Rolling Samples and Censuses

LESLEY KISH

ABSTRACT

Rolling censuses combine $F$ nonoverlapping periodic samples of $1/F$ each, so designed that cumulating the $F$ periods yields a complete census of the whole population area with $F/F = 1$. Intermediate cumulations of $k$ samples would yield samples of $k/F$ for more timely uses (annual or quinquennial censuses). Area sampling frames would cover the national territory for naturally mobile populations. These methods may often be preferable to other alternative methods for censuses, also discussed. Asymmetrical cumulations are also recommended to counter the problems of small sample cells for area domains (provinces, regions, states) common to most countries and to other population units. Split-panel-designs offer another use for cumulating periodic surveys by combining nonoverlapping portions $a - b - c - d -$ with panels $p$ for partial overlaps, $pa - pb - pc - pd -$, for multipurpose designs.

KEY WORDS: Periodic samples; Time sampling; Cumulations; Split-panel designs; Asymmetrical cumulations; Multipurpose designs.

1. INTRODUCTION AND DESCRIPTIONS

Several uses and methods for cumulating data from periodic samples are discussed below. This has been a rather neglected subject, as the literature on periodic and rotating samples has concentrated on the statistics for net changes and for current ("cross section") estimates; not on cumulations. The first concern here is on rolling censuses and samples, and let me attempt a definition of rolling censuses: a combined (joint) design of $F$ separate (nonoverlapping) periodic samples, each a probability sample with fraction $f = 1/F$ of the entire population, so designed that the cumulation of the $F$ periods yields a detailed census of the whole population with $F' = F/F = 1$. Intermediate cumulations of $k < F$ periods should yield rolling samples with $F' = k/F$ and with details intermediate between 1 and $F$ periods. We may appreciate that definition by looking at examples and counterexamples. We shall also examine possible variations that would satisfy the definition and conflicting needs that rolling samples can be aimed to meet.

Imagine a weekly national sample, each with $epsem$ selection rates of $1/520$, and so designed that in 520 weeks they are "rolled over" the entire population and the cumulation yields a complete census of the population averaged over ten years. Each year would yield national and local samples with selection rates of $52/520 = 1/10$. The design would combine weekly national samples into an averaged decennial complete census, and into sample censuses of ten percent each year.

The Health Interview Survey of the National Center for Health Statistics (1958) cumulate 52 weekly samples of about 1,000 households each. These samples select about $f = 1/80,000$ weekly; thus 520/80,000 represents cumulations of nonoverlapping periodic samples over ten years. But they are confined to a set of PSU's for reasons of cost chiefly, but also for better estimates of net change and for current estimates. However, rolling samples may better be reserved for samples designed for maximizing (increasing) the spread (representation) of the samples cumulated over national (or broad) populations. The words in the parentheses indicate that rolling samples constitute a special case of the more general cumulated periodic samples and that the boundary of the subset need not be precisely clear.

1 Institute for Social Research, The University of Michigan, Ann Arbor, U.S.A. 48106.
For overlapping between periodic surveys, the requirements for the selection of units of cumulated designs are diametrically opposed to the requirements for the objectives and substantive content of the interviews (the observations, variables). The content of the surveys must be as similar, standardized, identical as possible for the cumulations to be meaningful. Using periodic panels of the same elements for different contents could broaden the scope of surveys, but would not contribute to increasing the sample size for survey statistics. Most periodic surveys collect similar variables, though some may also have other contents attached at times. However, changes of methods, questions, and variables would cause conflicts and problems. Perhaps such changes should be introduced only with extended intervals of "splicing", using both the new and the old methods to study the differences. These problems are fundamentally similar to those faced when measuring differences from periodic surveys, but they seem more novel. I insist (Section 6) that solutions to such problems must be tailored to specific situations.

On the other hand, the cumulation of the same elements (persons, households) does not increase proportionately the sample size (base), and panels of the same elements would not help rolling samples. Many periodic surveys (e.g., labor force surveys of Canada, the USA etc.) have partly or largely overlapping fractions of segments (ultimate clusters), and those tend to contribute little toward increasing the sample size. Even in surveys with nonoverlapping segments (like the HIS of the NCHS (1985)), the segments are confined to the same first stage (and second-stage?) units; in these the positive correlations (clustering effects) tend to reduce the "effective" sample sizes for overall statistics. Furthermore, those periodic samples, confined to samples of primary units fail to meet the needs of rolling samples for spreading over the entire (national?) population.

