Journal of Official Statistics, Vol. 14, No. 1, 1998, pp. 31–46

Space/Time Variations and Rolling Samples

Leslie Kish[1]

Variations over spatial domains, like provinces, and over other domains, have gradually received some attention recently in sample designs. Similar problems are now considered in multinational designs also. However the variations in the temporal dimensions have been neglected and judgmental choice of reference periods for surveys and censuses is still the prevalent tradition. But probability selections should also be considered for time variations. Censuses yield spatial detail, but not temporal detail; and small area estimation is only a partial, imperfect answer for postcensal estimates. Cumulation of data from periodic surveys, existing or needed, is proposed as the method for dealing with these needs for spatial/temporal details. Furthermore ''rolling samples'' should be designed to yield these data. These general concepts need technical and specialized expertise to make them feasible and acceptable.

Key words: Periodic surveys; panels; rolling samples; small domain estimates; annual censuses.

1. Introduction

We must confront the triple aspect of our task here. First, we must lay the ground and construct the framework for the basic relations between two major sources of statistical variations in survey sampling: variations over space and over time. We must also add a third dimension to space and time: that involving diverse variables and populations, because surveys are typically multivariable and multipopulation in practice. Second, the triple sources of variations lead to a series of related problems that surveys must increasingly come to recognize and deal with; and to attempt to ''solve.'' The problems arise from expressed needs of institutions for public policy for data that are reasonably accurate and up to date (microtemporal) for small areas (microspatial), or other domains, and also rich in content (variables). Third, we shall suggest a general method of ''solution'' for those problems, for which I suggest a general name of ''rolling samples.''

To overcome variations in space and to yield statistics for small administrative areas serves as the chief justification for the complete decennial censuses. Because they are so large, censuses also yield data for many other kinds of small domains, e.g., by age, occupation, education, ethnic groups, etc. Because of their total national size, the total cost is very large, though not the per capita cost. Hence they are decennial, not annual; thus they lack temporal detail. Also because of their vast size, they are confined to counting persons and dwellings, and a few (say 7 to 17) ''simple'' variables. Therefore, in many countries, a ''census sample'' of from 5 to 20 per cent is selected at random from the complete census

[1] Institute for Social Research, The University of Michigan, Ann Arbor, MI 48106, U.S.A.

to obtain much richer data, of scores of variables, with much spatial detail. But only once in ten years.

On the other hand, for many variables – in employment, economics, health – much more temporal details are needed in order to measure fluctuations over time. Thus monthly and quarterly sample surveys were created in several countries, and these yield data with temporal detail for national statistics and for major regions and other major domains. But not for small areas or domains, because these samples of from 2,000 to 100,000 households cannot yield data in fine detail.

Thus neither decennial censuses nor sample surveys can provide both spatial and temporal details. However, public policy increasingly demands such twofold details, and now methods are being created to meet those demands. Administrative registers, created for and paid by other needs, can provide both timely and spatially detailed data – where they are accurate, reliable, up to date and publicly available. These registers generally exist in the Nordic countries and a few other places. Even there they provide only simple counts and a few other variables, but not rich data of many variables. And many registers are not good. I believe that good registers will spread, but only slowly; however, they can never be sources of the rich data of many variables, which are needed (Redfern 1989, 1997).

''Small area estimation'' method (or methods) are increasingly used to provide statistics that have both temporal and spatial details. It is a ''growth industry'' in statistics now, but it did not interest theoretical statisticians at first. Demographers developed ''postcensal'' and ''intercensal'' estimates, chiefly for population numbers, using demographic methods. Regression methods based on complete counts were first proposed (Schmitt and Crosetti 1949). Then I was associated with the first three Ph.D. dissertations that combined current sample data with complete census results (Ericksen 1973, in *Sociology*; Kalsbeek 1973; and Purcell and Kish 1979 in *Biostatistics*).

These methods have been greatly developed since then, and they have also been called ''small domain'' estimation, because the needs and methods can be applied not only to small administrative areas, but also to other small domains (Purcell and Kish 1979; Platek, Rao, Särndal, and Singh 1987). Furthermore I suggest here that it may also be viewed as ''small period'' estimation, when data are available for small areas from decennial censuses for example, and the method is used for estimates in post- (or inter-) censal small periods. The small area $\times$ small period cell may be viewed symmetrically in the space $\times$ time matrix.

Small area methods yield useful estimates, but their errors can be large, especially when the estimates are not closely related to (obsolete) census data and also when the sampling units of the sample data are few. These errors of estimation will probably tend to be even greater for richer, complex variables, because the relations in the estimation become more attenuated. Therefore, there exists a growing need for more frequent dates, especially for annual surveys with spatial detail.

I advocate designs of cumulated ''rolling samples'' that can provide statistics with adequate details both in temporal and spatial dimensions. The chief objective may most often concern good annual estimates with spatial detail based on cumulations of weekly and monthly samples. The monthly and quarterly samples can yield good national and major regional and domain estimates, as monthly and quarterly surveys in some countries

now yield, but not spatial details. We shall also discuss the possibilities of decennial cumulations, and this becomes a profound and contentious topic later (Kish 1981, 1990, 1997).

