term and can use mass-points; it allows building a mixture model of heterogeneity (in its present version it allows only logistic-normal) as well as other interesting models like Mover-Stayer model, lagged response model, and simple Markov model for two states. The data for this illustration is taken from the example data set (wemp.dat) provided with the manual, since it has as many as ten to 14 repeated measures.

This data set looks at the relationship between the employment status and a set of explanatory variables for 155 married women. For each woman, data are recorded annually, with a total of 1580 measurements (i.e. varying number of records for each woman). The variables in the data file are:

- 1) CASE - individual identifier
- 2) FEMP - wife's employment status; 1 = employed, 0 = unemployed
- 3) MUNE - husband's employment status; 1 = unemployed, 0 = employed
- 4) TIME - calendar time (year-1975)
- 5) UND1 - children aged < 1 year old; 1 = yes, 0 = no

- 6) UND5 - children aged < 5 years old; 1=yes, 0=no
 7) AGE - age (years-35)

Table 3.1 lists the variables in this data file along with the listing of a few cases. As seen from the insert, the first woman has four records, the second woman has 14, the third woman has 8 records.

Table 3.3. Logistic-Normal mixture model for women's employment

Deviance = 1237.1568, df = 1574

Parameter	Estimate	S. Error
int	2.0098	0.19833
fmune (1)	0.	ALIASED [I]
fmune (2)	-2.7033	0.42577
fund5 (1)	0.	ALIASED [I]
fund5 (2)	-2.6619	0.23273
scale	2.1082	0.19258

First, we fit a standard logistic regression to examine the probability of a woman being employed given the individual characteristics : whether the husband is employed and whether the woman has a child under 5 years. Table 3.2 displays the model and the parameter estimates.

As the results indicate, if a husband is unemployed, the woman is also less likely to be employed, and the presence of children under 5 lowers the probability of a woman's employment. Both the covariates are highly significant. As usual, we can use these estimates to find the probabilities of a woman being employed given the specific characteristics.

Table 3.2. Logistic Regression results for women's employment

Deviance = 1757.36 d.f. = 1577

Parameter	Estimate	S. Error
int	1.3068	0.074415
fmune (1)	0.	ALIASED [I]
fmune (2)	-1.7033	0.23584
fund5 (1)	0.	ALIASED [I]
fund5 (2)	-1.7335	0.12219

With the above information from the logistic fit, we move on to introduce the randomness by incorporating a mixture model for examining the likelihood of women's employment. SABRE allows only logistic-normal mixture, with end-points if desired, and the results from this model are displayed in Table 3.3. Compared to the standard logistic model, this model reduces the χ^2 value ("deviance" in the table) by $1757.36 - 1237.16 = 520.20$ for 3 degrees of freedom. The parameter estimates are now larger with the heterogeneity term included in the model. In addition, we have the scale parameter (2.1082) that is significant. This scale parameter is the unknown standard deviation of the mixing distribution of heterogeneity. With the assumed normal structure, this value means that 95% of subjects would fall within 4.2 (that is, $\pm 2\sigma$) logit units of the overall mean. This range on the logit scale translates into probabilities that range from 0 to 0.98, implying thereby that some women have little chance, while others have very high chance, of employment, given the two covariates - a very wide spectrum of heterogeneity present in the data.

We can proceed in this manner to fit other models that may be considered theoretically relevant; for example, a *lagged logistic-normal* mixture that would consider besides a term for heterogeneity the impact of the initial status on the measurements subsequently observed. This model is therefore related to some important aspects of life history analysis, particularly in a non-Markovian-type framework.

Table 3.4 presents the results obtained by fitting this lagged logistic-mixture model to the data, and the results are very enlightening. First, the model greatly reduces the deviance value thus making it the best model fit. Second, the parameter estimates are still stable even though reduced in their magnitude, except for the intercept which turns negative now. The lag parameter becomes conspicuously distinct and significant. The scale parameter is reduced by half (compared to the previous model), even though still significant. Thus, the initial status or the lag parameter becomes a very important variable; it reduces unobserved heterogeneity by 50%. This reveals to us the importance of doing a Markovian or non-Markovian type of analysis.

