

Linking public issues with private troubles: Panel studies in developing countries

Professor Trudy Harpham
South Bank University

Sharon Huttly
London School of Hygiene & Tropical Medicine

Ian Wilson
University of Reading

Professor Thea de Wet
Rand Afrikaans University

Reducing Attrition in Panel Studies in Developing Countries

Zelee Hill
London School of Hygiene and Tropical Medicine



CONTENTS

LINKING PUBLIC ISSUES WITH PRIVATE TROUBLES: PANEL STUDIES IN DEVELOPING COUNTRIES

ABSTRACT	2
1. INTRODUCTION: The Attraction of Panel Studies	3
2. EXAMPLES OF PANEL STUDIES IN DEVELOPING COUNTRIES	6
3. ADDRESSING PROBLEMS OF PANEL STUDIES IN DEVELOPING COUNTRIES	9
3.1 CONCEPTUAL FRAMEWORKS WHICH LINK MACRO AND MICRO CONTENTS	9
3.2 SAMPLING THE COHORT IN A COST-EFFECTIVE WAY	11
3.3. TRACKING INDIVIDUALS	13
3.4 ETHICAL ISSUES	14
3.5 DATA MANAGEMENT AND ANALYTICAL CHALLENGES	15
4. CONCLUSIONS	17
REFERENCES	19
TABLE I EXAMPLES OF PANEL STUDIES IN DEVELOPING COUNTRIES	6

REDUCING ATTRITION IN PANEL STUDIES IN DEVELOPING COUNTRIES

ABSTRACT	23
1. INTRODUCTION	24
2. IS ATTRITION A PROBLEM?	25
3. WHAT CAN WE DO TO REDUCE ATTRITION?	27
4. TRACKING LESSONS FOR DEVELOPING COUNTRIES	29
5. CONCLUSIONS	42
REFERENCES	43
TABLE 1 TOTAL ATTRITION AT END POINT OF SURVEY	28
TABLE 2 OUTLINE OF PANEL STUDIES IN DEVELOPING COUNTRIES THAT HAVE TRACKED RESPONDENTS	31

Linking public issues
with private troubles:
Panel studies in
developing countries

Abstract

Accompanying the call for increased evidence-based policy the developed world is implementing more longitudinal panel studies which periodically gather information about the same people over a number of years. Panel studies distinguish between transitory and persistent states (e.g. poverty, unemployment) and facilitate causal explanations of relationships between variables. However, they are complex and costly. A growing number of developing countries are now implementing or considering starting panel studies. The objectives of this paper are to identify challenges that arise in panel studies, and to give examples of how these have been addressed in resource-constrained environments. The main issues considered are: the development of a conceptual framework which links macro and micro contexts; sampling the cohort in a cost-effective way; tracking individuals; ethics and data management and analysis. Panel studies require long term funding, a stable institution and an acceptance that there will be limited value for money in terms of results from early stages, with greater benefits accumulating in the study's mature years.

I. Introduction: the attraction of panel studies

'...a new spectre is now haunting social scientists- the spectre of household panel surveys. Not content with cross-sectional surveys,... academic groups and national statistical offices have launched household panel surveys in most European countries as well as in North America. [They]... appear to have voracious appetites for severely constrained resources from national science foundation budgets and may suffer from a number of technical problems as well' (Duncan 2000 p. 54).

The above quote identifies the burgeoning number of panel studies in 'developed' countries. Reviews and lists of panel studies in developed countries are regularly produced by the UK Longitudinal Studies Centre of the Institute of Social and Economic Research at Essex University (www.iser.essex.ac.uk/ulsc). However, there is now an increasing call for panel studies in some developing countries that are moving towards evidence-based policy (see Garner et al 1998 for examples from the health sector). This paper outlines the strengths and weaknesses of panel studies and considers their implementation in developing countries.

The essence of the longitudinal survey is that it offers repeated observations of individuals over time. Such time-series design is often encountered under the generic term **panel study**. The unit of analysis is normally the individual and not (as in some cross-sectional surveys) the family or household. This is because the nature of families or households can change across time. 'Panel studies should use the individual rather than the household as the unit of analysis and map the relationship existing between the two at different points in time. One can use the household as the unit of *measurement* but ought to use the individual as unit of *analysis*, attributing to each individual the characteristics of the household in which he or she lives' (Laurie and Sullivan 1991: 122). In economic studies, the unit is often the household because many economic variables are not readily measurable at the individual level. This raises problems about defining what constitutes the 'same household'. Using Living Standards Measurements Survey (LSMS) data from Cote D'Ivoire, Grootaert and Kanbur (1995) called it the same

household if there was just one member from the previous round. This raises potentially serious issues of interpretation.

A **cohort study** is a form of panel study and is designed so that the cohort or sample members share the same initial condition such as being born in the same year (note that in epidemiology the terms cohort, longitudinal and follow-up are used interchangeably (Last 2001)). Because the same people are followed across time, data can be used to examine flows into and out of certain states (for example, poverty), thus opening up a wider range of possibilities in terms of causal analysis and inferences compared to cross-sectional surveys (Menard 1991).

Rose (2000, p.34) sums up strengths of panel studies as follows: ‘Essentially panel data allow us to distinguish between transitory and persistent aspects of phenomena such as poverty and unemployment. They allow us to examine gross change - flows as well as the stocks. As they mature, panels provide vital information on intergenerational issues, for example social mobility ... However, these advantages only emerge if panel surveys are well designed and are maintained so that the disadvantages inherent to panels - panel conditioning [respondents becoming atypical of the population because of their panel membership], wave non-response, attrition - are minimised’. ‘Transitory aspects’ can only be captured if the frequency of observation is high relative to the duration of the ‘transitory’ phenomenon.

Rose (2000) argues that the prime purpose of household studies is to provide both social scientists and policy makers with prospective micro data in order to improve our understanding of processes, causes and effects in relation to social trends and social change. He suggests that the most imaginative social science has sought to connect public issues and private troubles (Mills 1959) and thus to explore macro and micro interconnections. While cross-sectional studies can also examine macro-micro linkages, panel studies are particularly suitable because of the added time dimension. Panel studies are often aimed at the study of individual change within theoretical frameworks, which place micro-level changes in a macro-level context. To give

a concrete, sectoral example of a call for such research, consider the fact that epidemiology in the UK has demonstrated links between material inequalities experienced in childhood and inequalities of health outcomes in adulthood. Morrow (2000) points out that these studies cannot explain why such a link exists nor trace the processes involved as they failed to measure the macro-level context. She draws attention to Popay et al's (1998) call for studies that explore the complex interactive relationship between individual experience, social action and the way in which societies are organised at the macro level. In other words, the linkage between private troubles and public issues.

Panel studies are analytically strong, provide an opportunity to link macro-micro issues and are increasingly called for in the research recommendations of numerous projects. However, the disadvantages of panel studies need to be borne in mind: they are costly and complex; it takes a long time for results to become available; and determination of aims at the outset may restrict the ability to respond to emerging policy questions. The objectives of this paper are to identify challenges that arise in panel studies, and to give examples of how these can be addressed in a resource-constrained environment such as a developing country. Examples of panel studies in developing countries are presented then problems such as linking macro and micro contexts, sampling in a cost-effective way and tracking individuals are considered.