2. Statistical Variations Over Time and Space

We need two departures from our customary way of thinking, and these may cause difficulties for some, as discussions over twenty years have taught me. First, these two kinds of variations are not entirely similar in their essential structures, nor in practical survey design. However, I must emphasize certain basic similarities and base my proposals on those similarities.

Second, I propose different views of both time and space from those we normally use for our physical world. We tend to think of time as flowing forward evenly and unidimensionally. And we think of variations as occurring chiefly monotonically in a secular straight line; or along a logarithmic growth curve; or perhaps in a cyclical variation, governed either by the Earth's daily turns around its own tilted axis; or by its yearly path on that tilted axis around our sun. These diurnal and seasonal cyclical variations are seen in many cumulated averages. But in statistical and survey data taken over time intervals, we actually observe mostly irregular, random, or haphazard variations. This is true of individual blood pressure and blood counts, stock market averages, unemployment rates, air pressure and temperature, etc. The cyclical and secular trends are typically removed by either model-based adjustments, or by taking small time segments (like ''strata''), or by both. Thus the variations actually observed and used over time intervals are irregular, like the variations also measured in sampling over space.

Space also has a different meaning here for surveys than the three (or more) dimensional space of physics. It refers chiefly to partitions of the Earth's surface into administrative domains like provinces and districts, and into areal sampling units like strata, blocks, and segments. Furthermore, the same concepts can be applied to domains and partitions created by statistical analysis and treated similarly to spatial domains, such as age, sex, occupation, social economic classes, etc. So that, in contrast with the temporal dimension, ''space'' and ''spatial'' can stand as shorthand for other domains covered in cross-section surveys. For example, the teen-aged males and females, white and black, are among the most important domains for surveys of unemployment in the U.S.A. The cumulations for these other domains may be even more effective than for spatial domains (Kish 1994).

''Time'' and ''space'' are not entirely correct terms, but perhaps the simplest here. Furthermore, I have long made a distinction between ''proper'' and ''design'' domains and subclasses, used in sample designs (like provinces and districts in area sampling); and ''crossclasses'' (like age, sex, occupation, behavior, etc.) that cut across sample designs (Kish 1987, 2.3). Thus, despite the physical and philosophical differences between the temporal and ''spatial'' dimensions (and other domains) we can use the statistical similarities, in order to borrow from familiar principles of design for spatial aspects, also for the temporal aspects. However, we should examine those similarities from four distinct points of view. First, with regard to smooth continuity versus sudden discontinuity: each can exist in both the temporal and spatial aspects. Against the smooth temporal growth curves of peaceful nations, we can counterpose epidemics (influenza,

AIDS), stock market crashes, and sudden weather changes. Against the smooth spatial changes of the Midwest, we post drastic changes along the Andes and the Rocky Mountains, or the drastic social changes found when crossing the Rio Grande between Mexico and the U.S.A.

Second, most people seem to perceive a conceptual difference between temporal and spatial variations and especially cumulations. For example, adding regional, provincial statistics into national aggregates and averages appears ''natural,'' but rolling monthly samples into annual or decennial averages seems to run against perceptual walls. We may need a ''paradigm shift'' to hoist ourselves over that wall (see Scheuren's remarks in Kish 1990). I believe that this conceptual block is truly less philosophical than psychological and social, conditioned by our long acquaintance with the images of censuses and of monthly survey data.

Third, understanding the similarities may depend strongly on the time interval involved. For example, annual income is a readily accepted aggregation, and not only for steady incomes but also for occupations with high variations (seasonal or irregular). Averaging weekly samples for annual statistics will prove more easily acceptable than decennial averaging. Nevertheless, many investors in mutual stock funds prefer to rely more on their ten-year or five-year average earnings (despite their obsolescence) than on their up-to-date prior year's earnings (with their risky ''random'' variations). Most people planning a picnic would also prefer a 50 year average ''normal'' temperature to last year's exact temperature. There are many similar examples of sophisticated averaging over long periods by the ''naive'' public. That public, and policy makers, would also learn fast about rolling samples, given a chance.

Fourth, rolling samples will encounter formidable problems of feasibility. These will differ so much between countries, resources, and the nature of statistics, that I cannot discuss this topic both generally and usefully. One difficult example is the ''continuing measurements'' now being planned to begin for the U.S.A. in 2000 (Alexander 1993). On the other hand, designing rolling samples for annual statistics for most countries without monthly surveys, may be simple compared to its alternatives.

3. Major Surveys of the National Populations

Figure 1 lists the major types of population surveys now conducted in many industrialized countries. Some of these are also conducted in the ''less developed countries'' (LDCs), and decennial censuses cover almost all countries today. I refer here to population surveys of persons, families, households, and dwellings, and also to the characteristics and

1. Decennial census of BASIC data for SMALL AREAS (blocks)
2. Decennial socio-economic data: Census sample, ''long form'' for many variables
3. Annual basic data; mini-census; annual census?
4. Annual socio-economic surveys of diverse objectives
5. Monthly or quarterly labor force and population surveys
6. Monthly or quarterly surveys of, say, demographic and health data
7. Vehicle for diverse socio-economic surveys

Fig. 1. Major needs for survey data

variables based on them. The surveys listed here are mostly repeated and periodic, and mostly done by national statistical offices. They also conduct many surveys of other populations – of agriculture, industry, commerce – but we cannot deal with them here. Furthermore our discussion also concerns periodic surveys by nongovernmental centers. They are mostly national surveys, though others may be regional or even local in scope, or sometimes cover only specified subpopulations.