The package SABRE allows fitting of Markov logistic-normal mixture model as well. The results from this model are shown in Table 3.5 where the parameter estimates are classified by the initial status whether employed and unemployed. In the case of women who started their history in unemployed status,

having children under 5 is no longer significant and the scale parameter is larger than the one in the lagged model. In contrast, in the case of women who started in employed status, the scale parameter almost disappears and becomes non-significant. This interesting analysis clearly shows that unobserved heterogeneity, if any, is entirely due to the group of women who started their history in unemployed status and that including the initial status in the analysis eliminates unnecessary concerns about unobserved heterogeneity. The lesson from all these exercises is clear: Do not neglect past history in longitudinal analysis.

Table 3.4. Lagged logistic-Normal model

Deviance = 878.94
Deviance decrease from the model
Without the lag term = 220.72 for 1 d.f.

Parameter	Estimate	S. Error
int	-0.81251	0.22417
fmune (1)	0.	ALIASED [I]
fmune (2)	-1.6864	0.41477
fund5 (1)	0.	ALIASED [I]
fund5 (2)	-1.0872	0.23191
lag	3.5967	0.22898
scale	0.93090	0.21221

Table 3.5. Markov logistic-Normal model

Deviance = 846.49

1) Initial Status: Unemployed

Parameter	Estimate	S. Error
int	-1.3981	0.2906
fmune (1)	0.	ALIASED [I]
fmune (2)	-1.9934	0.7470
fund5 (1)	0.	ALIASED [I]
fund5 (2)	-0.2510	0.3512
scale	1.2878	0.2995

2) Initial Status : Employed

Parameter	Estimate	S. Error
int	3.0890	0.1802
fmune (1)	0.	ALIASED [I]
fmune (2)	-1.3567	0.4653
fund5 (1)	0.	ALIASED [I]
fund5 (2)	-1.8990	0.2643
scale	0.1160	0.2767

D. Conclusions

In this paper we have discussed or illustrated the following important points regarding longitudinal analysis. First, longitudinal data provide rich information on stability and change in the behaviour of systems as well as individuals, and these two dimensions should be examined together, not one or the other.

Second, longitudinal data introduce many complexities, and facing these complexities is a great challenge to researchers as well as to those who devise techniques of analysis. One such complexity is that of unobserved heterogeneity that has been much discussed in recent literature. Though its usefulness is under debate, research experience tells us that nothing is more important than a good research design backed by a solid theoretical framework. It is not enough to say, after doing an extensive analysis, that unobserved heterogeneity is important and significant in a study. Researchers have the obligation to show what it points to. Our experience shows that inclusion of past history, no matter how indirectly or how inconspicuously introduced into an analysis frequently eliminates the need for considering the heterogeneity term in an analysis.

Third, a point regarding the type of distribution that can be used in model building that was shown above: If any non-linear program is available or if it is included in any special packages like SABRE, we strongly recommend the use of beta-logistic distribution in contexts of longitudinal research. With beta-logistic, it is possible to examine repeated measures for “heterogeneity in persistence”. That is, with information on the same category identified j out of n times (waves), we include a beta distribution to allow for unobservable persistence over time. A case in point is voters’ persistence in the elections held in Canada; it is well-known that Catholics have been persistently choosing the Liberals in Canada since the beginning of the 20th century.

Fourth, we have intentionally included in the illustrations the case of lower-level variables - categorical and binary variables -, which many have come to consider as less useful than ratio-level variables. Far from it. Many longitudinal surveys in social research still collect much information in categorical or binary form. And the techniques are as “sophisticated” as, say, hazard models and can bring to light many hidden gems in longitudinal data.

Fifth, the dynamic analysis of life histories that examines the impact of past history is one of the best ways to analyse longitudinal data. The key concepts in life history, namely timing and sequence, are the most suitable to analysing any longitudinal data, provided proper measurements are made. We have not provided any illustration of life histories in this paper, as we hope that the readers are familiar with our earlier works on life histories.

Lastly, longitudinal data are the best sources for examining growth, particularly the so-called *intra-*

individual variability. As mentioned earlier, stability and change can be examined both in systems and in individuals. In practice, however, one is pitted against the other in empirical analysis. The main aim of collecting longitudinal data is centred on identifying the growth, development and changes taking place in individuals or groups. Thus, the “intra” aspect gains more importance in longitudinal analysis, this is very much so with repeated measurements. Individuals differ not only from each other (inter-) but also from one time point to another (intra-). The intra aspect can point to either stable or changing characteristics in individuals themselves, either short-term or long-term. In fact, it is intra-variability that contributes to inter-variability observed among individuals, and yet we do all our analyses focussing on inter-variability that is confounded with intra-variability. This topic will be dealt with in a later paper.