2. Examples of panel studies in developing countries

TABLE I. EXAMPLES OF PANEL STUDIES IN DEVELOPING COUNTRIES WITH MORE THAN 1,000 SUBJECTS, FOLLOWED FOR AT LEAST TWO YEARS WITH MORE THAN ONE FOLLOW UP

Place	Cohort	Current state	Tracked*	Notable Features	Reference
1. Johannesburg, S.Africa	4,029 children born April-June 1990	Ongoing	70%	Under-enrollment of white, middle class babies	Richter et al 1995
2. Pelotas, Brazil	1) 6,011 hospital births in 1982	Ongoing	1) 82% at age 12 months 87% at 20 & 42 months	3 follow ups at varying intervals	1) Barros et al, 1990
	2) 5,304 hospital births in 1993	Ongoing	2) 95% at 12 months		2) Barros et al, 2001
3. Yaounde, Cameroon	9,774 hospital births in 1978	?	88% after 1 month	Examining determinants of infant mortality,	Kuate-Defo B 1992
			65% after 2 years	7 follow ups at 4 month intervals	
4. Cebu, Philippines	3,080 mother-infant pairs born May '83-April '84 across 33 communities (rural and urban)	Ended 1999	By 1994, 75% were tracked (1991, 1994, 1999)	14 rounds in 1st 2 yrs, plus 3 others	www.cpc.unc.edu/cebu/
5. China	3,800 households across 8 provinces	?	?	4 rounds studying health, nutrition and family planning	www.cpc.unc.edu/china/
6. Guinea-Bissau	10,000 women across 5 northern regions	?	?	Visited 6-monthly for 6 years to measure maternal mortality causes	Hoj et al 1999
7. Rural Bangladesh	16,270 children aged 9-60 months	Ended 1985	?	Followed fortnightly from date of measles vaccination until: death/out-migration/attainment of 60 months of age/end of study (Oct '85)	Koenig et al 2001
8. Rural Senegal	4,852 children born March '83-Dec '89 to mothers resident in 22 villages in 1983	?	?	Yearly censuses 1983-87. From Jan '87, weekly visits (demographic surveillance area). Examined causes of child mortality.	Ronsmans 1996

? = Information not known

* It is not known whether this includes subjects who died before revisit or whether this is net of such mortality

Table 1 shows examples of large-scale (>1000 subjects) panel studies which have been conducted in developing countries. These studies have taken place in Asia, Latin America and sub-Saharan Africa and several have continued over many years. They all originated in the health and nutrition field. The table excludes 'population laboratories' or 'population observatories' which are prospective community studies that register the total population (typically 20,000) in a given area and follow up to monitor vital events such as deaths, births, and illness (Das Gupta et al 1997). Famous examples of these are the Matlab study in Bangladesh and the Narangwal study in India. Such studies have somewhat different aims and challenges compared to cohort studies and are not considered further here. The table also excludes some important panel studies from economics because the number of subjects was less than 1000 per sweep. Morduch (2001) points out that the Indian ICRISAT (International Crops Research Institute for Semi Arid Tropics) 10 village study, despite a cross section of just 120 households per year, still yields valuable new insights many years after the last wave was collected in 1983-84 (data were collected every year for ten years). The ICRISAT data have produced important work on risk-coping strategies, nutrition, crop choice, tenancy contracts, labour supply and poverty dynamics and have influenced safety net design world-wide. Few households stayed in the sample for the entire period (53 of the original 120 by 1983) but the survey was particularly rich in the variety of questions asked. This, combined with the fact that it managed 10 repeat rounds, made it particularly useful. Similarly some of the Living Standards Measurement Surveys (LSMS-www.worldbank.org/lsms) have had panel elements, for example, Cote D'Ivoire (which had a nationally representative cross section of households). Each year 1600 households were sampled; half of these were revisited the following year, the other half was replaced with new households. The result is four cross-sectional data sets and three two-year panels (for 1985-6, 1986-7 and 1987-8). Annual attrition (% of the base year sample) ranged from 5-13% for the 'poor' and 11-15% for the 'non-poor'.

Other examples of how results of panel studies have been used are:

- i) Birth to Twenty (previously known as Birth to Ten) is a birth cohort study of 4029 in Table children born during a seven-week period in 1990 in Soweto-Johannesburg, South Africa (<http://www.wits.ac.za/birthto20>). The first ten years of the study focused on the determinants of growth, health and development of the cohort. In the second 10-year period, new research areas such as sexual maturation, education, and the social risks and vulnerability of young people to contract HIV/AIDS are included. The Birth to Twenty study results have had a major impact on a number of areas of policy relating to children's health, well-being and education in South Africa. For example, children's perceptions and experiences of tobacco at 5 and 7 years of age contributed to changes in the South African tobacco control legislation passed in 2000. Birth to Twenty successfully uses policy briefs and newspaper articles to highlight findings of the study. Examples include the consequences of high blood lead levels in children living in the Johannesburg metropolitan area, the impact of tobacco advertising on children and the problem of bullying in schools.
- ii) Pelotas (Brazil) birth cohort studies (1982 and 1993). Data from these studies have been used to investigate inequities in child health and the implementation of new public health interventions (Victora et al 2000). The authors confirmed the hypothesis that new interventions and programmes initially reach those of higher socio-economic status and only later affect the poor, resulting in increases in inequity ratios for health indicators at the earlier stages of programme implementation. This has led to calls to policy-makers and programme managers to adjust their targeting strategies.

The principal challenges of conducting panel studies are addressed in the next section.

3. Addressing problems of panel studies in developing countries

3.1 Conceptual frameworks which link macro and micro contexts

‘In the search for effective policies to combat poverty, bridging the gap between macro-level policy analysis and micro-level livelihoods analysis is an essential task, but not an easy one’ (Shankland 2001: no page number). Because panel studies are well suited to linking macro policy changes with micro individual or household changes, it is imperative to have a conceptual framework that links the two. A conceptual framework requires hypothetical linkages between risk factors (or determinants, depending on what discipline one is from) and outcomes. For example, a child poverty study may have numerous outcomes such as a child’s physical and mental health, her nutritional status and developmental stage for age, numeracy, literacy, and the child’s own, subjective perception of well-being. These would be measured empirically, at each sweep, or round, of the study. In addition, there will be risk factors for such outcomes. Continuing the example of a study of child poverty, risk factors would include aspects of childcare (including use of services), work, play, education, and household structure including parental characteristics such as caregivers’ education status and physical and mental health. In addition, particularly in development studies, careful consideration needs to be given to the measurement of wealth, socio-economic status and livelihoods. A recently launched longitudinal study covering four developing countries, the Young Lives project (see Harpham 2002 and www.younglives.org.uk) for example, uses a livelihoods framework and tries to capture shocks and various assets (such as social capital) which may buffer the effect of shocks. As Dercon (2001:7) states: ‘panel data... are crucial for increasing our quantitative understanding of shocks and their impact’. But, all this is at the microlevel; what about the macro? In between the two levels are community structures and institutions. It is useful to capture contextual information by implementing a community (ecologic) questionnaire using a similar format to the community questionnaires implemented in the large-scale international household surveys such as the World Fertility Survey (WFS), the Demographic and Health Surveys (DHS) and the Living Standards Measurement Surveys (LSMS). The sampling procedure (see section 3.2) can facilitate or

hamper the collection of the most useful community-level data. We need enough communities to see variation in the effect of policy implementation (or to confirm its uniformity across the state). ‘Within’ communities we want all of a set of households to be part of the same ‘community’.

Measuring change at the macro level requires monitoring policy over time and then, in analyses, linking such change to micro level change at the household or community level. Studies may be concerned with specific sectors, such as health, or may be interested in broader issues such as poverty, in which case cross-sectoral policies have to be monitored (e.g. privatization). Evidence of policy implementation, not mere declarations or enactment, needs to be monitored. Thus, information is required on any differential implementation (between places (region, rural/urban), times or population groups).

If inter-sweep tracking visits are made to cohort members (see section 3.3 below), field workers can undertake checks of policy implementation at the local level during these visits. For example, if user fees are introduced at the national level, are local health facilities actually charging them? This confirmation of policy implementation at the local level is easiest for sectoral policies. In cross-sectoral policies it may be more difficult for the causal links to be made. For example, trade liberalization may be linked to rises in prices of agricultural inputs (e.g. fertilizers), which may result in changes in the productivity and ultimately the welfare of household members. However, changes in productivity may be associated with potential ‘confounding’ factors such as seasonal climatic change, gradual erosion of farmers’ capitals etc. Careful, step-by-step diaries of policy implementation are needed for robust analysis of macro-micro links. In addition, complementary longitudinal qualitative data may illuminate quantitative relationships observed (Laurie and Sullivan 1991).