Decennial censuses have been the first and foremost system of surveys in the world. They have been collected in the U.S.A. decennially since 1790, but French Canada (1665), Sweden (1749), and the Italian States (1770) are thought to have preceded them; and historical instances may be found as far back as 2 A.D. in China; and there are references in the Bible etc. Their principal aim is to count the population elements in geographical-administrative detail. The principal population elements are persons, families, households, and dwellings; at other times and in other countries they have been potential soldiers, taxpayers, or workers. The basis of geographic detail is association of persons (who are mobile) with ''usual places of residence,'' which are relatively stable; that is, with dwellings, villages, counties, blocks, towns, and cities. These provide *basic* data in geographic/administrative detail on numbers of persons by sex/age classes; also data on dwellings, families, and households (Hansen and Hurwitz 1946).

In addition, increasing requests arose over decades for socio-economic data of many varieties, to be also collected by the census enumerators along with the basic data. It seems cost-effective to use the census machinery to collect additional data also in great geographic detail. But collecting these added data increased the cost of censuses; then solutions were found in the last five decades by resorting to sampling the census households. These ''Census samples'' are usually 5, 10, or 20 per cent of the complete census (and may even be 1 per cent or less), but they are still much larger, hence provide more detail, than almost any national samples; they are now collected in many other countries. Typically, the complete census uses the ''short form'' with few questions, whereas the ''long form'' with many questions is reserved for the census sample. A sample of households may be assigned to all the enumerators, or a sample of special areas may be assigned to special enumerators trained in the complex long form.

However, decennial censuses fail to meet the need for current statistics. Long ago the ''Bureau of the Census has devised a plan for the conduct of a sample census of the population taken on an annual basis of small areas scattered through the country'' (Hauser 1942), but this plan has never been adopted in the U.S.A., and few countries have an annual sample census. Canada took a ''quinquennial'' census of 10 per cent in 1986 and again in 1996. The cost of sample censuses poses formidable obstacles: a 10 per cent census may cost half as much as a complete census, I understand. These sample censuses may consist of only brief and few basic data, or they may have the greater breadth of socio-economic survey data. We shall discuss annual samples in the next section.

Entirely different in scope are the monthly Current Population Surveys (U.S. Bureau of the Census, USCB, 1978) of the U.S.A. and its many monthly and quarterly equivalents in other countries. Their scope and aims have been entirely separated from those of the censuses, but I shall link them with rolling samples. With sampling fractions near 1/1,000, they have fewer than 100,000 households monthly, and based on clustered, multistage samples, ''partially overlapping segments.'' They are designed and used to yield

monthly and quarterly estimates of various national statistics, also for regions and for large subclasses. However, they are not designed nor used for small area statistics, nor for other small domains. Originally designed in 1943 for employment/unemployment and other labor force surveys, they have also been used for general population and socio-economic statistics. (Hansen, Hurwitz, and Madow 1953, 12B; Kish 1965, 10.4.)

These numbers and these descriptions are only approximate for the U.S.A. today. They have been evolving over time in the U.S.A., and they differ greatly elsewhere. The aims, scopes, methods, and resources vary between countries, and over the decades, and are much smaller or non-existent in most countries. Generally the samples are based on domiciles (dwellings), and the families, households, and persons who live there. The methods involve interviews in the homes, and increasingly by telephones in the U.S.A. and a dozen countries. But complete reliance cannot yet be placed on telephones, and therefore area segments are often used for frames, or as supplements.

The sampling frames and resources needed for these periodic statistics have also been used as resources and vehicles for other statistical needs, for example, for annual surveys of statistics of education, income, and crime victimization. Also *ad hoc* one time cross section surveys have been collected on many topics. Yet a great gap exists between the complete focus of decennial censuses on geographic/administrative and other domains, with great sacrifice of timeliness, and on the contrary, the complete focus of monthly samples on timeliness, with great sacrifice of spatial and domain details. Between these extremes, most statistical needs, which are now missing, could be filled with large annual samples. Cumulations and rolling samples are proposed in Section 4 to fill this gap.

4. Rolling Samples for Annual Statistics

Annual statistics seem neglected now by surveys which concentrate chiefly on decennial censuses at one extreme and on monthly labor force surveys on the other. This seems to be a historical curiosity, due to the success that those two great inventions have enjoyed in our times; and we placed our trust in them – more or less.

