KM 32(1989)
pag 41 - 63

REDUCING PANEL BIAS,

A REVIEW OF SAMPLING DESIGNS

Frank van de Pol*)

Summary

The sampling design of a panel is closely related to three problems that are inherent to panels: following a changing population, attrition and panel effects. First the design determines whether the panel sample is up to date, i.e. to which extent the problem of a changing population is overcome. Next the panel design determines to which extent the sample is maintained and whether control groups will be used. These features contribute to reducing the other two panel problems, selective attrition and panel effects. These problems are especially severe if the panel is also intended to provide descriptive statistics.

The sample may be updated by recruitment from the families of panel members or by another method of random recruitment. Selective recruitment in underrepresented strata (quota sampling) is a method for both updating and maintaining the sample, but will not reduce panel effects. Rotation is a way of updating and maintaining a panel sample by refreshing the sample continuously, thus confining panel effects. However, rotation cannot be applied if one wants to do long-term analysis. A split panel design is a combination of a panel without rotation and repeated independent cross-sections that serve as control groups. Attrition bias and panel effects are then not automatically restricted, but may be quantified.

The impact of the interview frequency is treated briefly. The final sections discuss some practical issues, such as the choice for stable units of measurement (persons or manufacturing sites, not unstable aggregates like households or firms), and financial considerations.

1. Introduction

Some of the most common reasons for starting a panel are getting information on micro level (respondent) changes, getting more efficient estimates of average change and reducing costs of data collection. Panels have many more benefits as well as drawbacks, which have been discussed in Duncan and Kalton (1985), Kish (1987), Cook and Campbell (1979), Davis (1978) and Van de Pol (1988). Related works are Nesselroade and Baltes (1979) and Goldstein (1979).

*) Centraal Bureau voor de Statistiek, Hoofdafdeling Statistische Methoden, postbus 959, 2270 AZ Voorburg

This paper focuses on the panel sampling design, which has a great impact on three panel problems. Firstly, the sample tends to become out of date; secondly, there may be a selective drop-out of panel members and, thirdly, panel membership may induce biased answers. These points are explained further below. Several sampling designs will be discussed that may reduce one or more of these problems (sections 2, 3 and 4). In section 5 the interview frequency is discussed briefly and section 6 examines the rules for following panel members. The paper concludes discussing some practical issues.

In order to choose an appropriate sampling design the purposes of the study should be made explicit. In official statistics an important goal is to obtain descriptive statistics, which implies an emphasis on reducing bias and variance. Bias may be reduced by the implementation (to some extent) of repeated independent cross-sections in the panel design (sections 3 and 4). Furthermore, a distinction may be made between current estimates, estimates of (net or gross) change and aggregates over time (for instance yearly totals), each having their own minimum variance design, as well as minimum variance estimators (Patterson, 1950, Cochran, 1977, Van de Pol, 1982, Kish, 1987). The main focus of this paper is on reduction of bias rather than variance.

Another goal with explicit consequences for the design of a panel is the study of long term dynamics on the micro level of respondents. Respondents have to be followed for a long time when the consequences of some circumstances or some sort of behaviour are expected to emerge only after several years. In such circumstances the renewal of the sample should be delayed as long as possible (section 2). Independent cross-sections may be added as control groups (section 4).

1.1. Population changes and updating the sample

In many cases panels are not (only) designed to observe individual (respondent) changes, but also to monitor changes in the characteristics of the population. A problem may emerge if the composition of the population changes and the panel sample does not change accordingly. In the Dutch population, for instance, family size has decreased dramatically in the eighties (AGB, 1986). The rules for following panel members would determine whether the change would be reflected in the panel.

It depends on the definition of the target population whether such a discrepancy between the current population and the sample is problematic. For a cross-sectional survey, the term population refers simply to a set of potential respondents with certain characteristics in a certain region at some *fixed* point in time. For panel sampling, three alternatives emerge.

The first alternative applies if a population is selected at a specific point in time (a cohort) and this 'fixed population' is to be followed through time. Then the target population relates to the time that the panel was started. This implies that people who immigrate after the start of this period are not members of the population, neither are those who come of age after that time. On the other hand, emigrants and other people who are no longer eligible will remain part of a fixed population (at least in theory). With this definition of the population the panel sample does not have to be updated.

The second alternative is that the panel sample should always be up to date. Then a population should be defined at every point in time that a panel wave is to take place. Some restrictions may be imposed on the current population, e.g. age range and region, the members of these consecutive populations will only partly be the same. Immigrants into the region enter the population, as do young people. On the other hand, there is an outflow of emigrants and some population members will die. The panel designs described in sections 2, 3 and 4, provide some way of keeping the sample up to date with the changing population.

A third alternative is to include both inflow and outflow in the consecutive target populations. Then at every wave a sample is obtained that includes a sample from a previous fixed population as well as a sample which is up to date. If a fixed population from a previous date is analyzed the panel members who were added for updating should be skipped and if the current population is analyzed those panel members that presently do not have the required characteristics should be skipped (for instance those who moved outside the relevant area).

As mentioned at the beginning of this subsection, the rules for following panel members are also linked up with the definition of the population. If the population is defined as having households as its constituent elements, then households that break up, for instance through divorce, should not be followed and newly formed households should be added to the sample in some

way. Firms that split up or that are taken over by other firms should be treated along the same lines. If, on the other hand, the population is defined as having persons or manufacturing sites as its elements, other following rules prevail. In section 6 some details on these rules will be discussed.