Another point to note about macro-micro linkages in panel studies is that a single cohort study is not able to inform us about trends. For life-course processes recorded for a single cohort,

there is no scope to examine variation from cohort to cohort. Only a comparison *between* cohorts allows one to establish whether there has been change across time, for example whether children born ten years later have better nutritional status (Dale and Davies 1994). However, this substantially adds to costs and a single, additional cross-sectional survey of children at a particular age is a compromise that will at least provide some comparative data, provided it uses the same sampling basis as the original cohort. In addition, the panel data can be compared to secondary data sources (e.g. previous national surveys) if population groups are sufficiently similar.

3.2 Sampling the cohort in a cost-effective way

Sample selection for a panel study needs very careful consideration. While a developed-country study might reasonably select children born in a hospital, a developing country sample would often be biased in favour of the wealthier and the urban by such a choice. In many cases, it may be impractical to take a sample that is representative of the relevant age-group, from the general population. It would need to be very large and expensive to contain enough of any given target group.

If it has the aim of illuminating macro-micro issues, a study needs to focus on geographical areas (“sites”) where it is meaningful to assess the local implementation of relevant policies and to collect a meaningful sample of micro-level information on individuals. The sampling strategy requires work at two levels: site selection and sampling of individual households within sites. Those individuals selected within sites studied will of course be personally unknown and anonymous to analysts concerned with the whole study, so the usual requirements of objectivity demand that sampling have the benefits associated with randomness, rather than being casual or opportunistic. Also, repeated follow-up makes it imperative that the selections made can be documented and traced, so the site sample can be reproduced without difficulty. In practice, this is likely to mean that a systematic form of sampling is used, carefully adapted to local conditions.

For example in a peri-urban area with moderate population density and no reliable long-term house-numbering system, one might sample households lying on transects, each beginning at an established road junction and heading directly towards a visible landmark. The target dwellings, the site description, and the instructions to later researchers who will follow the tracking rules need to be very well-described.

At the level of site selection, it can be anticipated that plausible entities that might become sampled sites will have numerous characteristics that are well known at least regionally. For example, one site may be a centre of a particular form of child labour in an ill-regulated industry, a market for some specific products, the centre of inter-communal religious violence, and so on. Sampling sites at random disregards this plethora of information, and will only work well if there are a huge number of selected sites so that most important features are adequately represented. To keep down the workload and the corresponding costs, developing country studies are likely to have a relatively small number of sites. Therefore, these must be sampled purposively, ensuring they serve to illuminate macro-micro policy linkages known to be important at the outset, e.g. individual or combined impacts of legislation on child labour, and an improved commitment to industrial regulation. A claim to have avoided subjective selection biases depends on a very clear qualitative description of the site selection, e.g. how it draws on published data, and was agreed by a consensus of named experts. The usual statistical approach to “cluster sampling” sets out to suit a one-off cross-sectional study; relatively many, relatively small clusters are chosen, often at random from a list of possibles, and the individual clusters are not regarded as individually interesting in their own right. In contrast, the panel study focuses its attention repeatedly on the same sites, has to maintain a reasonable sample at each site even though anticipating some degree of attrition as time goes by, and minimises attrition by good liaison with community members and bodies. Rather than following the precepts of cluster sampling, it is thus reasonable for a panel study to adopt the pattern of sentinel site surveillance – a relatively small number of carefully documented sites studied in a consistent way at intervals. Typical health sector sentinel

surveillance exercises entail quite frequent collection of relatively simple data in order to follow trends in public health indicators. The panel study is usually distinguished from such exercises by complex study instruments and a far greater need to retain the same sample members as time goes by.

The Young Lives study, in which the authors are involved, follows the above principles and regards it as important that the panel study sites provide excellent settings in which complementary research approaches can very effectively be implemented: given the necessary funding, this structure allows in-depth and perforce smaller-scale qualitative studies to be conducted in one or a few sites, and articulated with the broader, but shallower panel study information.

3.3 Tracking individuals

Minimizing loss of individuals over time is, perhaps, one of the greatest challenges facing any panel study. A review of the literature shows that refusal to continue participation is the main reason for loss in developed country cohort studies, whereas in developing countries the principal problem is failure to trace participants (Hill, 2002). Tracing rates, naturally, vary according to factors such as the inter-sweep interval and local conditions. The selection of strategies to maximize tracing is determined by resources and local circumstances. These strategies range from the gathering of pertinent information to the use of technology. Information on key friends, neighbours and family of a participant can lead to useful sources of help when a cohort member has moved. Another possibility is to employ a local community member to conduct occasional tracing checks on cohort members. Improvements in communications technology bring new possibilities. In an urban cohort in South Africa, the cheap availability of 'pay as you go' mobile phones has assisted in minimizing attrition by the researchers providing respondents with a cheap phone and the researchers regularly phoning to maintain contact with the respondents. Geographical Information Systems (GIS) provide exciting new opportunities for

tracking in remote areas. Inevitably resources will constrain the extent of the tracking that can be undertaken in a study. Decisions must be taken on, for example, whether or not to follow individuals who move outside the study catchment areas. Large-scale studies have used a central tracking operation for cases that move beyond a local area. In the Young Lives study, information will be gathered in the first sweep of the study to establish tracking strategies for the different sentinel sites, according to factors such as within-site mobility and migration rates.

3.4 Ethical issues

Ethical issues are particularly important in panel studies because of the increased burden on respondents. In some developing countries, appropriate ethical committees do not yet exist, and, therefore, they may have to be established before the implementation of a panel study. An important aspect of successful tracking of cohort individuals is that on entry to the study they are aware of its longitudinal nature and that they consent to being traced over time. This is one ethical issue that differs from single cross-sectional studies. Panel studies, naturally, share many ethical concerns with other study designs and these are important to address. These include issues such as informed consent, how to deal with cases of illness or abuse that are encountered, and how to interview vulnerable groups such as children. An issue which perhaps bears greater significance in panel studies, is that of incentives for participants. The question of whether to use incentives, in the form of cash or kind, particularly where there is no immediate benefit of participating in a study, is not straightforward. In the South African Birth to Twenty Project, for example, study participants receive simple tangible reminders of the project such as stickers, key rings, annual calendars, fridge magnets, pens and rulers, all with a prominent study logo. Participants are refunded for any transport costs incurred and a limited social and health service has been incorporated into the study. A toll-free number has been installed in the project office to enable families to contact the study for advice and information. Referral notes to local services are given to families when serious health or social problems were detected. Incentives in the form of communication to make participants feel involved have also proved helpful to keep participants interested and to remind them to inform organizers about change of address.

Examples include birthday cards, regular newsletters, reports on the study in the local media, and a website (<http://www.wits.ac.za/birthto20>).

3.4 Data management and analytical challenges

Improvements in both computing software and hardware have helped put panel studies within easier reach of researchers. Nevertheless, their complexities for data management and analysis should not be underestimated, and require considerable care in design and ongoing maintenance. Planning for data management and analysis at an early stage can also inform the design of data collection instruments, such as questionnaires - for example, by clarifying the coding of variables or by evaluating the utility of a question, in an analysis plan with dummy tables. These points hold for other types of study, but for panel studies they are perhaps of even greater importance, since the early planning stages need to consider the data management and analysis requirements of later sweeps. A good foundation will reap many benefits at later stages.