A few more remarks may help to broaden our frame of reference. (1) The discussion often assumes area sampling, but the concept can be generalized to other frames. (2) Equal selection rates for elements are often used, but cumulations may be modified to unequal selection probabilities. (3) The concept may be generalized from regular periodic samples to cumulations over less regular periods. (4) Cumulations over the entire time span (year or ten years) come most readily to mind, but we may envisage systematic sampling of the span; e.g., labor force surveys cover only single weeks of the months over the year.

2. ALTERNATIVE METHODS FOR CENSUSES

Rolling censuses would be expensive, and the reason for such an innovation should include the acknowledged relative weaknesses of the decennial censuses now widely used, and of sample surveys and administrative registers, which are proposed at times as possible alternatives. The chief reason for censuses is the need for detailed information, especially for small areas; and the chief weakness of decennial censuses is their obsolescence between censuses and their great total cost that prevents more frequent censuses. Sample surveys have many advantages for national statistics and for large regions, but they lack geographical and other details. Good registers are rare and they provide few variables beyond a few, bare demographic data.

Decennial censuses of population, housing, agriculture, industry and others, first and foremost, have spread into most countries in the last two centuries, and especially in the last two generations with the help of the United Nations State Statistical Office. In addition to detailed data for small domains, censuses often may obtain better coverage than samples, due to the concentrated publicity and the national "ceremony" connected with censuses; the Chinese census of 1982 is a good example (Kish 1979, 1989). The efforts of the census also yield lower unit costs (for short forms) than surveys, but much higher total costs than sample surveys, because of much greater size. At 2.6 billions, the 1990 censuses of the USA will cost $10 per
capita or $30 per household. That cost of about half to one hour of the median hourly wage per capita (once in ten years) seems to hold in international comparisons, though the number and complexity of census variables is one of the cost factors. Rolling censuses would probably be proposed and designed for surveys fairly rich in the numbers and complexity of variables. In Canada 260 weekly samples of 32,000 households would cumulate to the national population. In the USA 520 weekly samples of 160,000 would be needed by decennial cumulations to 80,000,000 households; the CPS surveys have 100,000 with state supplements.

No detailed comparison of decennial censuses with rolling censuses is possible here, but the issue of *timeliness* must be mentioned, because that is the chief issue in the comparison. Up to now the periods for using data from decennial censuses have varied from a start of 1-4 years to 14 year or more. Even with faster computers the start is slower for complex social statistics than for mere head counts; and the obsolescence over the ten intercensal years becomes worse with higher population mobility in our modern civilization. The biases due to obsolescence will be monotonic, if not linear, functions of elapsed time. The sizes of the biases will differ with variables, populations, *etc.*; but they will be present and considerable, I believe; often perhaps greater even than the famous biases due to under coverage (Kish 1981, 1979).

Increasing and rapid obsolescence of decennial census data should chiefly motivate the searches for alternatives, such as in *A Study on the Future of the Census of Population: Alternative Approaches* (Redfern 1987). "A serious weakness of the census is that it occurs relatively infrequently". About a "rolling census" it states: "The merit of this proposal is that . . . a much smaller, better trained organization and more experienced staff could be deployed both for the fieldwork and for processing . . . the public awareness of the rolling census would not be highly peaked. Whilst that might well lessen the risk of public protest, the reduced publicity would adversely affect the level of coverage achieved . . . (The method) would complicate the interpretation of the census results, especially comparisons between areas. Simultaneous national coverage, one of the virtues of the census, would be lost. The idea of a rolling census has not yet been developed and applied".

Most countries will probably still need censuses in 2000 AD. They are being replaced by population registers in the Nordic countries and still need to be introduced in some Third World countries in 1990. They have been stopped by opposition and by obstacles in a few. But most countries need and will have them in 1990. They have been a great and useful invention - like the steam locomotive, and at about the same time. However it is possible that the censuses also may be phased out gradually by some of the alternatives here considered.