1.2 Nonresponse, attrition and panel maintenance

If the panel is a survey, part of the sample will not respond at the recruitment wave. This is partly due to errors in the sampling frame. For instance some addresses do not exist or are not up to date. When the sample is reduced to its net size there will still be a part that does not respond. Some people refuse to cooperate, and others cannot be contacted because they are often out or because they are on holiday. Bias will occur if the *nonresponse* is selective; i.e. if these statistics of the nonrespondents differ from those of the respondents. Nonresponse bias may be reduced by poststratification and other methods of weighting (Bethlehem, 1985), or by applying special models.

Once the panel sample has been established there will be inflow and outflow in the population and therefore also in the sample (if this is provided in the sampling design). We shall call this *legitimate outflow* (and inflow). Legitimate outflow occurs because part of the sample dies or, for example in a panel of civil servants, because they leave the service. Apart from the legitimate outflow there is also unwanted, illegitimate outflow from most panel samples, just like there is nonresponse at the recruitment wave. Respondents may *drop out* and if they are lost permanently the size of the sample is reduced; this constitutes panel *attrition*. Apart from the reasons mentioned for nonresponse, dropping out may also occur because people move, even if movers are followed. If this dropping out from the sample is selective in some sense bias may occur in the statistics that are computed. Some evidence of selective dropping out can be found in Sobol (1959), Waterton and Lievesley (1987) and Van de Pol (1988).

One way to counteract attrition is to choose an appropriate sampling design, as will be explained in the next sections. Another, complementary, way is weighting the respondents with the inverse of their predicted probability to stay in the sample or a similar method. Weighting methods are reviewed by Kalton (1986), Hensher (1985) and Van de Pol (1986).

1.3. Panel effects and panel maintenance

Another problem in panel studies is that respondents may give systematically different answers on each occasion, creating the so-called *panel effects*. Some questions may raise the respondent's curiosity and therefore he may be better informed on the next occasion or even change his opinion on the subject (Vierkant, 1980). This panel conditioning effect is especially detrimental when changes in attitudes or opinions are the subject of study. When the focus is not on attitudes but on behaviour respondents may easily become tired of giving the same sort of information over and over again (Bailar, 1975, Golob and Meurs, 1986). On the other hand panel members will get trained in giving the correct answer, which is particularly relevant for complicated questionnaires, as used in budget surveys.

Fortunately respondents tend to forget the issues raised during an interview, although an interview on, for instance, war experiences may be less easily forgotten. Hence one way to reduce panel effects is to give the panel members enough time to forget the last interview (with the same questions). This may be realized by bringing down the interviewing frequency or by replacing (sections of) the questionnaire for every measurement. If, however, the questionnaire not only takes a "snapshot" of the present situation but also contains retrospective questions, a low interview frequency would be undesirable because of memory effects (Sudman and Bradburn, 1974, Sikkel, 1985). Another way to (partially) counterbalance panel effects is to make sure that at every wave at least part of the sample is a fresh sample. This kind of panel designs is treated in sections 3 and 4.

2. Updating and maintenance with permanent sample membership

A very simple panel design is to draw a sample and follow its members without replacing them at any time. This design seems appropriate if a fixed population (a cohort) is to be followed. In many cases, however, information on the current population is (also) required. It was pointed out in section 1.1 that it is a safe strategy to include both inflow (youngsters, immigrants) and outflow (emigrants) into the population.

In the next section some methods for keeping the sample up to date are described. Section 2.2 discusses quota sampling, which may not only keep the sample up to date, but may also reduce biases due to attrition. In section 3

rotating designs are treated, which not only reduce attrition bias, but also restrict panel effects. This is achieved at the cost of no data on long-term individual dynamics. The split panel design, which is considered in section 4, does not have the latter drawback. It does not restrict attrition bias and panel effects, but enables their quantification (assuming that a first measurement is unbiased).

2.1 Correcting for population inflow only

While attrition is absent in administrative panels, in panel surveys it always occurs. In an administrative panel the sample only has to be updated from time to time (unless one is not interested in the current population). The inflow into the population may be picked up by regularly drawing additional (random) samples from youngsters and immigrants. In section 3.2 a technique for drawing such additional samples from an official register is discussed. If the sampling frame does not identify these subpopulations, another way of sampling the population inflow is necessary, e.g. the *auto-rejuvenation* method.

If we have a random sample of adults (within certain strata) their children will be a random sample too, representing the population inflow of youngsters. So, adding the children of panel members to the sample is a method of updating the panel sample. The age limit for entering the sample does not have to be zero. Often children will enter the sample when they come of age. However, immigrants will not enter the sample in the first generation, neither will orphans. Hence in the long run, additional samples of immigrants and orphans should be added to the panel sample if these are sizable parts of the population.

When applying this strategy it should be clear what part of the population these new panel members stand for, or stated otherwise their *probability of inclusion* in the sample must be known. In a sample of households the auto-rejuvenation process is quite transparent. Within a household with inclusion probability π_h, child i has the same inclusion probability as the family he or she belongs to (when entering the sample), $\pi_i = \pi_h$.

In a sample of persons, however, things get slightly more complicated. A child who comes of age in a one-parent household has the same inclusion probability as the one parent that brought it into the sample. (This is in

accordance with the thusfar implicit rule that if a parent in the sample does not live in the same household as his children, then his children should not be included in the sample.) However, a child, i, in a two-parent household (with parents j and ℓ) has approximately a double probability of being selected, once via one parent and once via the other, $\pi_i = \pi_j + \pi_\ell$.