References

- Andersen, E.B. 1990. *The Statistical Analysis of Categorical Data*. Berlin:Springer-Verlag.
- Avison, W. 1996. Mental and physical health of mothers and children in single-parent families, London Ontario. The University of Western Ontario, London Ont., Canada.
- Bishop, Y.M.M., S.E. Fienberg and P.W. Holland, 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass.: MIT.
- Buffon, 1837. “Sur l’accroissement successif des enfants; Gueneau de Montbeillard mesure de 1759 a 1776”. *Oeuvres Completes*. Paris: Furne et Cie.
- Cain, L.D.Jr, 1964. Life course and social structure. In R.E.L. Faris (ed.), *Handbook of Modern Sociology*. Chicago:Rand McNally. pp.272-309.
- Cleveland, W.S., 1993. *Visualizing Data*. NJ:Hobart Press.
- Coale, A.J., 1973. The demographic transition reconsidered. Proceedings of the International Population Conference, Liege. Pp.53-72.
- Coleman, J.S., 1990. *Foundations of Social Theory*. Cambridge MA:Belknap Press of Harvard University.
- Elder, G.H., Jr. (1975). Age differentiation and the life course. *Annual Review of Sociology*, 1:165-190.
- Engel U. and J. Reinecke (eds), 1996. *Analysis of Change, Advanced Techniques in Panel Data Analysis*. N.Y.: Walter de Gruyter.
- Featherman, D.L. and R.W. Lerner, 1985. Ontogenesis and sociogenesis: Problematics for theory and research about development and socialization across the lifespan. *American Sociological Review*, 50: 659-676.
- Giele, J.Z. and G.H. Elder Jr., 1998. *Methods of Life Course Research, Quantitative and*

Qualitative Approaches. London:Sage.

Gottman, J.M., 1995. *The Analysis of Change*. UK: Lawrence Erlbaum.

Haberman, S.J. 1979. *Analysis of Qualitative Data*. N.Y.:Academic Press.

Jones, R.H. 1993. *Longitudinal Data with Serial Correlation: A State-space Approach*. N.Y:Chapman & Hall.

Kohli, M., 1986. The world we forgot: A historical review of the life course. In V.W. Marshall (ed.), *Later Life: The Social Psychology of Aging*. Beverly Hills CA: Sage. pp. 271-303.

Lesthaeghe, R. and C.Vandehoef, 1998. Ready, willing and able: A conceptualization of transition to new behavioural forms. Paper presented at the National Academy of Sciences Meetings on the "Social Dynamics of Fertility Change in Developing Countries", Washington DC.

Magnusson, D., L.R. Bergman, G. Rudinger and B. Torestad (eds.), 1991. *Problems and Methods in Longitudinal Research: Stability and Change*. N.Y.:Cambridge University Press.

Mannan, H.R., 1999. Contraceptive usage and method choice in Bangladesh. MA thesis, Department of Sociology, University of Western Ontario, London Ontario.

McCutcheon, A.L., 1996. Models for the analysis of categorically scored panel data. In Engel and Reinecke (eds.), 1996.

Menard, S. 1991. *Longitudinal Research*. London:Sage.

Nesselroade and Featherman, 1991. Intradividual variability in older adults' depression scores: Some implications for developmental theory and longitudinal research. In Magnusson et al. (Eds), pp. 47 seq.

Reinecke, J. and Schmidt, 1996. Model specification and missing value treatment in panel data: testing the theory of planned behaviour in a three-wave panel study. In Engel and Reinecke (eds.), 1996.

Uncles, M.D. (Ed.), 1988. *Longitudinal Data Analysis: Methods and Applications*. London Papers in Regional Science, London.

Vermunt, J. K.,1996. Causal log-linear modeling with latent variables and missing data. In Engel and Reinecke (eds.), pp.35-60.

Vermunt, J.K., 1997. LEM: A general program for the analysis of categorical data. Tilburg University.

Wall, W.D. and H.L. Williams, 1970. *Longitudinal Studies and the Social Sciences*. London: Heinemann Educational Books Ltd.

Waterton and Lievesley, 1988. Attrition, conditioning and attitude change: Some findings from the Social Attitudes Panel Study. In Uncles (ed.), pp.73-88.

Zazzo, R. 1967. Diversite, realite et mirages de la methode longitudinale: Rapport introductif au symposium des etudes longitudinales. *Enfance*,.20:131-136.