To capitalise on the primary strength of panel studies means looking at data longitudinally, that is linking data measured at one point in time with measurements at later points. This requires consideration of how stable or transient variables are in assessing their use as determinants or correlates of later outcomes. A key problem is that of missing data, either when an individual is lost to follow-up entirely or misses a sweep but returns later. This is a methodological field of enquiry which is currently attracting considerable attention and is likely to generate more sophisticated and accessible methods for dealing with missing data such as imputation.

Finally, a crucial issue for long-term successful use of panel data is comprehensive documentation of procedures and structures. Few researchers will be involved with a long-term study for its entire duration and as the complexity of datasets increases, the need grows for clear metadata - explanatory information associated with the actual data, e.g. as to what was measured or recorded by whom and about whom, when and where; and any changes to datasets during 'cleaning': These metadata should be securely locked on to the data themselves and all

documentation and definitions should be of sufficient standard that the provenance of early data can still be fully understood fifteen or thirty years thence. This requires rigorous attention to the archiving of metadata, in turn requiring a high degree of selectivity about the data and metadata items preserved, to avoid too large a bureaucratic burden.

4. Conclusions

Panel studies are important in developed countries due to their ability to inform policy (Office for National Statistics 1999, Institute for Social and Economic Research 2000). As the call for evidence-based policy increases in developing countries, the attraction of panel studies will grow. However, they are expensive (eg. the UK 1970 Birth Cohort Study and the UK National Child Development Study have each cost c. £1.5m) and complex. As experience of implementing panel studies in developing countries increases, more guidelines can be formed regarding the challenges considered in this paper: selecting a sample in the absence of a sampling frame; linking macro and micro contexts; tracking mobile populations; the need for ethics committee approval; and sophisticated data management and analysis. Panel studies should not be entered into lightly in any context, but in resource-constrained environments, their added value needs careful consideration. They certainly require quite long-term robustness of funding and a stable institutional setting, and they require their backers to have sufficient vision and patience to accept limited value for money in terms of results from the early stages, the greater benefits accumulating in the study's mature years.

Acknowledgements

The authors are grateful to Howard White of the UK Institute of Development Studies for his useful comments on a draft of this paper.

References

- Barros FC, Victora CG, Vaughan JP 1990 The Pelotas (Brazil) birth cohort study 1982-1987: strategies for following up 6,000 children in a developing country. *Paediatr Perinat Epidemiol* 4(2): 205-220.
- Barros FC, Victora CG, Vaughan JP, Tomasi E, Horta BL, Cesar JA, Menezes MB, Halpern R, Post CL, del Mar Garcia M 2001 The epidemiological transition in maternal and child health in a Brazilian city, 1982-93: a comparison of two population-based cohorts. *Paediatr Perinat Epidemiol* 15(1): 4-11.
- China health and nutrition study* <http://www.cpc.unc.edu/projects/china/sitemap.html>
- Dale A, Davies R 1994 *Analysing social and political change: a casebook of methods*. Sage: London
- Das Gupta M., Aaby P., Garenne M., and Pison G. 1997 *Prospective community studies in developing countries*. Clarendon Press: Oxford.
- Dercon S 2001 *Assessing vulnerability*. Oxford: Economics Dept., Oxford University.
- Duncan G 2000 Using panel studies to understand household behaviour and well being. In D Rose (ed.) *Researching social and economic change: the uses of household panel studies*. Routledge: London.
- Garner P, Rajendra K, Rumona D, Danis T, Salinas R 1998 Getting research findings into practice: Implementing research findings in developing countries. *BMJ* 317:531-535.
- Grootaert C, Kanbur R 1995 The lucky few amidst economic decline: distributional change in Cote D'Ivoire as seen through panel data sets 1985-88. *Journal of Development Studies* 31(4): 603-619.

- Harpham T 2002 Measuring child poverty and health: a new international study. *Journal of Tropical Paediatrics* 48: 128-129.
- Hill Z, 2002 Tracking in panel studies. *Young Lives working paper series*. London.
- Hoj L, Stensballe J, and Aaby P 1999 Maternal mortality in Guinea-Bissau: the use of verbal autopsy in a multi-ethnic population. *International Journal of Epidemiology* 28: 70-76.
- Institute for Social and Economic Research 2000 *The future of UK longitudinal studies*. ISER: Colchester.
- Koenig M, Bishai D, and Khan M 2001 Health interventions and health equity: the example of measles vaccination in Bangladesh. *Population and Development Review* 27(2):283-302.
- Kuate-Defo B 1992 Mortality and attrition processes in longitudinal studies in Africa: an appraisal of the IFORD surveys. *Population Studies* 46(2): 327-348.
- Last JM 2001 *A dictionary of epidemiology* Oxford University Press: Oxford.
- Laurie H, Sullivan O 1991 Combining qualitative and quantitative data in the longitudinal study of household allocations. *Sociological Review* 39(1): 113-130.
- Menard S 1991 *Longitudinal research*. Sage London.
- Mills C 1959 *The sociological imagination*. Oxford University Press: Oxford.
- Morduch V 2001 Consumption smoothing across space: testing themes of risk-sharing in the ICRISAT study region of South India. Paper presented at UNU meeting on Insurance against Poverty, Helsinki, June 15.

Morrow V 2000 'Dirty looks' and 'trampy places' in young people's accounts of community and neighbourhood: implications for health inequalities. *Critical Public Health* 10: 141-152.

Office for National Statistics 1999 *Tracking people: a guide to longitudinal social sources*. ONS: London.

Popay J, Williams G, Thomas C and Gatrell T 1998 Theorising inequalities in health: the place of lay knowledge. *Sociology of Health and Illness* 20 (5): 619-644.

Richter L, D Yach, N Cameron, R Griesel, T de Wet 1995 Enrolment into birth to ten: population and sample characteristics. *Paediatric and Perinatal Epidemiology* 9(1): 109-120.

Richter L, and T de Wet Almost there...Birth to Ten Children are 9 years old (Forth coming)
<http://www.wits.ac.za/birthto20>

Ronsmans, C 1996 Birth spacing and child survival in rural Senegal. *International Journal of Epidemiology* 25(5):989-997.

Rose D 2000 Household panel studies: an overview. In D Rose (ed.) *Researching social and economic change: the uses of household panel studies*. Routledge: London.

Shankland A. 2001 *Matching macro and micro: building bridges between policy and livelihoods*
www.id21.org.

Victora CG, Vaughan JP, Barros FC, Silva AC, Tomasi E. 2000 Explaining trends in inequities: evidence from Brazilian child health studies. *Lancet* 356(9235): 1093-1098.

Reducing Attrition in Panel Studies in Developing Countries

Abstract

In panel studies in developing countries attrition or loss to follow-up is mostly due to respondents moving. Attrition can cause bias if it is selective, and efforts should be made to track respondents. This can be costly and difficult as populations in developing countries are often highly mobile, infrastructure is poor, structures frequently change and formal address systems and population records rarely exist. In this paper, the experiences from panel studies in developing countries are reviewed in terms of the importance of attrition, the importance of tracking respondents on reducing attrition and then makes recommendations for setting up systems to track respondents in developing county settings. Tracking can reduce attrition by up to 45%, and is feasible if procedures are locally appropriate, well planned, involve the community, collect as much locating data as possible, criteria are explicit, tracking is done and at regular intervals, and interviewers are well trained, supervised and motivated.

I. Introduction

In panel studies, respondents are lost when they die, decide to drop out of the study or when they are not relocated. In their review of important issues to consider in the design of panel studies in developing countries, Harpham et al (2002) identify the loss of individuals over time (attrition) as one of the greatest challenges. Much has been reported about attrition in developed countries, but literature from developing countries is scarce. The experiences from developed countries may not be relevant for developing countries, as the main reasons for attrition differ. In developed countries, attrition is mostly due to refusals which studies have found are often situational, for example because the interviewer called at a bad time (Singh 1995, Buck et al 2002), whilst in developing countries, attrition is mostly due to moves (Barros et al 1990, Kuate Defo 1992, Haaga et al 1994, Alderman et al 2001, Thomas et al 2000). This paper reviews the limited literature from developing countries and explores the importance of attrition, the importance of tracking respondents on reducing attrition and makes recommendations for setting up systems to track respondents in developing county settings.