2.2 Quota sampling: correcting for both population changes and attrition

When attrition is selective, the composition of panel samples tends to deviate from that of the population after some time. This is especially so if the response burden is high, e.g. in consumer panels that have waves every week. Random recruitment of new panel members will not mend the sample, nor will it be a solution to add equal proportions of the population in every stratum, whether the stratum has a high attrition rate or not.

A good method to maintain the panel would be to add population members that have the same scores on target variables as the drop-out respondents. Unfortunately, however, there is no sampling frame available with target variables in it (if it existed no further research would be necessary). But what can be done is to select new population members based on relevant auxiliary variables, such as age, sex, marital status and the socio-economic status of the neighbourhood.

Using these auxiliary variables for stratification it may be assured that in each stratum about the same fraction of the population is a respondent in the panel sample. The rationale behind this method, which is known as *quota sampling*, is the same as the rationale behind poststratification. It is based on the assumption that respondents and drop-outs have the *same (multivariate) distribution* on target variables *within a stratum* (Bethlehem and Kersten, 1985). If this assumption is valid, drop-outs may be replaced by other units from the same stratum without any resulting bias. In some cases, however, nonrespondents and panel drop-outs are *not* similar to the respondents (Leeflang and Olivier, 1983, Waterton and Lievesley, 1987).

In order to fill up the strata to their quota, additional respondents should be sampled with probabilities that are different for every stratum (but the same within every stratum). So one should have a sampling frame with information on the strata for every member of the population. Four sorts of sampling frames can be distinguished for this purpose.

The simplest strategy is to employ an official register, and one is confined to the (limited number of) characteristics in the register.[1)]
This limitation is less severe if the data of a recent, large survey are used as a master frame. In many countries, large scale surveys on the labour force, on use of the media, on medical care are carried out which may be used as a sampling frame.[2)]
A third possibility is to carry out a survey dedicated to the selection and recruitment of panel respondents. In the Netherlands, a number of marketing bureaus operate a 'mini-census' for this purpose. The fourth strategy is to extend the data from an official register with characteristics that are recorded in the field, by the interviewer.[3)]

The last strategy has been commonly used in a way that has given quota samples a bad reputation. Cochran (1977) writes "To expedite the filling of quotas, some latitude is allowed to the enumerator regarding persons or households to be included. ..... The quota method seems likely to produce samples that are biased on characteristics such as income, education, and occupation, although it often agrees well with the probability samples on questions of opinion and attitude." If, however, the quotas are not filled in the field but from a good sampling frame the method is no worse than post-stratification.

1) Several sampling frames for the Dutch population are generally available; all have some drawbacks. Van der Eijk et al. (1981) mention four sampling frames for the Dutch population. The municipalities keep two registers of persons, the electoral register and the population register. Both are becoming less suitable for national samples of individuals because municipalities increasingly refuse to cooperate or ask too high a price. Therefore nowadays mainly household samples or address samples are drawn. These may be obtained from the registers of dwellings which are kept by the municipalities or from the mail delivery register which is exploited by the postal services. However care must be taken that the inclusion probabilities do not depend on the size of the household. For sampling firms one may think of the legal administration of firms, in Holland the chambers of commerce. The Netherlands Central Bureau of Statistics (CBS) has its own sampling frames, one for firms and another one for addresses.

2) One may also gain advantage from an existing survey by considering it as the first wave of a panel survey. This is sometimes called a follow-up study.

3) The instruction may be for instance to find a respondent in a specific street with certain characteristics on age, sex and marital status. Although interviewers may be instructed to work through the street in a systematic sampling procedure, some will probably look for people with the desired characteristics in a selective manner. Neighbours who are willing to help will probably not send the interviewer to people who are known to be seldom home or who they suspect will not be interested in the subject. So people with a low response probability may be avoided, thus aggravating the biases due to selective nonresponse and attrition. A less important consequence of asking for people with special characteristics is that response figures may be given an unrealistic face-lift. In fact for many quota panels no reliable response figures are known.

2.3. Permanent panel membership: summary

Figure 1 summarizes the effects of non-rotating panel sampling designs on three panel sample deficiencies. Only outdatedness can be prevented to some degree with a nonrotating design. Practical alternatives for random entry of new population members are auto-rejuvenation and quota sampling (if there is a separate stratum for new population members or young people).

Figure 1. Ways of introducing new respondents in the (non-rotating) sample and their effects on three panel sample deficiencies.
(- none, . small , + considerable)

way of introducing new respondents in sample	section	outdated-ness	selective attrition	conditio-ning
no entry	§ 1.1	-	-	-
random entry	§ 2.1	+	-	-
auto-rejuvenation	§ 2.1	+	-	-
filling up quota	§ 2.2	-[1)]	.	-

[1)] If new population members form one of the strata: +

Selective attrition and panel effects (if present) can hardly be reduced. Only quota sampling can counteract bias due to attrition to some extent. But if this sampling design is carried out without a good sampling frame new biases may be introduced.

3. Updating and maintenance by renewing the sample

Panel samples that are either not updated or not maintained have several deficiencies. The sample tends to grow out of date, and strata with a high attrition rate become underrepresented. Moreover, panel members may give systematically different answers in different waves (panel effects). These flaws are especially detrimental to the validity of descriptive statistics.

The present section describes panel sampling designs that seek to neutralize these deficiencies mainly by renewing the sample systematically. Such *rotating panel samples* are especially suitable for achieving a sample that is up-to-date with the population. Also, the bias that may be caused by attrition and panel effects is probably reduced (but not eliminated). If, for instance, the first measurement is considered least biased, then a model may be conceived to correct the panel estimates for bias.

