distribución restringida

robert o. carleton

ACCURACY IN THE MEASUREMENT OF FERTILITY CHANGES WITH ONLY SMALL SAMPLE DATA

Serie A, n° 54

SUMMARY

Retrospective sample data permit the measurement of fertility to be made with exactly the same women in two successive time periods. Were it not for the fact that age (and certain other characteristics) change with time, the use of the same women would, without any increase in sample size, provide a control equivalent to simultaneous standardization by all characteristics in which sampling variation could affect the measurement of change. The present paper proposes a procedure for achieving this objective while at the same time circumventing the difficulty presented by change in age and other characteristics.

The underlying principle is a) first to reduce the change in characteristics to a minimum by observing the change in each quinquennial group in two successive intervals of only one year instead of five years, i.e., the change in the age group 25 to 29 is observed as it passes on to age 26 to 30 instead of to age 30 to 34, and b) then to compare the changes in fertility occurring to the successive partially different cohorts observed in the retrospective data. The comparison in terms of changes of fertility instead of levels of fertility not only eliminates the effect of the one year of aging, but also at the same time achieves the simultaneous standardization by all other relevant characteristics in so far as sampling variations in these characteristics lead to <u>differential levels</u> of fertility.

The simultaneous standardization, however, is not achieved in so far as these characteristics are related to <u>differential predisposition to change</u> fertility. Strong protection against this type of distortion can perhaps be derived from the proper use of the 20 percent rotation of successive cohorts.

As opposed to the standard procedure for testing the effectiveness of an action program by comparing the levels of fertility in the base and treatment periods in order to note the change occurring, the proposed procedure has the advantage of being able to compare the change occurring in the treatment period with whatever change was occurring before in the base period.

1. Until very recently the field of family planning has been an area of demography in which methodology has of necessity fallen somewhat short of the rigorous standards which characterize the discipline in general. The principal subject-matter --testing the efficacy of various contraceptive methods-- together with the usually low-cost budgets of its projects obliged working with clinical populations, ordinarily highly selective samples representative of some unspecified universe of persons in all probability significantly different from the general population.

2. In the last few years the population problem has come to occupy a prominent position in the public attention. The United Nations has adopted programs of assistance to those governments which find that rapid population growth is a serious obstacle to the achievement of the economic and social development of their countries; many of the industrialized nations are providing financial and technical support in order to encourage underdeveloped countries to adopt and carry out programs aimed at the reduction of fertility. With the rediscovery of the IUD new optimism has been imparted to the field of family planning. Large sums of money are being contributed to support its activities. Action programs designed to affect the fertility patterns of either the general population, or at least of wide sectors thereof, have come to the forefront. It has become increasingly possible to work with randomly selected samples representative of clearly designated populations. At the same time, demographers with experience in the more rigorous methodology of the other branches of demography have become interested in the problems of family planning programs.

3. A noteworthy sharpening of analytical tools can be observed. This was already evident in the Princeton Study. Then at Belgrade Potter presented a paper with an ingenious proposal for eliminating the bias in testing contraceptive methods which results from the tendency of high-risk women to drop out of studies with the consequence that data are excessively weighted by the experience of low-risk women.1/ At the August 1965 International Conference on Family Planning Programs in Geneva Freedman reported on methodological progress in family planning and strongly advocated the development of "means for estimating age-specific birth rates as well as crude birth rates ... in countries with poor vital records ... since the crude rates may be quite misleading in some cases".2/

4. For the same Geneva Conference Bogue contributed an important book3/ containing a series of notable suggestions for significantly improving the evaluation of family planning programs. He suggested the use of retrospective data as a means of reducing the sampling error without increasing the number of women surveyed; a survey conducted five years after the initiation of the program could obtain retrospective data covering the five-year periods before and after the program.4/ Another Bogue refinement for increasing the accuracy of measurement involves the utilization of computers to calculate age-specific fertility rates using women-years of exposure in each age group based on month as well as year of birth. Perhaps most promising of all is Bogue's idea of using the same sample of women in the before and after comparison in

---

1/ Robert G. Potter; Application of life table techniques to measurement of contraceptive effectiveness, United Nations World Population Conference 1965, (B.13/I/E/301).

2/ Ronald Freedman: "Family Programs Today", rough draft published in Studies in Family Planning, N° 8 (supplement), October 1965, Population Council.