2. Is attrition a problem?

Attrition in panel studies results in a diminishing number of study respondents. Because the loss of respondents is cumulative, what appears to be a small loss in each survey round can add up to a considerable loss over time. This loss of study respondents can reduce the statistical power of the study (through a reduction in sample size) and, if it is selective, can cause bias as the remaining sample may not be representative of the population it was originally selected from. In general, attrition in panel studies is regarded as problematic, but its actual importance varies depending on the study aims and the study population. For example, loss of respondents through moves can cause particularly large bias in studies which aim to explore transition and change linked to migration.

Four studies were located exploring the impact of attrition bias in panel studies in developing countries. In Cameroon (Kuate Defo 1992), the probability of children being lost through attrition was associated with the mother's marital status and mother's education level. However the attrition of 40% was not found to affect the outcome measure of infant mortality. In Malaysia (Haaga et al 1994), selective attrition of 30% meant that younger, better educated, urban and ethnic Chinese were under-represented in the sample, but it was concluded that meaningful analysis could be conducted for the outcome variables by re-weighting the sample and using other corrective statistical methods. In a Brazilian panel, attrition was 13%, but there was no marked variation in birth weight or income by whether follow up was successful or not (Barros et al 1990). In a review of three panel studies in Bolivia, Kenya and South Africa (Alderman 2001), attrition was 35%, 28% and 15% respectively, and was found to be associated with several background variables including education, age and assets and with outcome variables in some multivariate models. The general conclusion of the review was, however, that attrition was not a pervasive problem for obtaining consistent estimates of coefficients in any of the three sites, a further evaluation of the South African data drew similar conclusions (Maluccio 2000). These encouraging conclusions are similar to those from panel studies in developed countries (Fitzgerald et al 1998, Lillard & Pannis 1998, Zabel 1998, Ziliak et al 1998, Falaris and Peters

1998, Twisk and de Vente 2002). However, they must be interpreted with some caution as the studies are not conclusive and the results can not be used to justify relaxed efforts to reduce attrition, particularly in studies that explore transition and change. The four studies discussed above did find that some variables were associated with attrition and, as it is impossible to predict which and how many variables will be affected, efforts to limit attrition are important. Studies with high attrition should not, however, assume their results are invalid.

3. What can we do to reduce attrition?

Most studies recognise that attrition is an important problem and make efforts to reduce it. In developed countries, this often involves trying to reduce refusals but in developing countries it mostly involves dealing with respondents who move. There are three main ways of dealing with respondents who move: 1. They can be excluded from the sample, 2. They can be replaced and 3. They can be tracked. Which of these three strategies is appropriate depends on the aim of the study, but in most cases tracking respondents is the most desirable strategy for minimising bias.

Keeping track of cohort members over time involves considerable effort, particularly in developing countries where mobility is high and researchers face logistical problems such as poor infrastructure, structures frequently changing, a lack of formal address systems and limited population records. Some studies make great efforts to track panel members, whilst others perform no tracking at all. Table 1 shows attrition and the impact of tracking on attrition for a number of large cohort studies (>1000 respondents followed for at least 2 years) conducted in developing countries. As Table 1 shows, tracking does not eliminate all attrition but it can significantly reduce it. Studies with no tracking had between 10 and 20% attrition per year compared to only 1 to 7.5% for those with tracking. The number of people located by tracking, and who would have otherwise been lost, ranged from 5 to 45%.

TABLE I: TOTAL ATTRITION AT END POINT OF SURVEY

Study	Reference	Survey length (years)	% Attrition per year	% Attrition at end of survey	% study respondents found through tracking
IFORD Yaounde Survey (Cameroon 1978-1980)	Kuate Defo 1992	2	20	40	No tracking
Bolivia (1996-1999)	Cited in Alderman et al 2001	3	11.7	35	No tracking
Peru (LSMS) (1985-1990)	World Bank 1993	5	10	50 ¹	No tracking
1982 Birth Cohort Study in Pelotas (Brazil 1982-1987)	Barros et al 1990	2	7.5	13 ²	45
Birth to Ten (South Africa 1990-present)	Richter & De Wet (forth coming)	8	3.8	30 ³	Unavailable
Cebu longitudinal health and nutrition survey (Philippines 1983-present)	Glewwe et al 1999	8	3.6	29 ⁴	Unavailable
South Africa KIDS/PSLSD (KwaZulu-Natal 1993-1998)	May et al 2000	5	3	15	5
Malaysian family life study (1977-1989)	Haaga et al 1994	12	2.5	30	37
China health and nutrition survey (1989-1993)	Cited in Thomas et al 2000	4	2.3	9	Unavailable
Indonesia family life study (1993-present)	Thomas et al 2000	5	1	5 ⁵	19
Vietnamese Longitudinal Study (1995-present)	VLS web page	36	1	3	Unavailable

1 As households were replaced, the attrition figures refer to the number of households that had to be replaced, i.e. houses not occupied by the same family, unoccupied or demolished. It does not include refusals

2 Figures refer to 1984

3 Figures refer to 1998

4 Figures refer to 1991

5 Figures refer to 1995

7 Figures refer to 1997

4. Tracking lessons for developing countries

This section uses lessons learnt from previous large-scale panel studies in developing countries (outlined in table 2) to describe some of the basic principles of tracking. Generally, the success of tracking is linked to population mobility, resource and infrastructure availability, study length, tracking procedures, time between contact and interviewer motivation.

TABLE 2: OUTLINE OF PANEL STUDIES IN DEVELOPING COUNTRIES THAT HAVE TRACKED RESPONDENTS

Study name and reference	Study design	Tracking procedure
<p>Malaysian Family Life Study (1977-1989)</p> <p>Haaga et al 1994</p>	<p>The panel consisted of 1262 ever married women resident in peninsular Malaysia, aged less than 50 in 1976. The women were interviewed between 1976-77. One follow up round was conducted 12 years later. Data collection collected includes: demography, fertility events/desires, education, time allocation, income and intergenerational transfers.</p>	<p>Follow up not planned, thus tracking relied solely on the addresses from the 1977 questionnaires. Three teams covered different regions, they started in remote areas and finished in the cities where rural migrants were likely to have gone. The teams visited the recorded address, if the respondent was no longer there they tried to get the new address, the best informants were neighbours and the current tenants of the property. Village headmen, postmen, employers & shop-owners were also useful. The teams regularly sent lists of names & locations of movers to relevant team, thus tracking was countrywide.</p>
<p>1982 Birth Cohort Study in Pelotas (Brazil 1982-1987)</p> <p>Barros et al 1990</p>	<p>The panel consisted of all children resident and born in hospitals in the Pelotas urban area in 1982 (n=6011). Three follow up rounds were conducted at 1, 2 and 4 years. Data collected includes: mortality, morbidity, growth & development & health care utilisation.</p>	<p>Addresses were obtained from the mothers' at the hospital. In the first follow up interviewers visited the recorded addresses, if the respondent was not located relatives, employers and credit sections of department stores were used to try and locate them. Children were not tracked outside the study area. For the second and third rounds a census of Pelotas was conducted and records matched, for children who were not located this way the method for round one was followed. A mass media campaign proceeded each follow up. Special permission was obtained to look at adoption records to trace adopted children.</p>
<p>Cebu Longitudinal Health and Nutrition Survey (Philippines 1983-present)</p> <p>CLHN 1989</p>	<p>The panel consisted of a sample of 2600 women pregnant between 1983-1984 and resident in the Cebu metropolitan area. 14 interviews were conducted in the first 2 years and three follow-up rounds were conducted at 8, 11 and 15 years. Data collected includes: feeding practices, health, demographic, socio-economic characteristics and time allocation.</p>	<p>No information located except that respondents were traced within the study areas but long distance migrants were excluded.</p>
<p>CEP-CPC Study of Social Change in Nang Rong (Thailand 1984-present)</p> <p>CEP-CPC Nang Rong Project website and Migrant follow up handbook</p>	<p>Census of 51 villages in Nang Rong (n=42,219) conducted in 1984. The census was repeated in 1994 and out migrants from 22 villages tracked. Data collected includes: migration processes, fertility and contraceptive behaviour and life choices.</p>	<p>Local migrants were located through census. Out migrants from 22 villages (n=2380 of which 1781 were located) were tracked. Interviewers tried to find the migrants new address when they were working in the study area. Once data collection in the study area was completed interviewers tracked the migrants if they had moved to Bangkok, Buriram town or the eastern seaboard. The interviewers made several attempts to locate the respondent and recorded details about each attempt, they were encouraged to use their ingenuity in their tracking attempts and utilised family, friends and employers. When they located migrants they asked each migrant if they knew about any other migrants from Nang Rong living in their location.</p>