The same may be done if a *split panel*, a combination of a panel without rotation and repeated cross-sections, is applied (section 4, Kish, 1987). Similar designs have been developed in other fields. In primarily bio-medical research on child development *mixed-longitudinal designs* combine a cohort study with repeated cross-sections in much the same way as a split panel does (Prahl-Andersen et al., 1979). Cook and Campbell (1979) deal with a great many of *quasi-experimental designs* and focus among other things on which inferences are valid for each design.

3.1. Rotation

In a rotating panel design a new group of panel members is recruited at every wave. The size of the sample is controlled by restricting the duration of panel membership to a fixed period of for instance 6 waves or two years. Exaggerating somewhat, it might be said that the sooner the respondents are dropped from the sample the better the sample is. If respondents are interviewed only once, there will be minimum outdatedness, no attrition and respondents will not be conditioned by previous interviews. But then we have repeated independent samples, not a panel sample.

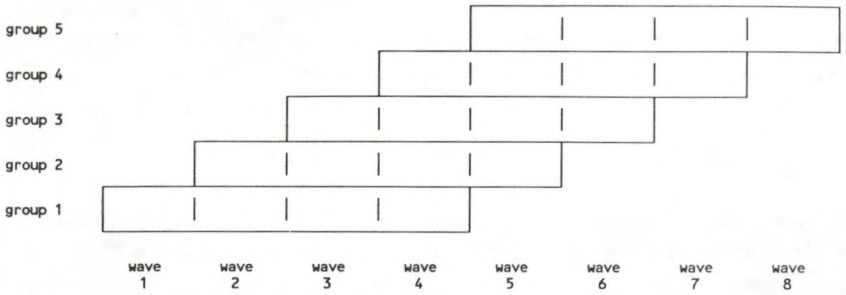
Of course the benefits of repeated independent samples should not be overrated. They too have a nonresponse problem and they lack experienced respondents, which may be a drawback in the case of difficult questions. Also, if scores on consecutive panel measurements are positively correlated (as they usually are), trends are estimated with a higher variance by repeated independent samples than by a panel (Kish, 1987, Van de Pol, 1988).

If, on the other hand, respondents are never dropped and no other method of updating or maintaining the panel is applied, the panel sample will become less and less up to date, and will deteriorate more and more if selective attrition and panel effects occur. By applying rotation, the deterioration of the sample is bounded because the numbers of new respondents, respondents with one interview experience and respondents with more experience in the sample are fixed (apart from fluctuations in the attrition rate). Thus the amount of bias resulting from outdatedness, attrition and panel effects is virtually constant once the rotation process has started.

Building up a rotating panel may be done group-wise. Suppose it is decided that respondents should not remain in the panel longer than W waves

and that the panel should have n respondents once the process is started up. Then, neglecting attrition, at the time of every wave a new panel group is added to the sample of net size n/W. At wave W the full sample size n is reached and the first panel group is interviewed for the last time. In figure 2 this is illustrated for W = 4.

Figure 2. Rotation with replacement of respondents after 4 waves



In practice, samples larger than n/W will have to be drawn, as considerable losses are to be expected due to initial nonresponse and attrition.

Rotation is especially useful for "measurement panels", panels that are mainly used to obtain descriptive statistics from separate panel waves (used as cross-sections). A simple example is the Private Car Panel (PCP) of the Netherlands Central Bureau of Statistics. This used to be a survey of independent monthly samples with a retrospective question about the mileage driven in the last month. Experiments showed, however, a 15% downward bias in these retrospective mileages. This bias does not exist in a panel design where the odometer is read at the beginning of every month.[4)]

PCP respondents are requested to participate in 4 waves. The first wave is conducted by an interviewer. After this basis measurement three monthly waves are conducted by telephone and then the respondent is dropped from the sample.

Mean changes are usually measured more efficiently with panel data than with independent longitudinal samples because of the overlap of the panel samples. So, for measuring the mean month-to-month change, the PCP panel is more efficient than the (longitudinal) cross-sectional design that preceded

4) This was investigated at the Netherlands CBS by Peeters in 1980.

it. However, for year-to-year comparisons this gain in efficiency was not achieved, because there is no overlap between the rotating panelsamples that are twelve months apart.

This is one of the reasons why the U.S. Current Population Survey (CPS) comes back to his respondents after a year. In the CPS respondents are interviewed in four consecutive months, then for eight months they leave the sample, and finally they are interviewed again in the same four months as the year before (Hansen et al., 1955). This design will not only lower the variance of year-to-year comparisons somewhat, but will also improve the comparison of the same month in consecutive years. However, nonresponse rates nowadays are higher than they used to be. Nowadays the gain in efficiency achieved by coming back after a year may be offset by bias due to attrition.

It was stated above that rotation will reduce bias due to outdatedness, attrition and panel effects. These biases appear to be dependent on the time the respondent remains in the sample (or the interview frequency, which is directly related to the former parameter for a fixed number of interviews per respondent.) If the respondent is to stay in the sample for several years, the rotation groups may be updated as described in the next subsection. The relation between the interview frequency and other panel biases is described in section 5.

3.2. Updating rotation

In the rotating design of the previous subsection the only inflow of new panel members is the inflow of complete panel groups. The panel group samples are not updated. However, when the intervals between panel waves are large, it may be necessary to update the panel group for inflow in the population (youngsters and immigrants). A system to incorporate population changes in the panel sample was described by Bäcklund (1975) and, with modifications, by De Ree (1983).