*Quinquennial annual censuses* have been proposed, and quinquennial censuses have been initiated or carried out in a few countries, including Canada and Turkey. But these are not destined for quick acceptance, I suspect. They seem too costly: ten percent samples in two countries had half of the costs of complete censuses. Also they still leave a great deal of obsolescence. On the other hand, much smaller (*e.g.,* 5 or 1 percent) yearly sample censuses would fail to offer enough geographic detail. The one percent "microcensus" of West Germany provides yearly sample data. China had a one percent census in 1987; their yearly samples of 1/2,000 (also about 500,000 people) collect chiefly fertility data only (State Statistical Bureau 1987; Kish 1989). Quinquennial censuses are not frequent enough and yearly censuses would be too costly.

*Administrative registers* provide a great deal of diverse data in many countries, and they are likely to spread in the future. Excellent *population registers* exist in the Nordic countries of Sweden, Norway, Denmark, and Finland, and perhaps in some other countries of Northern Europe. Their completeness is based on cooperation, motivation (with social incentives), and literacy; in a few cases they are replacing censuses with data from the population registers. In other situations their coverage, quality, and updating are far from adequate. We can expect
future improvements in the quality, spread, and use of population registers but not quickly and not widely. We should not expect them to replace censuses even in developed countries like the USA and Canada, and their use in less developed countries soon is even less likely (Redfern 1989).

Furthermore, even after population registers become adequate in quality and coverage, they will contain and supply only a few, bare demographic variables: head counts, age, sex and little more. Thus, they will fail to meet the demands of modern society for richer sources of statistics. For these the registers will serve only as auxiliary variables.

Synthetic, ratio regression, and raking estimators are being used increasingly for small area statistics (Platek et al. 1987; Purcell and Kish 1980). Census data are usually obsolete, data from registers inadequate, and sample data lack details for small areas. The weaknesses and strengths of the three methods are complementary, hence combining the advantages of the three methods seems like good strategy. This is the common purpose of the several methods of small area estimation: to provide estimates for small areas and for other small domains that are current, accurate, and relevant.

These methods are now being used for local area estimates of population counts for the intercensal years, in order to compensate for the obsolescence of the decennial censuses, thus sometimes called postcensal estimates. They also have other uses in increasing numbers, e.g., they have been proposed to compensate for undercount biases. However, those methods have all combined censuses with sample surveys and registers. Therefore, they should not yet be considered as alternatives to censuses. Nevertheless, we may raise the question whether rolling censuses would perform better or worse overall than decennial censuses in those combinations. The answer is uncertain, but I believe that the balance of variance components would favor rolling censuses in most cases. However, theoretical as well as empirical investigations will be needed to decide this question as well as several others here.

Partially overlapping samples from multipurpose designs must be considered because they exist in many countries for several purposes and they absorb some of the funds available for national statistics. These multipurpose surveys often provide labor force statistics and other valuable data. They vary in parameters between countries but they also have several basic features in common with those of the USA and Canada. They are periodic samples with overlaps that are constant and for fixed periods (but all three parameters differ between countries). They use area segments for bases, but not panels of households (movers are not followed). The overlaps are usually large and these are generally justified with references to reductions of variances from positive correlations in the overlaps. But an even greater advantage of overlaps may be the lower costs of interviewing in later calls, especially where telephone calls follow first calls on foot. These “rotation designs” have dominated practice and literature and they represent an important innovation (by H.D. Patterson 1950 and R.J. Jessen 1942). They are designed for measuring net changes and current (level) statistics, but not for cumulations. However, the variances (per household) would not be greatly increased for overlaps of even a small fraction (< 0.3), when compared to the large overlap (> 0.7) commonly used. This is particularly true for many variables like being unemployed, which have low correlations between periods. Furthermore the overlaps could be changed in other ways (Section 5). Therefore it is possible that these surveys could be combined with the cumulations needed for rolling samples and censuses.