3/ Donald J. Bogue: Inventory, Explanation and Evaluation by Interview of Family Planning, Motives -- Attitudes -- Knowledge -- Behavior, document prepared for discussion at International Conference on Family Planning Programs, Geneva, Switzerland, August 23-27, 1965.

4/ Or alternatively, as Bogue notes, there could be two surveys -- one at the time the program is started and covering the previous five-year period, and another five years after the program and covering the five-year period of exposure.

order to solve the thorny problem of controlling all the extraneous factors[5/] such as differences in duration of marriage, education, economic activity, etc. that are likely to distort the magnitude and possibly even the direction of any observed change in fertility when two different samples are used for measuring the level of fertility before and after the program. Otherwise, Bogue notes, the size of sample needed to standardize by all the relevant differences in characteristics "would be intolerably large".

5. Unfortunately, as Bogue has observed, the objective of controlling all the factors affecting fertility by using the same sample of women is to a large extent obstructed by the fact that some of these factors -- such as age and duration of marriage -- automatically change with the passage of time. Even the same women do not stay the same with respect to these characteristics. Nor can other factors such as labor force participation and even level of education be assumed to remain constant in time. If one studies, for example, the fertility of the cohort of women in the ages 25 to 29 in the five-year base period before the program, the comparison with the five-year treatment period after the initiation of the program -- if made with the same women -- will necessarily refer to the fertility of women in the ages 30 to 34 (see col. (2) of Table 1). As the recent fertility study[6/] of the United Nations has shown, in almost all countries of the world age-specific fertility is significantly lower at age 30 to 34 than at age 25 to 29. These same women in the sample, therefore, will probably have lower fertility after the program than before; it is clarly not permissible to interpret this lower fertility as a decline in the trend of fertility.

5/ Unless otherwise specified, the expression extraneous factors affecting fertility is used in this paper to refer to those characteristics of women differences or changes in which affect the measurement of fertility, It is this type of factor which can be controlled by using the same sample of women. In a simulated controlled experiment extraneous factors could also be used to refer to events whose effect on fertility patterns is observed simultaneously with the effect of the variable under study -- such as a family planning action program. Using the same sample of women would be of no assistance in isolating this kind of extraneous factors.

6/ United Nations: Population Bulletin of the United Nations, N° 7 - 1963, with special reference to conditions and trends of fertility in the world. ST/SOA/Ser.N/7.

Table 1

Period	Age during base and treatment periods	
	Same women in each period	Same age in each period
(1)	(2)	(3)
Base Period	25 to 29	25 to 29
Treatment Period	30 to 34	25 to 29

6. On the other hand, when age is controlled so that the base period fertility of women in the ages 25 to 29 is compared with the treatment period fertility of women in the same ages (see col. (3) of Table 1), then the comparison is no longer made between the same women even though taken from the same sample of women. While the fact that they are from the same sample of women does undoubtedly control to a certain extent for characteristics such as level of education and labor force participation,[7/] Bogue recognizes that the women from two different age groups in the same sample are different women and essentially different samples of women as far as the control of factors related to fertility are concerned.

7. The purpose of this paper is to suggest a method for exploiting still further Bogue's idée maitrêsse of using the same sample of women before and after the initiation of a program in order to control for differences with respect to the factors that affect fertility. Although the method suggested has yet to be tested in practice and many of the details of its application remained to be worked out, the important advantages of its underlying principle seem clear enough to justify throwing it open to general discussion and comment.

7/ If the clusters of a sample, for example, contain an over-representation of women with a low educational level, this over-representation will ordinarily affect all or most of the age groups rather than be concentrated in only one age group.

8. The exposition of the method is presented with special reference to measuring the changes that occur in a high fertility population after exposure to a family planning action program. As the title suggests, however, it is thought to be generally applicable, with certain, usually obvious, modifications (i.e., such as the terms base and treatment periods) to the measurement of fertility changes in malthusian as well as non-malthusian populations and without reference to an action program or to any other factors intervening between the two time periods being compared.