Suppose the members of a panel group are put together in the sampling frame and that nobody enters or leaves the sampling frame. Then the same panel group can be sampled at every wave by picking out the same part of the sampling frame. Let us now suppose that the sampling frame is continually updated according to population changes and that the new population members are brought into the sampling frame on a random position. Then picking out

the 'same part' of the sampling frame at every wave will result in slight changes in the composition of the panel groups. New population members will have the same probability of inclusion in the sample as the old ones (in the same stratum).

This method could only be used to update the sample (subsection 2.1), but once we are in a position to update the sample we might rotate it as well. Adding a rotation group may be realized by, at every wave, slightly shifting the part of the sampling frame to be picked out. So one may also say that this *updating rotation* is achieved by drawing largely overlapping samples for consecutive rounds.

Updating rotation can be applied if there is a procedure to add new population members randomly into the sampling frame (or, even better, randomly within relevant strata). If such a procedure is available, updating rotation will be implemented by assigning a random key number to every population member. Rotating through such a sampling frame is easily performed by selecting key numbers within a certain range for wave 1 and selecting key numbers within a somewhat different, but largely overlapping, range for wave 2, etc..

In the design of figure 3 the random numbers 1 - 2000 are selected in the first wave, the numbers 1000 - 3000 in the second wave, etc. So this design will keep respondents in the sample for only two waves. Of course a design with a longer panel membership is also feasible.

People who leave the population, because of death or emigration, will not be sampled on the next occasion, provided the sampling frame is up to date (c.f. respondents Nassau and Ceased). The main advantage of this design, however, is the automatic sampling from new population members. Respondent Young was born into the population between wave 2 and wave 3. He received a random number (2410) that, by chance, automatically introduced him in the sample. Respondent Born on the other hand was assigned a number (4163) that did *not* introduce him in the sample at wave 3.

In practice, most databases are not accessible for outsiders because of privacy reasons. However, it is sufficient if the database supervisor is willing to run a job for you from time to time, as it is not necessary to sort the file and add a random key number. First a sorted file as displayed in figure 3 is not necessary, because what has to be done is just retrieving records with a key number within a certain range. The only advantage of real

Figure 3. Population members sorted on a random key number in a continually updated sampling frame. In the first panel wave numbers 1-2000 are sampled, in the second wave numbers 1000-3000, etc.

name	wave 1	wave 2	wave 3
N.O. One	**0001***	0001	0001
J. Bond	**0007***	0007	0007
S.O.M. Body	**0987***	0987	0987
S.T. Elizabeth	**1421***	**1421***	1421
W. Nassau	**1584***		
C.B. Statistics	**1299***	**1299***	1299
S.O. Sealand	**1953***	**1953***	1953
D.E. Ceased	2112	**2112***	
O. Range	2345	**2345***	**2345***
J. Young			**2410***
E. Psilon	2712	**2712***	**2712***
P.I. Circle	3142	3142	**3142***
P. Erson	3745	3745	**3745***
P.O.P. Member	3811	3811	**3811***
I.N. Habitant	3951	3951	**3951***
N.E.W. Born		4163	4163
E.A.U. de Cologne	4711	4711	4711

*: selected in the present wave

sorting is that the sample size can be fixed exactly, whereas when selecting records within a certain range one can only fix the sample size approximately.

Secondly the random key number does not have to be entered but can be deduced from variables that are already present in the file. Sunter (1986) points out that a pseudorandom number may be defined as some function of fixed characteristics like date of birth, telephone number, tax number, last name, or whatever is available in the sampling frame. So the same pseudo-random key number can be computed on every occasion that the sampling frame is passed through to pick out a sample.

However, the sampling frame used must match the target population. People who move will be lost if this updating method is applied to a sampling frame of addresses (unless the target population consists of addresses or houses). At the Netherlands CBS updating rotation is succesfully applied for sampling firms from the general firm register.

Updating rotation is not only a good method for keeping the sample up to date, but is also a suitable means to *spread the response burden in small*

populations. Suppose a sample of size 850 must be drawn from for instance a population of 8500 primary schools. Then updating rotation will prevent some schools from never being sampled while other, perhaps neighbouring, schools are sampled every year.

Updating rotation will add a random sample of new population members to every panel group at every panel wave. However, if no suitable sampling frame is available, then new population members may be introduced to the sample from the families of the panel members themselves (*auto-rejuvenation*, subsection 2.1). A somewhat more coarse method is to split up the rotation groups in age strata and to make sure that the number of panel members in these strata is correct at every measurement (*quota sampling*, subsection 2.2).

3.3. Rotation: summary

Figure 4 summarizes the effects of the rotating panel sampling designs on three panel sample deficiencies. Rotation is a good method to keep the sample up-to-date. If, however, panel groups remain in the sample too long to represent the population adequately, an additional updating method will increase the quality of the sample. As stated before, auto-rejuvenation is one practical alternative for random entry of new population members, and quota sampling is another (if there is a separate stratum for new population members or young people).

Figure 4. Ways of introducing new respondents in the rotation groups and their effects on three panel sample deficiencies.
(- none, . small , + considerable)

way of introducing new respondents in sample	section	outdated-ness	selective attrition	conditio-ning
no entry	§ 3.1	.[1)]	.	.
random entry	§ 3.2	+	.	.
auto-rejuvenation	§ 2.1	+	.	.
filling up quota	§ 2.2	.[2)]	.	.