Annual statistics play leading roles in many endeavors; in economic data, in accounting practice, in weather reports, in demographic reports, etc. There are annual social and demographic statistics released in some countries, based on the last decennial censuses with ''postcensal'' adjustments based on vital and other registers. There are annual fertility and population samples of 1/2,000 in China (Li 1985); and Germany had annual 1 per cent counts of the population; but I have made no study of these efforts. An annual sample of 1 per cent was advocated for the U.S.A. long ago by Hauser (1942). However, I believe that these yearly snapshots, taken on some fixed date, would be more costly, less useful, and feasible than rolling samples. The yearly data from 52 weekly samples of 1,000 dwellings in the National Health Interview Surveys come somewhat closer to rolling samples, but are not quite that (National Center for Health Statistics, NCHS, 1958).

To avoid confusion with other methods, I define rolling samples as: *a combined (joint) design* of k *separate* (non-overlapping) periodic samples, *each a probability* sample with the selection fraction $f = 1/F$ of the *entire population, so designed* that the cumulation of k

periods yields a *detailed sample* of the whole population with $f = k/F$. Several feasible modifications can be accommodated within the definition (Kish 1990).

a) When $k = F$, the cumulated sample yields a complete census with $f = F/F = 1$; perhaps decennially. We defer problems of weighting and inference from cumulations, especially over longer periods, such as decennials.
b) The fixed, constant sampling fraction can be changed from $1/F$ to P_h; perhaps to accommodate with larger P_h the smaller domains, or because of frame problems, etc.
c) Changing the periods and the sampling fractions $1/F$ between periodic waves are both possible, but the population weights for the periods must be considered.
d) It is implicitly assumed that the reference periods of the waves are ''mutually exhausting,'' so that weekly (or monthly) samples refer to the entire weeks (or months). But the reference periods can also be only systematic samples of the periods; for example, one week in the month, as in the CPS sample (Kish 1987, 6.1).
e) For simple and efficient combining we assumed separate samples that are ''mutually exclusive'' (not overlapping), but overlapping designs can be accommodated with special care and methods. It may be most convenient to design an annual sample, then partition that into 52 weekly national samples.

I propose rolling samples to be collected weekly (or perhaps monthly) to serve simultaneously several major objectives.

I. They can replace the present monthly and quarterly surveys of labor force and/or current population surveys. Countries that have not yet adapted these may now have added incentives for starting them. The multiple objectives of rolling samples can be built into the designs from the start. Countries that have good, large surveys can use those budgets, but may face problems of conversion, because of two main obstacles. Some have large month-to-month overlaps, which may yield variance reductions for some change statistics, such as changes in unemployment. Furthermore in countries with many telephones, later interview may be cheaper than the first doorstep interviews.
II. Annual statistics based on 52 weekly rolling samples may be the chief product.
III. Decennial (and quinquennial) samples will be based on combinations of annual samples.
IV. Panel studies may also be attached, as discussed in Section 6.
V. The entire operation can also serve as a basis for other periodic or one-time surveys.

Thus, the budget of the rolling samples should be compared to the combined cost of all these operations, rather than to the cost of only one of these, such as the monthly surveys alone, or the complete census alone.

I read about three important, excellent samples of a kind, which I propose to call ''cumulated representative samples'' or ''cumulated representation.'' Each has the following characteristics:

a) Representative of a national (or large) population.
b) Probability selection methods.
c) Mostly nonoverlapping periodic samples.

d) Cumulated periodic samples: weekly, monthly, or quarterly.
e) Annual (or longer) cumulations.

The Health Household Interview Surveys collect weekly samples of about 1,000 U.S. households, and their 2,500 occupants, cumulated to 52,000 households and 130,000 persons yearly for multipurpose health interviews (NCHS, 1958). The Australian Population Survey Monitor is designed to cumulate nonoverlapping distinct quarterly samples into annual data for better domain statistics, but with 0.5 overlaps from year to year (Australian Bureau of Statistics, ABS 1993). In the U.K. also quarterly surveys are cumulated annually but these still have overlaps between quarters (Steel 1997).

There may be others like these, each confined to a set of primary sampling units. They can give good estimates for national estimates and major regions. The yearly cumulations can give improved estimates for many domains. But because they are not designed for ''rolling'' over all PSUs, they are not designed for spatial detail. I propose that we should clarify this distinction and assign ''cumulated representative sample'' to this broader class which deserves a special name. This would also help to distinguish what rolling samples are. Rolling samples are a special type that are ''rolled'' over all the primary units in the cumulated sample, and a rolling census refers to complete coverage of the population. (I have used them for non-national populations: Mooney 1956; Kish, Lovejoy, and Ruckow 1961.)

5. Rolling Samples for Decennial Censuses

Now we face a dilemma. On the one hand, decennial censuses are the most costly and the most important surveys worldwide. The need for them is widely felt by the public and by its funding bodies. They often have tradition and even legal force behind them. However, they often encounter vast problems and deficiencies. And there exist vast differences in the way they have been collected in different countries and ages. A long time ago I wrote that ''I think that some of the activity of the decennial census, now focused on a single day every 10 years, could better be split into separate periodic samples, covering the country in 120 monthly surveys, or 40 quarterly surveys'' (Kish 1965, p. 475). Now I recognize better than in several papers (Kish 1979b, 1981, 1990; Kish and Verma 1983) the difficulties and delays caused by tradition, habit, and law. Decennial censuses were great inventions, but perhaps if they had not existed for 200 or so years, they would not be invented now. (Some scientists said that about the internal combustion autos driven by petrol explosions.) This may be true but not immediately convincing now. It seems fairly clear that the adoption of rolling censuses will need different specific techniques in different countries. Several countries are considering rolling samples and censuses, but the most specific and most advanced are the ''continuous measurements'' (or American Community Surveys) being planned by the U.S. Census Bureau (Alexander 1993; Herriot et al. 1988; Bounpane 1986).