TABLE 2: OUTLINE OF PANEL STUDIES IN DEVELOPING COUNTRIES THAT HAVE TRACKED RESPONDENTS

Study name and reference	Study design	Tracking procedure
China health and nutrition survey (1989-1993) Thomas et al 2000 and CSHN 1999	The panel consisted of 3795 households selected from 9 provinces. Three follow-up rounds were conducted at 2, 4 and 8 years. Data collected include: time allocation, economic activities, health care utilisation, nutrition and health.	No information located except that respondents were traced within the study areas but long distance migrants were excluded.
Birth to Ten (South Africa 1990-present) Richter & De Wet (forth coming), De Wet personal communication and Richter et al 1995	The panel consists of all children born and resident in Johannesburg-Soweto during a seven week period in 1990 (n=3275). Six follow up rounds were conducted at 6 months, 1, 2, 4, 5 and 7/8 years. Data collected includes: prenatal risk-factors, mortality, morbidity, growth	and development, pollution and child care. Contact addresses were collected at each round. Follow up contacts began with a phone call, home visit, inquiries among the neighbours, letters left at home and further inquiries among the contact addresses, these activities were done by a designated tracer. Between rounds contact was maintained by tangible reminders (e.g stickers, newsletters, radio reports, birthday parties for the children). A toll free number was installed to enable families to contact the project.
Indonesia Family Life Study (1993-present) Thomas et al 2000	The panel consisted of a sample of 7224 households from 13 provinces (urban areas over sampled). Two follow up rounds were conducted at 4 and 5 years. Data collected includes: demographic data, consumption, employment, health status, service utilisation and income.	Relocation sheet filled out for each subject containing relevant detailed information from the first interview. Interviewers returned to the 1993 houses, if no household member was located interviewers tried to obtain information about their location from neighbors, relatives, friends, employers and local community leaders. Local movers (1/2 an hour by public transport) were then traced during main data collection, far movers were traced in a separate, centrally co-ordinated, exercise with bonuses for interviewers for every subject traced. Moves outside the 13 study provinces were excluded.
South Africa KIDS/PSLSD (KwaZulu-Natal 1993-1998) May et al 2000	The panel consisted of a sample of 1393 households in KwaZulu-natal in 1993. One follow up round was conducted 5 years later. Data collected includes: health demographics, household environment, expenditure, employment and income.	Interviewers went to the original location of the 1993 house. If they learnt the household had moved they sought the new address from other family members, schools, employers etc. If a new address was found and was sufficiently detailed at the end of the interview round the interviewers followed the household.
Vietnamese Longitudinal Study (1995-present) VLS web page	The panel consists of a sample of 1855 households from 10 communes in Nam Ha and Ninh Binh provinces in 1995, three follow up rounds were conducted at 1, 2 and 3 years. Data collected include: social and economic activities and demographic behaviour.	No information located

There are several lessons that can be learnt about tracking in developing countries from the studies outlined in Table 2. First, studies must consider all of the different options for tracking and use those that are locally appropriate. Second, protocols need to include rules and tracking limits. Third, the shorter the gap between contacts the better. Fourth, respondents should feel part of the study, and fifth, choice, training, supervision and motivation of interviewers are essential to consider.

4.1 Studies must consider all of the different options for tracking and use those that are locally appropriate

A tracking protocol that is feasible and effective in one country may not be appropriate in others because of differences in geography, infrastructure and mobility. Mobility can differ both in quantity and destination: In the Indonesian and Malaysian Family Life Studies (Thomas et al 2000, Haaga et al 1994) attrition was highest in urban areas where residents were more transient. In contrast, in the KwaZulu-Natal Income Dynamics Study, attrition was highest in rural areas because of the high migration between rural and urban areas (Alderman et al 2001). In the Malaysian study, the tracking protocol utilized the fact that the destination of rural migrants was predictable by starting the follow-up rounds in rural areas and working towards urban areas. This facilitated tracking as the interviewers were moving in the same direction as most of the rural migrants (Haaga et al 1994).

Infrastructure varies greatly between developing countries. The Birth to Ten setting (urban South Africa) has the advantage of good infrastructure and the majority of study respondents have access to a phone, and thus tracking respondents by telephone was feasible (De Wet personal communication). In the Indonesian Family Life Study, only 20% of households could provide a telephone number at which they could be contacted (Thomas et al 2000). For such settings, other more appropriate means of contact must be used. In some settings, there are no formal address systems (e.g. urban slums or remote rural dwellings), and in such cases, it can be

difficult for interviewers to relocate the respondent's house. Relocation can be facilitated by drawing maps as in the Malaysian Family Life Survey (Haaga et al 1994), by using Geographic Positioning Systems to obtain coordinates as in the Young Lives Study in Peru (Lanata personal communication) or by recording detailed descriptions of permanent land marks and directions to the house.

Infrastructure, resources, stability and mobility also vary within countries and tracking protocols need to take this into account, for example they may need to be different in rural and urban areas. Panel studies can reduce attrition by identifying areas with high mobility or with special tracking needs (e.g. communities where people do not know each other's movements, such as those that are not closely-knit). Once these areas are identified, special tracking arrangements can be made, such as more frequent visits by interviewers or employing a community member to help with tracking. In both the Malaysian and Indonesian Family Life Studies, shantytowns were a problem area for tracking as the redevelopment or demolition of shanty towns eliminated entire clusters (Haaga et al 1994, Thomas et al 2000). Where demolition is frequent, studies could make checks on redevelopment and demolition plans and visit the study respondents prior to their removal. Other examples of highly transient respondents or places where communities are not closely knit include people living in government staff quarters, agricultural estate workers (Haaga et al 1994) and people living in areas around training schools, universities or markets (Thomas et al 2000).

One of the best ways to identify mobile groups and other tracking problem areas is to conduct formative research before the start of the study. This should include working with the community to explore how people are commonly identified in that community and the best tracking approach. Formative research can also be used to identify any tracking or linking information (such as names, addresses and dates of birth) that may be reported incorrectly or inconsistently by respondents. The Pelotas study (Barros et al 1990) enrolled children at birth in

hospitals, and found that follow up rounds were hindered because some mothers had given false social security numbers and incorrect dates of birth for their children. This made linking respondents between rounds difficult. It was in the mothers' interest to do this, as their social security number determined the quality of care they received at the hospital and to avoid fines for the late registration of births. In South Africa, names can be spelt using two spelling systems, and, as a result, in the Birth to Ten study linking children between rounds on the data base was difficult because names were not recorded consistently using one spelling system (Anderson and Richter 1994). In both of these examples, well-conducted, formative research could have prevented these problems, as could have issuing each respondent with a unique ID card and number.