3. CUMULATIONS OVER TIME AND SPACE

Changes in populations and in their variables are often recognized as of three kinds: “secular” trends, which are more or less smooth and monotonic, like “growth”; periodic and “cyclical”, such as seasonal fluctuations; and irregular variations which are difficult to describe
and often treated as “random”. Designs for cumulating, averaging, and sampling over temporal variations face psychological obstacles that differ from our acceptance of designs for variations over spatial variations. Spatial variations can be large and sometimes accountable, but more often irregular. However, we have learned to accept samples, averages, and cumulations over them in population (national) aggregates and averages.

The psychological blocks still facing rolling samples and censuses may be countered with both theoretical and pragmatic arguments. The theoretical and philosophical arguments are hinted at above and in later discussions of alternatives (Kish 1987, 6.1B). The pragmatic and empirical arguments may be buttressed with several types of uses we recognize as common and successful. The same periodic samples for obtaining current data and for measuring changes can also be used for aggregates needed for spatial and domain details. Furthermore, by averaging (over a year or longer) the temporal variations (seasonal or cyclical or erratic) are smoothed over in the moving averages.

Retrospective data. “Children ever born” to women who completed fertility over the entire fertile span of 30 years may represent an extreme for retrospective spans; but other individual interview data aggregated over life spans include serious diseases, education, etc. Interviews aggregated over yearly spans include farm production, work history, income, home and auto purchases. Of course, all these data have imperfections, which differ across variables, respondents, methods, etc. But even cumulations over a week or over a day (such as purchases of bread or cigarettes) have errors. Multiround surveys are used for cumulating short term data; for example, births during the past month have been cumulated from 12 monthly samples over the year.

Cumulating rare elements from periodic surveys has often been used to deal with these difficult and expensive problems. The topic has been dealt with and illustrated in publications on rare items (Kish 1965 11.4; Kalton and Anderson 1986). Statistics for small domains may also benefit from cumulations, and single years of birth may exemplify such small domains, which consist of “crossclasses”. But geographical and administrative units are “proper domains”; for these the periodic samples are not adequate, because those domains need the designs of rolling samples or censuses.

Cumulations from periodic samples. The Health Interview Survey (NCHS 1958), described above, may be the best known example with yearly cumulations of weekly samples of about 1,000 households from nonoverlapping area segments. It is designed for multipurpose objectives (like most periodic surveys) including cumulations for some rare diseases, but also estimates of current levels and net changes. It provides some estimates for larger domains, as well as national estimates for the common diseases. To convert it into a rolling sample, by increasing the spread of the yearly samples, would increase field costs, especially in that portion (about 30 percent only) where the PSU’s are counties (not self-representing).

A traffic survey provides an interesting example of cumulations, because the population is very mobile within the sampling frame of sampling units of locations x hours (Kish, Lovejoy and Rackow 1961). The general concept is applicable to nomads and other mobile populations. It may also serve less mobile general populations over a longer period, such as the decennial spread.

The earliest cumulation I found is for a sample of California in 1952 (Mooney 1956). “The samples were selected in such a manner that they resulted in a uniform overall sampling rate of 1 in 385. For purposes of enumeration, the sample was divided into 52 equal subsamples, and a different subsample was enumerated during each week of the survey year. Consequently, each week’s enumeration was based on a sample of 1 in 20,020”. For smaller states (populations) and/or larger samples one may imagine weekly samples of 1/520, and complete rolling samples in the 520 weeks of the decennial census period. It is likely that such rolling samples have been designed for smaller populations.
The above examples refer to nonoverlapping periodic samples. Cumulations from partially overlapping samples have been used, but with the "effective sample sizes" reduced by the amount of the overlap (Ericksen 1974). Furthermore, this paper concerns cumulations of individual cases, but periodic or repeated surveys may also be used for combining statistics from them (Kish 1987, 6.6) as in "meta-analysis".

4. ASYMMETRICAL CUMULATIONS

This term denotes a proposed method of cumulation for problems that arise because "natural" subpopulations generally vary greatly in size. For example, I have been faced within the past few years with ranges of 50 or even 100 to 1 among the provinces (or states) of Canada, USA, Australia and China; and those ranges of relative sizes are similar for the provinces of most countries. Those inequalities arise because administrative units tend to be created roughly equal in areas, but spread over lands with highly unequal population densities. They also exist for districts, counties, etc. within most provinces. They also arise for other social units and social organizations, like firms, hospitals, universities. But not for all: military units, census enumeration districts and elementary schools are created roughly equal.