We should separate the two functions of censuses which are represented by the ''short forms'' and the ''long forms.'' The short forms are designed to obtain complete counts of dwellings, households, families, and persons by age and sex and little more. This function is performed in very different ways and with different success in different countries. In the Nordic and a few other countries, administrative registers do this so well that some wonder

whether a complete count or a rolling census can do it better or cheaper. Therefore, the possible functions and advantages of a rolling census must be assessed country by country.

The long form of many socio-economic variables presents a different situation generally, also in relation to the rolling census. It is often based on a ''census sample,'' selected from the complete count, perhaps 5, or 10, or 20 per cent. It is still much larger than sample surveys, hence it provides geographic, spatial, and domain details. But attached to the decennial census it has the same disadvantage of obsolescence in comparison with annual samples.

These days in a growing number of countries there are increasing demands for timely geographic/administratively detailed data of socio-economic content. These are wanted not only for research purposes but also for public policy actions. Also with increasing mobility, changes become more rapid and frequent, thus decennial census data become obsolete. But the census sample often has a great cost advantage in addition to traditional acceptance: its true cost may be hidden when attached to a compulsory complete simple count.

With rolling samples, national and large domain estimates may be based on the latest annual cumulations. On the other hand, for small domains the rolling samples can provide 10 year cumulations. I assume these need not suffer the discontinuities of the present practice of decennial censuses; that 10 year cumulations can be presented annually. These topics are discussed later under ''asymmetrical cumulations,'' also here under weighting.

I cannot evade the tasks of saying something about weighting the cumulated annual samples and to make it as simple as possible. Assume that the periodic annual samples are about the same size and design. Thus we can concentrate on obsolescence versus cumulated sample size as the main factors. Assume there are 10 yearly means $\bar{y}_i$, whose mean is $\bar{y} = \Sigma W_i \bar{y}_i$, where W_i are the weights assigned to the 10 years, and $\Sigma W_i = 1$. Consider now four models:

$\bar{y}_c = \bar{y}_0$, with $W_0 = 1$, and all other $W_i = 0$. This resembles the current practice of assigning the full weight to be base census year $\bar{y}_0$.

$\bar{y}_f = \bar{y}_9$ with $W_9 = 1$ and all other $W_i = 0$. This would utilize only the final, current year $\bar{y}_9$. This could be done for national and large domain estimates, where timelines dominates sampling precision, because the sample size is large enough. This may also hold for epidemics, stock prices, unemployment etc, where irregular fluctuations dominate stable changes.

$\bar{y}_e = \Sigma W_i \bar{y}_i$, with all $W_i = 0.1$. Equally weighted years may be good for (relatively) changeless stability. Some could argue that this stable model is even more needed for the present use of obsolescent decennial censuses.

$\bar{y}_w = \Sigma W_i \bar{y}_i$, with $W_0 \leq W_1 \leq W_2 \ldots \leq W_8 \leq W_9$, with monotonically nondecreasing W_i. The curve of increase may be determined by some model, or with empirical data, or their combination. The means $\bar{y}_f$ and $\bar{y}_e$ are trivial cases, but they all seem better than $\bar{y}_o$. Some monotonic increase like

$W_{i+1} = W_i + k$ (with $k > 0$), or even better $W_{i+1} = cW_i$ (with $c > 1$), seem better than $W_i = 0.1$.

Important questions remain for further discussion and research. Perhaps forever, and this can become a ''growth industry.''

a) Should the statistical source office develop and provide the one ''optimal'' estimate?
b) Or two or three estimates?
c) Or all 10 yearly data and let the consumer (beware and) belabor? Or both a) and c)?
d) If a), should the office use different estimates for small and large domains? (Asymmetrical cumulation is suggested later.)
e) Should there be ''overlaps'' within the year? Between years? Using ''composite estimates?''

We have emphasized the cumulation of periodic surveys by government offices. However, academic and other non-governmental centers have also started to collect increasing numbers of annual surveys in diverse fields. The main emphasis has been on changes and comparisons of annual results. But cumulation of cases from the periodic surveys would be useful for strengthening the sample base domains, especially for small domains, particularly for variables with stabilities that encourage cumulations. Cumulations of cases may be more efficient than weighted combination of statistics for rare items, from changing domains, and from small samples (Kish 1987, Sections 6.2–6.3).