1) For rotation schemes that drop respondents soon: +
2) If new population members form one of the strata: +

Rotation will confine bias that may be caused by selective attrition and by panel conditioning. A slight additional benefit may be obtained if the

panel groups are maintained in a selective manner, i.e. if, in strata with a high attrition rate, more new respondents are recruited (filling up quota). However, rotation will cut off the time series of the respondents after a relatively short period.

4 Updating and maintenance by control groups

When panels are designed primarily for the study of long-term dynamics on the micro level of respondents, they often are also used to obtain current estimates and trends of descriptive statistics. However, for the first objective long-term panel membership is necessary, and for the second refreshing the sample is to be recommended. Kish (1987) advocates a 'split panel' design that satisfies both prerequisites.

A split panel sample consists for one part of new respondents and for the remainder of long-term panel members. It differs from a rotating design in that the new respondents will, in principle, not be reinterviewed. Kish stresses, among other things, reduction of variance: the panel part of the sample is efficient for measuring (average) change and the other, repeated cross-sections, part is efficient for measuring aggregates over time, for instance an annual total.

The new respondents may also be viewed as a control group, which is not exposed to the reinterview condition and hence does not exhibit panel effects. Moreover, the size of the control group is not reduced by attrition and statistics obtained from this group cannot be biased by that factor. If statistics from the panel sample turn out to be systematically different from those from the control group samples, correction procedures may be devised for the panel part.

In practice, adding control groups for every panel wave will be inhibited by the fact that the costs of the panel will be approximately doubled in this way. Attrition bias and panel effects may also be traced with the aid of reference data from existing longitudinal surveys, which use repeated independent cross-sections, or other reliable official data on target variables.

5 The interview frequency

The interview frequency of a panel is largely determined by the type of data to be collected. When for instance data on daily expenditures are required, a high interview frequency is necessary, but if the focus is on the labour market or on attitudes the interview frequency can be lower. In the present section the relation of the interview frequency with some other factors is taken into account.

Manipulating the interview frequency cannot make the sample more or less up to date. There is, however, a relation between the interview frequency and the attrition rate, which seems to have two components. On the one hand there should be a negative association because most removals will be observed when respondents are interviewed frequently. Then the interview staff will not easily loose track of respondents. On the other hand there should be a positive association as a high interview frequency usually implies a high burden for the respondent. And a high response burden may lead to high levels of attrition. (If the respondents are fully informed in advance how time consuming panel membership will be, the initial response will be low instead, as is the case for instance with the budget survey at the Netherlands CBS).

Little is known about the net effect of both components together. A weekly interview will generally have either high nonresponse rates or high attrition rates, and a reinterview after say ten years will generally have a high attrition rate (although exceptions are known to exist, c.f. Saris and De Pijper, 1986, Freedman et al., 1980). Our educated guess of the interview frequency with the lowest attrition rate (for a fixed interval of time) is about once a year.

As mentioned in section 1.3, some panel effects depend on what respondents remember about the last interview. Hence, if the questionnaire aims at making a snapshot of the present situation, one should "stretch" the interval between interviews, thus giving the respondents the opportunity to forget the last interview. Interviews on matters like children, health and housing are usually experienced as interesting (as reflected by lower nonresponse rates and lower attrition rates) and therefore may be remembered somewhat longer than interviews on other subjects. Of course forgetting about the last interview is not the same as not being influenced by the last interview. Panel effects can also emerge if the last interview was

forgotten.

On the other hand, a panel design may be used to cover all events of some kind during the period between two interviews (retrospectively). Professional careers may be reconstructed successfully using retrospective questions, even without a panel design. For the Netherlands CBS labour force survey (Enquète Beroeps Bevolking), for instance, a retro-period of 12 months is used. Contacts with physicians, however, are known to be sensitive for memory effects (Sikkel, 1985). A respondent may situate an event after the last interview (due to telescoping), although it actually took place before that interview. Some of these errors may be mended if the last interview's questionnaire can be inspected by the interviewer. The best way, but also the most expensive, to suppress memory effects on retrospective questions is, however, to shorten the period between two interviews.

The last conclusion is at variance with the conclusion from the preceding paragraph. This is a problem if events as well as attitudes or behaviour prone to panel effects are to be studied in the same panel survey. In that case memory effects may also be reduced by using diaries (at the cost of more attrition) or by placing a computer at the home of the respondent with a diary-like program.

6. Following rules for panel members and implications for weighting

Some panel surveys are aimed at monitoring individual changes and others focus primarily on household characteristics. Working conditions and attitudes, for example, will often be analyzed on the level of individuals, whereas housing conditions and budget details are better suited to study on household level. It seems sensible that the panel sample should follow the subject of the analysis in time, i.e. follow persons if the analysis concerns a population of persons and households if not. In practice, however, it turns out that family composition changes so fast that a panel based only on those households that maintain the same composition in consecutive waves has such high rates of legitimate outflow (and inflow, if provided in the design), that valuable information is lost.

The same problem occurs when firms take over or sell off divisions or manufacturing sites. Production figures of such firms may be shaken up or down. The analysis should not concentrate on that part of the population

which had a stable composition over the period under consideration, because some of the changes that take place concern divisions or manufacturing sites that split up or combine. Hence in most cases individual persons or manufacturing sites should constitute the sampling population, even if the target population consists of an aggregate of these units.