For many other frequency distributions rough equalities of classes are created with traditionally accepted cumulations over roughly logarithmic scales; e.g., income, city size, etc. are often tabulated in classes like 10-25, 25-50, 50-100, 100-250, 250-500, 500-1,000, 1,000-2,500, etc. This shows a sensible method of cumulation that creates roughly equal cells on a roughly logarithmic scale, and they are traditionally accepted and understood, although highly asymmetrical.

Note also that cells in tables for sample data are generally cumulated over both space and time. For example, monthly surveys of labor force often show labor force statistics cumulated over the month (or over a week as a "sample" of the month), and also over the provinces (from a sample of sampling units). Quarterly and yearly statistics show further cumulations, as do the national statistics. The spans of cumulations must balance three parameters of restraints: the span of the reference period that may be relatively flexible; the domains of subpopulations, which may be more rigid, like provinces; and the sample size expressed in sampling units and variance components. Other variables, such as cost factors and "required precisions", tend to be expressed through the basic three parameters of cell size.

Decennial censuses of the population counts represent extremes by emphasizing locational detail: persons are placed in homes as of the reference date (April 1 in the USA). But yearly and longer cumulations are possible for income, etc. Time gets sacrificed in obsolescence, and sample sizes and costs in complete coverage. At the other extreme are monthly sample surveys for labor force and health variables, and myriad other variables, where the emphasis is placed on timelines and reduced costs, but at great sacrifice of spatial detail.

Population inequalities between provinces impose severe restraints on timeliness and sample sizes. Often higher sampling rate are introduced for the smaller provinces, but such "optimal" selection rates bring disadvantages in increased variances both overall and for cross-provincial "crossclasses" (age, sex, etc.) (Kish 1988, Section 5; Trewin 1987). Thus those mildly unequal rates fail to solve conflicts in provincial sizes of 50:1 or 100:1.

Because of those conflicts the tables for monthly surveys commonly present cells for small provinces with inadequately small sample sizes. Two alternative procedures have been advanced and practiced for such small cells. A. Release the same data for small cells as for large cells, and let the reader (user, consumer) beware, caveat emptor, with perhaps warnings posted
to appendixes to sampling errors. B. Don’t release, but suppress small cells, leaving them blank, after applying some declared curtailing limits. Readers may be directed to other released publications, based on cumulated data (quarterly, annual).

Asymmetrical cumulation proposes a compromise between symmetrical releases (A) and asymmetrical suppression (B).

C. Asymmetrical cumulation proposes to release for small cells the specified cumulations of periodic data. These cumulations may be flexible: for example, quarterly for small cells and yearly for very small cells, instead of the monthly data for large cells. The readers may be notified (with * or italics or other signs); thus they may choose either C (cumulation) or B (disregard).

AC. This procedure would allow readers to choose either A or B or C by publishing both the current monthly data A and the cumulated C data.

Procedures B and C have the disadvantage that the cells do not sum to the marginals. But AC like A do sum to the marginals. Some iterative method could overcome these disadvantages of B and C.

5. MULTIPURPOSE SPLIT PANEL DESIGNS (SPD)

In order to find adequate funds for rolling samples and censuses it is desirable to consider how they could be combined with the periodic surveys now being funded and conducted in many countries. These are either monthly or quarterly surveys (sometimes yearly or weekly). They are typically partially overlapping samples designed for improved estimates for current level and net changes. However they are not designed either for cumulated rolling samples, or for panel studies based in the overlaps. I proposed SPD as the design for providing data for all those four purposes; and also for some fringe benefits (Kish 1987, 6.5).