6. Partial Overlaps, Cumulations, and Panels

Monthly and quarterly surveys of the population and of the labor force typically use some kind of rotation design based on partially overlapping sampling units of blocks, area segments, or dwellings. These were great inventions about fifty years ago, chiefly for improving neighboring periodic estimates with reduced variances due to covariances in the overlaps (Jessen 1942; Patterson 1950; Cochran 1977, 12.9–12.12; Kish 1965, 12.4). These were further improved with ''composite estimators'' not only for two overlapping periods, but also ''chained'' for more periods, and developed chiefly at the U.S. Census Bureau and Statistics Canada (Singh and Merkouris 1995). With high correlations between periods they can obtain great variance reductions for net changes; and modest gains (amazingly) even for current estimates. Of course, with low correlations for many statistics the gains are more modest (Binder and Dick 1989; Binder and Hidiroglou 1988).

The reduced cost of reinterviews reinforces the advantages of overlaps. It is difficult to estimate the relative total costs of new versus reinterviews; they have three components: initial contact, interview, processing. The first two can be considerably lower when conducted by telephone. Thus the relative costs can differ greatly between countries.

Overlaps and reinterviews are common to most (not all) periodic surveys, but the patterns of rotation show a great variety of choices between countries. It would be difficult to separate, I fear, decisions based on diverse country situations and needs, from just idiosyncratic choices and traditions. In the U.S.A. the segments are in for 4 months, out for 8, then in for another 4, for a total of 8 months of interviews (USCB 1966, 1978; Kish 1965, 10.4). The greatest overlaps (6/8) are for successive months, and for successive years (4/8), and these get the best gains when correlations due to overlaps are high.

For designing periodic studies we face a conflict, so familiar to multipurpose designs.

a. For cumulations of rolling samples it is best to have no overlaps.

b. For monthly and yearly net changes it is best to have high overlaps, especially with composite estimators.
c. For current estimates, using composite estimates, modest overlaps (1/4 or 1/3, or $\frac{1}{2}$) may be best. This may also hold for ''time series'' uses.

Justifications for these statements can be found in the references above and in the sampling literature. However, the advantages of overlaps are greatly weakened by other considerations, which are usually neglected.

1. The gains are computed for months that are fixed arbitrarily, as is shown by the country differences. The needs of public policy may often vary over time and be quite different from the fixed design.
2. The gains in variance for composite estimates in band c above may be much lower for low and for variable correlations.
3. The gains in variance are shown against a single period's statistics, but cumulated periods will have much lower variances.
4. Biases of reinterviews are often mentioned, sometimes as ''rotation group bias.'' I am not an expert, and some reinterviews may be better than the first. But problems of reinterviews should be considered.

It would be possible to arrive at a good compromise by having, in addition to a non-overlapping sample of segments, another sample in the ratio of p (perhaps 1/5 or 1/4) of overlapping segments, with some rotation scheme. This would overcome my objection #1 by having overlaps for more comparisons, net changes. However I would prefer to have the overlaps also provide gross changes, with the use of panels which are so lacking today in periodic surveys.

Panels denote samples in which the same elements (persons, families, households) are measured on two or more occasions for the purpose of obtaining *individual* changes. From the mean of these individual changes the net mean population change can be estimated. However, from the net changes of means we cannot estimate (directly) the gross change of individuals. This contrast of population/element change has been variously designated by individual/mean, or gross/net, micro/macro, or internal/external.

Only panels can reveal the gross changes behind the net changes generally (exceptions can be found with strong models) (Kish 1987, 6.2D, 6.4–6.5). The periodic labor force surveys fail to yield it, because the samples are rotated, and also because households and people change and move.

''*Split Panel Designs*'' (SPD) may be added to rolling samples, as I have proposed (Kish 1981, 1987, 1990). This would displace partial overlaps with two samples: a panel p added to the independent rolling samples $a, b, c, d \ldots$. Thus the periodic samples will consist of *pa-pb-pc-pd* etc. The size of the panel p relative to the independent samples can be varied, but a small ratio, $p/a < 1/3$, will usually suffice. This SPD has two critical advantages over the classical partial overlaps. First, it provides true *panels* of elements (e.g., persons or households), which are missing for the moving elements in designs of mere overlaps. However, panels involve following the movers, and thus they can uniquely yield most valuable statistics, which mere overlapping samples of sampling units (e.g., segments, PSUs) fail to yield. Second, in SPD the correlations are present for *all* periods,

1. Overlaps between samples. Excluded from cumulations?
2. SPD-Split Panel Design. Panels and overlaps for all periods
3. Oversampling some small domains
4. Undersampling some expensive domains
5. Weighting, e.g., moving averages to favor recent data
6. Over (under) sampling for some periods
7. Synthetic estimation for small areas and periods (SPREE)
8. Other cumulations of F periods; e.g., 52 weeks = 1 year.

Fig. 2. Possible modifications of rolling samples

not only for the pairs arbitrarily designed in the classical symmetrical rotation designs. These overlaps are mostly designed for successive monthly and yearly changes. However, often the most desirable comparisons may not be foreseen in the design, hence the benefits of correlations are absent for them. These comparisons would benefit from the correlations of SPD designs.