9. The basic principle of the method here proposed consists in the comparison of changes instead of levels of fertility in the base and treatment periods. First of all, however, changes in fertility are observed by comparing successive levels of fertility after one-year (instead of five-year) intervals. In this way, cohorts age 25 to 29 in one year are only one year older (i.e., age 26 to 30) the following year. Although exactly the same women, the difference in level of fertility from one year to the next due to aging will be minimal.[8/] Then the comparison of changes in fertility is made among the different base and treatment period cohorts as they pass from age 25 to 29 to age 26 to 30. Comparing changes has the effect of standardizing by age since the small remaining difference due to aging is approximately the same in all the cohorts being compared. Far more important, the comparison of changes is equivalent to standardizing simultaneously by all other relevant characteristics in so far as differences in these characteristics among different samples of women would ordinarily result in different LEVELS of fertility.

10. It is necessary at this point to introduce a distinction between differences in characteristics related to differential levels of fertility and those related to a differential predisposition to change fertility patterns. As will be explained presently, the comparison of changes in fertility accomplishes the simultaneous standardization of all characteristics only in the first sense of their relation to differential levels; the standardization is not achieved in the second sense of differential predisposition to change.

8/ Changes in other characteristics such as duration of marriage also assume less importance when the fertility of the same women in two successive years is compared.

11. In table 2, including the year Z which denotes the 12 month period[9/] after the initiation of the program, there are eleven years of observation (from year Z-5, the fifth year before the program, to year Z+5, the sixth year after the initiation of the program and the fifth year after the program could be supposed to have had some effect) in order to get ten years of observed change: five observed changes during each of the base and treatment periods (from Z-5 to Z-4, from Z-4, to Z-3, etc., and from Z-1 to Z during the base period and from Z to Z+1, from Z+1 to Z+2, etc., and from Z+4 to Z+5 during the treatment period). All ten observed changes, it should be stressed, refer to cohorts age 25 to 29 in one year and age 26 to 30 in the next year.[10/] In this way the comparison of changes is standardized by age.

12. In comparing (col. (1) of Table 2) the base period change in fertility of the women age 25 to 29 in year Z-5 (δ_1^B) with the treatment period change in the women age 25 to 29 in year Z (δ_1^T), one is comparing the changes of fertility in entirely different cohorts of women. However, by virtue of the comparison being between changes instead of between levels of fertility, the effect of even this one year of aging is eliminated since it can be supposed approximately the same for both cohorts.

13. This procedure designed to standardize by age manages in one fell swoop to standardize simultaneously by all other relevant characteristics in so far as differences in these characteristics will lead to differential levels of fertility. Because the two cohorts of women -- those age 25 to 29 in years Z-5 and Z respectively -- are different women and probably differ

---

9/ Although the treatment period begins at the start of year Z, at least part of this year must have the characteristic of base period fertility in the sense that its fertility is unaffected by the program. Because of the prolonged gestation period characteristic of human fertility, any effect the program has on reproductive behavior could not be noticeable in terms of births until at least nine months after the launching of the program. To round out a full year, an additional three months has been added somewhat arbitrarily to show that some time must be allowed between the initiation of the program and the moment when it can be supposed to exercise some effect on family planning attitudes and behavior. Year Z+1, therefore, is the first year of the treatment in which treatment might have some effect on fertility.

10/ For illustrative purposes the discussion throughout this paper is in terms of the age group 25 to 29. In actual practice, of course, all the age groups in the reproductive age span would normally want to be studied in similar fashion.

Table 2

MEASUREMENT OF CHANGES (δ_i) IN FERTILITY IN BASE (δ_i^B) AND TREATMENT (δ_i^T) PERIOD AGE 25 TO 29

Cohorts used in each year of Observation
(Not arranged to show rotational sequence)

Year of observation[a/]	Cohorts by age in year of survey (Z+6)				
	36-40	35-39	34-38	33-37	32-36
	Age during years of observation				
	(1)	(2)	(3)	(4)	(5)
	Base Period				
Z - 5	25-29				
	δ_1^B				
Z - 4	26-30	25-29			
		δ_2^B			
Z - 3		26-30	25-29		
			δ_3^B		
Z - 2			26-30	25-29	
				δ_4^B	
Z - 1				26-30	25-29
					δ_5^B
Z					26-30

Year of observation[a/]	Cohorts by age in year of survey (Z+6)				
	31-35	30-34	29-33	28-32	27-31
	Age during years of observation				
	Treatment Period				
Z	25-29				
	δ_1^T				
Z + 1	26-30	25-29			
		δ_2^T			
Z + 2		26-30	25-29		
			δ_3^T		
Z + 3			26-30	25-29	
				δ_4^T	
Z + 4				26-30	25-29
					δ_5^T
Z + 5					26-30

a/ The year Z refers to the 12 month period after the initiation of a treatment program.