a. Combining two separate periodic samples forms the basis of SPD: to add a panel p to a parallel series of nonoverlapping samples a — b — c — d etc., with the combination then denoted as pa — pb — pc — pd etc. The panel p provides individual (micro) changes and the nonoverlaps can be cumulated into larger samples and rolling samples. The combined samples provide the partial overlaps best for current estimates and for net changes; thus they can replace the usual rotating samples. This combined use is a main feature of SPD, together with the provision of a flexible and potentially large sample of nonoverlapping portion for use in cumulating samples.

b. The designs for p and for a — b — c can be separate and distinct, each “optimized” for its own objective. But they must also be combined for joint estimates of net changes and current levels; and for that purpose the populations covered and the measurements used must be similar enough for the combination.

c. SPD has considerable advantages because its overlaps exist for all periods, whereas they are rigidly fixed in classical rotation designs. This advantage is clear and important for net changes because it exists for all desired comparisons. But it also exists for current levels, because the correlations may differ among variables.

d. Including proper panels p of elements necessary for measuring individual (micro or gross) changes would be a great advantage for SPD over partial overlaps now used. However, the other features can be satisfied with overlaps p' of area segments as at present. Furthermore a modest and slow rotation can be built into the design of either the panel p or the overlap p', so as to retain most of the gains from covariances and from panel information. Perhaps some alternation may be introduced to reduce panel fatigue or deterioration. Several surveys have used both the overlap p' and panel p by following as many movers as possible. Most
households belong to both samples. The extra cost for the panel depends on the proportion of movers and their cost (Kish 1987, 6.2, 6.4).

e. The advantages and problems of panel interviewing pose difficult problems, with a large and varied literature and conflicting results (Kish 1987, Sections 6.4, 6.5). The number and spacing of reinterviews that are possible, desirable, and reliable need to be established.

SPD has an advantage in separating the panel $p$ whose cumulated data may be checked against the nonoverlaps for “panel biases”, and perhaps even for adjustments of biases when those are measured adequately.

Another useful modification may be to recruit sampling units into the panel by different (“optimal”) selection rates on the basis of their being “screened” in the nonoverlaps.

f. The size of $a - b - c - d$ need not always be the same; this flexibility of SPD, which differs from the rigidity of rotating designs, may be used for needed sample enlargements or for cost retrenchments. Such changes would raise weighting problems (solvable) for cumulations.

g. The relative size of the panel $p$ against the nonoverlap $a - b - c - d$ portions depends on feasibilities and costs and needs study (Section 6). For individual changes we need larger $p$, but for cumulations larger $a - b - c - d$. The larger $p$ portions now common may be favored by lower field costs for telephone reinterviews.

Lower values of $p$ than are now common are good enough for current levels and for net changes with weighted estimates; the optima are insensitive and $p$ between 1/4 and 1/2 are all nearly best; lower $p$ may also be used where the emphasis lies in nonoverlaps $a - b - c - d$ for cumulations.

6. CONCLUSIONS AND QUESTIONS

Cumulated samples provide the bases for four new methods proposed here: rolling samples, rolling censuses, asymmetrical cumulations, and split panel designs. Rolling samples have been designed, but the other three still await practical applications. Meanwhile we should welcome methodological developments that would outline the parameters of feasibility.

However, the chief tasks for these methods must be found in the details of specific situations rather than in theoretical generalities. The factors of costs, variances, biases, feasibilities, and public acceptance for novel procedures must be worked out specifically for each situation. We can do no more than raise a few questions as examples, in addition to those raised implicitly or explicitly in the preceding sections.

1. For rolling samples and censuses what kinds of moving averages may prove most useful? For national aggregates the last six months (or quarter or year) may receive the full weight. But for small local areas the data may be cumulated over ten years; with equal or with increasing weights? Are “shrinking” (Stein-James) estimators useful?

2. How to deal in the aggregates with changes in the population, in methods, in variables?

3. For asymmetrical cumulation similar questions arise. Should the latest monthly estimates (A) be printed together with the cumulated (C)? Methods are needed to make the cells and the marginals consistent.

4. For the split panel design, how large should the overlap ($p$) be? Can it be a panel or merely overlapping segments? Or must we, can we, have both? How does it depend on the correlations for diverse variables? How do we balance the four chief purposes of periodic surveys?

There will be other interesting questions but this essay must come to an end before they do.
REFERENCES