7. Asymmetrical Cumulations

This topic may serve to best distinguish the rational statistical designs that rolling samples can offer, from the traditional designs that pass for ''common sense.'' However, I want to emphasize that asymmetrical cumulations do not depend on rolling samples, and can be applied to other sample designs (Kish 1986). I refer mainly to the strategy of balancing sampling errors against biases due to obsolescence of data from temporal changes. Take for example the justly famous Current Population Surveys of the U.S.A., with monthly samples of about 60,000 households, with twice as many adult persons. Many judge that sample too large because its sampling precision is swamped by structural, temporal, nonsampling errors, due to the vagaries of the weather, or the calendar, or other haphazard factors that appear in its monthly news releases. On the contrary, for the statistics of important small domains the sample is too small and the sampling variability is much too large for reliable statistics. Small domains may be either geographical-administrative, such as a state; or they can consist of ''crossclasses,'' such as the Black teenage girls and boys in the labor force. Sampling variability of the statistics is even larger for the many comparisons between statistics of small domains. This is a general problem with applications in many countries and for many subjects and variables (Kish 1987, 2.1–2.3).

The same periodic surveys must serve both for overall (national) statistics and for domain statistics. Asymmetrical cumulations can best satisfy both needs: frequent (monthly) statistics for the total (national) statistics, but less frequent (e.g., quarterly or annual) statistics for smaller domains. And for these multipurpose aims, rolling samples can serve best.

Three main reasons should lead to asymmetrical cumulations. 1) The principal divisions of most countries tend to vary greatly in size, with ranges of 50 or even 100 to 1; e.g., the states of the U.S.A. and Australia, the provinces of Canada and China. Similar variations also exist for other social organizations, like firms, universities, and hospitals. 2) Below the level of the principal divisions, statistics are also wanted for their subdivisions; e.g., counties, districts, etc., which are much smaller and more numerous. 3) Cumulations

are often needed for rare items, which can be of three kinds (the size M or proportion M/N of a rare population; the variable Y or Y/M in M; or Y/N in the whole population; Kish 1965, 11.4).

I recognize the practical problems of inconsistencies that arise from asymmetrical cumulations, and they have been pointed out to me. Also, we have not been trained for these by statistical models, which always assume a single population, and fail to prepare us for the multipopulation nature of surveys (Kish 1994). The spatial (and other) domains form the multipopulation nature of surveys, and the cumulation of rolling samples provides the possibility of rational means of estimation.

I also recognize the difficulties of balancing variances, that depend on the sizes of sample bases, against biases of obsolescence, which differ greatly between variables. Nevertheless, if ''total survey errors'' has any meaning at all, the advantages of asymmetrical cumulations must be considered. Developing valid and useful methods will take more effort, talent, and time than I can muster.

8. Conclusions

Periodic surveys – quarterly, monthly, weekly – have become much more widely used and valued for the last fifty years. They are used in more countries, more fields, and collected not only by national statistical offices, but also by other institutions and centers, sometimes as yearly surveys, and I see them as the wave of the future.

These are designed to detect temporal variations in variables like unemployment, epidemics, economic fluctuations, etc. They provide data for national trends and for major domains, but not for small areas or domains.

However small domain statistics are in ever greater demand not only by social scientists but also by institutions responsible for public policy. For small area and domain details we have depended on decennial censuses that have been introduced gradually into most countries in the last two centuries. Not restricted to simple counts of persons and households, many censuses now also provide spatial detail for diverse socio-economic variables. But decennial statistics do not satisfy demands for more timely data, because of greater mobility, and also because new methods and resources now appear available.

I propose rolling samples to be collected weekly (or perhaps) monthly to serve simultaneously several major objectives.

I. They can replace the present monthly and quarterly surveys of labor force and/or current population surveys. Countries that have not yet adapted these may now have added incentives for starting them. The multiple objectives of rolling samples can be built into the designs from the start.

   Countries that have good, large surveys can use those budgets, but may face problems of conversion, because of two main obstacles. Some have large month-to-month overlaps, which may yield some modest gains for some change statistics, such as changes in unemployment. Furthermore in countries with many telephones, later interview may be cheaper than the first doorstep interviews.

II. Annual statistics based on 52 weekly rolling samples may be the chief product.

III. Decennial (and quinquennial) samples will be based on combinations of annual samples.

IV. Panel studies may also be attached, as discussed in Section 6.
V. The entire operation can also serve as a basis for other periodic or for one-time surveys.

Thus, the budget of the rolling samples should be compared to the combined cost of all these operations, rather than to the cost of only one of these, such as the CPS or the complete census.

Let me end, surprisingly, by noting the difficulties facing the cumulations that I advocated here and elsewhere. There are technical aspects to cumulations over time, which may be formidable and may differ greatly between countries and between situations. The idea of ''rolling samples'' must be translated into technically valid operations and ''the devil is in the details,'' as the saying goes. In some situations spreading the sample from a set of primary areas to the whole population may involve considerable expense in some of the strata. Finally, inferences over a span of time face obstacles that are statistical, philosophical, psychological, and traditional. They may require a new view, a new ''paradigm.''

9. References

Alexander, C.H. (1993). A Continuous Measurement Alternative for the U.S. Census. Report to USCB, also presented to the 1993 meeting of the American Statistical Association.