The need for context specific tracking protocols is illustrated by those drawn up for a new multi-site panel study called 'Young Lives'. This study aims to explore what happens to children born into poverty and will collect panel data on 8000 children from four developing countries. Each of the four countries (India, Ethiopia, Peru and Vietnam) used the same principals to produce their tracking protocol, but has ended up with unique and individual tailored protocols; reflecting their unique and individual contexts.

Although tracking protocols must be context specific, there are some general 'rules' that should be applied in all settings. Most importantly, the more information you can collect about the respondent the better. A good strategy is to try and identify a network of people associated with the respondent. These people can be invaluable when trying to locate respondents as they are most likely to know of their new location, they can also be tracked themselves if the entire household has moved. It is useful in most settings to record the full names of all adult members of the house, the name of the children (checking their relationship to the adults in the household and that the surname is the same), the names and addresses of external family members and other people who respondents think will know their whereabouts (e.g. boyfriends/girlfriends).

All this must be done accurately and completely for formal names and names people are known by in the community. In Peru, initial experiences from the Young Lives study (Lanata personal communication) have been that recording formal contact addresses is difficult; respondents often describe how to get to the contact address (e.g. 'take the yellow bus from the corner by the bakery and then get off at the drug store'), but can not say exactly where the address is. Studies in such settings need to make sure they can utilize these informal directions.

It can also be useful to record other information about the respondent and their network such as place of work, ethnicity, date of birth, place of origin, the location of any property owned and whether they have any plans to move in the future. In the Indonesian Family Life Study (Thomas et al 2000), a relocation sheet was produced for each respondent, which was updated at each round and issued to the relevant interviewer. The sheet contained data on the household, the name of a person who may know how to contact the respondent, economic and demographic information such as employment, place of birth, all of the places the respondent had lived and the names of their family. Collecting information about the respondent and their networks can generate a lot of data, and plans need to be made for the handling of all the tracking information gathered. If a respondent has no network, they may be difficult to track and special tracking arrangements should be made.

As well as collecting information about the respondent and their networks, studies should explore who else would be useful locating resources in their setting. Whilst formal registration and documentation systems are rare in many developing countries, they can be a useful resource in some settings. For example, the Young Lives Study in Vietnam is able to track respondents using the commune registration systems (Tuan personal communication). The following may be a useful source of addresses (if available) of those who have moved: birth, marriage or medical records, health or welfare registers, city or municipal records, voter or census lists, phone books or postal records, land titles, tax or driving license records, school or church records, prison or

labour union records, organisation membership lists. In the Pelotas study (Barros et al 1990) credit records of department stores were successfully used to locate respondents and in the Malaysian Family Life Study electricity meter readers were useful in relocating households (Haaga et al 1994). The Nang Rong project (CEP-CPC Study) uses the located movers themselves as a tracking resource in a method they call 'snow balling'. They anticipate that migrants from the same village in the same destination will know one another, thus once the interviewer locates a respondent through tracking they ask them if they know of any one else from the study site living in their location.

Whoever the appropriate tracking contacts are, respondents should be informed as part of the consent procedure at the start of the study that they will be tracked and who may be contacted for this purpose. Without this information the respondent and the tracking contacts may be surprised, suspicious and upset by the tracking efforts. Tact and discretion must be used in tracking especially when the respondent has left the study area because of difficult circumstances such as owing money.

It is also important not to assume that respondents lost in one round are lost for all future rounds. In some settings, circular migration is common and respondents return to the community. In the Birth to Ten Study (Richter and De Wet forthcoming), between one third and a half of all cases lost in one round were found in the Johannesburg-Soweto metropolis in later rounds and interviewed again. Similarly, in the Indonesian Family Life Study (Thomson et al 2000), 60% of households that were lost in the first follow-up round were relocated in the second round. In the Cebu health and nutrition study (CLHN 1989), 8% of women classified as 'lost' were later found at their original address.

It is best to start tracking with the simplest and cheapest tracking options, which will locate the most respondents. Studies should, however, consider all of the tracking options. What sounds expensive or unfeasible at first may save time and money in the long run. The Pelotas study

successfully relocated respondents by performing the huge task of conducting a census of the 68,590 households in urban Pelotas (Barros et al 1990), and in the Birth to Ten study mobile phones were given to some respondents, so they could be easily contacted (De Wet personal communication).

The timing of tracking is important; households should be visited when respondents are most likely to be home and tracking should be conducted in the most appropriate season. Seasons associated with difficult weather conditions and when seasonal migration is occurring should be avoided.

4.2 Tracking protocols need to include rules and tracking limits

Some respondents can not be tracked and for others, tracking will be exceedingly costly. The cost of tracking increases dramatically for hard-to-reach cases. In the Indonesian Family Life Study (Thomas et al 2000), tracking outside the local enumeration areas cost 20% more than local tracking.

Generally the studies reviewed tracked respondents within their study areas. The studies that covered whole regions or the entire country had the advantage of having a network of interviewers in several locations and could track respondents who moved from one study area to another. Studies need to set criteria and limits on the time, money and resources they will put aside for tracking.

When deciding the tracking limits, studies need to consider where people move. In the Indonesian Family Life Study (Thomas et al 2000), 40% of respondents found by tracking were located within their study area and 60% had moved to another study area. In the Malaysian Family Life Study, 28% of respondents found by tracking were located in the same district, 4% in a different district in the same state and 5% in different states (Haaga et al 1994). In the

Pelotas study (Barros et al 1990) and the Cebu study (CLHN 1989), 6% and 27% of their respective respondents moved outside the study area (the metropolitan limits in both cases).

If resources allow, tracking outside the study area should be considered but with a time or distance limit. The Indonesian Family Life Study (Thomson et al 2000) set a limit for local tracking of half an hour by public transport. Regional and national studies should consider tracking respondents who move to other study areas. This requires that studies make provisions for handling information gathered, keeping information updated and getting it to those who need it. A successful approach in several national studies (Haaga et al 1994, Thomson et al 2000) has been to have a central 'tracking shop'. In the Indonesian Family Life Study (Thomson et al 2000), tracking was divided into two distinct phases:

1. Local tracking: This was defined as tracking people who could be reached by interviewers from the study area by public transport in half an hour.
2. Tracking phase: Details of respondents who had moved further away were sent to the study's central office who assigned their tracking to interviewers in the relevant areas once the main phase of fieldwork was completed. Each interviewer then planned the most efficient visit route. The interviewers reported to the central office several times a week, where they were informed which respondents they should prioritise and when tracking should end. This centrally co-ordinated tracking allowed better planning and a more efficient use of time and personnel. This was important, as respondents who had moved far away needed effort and organisation to track.

4.3 The shorter the gap between contacts the better

Tracking respondents who have moved becomes more difficult the longer the length of time that has elapsed since the move. Contacting respondents at regular intervals allows records to be regularly updated and respondents to be tracked as quickly as possible after they have moved. These contacts can either be survey rounds themselves, or contacts solely for tracking purposes.

4.4 Respondents should feel part of the study

Respondents who feel part of the study are more likely to make an effort to inform the project they have moved or leave an address with neighbours. A good respondent-study relationship can be encouraged by keeping respondents informed about the study and about tracking, and by using tangible reminders such as radio publicity (Barros et al 1990), stickers, fridge magnets, newsletters, birthday cards, toll-free numbers and change of address cards (Richter and De Wet forthcoming). The Birth to Ten Study managed to create a real sense of ownership among respondents, and recommends producing simple materials for respondents such as 'study passports' with a photograph and ID number (De Wet personal communication).

4.5 Choice, training, supervision and motivation of interviewers

Studies have found that a careful choice of interviewer is essential to ensure successful tracking (Thomas et al 2000, De Wet personal communication). The Indonesian family life study (Thomas et al 2000) found that success in tracking was associated with interviewer quality (as measured by mathematics scores and salary in their previous job) and contact with the supervisor.