If this is opted for, another question arises. Which part of a household that has split up is to be followed? The safest strategy is to postpone a decision and follow all panel members, as is done in the panel study on income dynamics (PSID, Morgan, Duncan et.|al. 1984), the Netherlands Socio Economic Panel and similar panel surveys. Adopting this policy a database is created that may be analyzed on the individual as well as household level.

Following all panel members in a panel analyzing households implies that every respondent who starts cohabitating with a panel member should become a panel member too. Hence some people marry into the panel and children of panel members enter the panel when they come of age or attain whatever age defines the target population. The inclusion probability of a child was dealt with in section 2.1.

The inclusion probability of a nonpanel spouse, i, is determined by two factors (Morgan, Duncan et al., 1984). First he (or she) is included because he married a panel member, j, who originally had inclusion probability π_j. Therefore this part of his inclusion probability is equal to the original one of "his" panel member. Secondly he had an extra inclusion probability at the time that sampling took place. This inclusion probability depends on the spouse's characteristics concerning stratification variables at that time. This value is unknown, but may be in the same order of magnitude as the original inclusion probability of his panel member, about π_j. Hence one may set the inclusion probability of nonpanel spouses, π_{ij}, to twice the original one of the panel member he is affiliated with, $\pi_{ij} = 2\pi_j$.

The present inclusion probability of their panel member was also affected by their marriage (or cohabitation). Now he (or she) has an extra probability of being included via the spouse, because of the spouses probability of being sampled, which was supposed to be about π_j. Thereforethe probability of inclusion of panel members changes after marriage to $\pi_{ji} = 2\pi_j$.
Weights for the inclusion probability should be corrected accordingly.

However the rule of following all panel spouses and children is not efficient when only individual data are to be analyzed. After some time a

large part of the sample will consist of respondents with halved inclusion weights. If weighting for nonresponse and attrition is applied and these weights do have a considerable variance, it may be decided to ignore inclusion probabilities and treat the sample as a quota sample.

7. Discussion

Considering which panel design best fits the purposes at hand is one thing, but carrying out such an ideal design seems to be quite another. Many panel samples are neither updated nor maintained, for various reasons. Firstly, panel studies are rarely designed to be permanent from the very start. *Later, new plans may arise* and funds may become available to extend an already existing database with another wave. For each single wave it may be argued that the inflow into the population can be neglected. The accumulated inflow over several years, however, cannot be neglected, at least not for the estimation of means and frequency distributions.

Secondly, there is a lack of *sampling frames* which are regularly kept up to date, especially if the target is some specific subpopulation. So it is often quite expensive or virtually impossible to fill up the sample with members of a specific subpopulation. Thirdly, it is not easy to perform a long-term analysis with a *datamatrix which is not rectangular* and which has holes in it. Therefore, long-term analysis itself being far from simple, many researchers will avoid the complicated sampling procedures that produce a non-rectangular datamatrix.

In the fourth place, it is sometimes argued that a sophisticated sample is not needed for the purposes at hand, especially for structural models. Measures of association, regression coefficients and structural parameters (Hartog, 1986, Hoem, 1985, Schirm a.o., 1982) are generally less sensitive to selective attrition than means and proportions (BGC, 1988).

Finally, *financial considerations* will also influence decision making about the design. A panel will be more expensive than repeated cross-sections, due to a higher nonresponse rate and keeping up the address file. Also, matching consecutive measurements of the same respondents and carrying out longitudinal controls are expensive, but it is not fair to count these costs if repeated cross-sections are the alternative. On the other hand, repeated cross-sections are more expensive due to more first interviews.

Furthermore the interview frequency is a determinant. A panel wave of interviews is cheaper if the interview frequency is high (every week). In that case a panel may be even cheaper than repeated cross-sections, especially if there are high costs per respondent, such as training the respondent for a difficult questionnaire or installing a home computer. Another cost affecting factor that may be mentioned is the effort put into following panel members who have moved.

Morgan and Duncan (1983) claim on the basis of their PSID experience that panels are "not appreciably more costly than separate cross-sections". However the PSID has very little attrition. Applying autorejuvenation only (no supplements to fill up quota or whatever), the size of the PSID sample has grown. Moreover, autorejuvenation is probably the cheapest way of recruiting panel members. Recruitment in other sampling designs is more expensive, especially in a quota sample, where not all respondents that are recruited are actually added to the panel sample. This cost difference between quota samples and rotating samples may be outdone by the fact that for a rotating sample more first interviews have to be conducted.

A more important cost determinant than the choice between designs is the interview method. Mail surveys are cheapest but will induce more drop-out unless additional measures are taken. Telephone interviews come second, and face-to-face interviews come third. The last two alternatives can be carried out with the assistance of a home computer at the respondent's home, a CATI system, or hand-held computers, carried by the interviewer. Of course, such an investment requires regular use of these computers. Finally, costs depend on response and attrition rates, which, in turn, are affected by the subject of the study, the difficulty of the questionnaire and the capability of the interview staff.

REFERENCES

AGB, 1986, AGB Mini-census 1986, een trendrapportage (Netherlands AGB, Dongen).

Bäcklund, S., 1975, Coordinated samples with the JALES-technique, contributed papers of the 40th session of the International Statistical Institute (Warschau).

Bailar, B.A., 1975, The effects of rotation group bias on estimates from panel surveys. Joural of the American Statistical Association 70, pp. 23-30.

Bethlehem, J.G., 1985, LINWEIGHT user manual, CBS report (Netherlands Central Bureau of Statistics, Voorburg).

Bethlehem, J.G. and H.M.P.Kersten, 1985, On the treatment of nonresponse in sample surveys. Journal of Official Statistics 1, pp. 287-300.