Binder, D.A. and Hidiroglou, M.A. (1988). Sampling in Time. In Handbook of Statistics, Krishnaiah, P.F. and Rao, C.R. (eds.), Vol. 6, Amsterdam: Elsevier, 187–211.

Binder, D.A. and Dick, J.P. (1989). Modeling and Estimation for Repeated Surveys. Survey Methodology, 15, 29–45.

Bounpane, P. (1986). How Increased Automation Will Improve the 1990 Census. Journal of Official Statistics, 4, 545–553.

Cochran, W.G. (1977). Sampling Techniques. Third Edition. New York: John Wiley and Sons.

Ericksen, E.P. (1973). A Method for Counting Sample Survey Data and Asymptotic Indicators. Demography, 10, 137–160.

Hansen, M.H. and Hurwitz, W.N. (1946). Sampling Methods Applied to Census Work. In U.S. Bureau of the Census, The History, Operations and Organizations of the Bureau of the Census, Washington: Government Printing Office, 83–94.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). Sample Survey Methods and Theory, Vol. I. New York: John Wiley and Sons.

Hauser, P.M. (1942). Proposed Annual Census of the Population. Journal of the American Statistical Association, 37, 81–88.

Herriot, R., Bateman, D.J., and McCarthy, W.F. (1988). The Decade Census Program. U.S. Census Bureau, Internal draft.

Jessen, R.V. (1942). Statistical Investigation of a Sample Survey for Obtaining Farm Facts. Ames, IA: Iowa Ag. Exp. Station, Bulletin 304.

Kalsbeek, W.D. (1973). A Method for Obtaining Local Postcensal Estimates for Several Types of Variables. Ann Arbor: University of Michigan, Ph.D. Theses.

Kish, L. (1965). Survey Sampling. New York: John Wiley.

Kish, L. (1979a). Samples and Censuses. International Statistical Review, 47, 99–109.

Kish, L. (1979b). Rotating Samples Instead of Censuses. Asian and Pacific Census Forum, 6, 1–2, 12–13.

Kish, L. (1981). Using Cumulated Rolling Samples. U.S. Government Printing Office, No. 80-52810.

Kish, L. (1986). Timing of Surveys for Public Policy. Australian Journal of Statistics, 1, 1–12.

Kish, L. (1987). Statistical Research Design. New York: John Wiley and Sons, Chapter 6, Sample Designs Over Time.

Kish, L. (1990). Rolling Samples and Censuses. Survey Methodology, 16, 63–79.

Kish, L. (1994). Multipopulation Survey Designs. International Statistical Review, 62, 167–186.

Kish, L. (1997). Periodic Rolling Samples and Censuses. In Statistics and Public Policy, Bruce D. Spencer (ed.), New York: Oxford University Press, Chapter 7.

Kish, L., Lovejoy, W., and Rackow, P. (1961). A Multistage Probability Sample for Continuous Traffic Surveys. Proceedings of the American Statistical Association, Social Statistics Section, 227–230.

Kish, L. and Verma, V. (1983). Censuses Plus Samples: Combined Uses and Designs. Bulletin of the International Statistical Institute, 50, 66–82.

Li, C. (1985). Statistical Reporting on the Impact of China's Economic Policies. Bulletin of the International Statistical Institute, Vol. LI, Book 1, 1–17.

Mooney, H.W. (1956). Methodology in Two California Health Surveys. U.S. Public Health Monograph No. 70.

National Center for Health Statistics (1958). Statistical Design of the Health Household Interview Survey. Public Health Services, 584-A2, 15–18.

Patterson, H.O. (1950). Sampling on Successive Occasions with Partial Replacement of Units. Journal of the Royal Statistical Society (B), 12, 241–255.

Platek, R., Rao, J.N.K., Särndal, C.E., and Singh, M.P. (eds.) (1987). Small Area Statistics. New York: John Wiley and Sons.

Purcell, N.J. and Kish, L. (1979). Estimation for Small Domains. Biometrics, 35, 365–384.

Redfern, P. (1989). Population Registers. Journal of the Royal Statistical Society, A, 152, 1–41.

Redfern, P. (1997). Numbering the People. In Spencer, B.D., (ed). Statistics and Public Policy, Oxford University Press, Chapter 8.

Scheuren, F. (1990). Comments on ''Rolling Samples and Censuses.'' Survey Methodology, 16, 72–79.

Schmitt, R.C. and Crossetti, A.H. (1954). Accuracy of the Ratio-Correlation Method of Estimating Postcensal Population. Land Economics, 30, 279–280.

Singh, A. and Merkouris, P. (1995). Composite Estimation by Modified Regression for Repeated Surveys. Proceedings of the American Statistical Association, Section on Survey Research Methods, 420–425.

Steel, D. (1997). Producing Monthly Estimates of Unemployment and Employment. Journal of the Royal Statistical Society (A), 160, 5–46.

U.S. Bureau of the Census (1966). Methods of Population Estimation. Current Population Reports Series P-25, No. 339 and No. 427.

U.S. Bureau of the Census (1978). The Current Population Survey: Design and Methodology. Technical Paper 40. U.S. Government Printing Office.

Received November 1995
Revised June 1997