significantly with regard to characteristics related to fertility,[11] their levels of fertility will in general be different quite apart from and independent of any changes in fertility that are occurring. If the cohort age 25 to 29 in the year Z (the treatment period cohort) has a greater proportion, for example, of economically active women or of women with more education than the cohort age 25 to 29 in the base period year Z-5, it can for this reason alone be expected to have a lower level of fertility than the base period cohort in both the first and the second of the two years under observation with regard to change. If the levels of fertility of the two different cohorts were being compared, the lower fertility of this cohort would be indistinguishable from and, therefore, liable to be interpreted mistakenly as a decline in fertility. If, however, the changes in fertility of the two cohorts are being compared, the lower fertility of the treatment period cohort gets washed out because it is equally present in both years under observations: it, therefore, disappears in the calculation of the difference between these two observed levels.

11/ As was noted above, these differences are reduced only to a limited extent by the fact that the two cohorts of women come from the same sample of women.

14. It should be clearly stated that the method proposed here makes no pretense at a causal evaluation of family planning programs. Fertility is a highly complex phenomenon simultaneously affected by a wide variety of different factors, some of which may well be acting to increase fertility, while the effect of others is to decrease it. Observed changes are net changes, the resultant of the combined effect of all factors in operation. Causation is most appropriately studied through the use of a control group as similar as possible to the experimental group in all relevant respects except exposure to the action program.12/ It is nevertheless possible to some extent, without the use of a control group, to evaluate the causal effect of a particular factor by recourse to independent data on the behavior of other relevant factors during the base and treatment periods. It would, for example, be more legitimate to infer the causal effect of an action program in a population where social change was very slow than in a dynamic society where many other active factors were abounding. The present proposal, however, aims merely at comparing the net changes in fertility patterns in the base and treatment periods before and after the initiation of a family planning program and expects to achieve this with more accuracy than has been possible heretofore.

15. It perhaps also should be pointed out that comparisons are made always in terms of period (cross-sectional) age-specific fertility rates during two successive time periods of limited duration. In interpreting any changes observed in these period fertility rates, it will usually be impossible to distinguish genuine changes in fertility patterns from fluctuations due merely to changes in the timing or spacing of children.

16. An extremely important advantage of comparing changes instead of levels of fertility in the base and treatment periods lies in that prevision is made for the by no means impossible contingency that the level of fertility has not

---

12/ Another advantage of the use of retrospective data is that the validity of the control group can be tested by comparing the control and experimental groups not only with respect to their composition in terms of characteristics related to fertility, but also with respect to their fertility patterns during the base period, i.e., before exposure to the program. To the extent that the control group is similar to the experimental group with respect to all relevant characteristics, the pre-program fertility of the two groups should be similar. The non-exposure of the control group to the action program cannot, of course, be tested in this way.

been constant during the base period prior to the program. When levels of fertility are compared, the most that can be affirmed is whether fertility has declined since the inception of the program. But if fertility was declining before the program began, the question the sponsors of a program want to investigate is whether it has been declining faster since the program started. Similarly, if fertility was increasing in the base period, the investigators will want to determine whether the increase has decelerated during the treatment period. The comparison in terms of fertility changes during the base and treatment periods is directed precisely to the investigation of this aspect of the fertility patterns of the two periods.

17. Furthermore, by measuring the change in fertility occurring to a base period cohort in two successive years and comparing this change with that occurring to a corresponding treatment period cohort, a potentially serious source of bias in retrospective data — the increasing forgetfulness of births as the reference period is farther in the past -- is largely averted. When, as in col (3) of Table 1, the fertility of the cohort age 25 to 29 in the base period is compared with that of the 25 to 29 cohort in the treatment period on the basis of retrospective data obtained at the end of the treatment period, the fertility of the treatment period can be expected (because of fewer unreported births) to be underestimated less than base period fertility. The bias will give the impression of increasing fertility and would tend to obscure partially or totally any decrease caused by the family planning program. Bias of this kind would be most damaging among illiterate populations of high fertility and probably also more so among older cohorts than in the 25 to 29 age cohort.

18. When fertility change is measured over two successive years (such as Z - 5 and Z - 4) the difference in the forgetfulness of births will be very small in comparison with that resulting when the change is measured over two five-year periods. Even this very small difference is mostly eliminated by comparing changes instead of levels of base and treatment period fertility. Both changes will have a small upward bias which largely washes out when the difference in the changes is compared.