BGC, 1988, Meetfouten in het panel, een concept-onderzoeksverslag (Bureau Goudappel Coffeng, Deventer, Netherlands)

Cochran, W.G., 1977, Sampling techniques, 3rd. ed. (Wiley, New York).

Cook, T.D., and D.T.Campbell, 1979, Quasi-experimentation, design and analysis issues for field settings (Houghton Mifflin, Boston).

Crider, D.M., F.E.Willits and R.C.Bealer, 1973, Panel studies: some practical problems. Sociological Methods & Research 2, pp. 3-19.

Davis, J.A., 1978, Studying categorical data over time. Social Science Research 7, pp. 151-179.

Duncan, G.J. and G.Kalton, 1985, Issues of design and analysis of surveys across time, Bulletin of the International Statistical Institute, Proceedings of the 45th session, invited papers, vol.51, book 2, 14.1 (Amsterdam).

Eijk, C.van der, B.Niemöller and A.Th.J.Eggen, 1981, Dutch parliamentary election study 1981 (University of Amsterdam).

Freedman, D.S., A.Thornton and D.Camburn, 1980, Maintaining response rates in longitudinal studies. Sociological Methods & Research 9, pp. 87-98.

Goldstein, H., 1979, The design and analysis of longitudinal studies, their role in the measurement of change, (Academic Press, London Etc.).

Golob, Th.F. and H.Meurs, 1986, Biases in response over time in seven-day travel diary. Transportation 13, pp. 163-181.

Hansen, M.H., W.N.Hurwitz, H.Nisselson and J.Steinberg, 1955, The redesign of the Census Current Population Survey. Journal of the American Statistical Association 50, pp. 701-719.

Hartog, J., 1986, Survey non-response in relation to ability and family background: structure and effects on estimated earnings functions. Research memorandum nr. 8620 (Department of Economics, University of Amsterdam).

Hensher, D.A., 1985, Issues in the pre-analysis of panel data. (Netherlands Ministry of Transport and Public Works)

Hoem, J.M., 1985, Weighting, misclassification and other issues in the analysis of survey samples of life histories. In: J.J. Heckman and B. Singer, eds., Longitudinal analysis of labor market data (Cambridge University Press, London), pp. 249-293.

Kalton, G., 1986, Handling wave nonresponse in longitudinal surveys. Journal of Official Statistics 2, pp. 303-314.

Kish, L., 1987, Statistical design for research (Wiley, New York).

Leeflang, P.S.H., and A.J.Olivier, 1983, Bias in consumer panel and store audit data (University of Groningen, Economic Faculty).

Morgan, J.N., and G.J.Duncan, 1983, Experience with panel study of income dynamics. paper presented at the symposium on micro-analytic simulation models to support social and financial policy (Bonn, June 1983).

Morgan, J.N., G.J.Duncan et. al., 1984, User guide for the panel study of income dynamics (Survey Research Center, University of Michigan).

Nesselroade, J.R., and P.B.Baltes, 1979, Longitudinal research in the study of behavior and development (Academic Press, New York).

Patterson, H.D., 1950, Sampling on successive occasions with partial replacement of units, Journal of the Royal Statistical Society B, 12, pp. 241-255.

Pol, F.J.R.van de, 1982, Efficient estimation of descriptive statistics from panels with partial replacement. Proceedings of the ESOMAR seminar on the effective use of panels, Lucerne (ESOMAR, Amsterdam) pp. 39-58.

Pol, F.J.R.van de, 1986, Herwegen van panels, een inventarisatie van methoden. Internal CBS report (Netherlands Central Bureau of Statistics, Voorburg).

Pol, F.J.R.van de, 1988, Better measurement with a panel design? In W.E.Saris and I.N.Gallhoffer, eds., Sociometrics today, vol. 2 (MacMillan, New York) pp.24-60

Prahl-Andersen, B., C.Kowalski and P.H.J.M.Heyendael, eds., 1979, A mixed-longitudinal interdisciplinary study of growth and development (Academic Press, New York, San Francisco, London).

Ree, S.J.M.de, 1983, A system of coordinated sampling to spread the response burden of enterprises, Contributed papers of the 44th session of the International Statistical Institute (Madrid), pp. 673-676.

Saris, W.E., and M.dePijper, 1986, Computer assisted interviewing using home computers. European Research 14, pp. 144-150.

Schirm, A.L., J.Trussell, J.Menken and W.R.Grady, 1982, Contraceptive failure in the United States: The impact of social, economic and demographic factors. Family planning perspectives 14, pp. 68-75.

Sikkel, D., 1985, Models for memory effects. Journal of the American Statistical Association 80, pp. 835-844.

Sobol, M.G., 1959, Panel mortality and panel bias. Journal of the American Statistical Association 54, pp. 52-68.

Sudman, S. and N.M.Bradburn, 1974, Response effects in surveys (Aldine Publishing C

Sunter, A.B., 1986, Implicit longitudinal sampling from administrative files: A useful technique, Journal of Official Statistics 2, pp. 161-168.

Vierkant, P., 1980, Het functioneren van het telefoonpanel van de Nederlandse Omroepstichting, seizoen 1977-1978, Massa Communicatie 8, pp. 279-289.

Waterton, J., and D.Lievesley, 1987, Attrition in a panel study of attitudes. Journal of Official Statistics 3, pp. 267-282.

Ontvangen: 03-09-1988
Geaccepteerd: 03-07-1989