Interviewers should have certain personal characteristics: enthusiasm, flexibility e.g. in scheduling appointments, ingenuity and perseverance are all-important. Tracking is made easier if interviewers know the community, as this will help them locate households and appropriate tracking contacts. It is also desirable to use the same interviewers in each survey round, as they are familiar with the study area and it helps foster a good relationship between the interviewers and the respondents.

Studies need to equip interviewers with all the skills they need for successful tracking.

Interviewers must be well trained, motivated and supported and their workload kept realistic.

Interviewer training should include techniques of tracking and the legal and ethical implications

of tracking (these will vary depending on the context). The Nang Rong project (CEP-CPC Study) provides excellent instructions in their field manual for tracking out-migrants; they clearly describe the process (see box 1) and record information about each interviewer's attempts to try and find the migrant.

BOX 1: AN EXAMPLE OF GOOD INSTRUCTIONS FOR INTERVIEWERS REGARDING TRACKING TAKEN FROM THE NANG RONG PROJECT 'MIGRANT FOLLOW UP FIELD MANUALS' (CEP-CPC STUDY)

DEFINITION OF A TRACKING ATTEMPT

"The best way to think about an attempt is to imagine the attempt as one path where there are a series of clues leading to an interview with a migrant. Each clue would be linked to the next. If the path or trail of clues does not lead to the location of the migrant and a new set of clues is needed then this would be the beginning of a new attempt.

For example, an interviewer may have information from a respondent at the migrant's origin house in Nang Rong. That information may be that the migrant works in factory A in Bangkok. Then the interviewer finds the telephone number for factory A, calls factory A, talks to a personnel manager, and then the personnel manager tells the interviewer the migrant's home address and telephone number. The interviewer calls the home telephone number and talks to the migrant, makes a date for the interview and conducts the interview. This would be considered one attempt, which uses a number of different ways to locate the migrant.

Thinking about the above example, if the telephone number is a wrong number and then the interviewer goes to the home address and cannot find the migrant and must return to factory A to find someone else to ask for better directions this would be the end of the first attempt and the beginning of the second attempt.

Sometimes an attempt will take only one day's time and sometimes an attempt will take several days' time. So, remember that an attempt is a series of linked clues that brings the interviewer closer to finding the migrant."

Facilitating communication between interviewers can be both motivational and an opportunity for interviewers to learn from each other and should be encouraged. Methods of motivation used in the Indonesian Family Life Study (Thomas et al 2000) include interviewers working in pairs when tracking respondents who had moved far afield and giving rewards per respondent located.

5. Conclusion

Tracking in panel studies in developing countries can significantly reduce attrition. Attrition can cause a reduction in sample size and potentially causes bias. Tracking in developing countries can be difficult because populations are highly mobile and infrastructures are often poor. We have shown, however, that several panel studies in developing countries have had good follow up rates as a result of tracking respondents. Although tracking procedures need to be context specific, there is much to learn from these studies. Studies must consider all of the different options for tracking and use those which are locally appropriate, have a well defined tracking protocol with the rules and limits of tracking, make the gap between respondent contact as short as possible, make respondents feel part of the study and select, train, supervise and motivate the interviewers well.

References

- Alderman H, Behrman J, Kohler H, Maluccio JA, Watkins SC. 2001. Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples. *Demographic Research* 5(4). Available at: [http://www.demographic-research.org/? http://www.demographic-research.org /volumes/vol5/4/](http://www.demographic-research.org/?http://www.demographic-research.org/volumes/vol5/4/)
- Anderson A, Richter LM. 1994. Birth to Ten: Error detection and correction in a longitudinal address database. *S Afr J of Epi & Infect* 9 (1):16-19.
- Barros FC, Victora CG, Vaughan JP. 1990. The Pelotas (Brazil) birth cohort study 1982-1987: strategies for following up 6,000 children in a developing country. *Paediatr Perinat Epidemiol.* 1990 4(2):205-220.
- Buck N, Burton J, Laurie H, Lynn P. 2002 DRAFT. Quality Profile British Household Panel Survey Waves 1 to 10: 1991-2000. Available at: [http://www.iser.essex.ac.uk/bhps/ index.php](http://www.iser.essex.ac.uk/bhps/index.php)
- Cebu Longitudinal Health and Nutrition Survey (CLHN). 1989. *Survey Procedures*. Available at <http://www.cpc.unc.edu/projects/cebu/docs/clhn1.pdf>
- CEP-CPC Study of Social Change in Nang Rong, Thailand. *Information and Migrant follow up field manual*. Available at: [http://www.cpc.unc.edu/ projects/nangrong/intro.html](http://www.cpc.unc.edu/projects/nangrong/intro.html)
- China Health and Nutrition Survey (CSHN). 1999. [http://www.cpc.unc.edu/projects/china/china _ home.html](http://www.cpc.unc.edu/projects/china/china_home.html)
- Falaris E, Peters E. 1998. Survey Attrition and Schooling Choices. *Journal of Human Resources.* 33(2): 531-554.
- Fitzgerald J, Gottschalk P, Moffitt R. 1998. An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources* 33(2):251-299.

- Glewwe P, Jacoby H, King E. 1999. Early childhood nutrition and academic achievement: A longitudinal analysis. *FCND discussion paper No. 68*.
- Haaga J, DaVanzo J, Peterson C, Tey Nai Peng. 1994. Twelve Year Follow-Up of Respondents and Their Adult Children in a Panel Survey in Peninsular Malaysia. *Asia-Pacific Population Journal* 9(2): 61-72.
- Kuate Defo B. 1992. Mortality and attrition processes in longitudinal studies in Africa: an appraisal of the IFORD surveys, *Population Studies* 46(2): 327-348.
- Lillard L, Panis S. 1998. Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality. *Journal of Human Resources* 33(2):437-457.
- Maluccio J. 2000. Attrition in the Kwazulu natal income dynamics study, 1993-1998. *FCND Discussion paper no. 95*. Available at: <http://www.ifpri.cgiar.org/srstaff/maluccij.htm>
- May J, Carter M, Haddad L, Maluccio J. 2000. KwaZulu-Natal Income Dynamics Study (KIDS) 1993-1998: A longitudinal household data set for South African policy analysis. *Development Southern Africa* 17(4): 567-581.
- Richter L, Yach D, Cameron D, Griesel D, de Wet T. 1995. Enrolment into birth to ten: population and sample characteristics. *Paediatric and Perinatal Epidemiology* 9(1): 109-120.
- Richter L, T de Wet. Forth coming. Almost there...Birth to Ten Children are 9 years old.
- Singh C. 1995. A Comparative Analysis of Attrition in Household Panel Studies. Comparative Research on Household Panel Studies. *Document no 10. CEPS/INSTEAD B.P. 48 L-4501*. Available at: <http://www.ceps.lu/paco/documents/pacwp10.doc>

Thomas D, Frankenberg E, Smith J. 2000. Lost But Not Forgotten: Attrition and Follow Up in the Indonesian Family Life Survey. *RAND Working Paper DRU-2059-1-NICHD*.

Twisk J, de Vente W. 2002. Attrition in longitudinal studies: How to deal with missing data. *Journal of Clinical Epidemiology* 55 (4): 329-337.

Vietnam Longitudinal Survey: <http://csde.washington.edu/research/vietnam/documents.html>

World Bank. 1993. *Basic Information Peru Living Standards Survey (PLSS), 1990 Lima, Peru*. Available at: <http://www.worldbank.org/html/prdph/lms/country/pe90/docs/pe90.pdf>

Zabel J E. 1998. An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior. *Journal of Human Resources* 33(2):479-506.

Ziliak JP, Kniesner TJ. 1998. The Importance of Sample Attrition in Life Cycle Labor Supply Estimation. *Journal of Human Resources* 33(2):507-530.