19. At least two difficulties should be mentioned in connection with this approach. First, the comparison of changes in fertility does not completely

solve the problem of controlling for differences with respect to the factors affecting fertility. Women of different characteristics are undoubtedly differentially predisposed to adopt family planning practices. Should a decline in fertility occur, its magnitude will tend to be over- or understated if in the base and treatment period cohorts there are different proportions of women with characteristics that predispose them to family planning.

20. Secondly, the comparison between changes of fertility after only one year of aging reduces substantially the number of women-age-years of experience involved in the measurements and would require a larger sample of women if the sampling error is not to be materially increased. The Bogue comparison in column (3) of Table 1 of women age 25 to 29 in the base and treatment periods takes in altogether 25 women-age-years of experience per woman in each of the two periods (each of the five ages --25, 26, 27, 28 and 29 -- in each of the five years of the base period and in each of the five years of the treatment period). The present proposal, in comparing the changes occuring to one base period and one treatment period cohort each age 25 to 29 in one year and age 26 to 30 in the next, uses only ten women-age-years in each period.

21. The second difficulty is readily resolved because in both the base and treatment period are five (partially)[13/] different cohorts age 25 to 29 whose changes (δ_i^B and δ_i^T) in fertility are recorded as they pass on to age 26 to 30. It is possible to combine the experience of each of the five base period cohorts into one single average figure ($\bar{\delta}^B$) and to compare this average change with the corresponding average ($\bar{\delta}^T$) of the changes in the five treatment period cohorts. In this way, the entire experience of women in these ages will be used in the comparison.

22. This solution also partially resolves the first difficulty of controlling for differences in factors relating to fertility when the cohorts being compared are entirely different women and in general with characteristics somewhat differently predisposing them to a voluntary limitation of family size. By using the average change in each of five only partially different cohorts in order to compare the changes in fertility among the base and treatment period cohorts,

13/ The five different cohorts, of course, are only partially different since in each successive year the women age 25 to 29 are exactly the same women in four of the five ages. In each successive year there is an approximately 20 percent rotation, with the women age 29 passing out as they reach age 30 and being replaced by the women age 24 who reach age 25.

it becomes possible to take advantage of the 20 percent rotation in the cohorts each year; in the women-age-years used in calculating the average changes for the base and treatment periods, approximately 40 percent of the age-years in the two periods refer to the same women. This is seen more clearly in Table 3 where the entries in Table 2, in order to represent better the rotational sequence involved, are re-arranged in one long diagonal covering both of the two periods instead of two separate diagonals, one for each period. Of the five treatment period cohorts, the following approximate percentages of women age-years refer to women also in the base period: 80 percent of cohort age 25 to 29 in year Z, 60 percent of those age 25 to 29 in Z + 1, 40 percent in Z + 2, 20 percent in Z + 3 and none at all of the cohort age 25 to 29 in Z + 4.

23. To what extent this 20 percent rotation of women in each successive cohort reduces the minimum sample size required to control for differences in characteristics related to differential predisposition to change fertility is a matter for investigation by sampling specialists. Attention is called, however, to the apparent similarity of this procedure to the 25 percent rotation used by the U.S. Bureau of the Census in its Current Population Survey in order more effectively to measure changes.

24. Although the principle of comparing changes instead of levels of fertility has been shown effectively to standardize by all the characteristics differentials in which result in <u>differential levels</u> of fertility, the 20 percent rotation of women is the only device mentioned up to now for isolating in a genuine decline in fertility the effect of a differential decline among women with characteristics <u>differentially predisposing</u> them to change their fertility. To what extent can the differential composition of the base and treatment period cohorts (when caused by sampling variation) distort the analysis of the changes that have occurred? To answer this question, it is necessary to considerer the different ways in which fertility differentials can act in the course of a fertility reduction. For the sake of simplicity and concreteness, let us take as an example a dichotomous educational differential; during the base period women with a "high" level of education have lower fertility than women with a "low" level of education:

Table 3

MEASUREMENT OF CHANGES (δ_i) IN FERTILITY IN BASE (δ_i^B) AND TREATMENT (δ_i^T) PERIODS, AGE 25 TO 29

Year of observation	Base Period 36-40	35-39	34-38	33-37	32-36	Treatment Period 31-35	30-34	29-33	28-32	27-31
	Cohorts used in each year of observation (Arranged to show rotational sequence) — Age during years of observation									
Z - 5	25-29									
Z - 4	26-30 δ_1^B	25-29								
Z - 3		26-30 δ_2^B	25-29							
Z - 2			26-30 δ_3^B	25-29						
Z - 1				26-30 δ_4^B	25-29					
Z					26-30 δ_5^B	25-29				
Z + 1						26-30 δ_1^T	25-29			
Z + 2							26-30 δ_2^T	25-29		
Z + 3								26-30 δ_3^T	25-29	
Z + 4									26-30 δ_4^T	25-29
Z + 5										26-30 δ_5^T

a. Fertility may decline only among the high educational women; as a consequence the differentials widens.[14]

b. Fertility may decline by proportionately the same amount among both high and low educational women; as a consequence the differential remains unchanged.

c. The composition of the population changes with women shifting from the high fertility category (i.e., low educational level) to the low fertility category (i.e., high educational level); as a consequence over-all fertility declines without any change occurring either in the differential or in the levels of fertility of the educational categories.

25. Any significant decline in fertility during the treatment period may very well be the result of some combination of all three modes of acting.[15]

14/ It is also logically possible that the differential narrow as a consequence of a decline in fertility only among the low educational women. In non-malthusian populations of very high fertility this possibility can be neglected as extremely unlikely. When the model is applied to countries of moderate or low fertility, this mode of behavior would have to be taken into consideration.

15/ It would be difficult to specify in general the precise proportion of each of the three. Undoubtedly, this will vary both from country to country as well as from one type of differential to another. So far as the writer can determine, no systematic study has yet been made on the relative importance of the three modes. Ryder (in his chapter on fertility in Hauser and Duncan's Study of Population, page 412) mentioned this among the questions in differential fertility analysis which are "both important and relatively unanswered".

The writer believes he has demonstrated that the third mode (changing population composition without any change in fertility) cannot be of great significance in a short-run fertility decline unless the magnitude of both the fertility differential and of the change in population composition are very great (Robert O. Carleton, "Fertility Trends and Differentials in Latin America", The Milbank Memorial Fund Quarterly, Vol. XLIII, N° 4, Oct. 1965, pp. 15-29).

In the same work he found among five Latin American countries around 1950 the urban-rural fertility differential to be smallest in the countries with highest fertility. This finding suggests that a widening differential is associated with decreasing fertility at least in some countries (mode #a). Finally, not very comprehensive data selected from countries with different levels of fertility show that rural fertility as well as urban fertility is usually lower in the low fertility countries; the implication would seem to be that mode #b is operative to some extent also.

To the extent that mode #b (an even decline in fertility among the different categories of each differential) is operative, no control need be established for differences in composition of the cohorts in the base and treatment periods; the composition of the cohorts is irrelevant in this kind of decline.

26. It must be admitted that the method under discussion, by the very nature of its design, is completely ineffective in detecting mode #c changes in fertility (a shift of population composition from the high to the low-fertility categories). The objective of discounting differences in characteristics relevant to fertility by comparing changes instead of levels of fertility backfires in the case where the differences are not due to sampling variation, but represent genuine changes in the compositional structure of the population.16/ The importance of this limitation cannot be very great, however, because short-run changes in population composition are seldom of sufficient magnitude to have a marked effect on fertility levels. In the exceptional cases where abrupt and drastic changes do occur, their revolutionary character will render them highly conspicuous as a consequence of which the danger of unconscious bias will be very

16/ At first glance, it might appear that the measurement of both base and treatment period fertility with data from a sample of women based on one, single survey taken at the end of the treatment period would attribute post-treatment composition to the base period cohorts and, therefore, also fail to take into account changes in population composition. These changes, however, are partly taken into account by a parallel evolution of the composition of the women in the sample. The change in their characteristics between the first year of the base period and the year after the treatment period should reflect to some extent the structural changes occurring in the general population.

The 20 percent rotation of women whereby each successive cohort was born on the average one year later provides even greater assurance that changes in the composition of the population will be reflected in the characteristics of the base and treatment period cohorts. As Ryder has shown so eloquently (Norman B. Ryder, "The Cohort in the Study of Social Change", American Sociological Review, December 1965, pp. 843-61), changes in population composition tend to occur successively in the younger cohorts rather than to appear all at once in all the age groups. The 20 percent rotation provides a built-in protection against population composition changes of this kind. The base period cohorts were born on the average five years before the treatment period cohorts and, therefore, should have an appropriately smaller proportion of women in the low-fertility categories.

small. In such instances special controls could be instituted; the sample size might have to be increased so that cross-tabulations can be made.

27. In the case of mode #a (a widening differential as a result of declining fertility exclusively in the low-fertility categories), the procedure proposed does not provide altogether satisfactory controls for differences in the composition of the base and treatment period cohorts with respect to factors affecting fertility. To the extent that the treatment period cohorts contain a greater proportion of women in the low-fertility categories (i.e., urban rather than rural, high rather than low level of education, economically active rather than inactive), then -- provided that this difference in composition is a result of sampling error and does not represent genuine changes in the population composition -- the declining fertility of these low-fertility categories will have a greater impact on the changes in fertility of the treatment period cohorts (the δ_i^T) than on the changes in the base period cohorts (the δ_i^B) in part because of a real decline in fertility, but in part also because sampling error gives greater weight to these low-fertility categories in the treatment period cohorts. The decline that has in fact occurred is exaggerated by the failure to control for differences in composition. Conversely, if the treatment period cohorts contain a smaller proportion of women in the low-fertility categories, the decline will be understated because of the failure to control for differences in composition.

28. Since this kind of bias occurs only when the differences in base and treatment period cohorts are due to sampling variation, the importance of being able to distinguish between differences arising from changes in the composition of the population and those due to sampling error will be readily appreciated. It should be possible to make this distinction (at least when the differences in composition are very pronounced) by an analysis of the pattern of changing composition among the ten successive partially different cohorts in the base and treatment periods. The pattern of change will tend to be more orderly when it reflects a change in the population at large, and will have a random character when caused by sampling variation.

29. Furthermore, it is worth noting that a bias of this kind occurs only when a decline in fertility does in fact occur whose type is that of mode #a. If there is no decline in fertility, the presence of more low-fertility women in the treatment period cohorts has no effect on the comparison of base and treatment period cohorts since the comparison is being made in terms of changes and not of levels. As was observed above in paragraph 11, differences in level of fertility disappear when changes in fertility are compared. In other words, although the method can exaggerate a decline that has taken place, it cannot manufacture a decline out of whole cloth and misleadingly give the impression of a decline when in fact there has been none.

30. Among the many aspects that could benefit from discussion and further investigation the following have been selected somewhat at random:

a. What sample size is required in a quinquennial group in order to determine whether a specified difference in the average changes of the base and treatment period cohorts is significant?

b. In addition to comparing the average changes in the base and treatment period cohorts, it would be useful to study the sequence of changes in time for the successive base period and treatment period cohorts in order i) to learn whether observed trends are gathering momentum and ii) to detect signs of random variation which indicate insufficient sample size. How can this best be done?

c. It has been proposed somewhat arbitrarily to study changes in fertility with five base and five treatment period cohorts, all of these during two successive one-year intervals, i.e., at age 25 to 29 and then at age 26 to 30. One could have selected either shorter intervals with more cohorts (e.g., six-month intervals with ten base and ten treatment period cohorts, each at age 25 to 29 and then at age 25.5 to 29.5) or longer intervals with fewer cohorts (e.g., the extreme case would be $2\frac{1}{2}$-year intervals with one base and one treatment period cohort, each at age 25 to 29 and then at age 27.5 to 31.5). Which of the alternatives would be more efficient and require the smaller sample size is not immediately obvious. Even if the present one-year intervals of observation should be maintained, it might be advisable, for the sake of comparability with other data, to take the quinquennial groups one-half year younger in order to make them more representative of the conventional age groupings. For example, the age group 24.5 to 28.5 could be observed as it passed to age 25.5 to 29.5. Its average age during the two years of observation would be exactly 25 to 29.

d. The distinction introduced here between characteristics related to <u>differential levels</u> of fertility and those related to <u>differential predisposition to change</u> fertility, are they new concepts, or are they old concepts which I, having failed to recognize, have dressed up in new terminology?

e. The analysis of differences in terms of numbers of persons requires an unusually large total sample because the number of cases remaining in residual categories will otherwise be too small. What are the sample requirements in the present instance where the differences under analysis are changes in <u>rates</u>?