

ON THE VALIDITY OF FERTILITY DATA COLLECTED THROUGH INTERVIEWS

By J. M. SENGUPTA

Indian Statistical Institute

SUMMARY. In a city like Calcutta, where the general literacy is high, we have found no reason to suspect that lapses in recalling every child birth over a period of several years, is of serious dimensions. It appears that reasonably dependable results may be expected, when the investigation is carried out with a specific emphasis on the fertility data, and not as one out of many topics covered in a multipurpose survey.

In sample surveys spread over a number of months, one has to consider carefully, if in the recall of earlier births, all the twelve months of the year have been equally represented and the seasonal effect is duly balanced. Due to a pronounced seasonal pattern in the natural rates of birth, a sample not duly balanced over time may lead to biased conclusions.

The vital statistics compiled by the Public Health Department based on the municipal records in Calcutta are collected by Wards, i.e., Constituencies. From our present study, it has been observed that the major proportion of child births is being reported from the hospitals and similar institutions. The element of under-reporting of birth for the vital statistics of the city is thus becoming less and less important.

1. *Object of the study.* The ultimate measure of success in any family planning programme is in the effective reduction of live births within a stated interval of time. It is of paramount importance therefore that the number of live births should be capable of accurate assessment. With large populations covered by such programmes this assessment has to be made usually on the basis of sample investigations made at reasonable time intervals. It is hardly possible to maintain a day-to-day contact with the individual couples under observation, so that not a single incidence of live birth is missed or non-reported. The informant when interviewed after a given interval of time has therefore to recall each incident occurring in the interval. The shorter interval, the easier it would be to recall, without serious lapses. It is generally believed, that the lapse in recall is considerable and data thus collected is apparently not so reliable. The interval between two successive enquiries should, therefore, be as small as possible and since repeated enquiries involve expenditures in time and money, an optimum length of interval has to be worked out to guarantee data which would be practically free from recall lapse compatible with the cost of enumeration.

The object of the present study is to examine the extent and magnitude of this phenomenon of memory lapse and to ascertain up to what length of interval, this factor does not assume serious proportions. The data at our disposal is the material collected by an initial bench-mark survey conducted by the Indian Statistical Institute in two selected centres of Calcutta in 1963 covering a total population of 10700 persons with 1405 active couples. A couple has been defined as an active one, when the husband's age is above 18 years and the wife's age is between 15 and 45. The initial enquiry according to the scheme, will be followed

up by an action programme on mass communication lines. The effect of this drive will be measured by a resurvey of all active couples in the area successively at intervals of every six months. This investigation should enable us to measure the effective change in birth rates resulting from the family planning drive, which is its direct object, and would also furnish valuable material to study the fertility pattern in the different socio-economic classes.

2. *Recall of child births in their chronological sequence.* In the present enquiry, the couples were asked to recall each incidence of child delivery even if it resulted in a still birth or abortion since the year of marriage. The present age of each living child or the age at death of children who have subsequently died, were also recorded. The recording of each incidence of delivery in a chronological sequence is expected to reduce lapses in recall, i.e., skipping over intermediate child birth, if any, between two reported births. For whenever the interval between successive child births was large, the investigator could put in additional probing questions which would help to refresh the informant's memory. Further questions were made to ascertain the place of each delivery, whether at one's own residence, in a relation's house, or in a hospital or clinic. Ascertainment of such details directly associated with the incidence of birth has been considered to be an aid to the informant's memory.

The enquiry was confined to information relating to the growth of family size and allied topics in a short questionnaire and not in the form of a multipurpose schedule, where the investigator in his hurry to cover up all the blocks in his schedule of considerable dimensions, may have to deal with the individual items superficially. On *a priori* grounds, there is every reason to believe that the quality of the data so collected should be reasonably good, if we are permitted to assume that the parents would rarely forget events of such importance in their life, like the birth of a child. For lapses, if any it would perhaps be more reasonable to suspect an weakness in the interviewing procedure itself, than a natural failure of memory on the part of the couple.

3. *Age specific rates of live birth per active couple from year to year.* Table I gives the age-specific rates of live birth as in the different years from the current year backwards up to 1951, based on data from both the centres combined. In each year, the couples were classified under six age-groups, according to current age attained by the woman in that year, and rates of live births were computed separately for each age-group.

It will be seen that the rates of live birth have remained more or less steady in the younger age-groups up to 25 years. In the higher age-groups, there is a sharp fall in birth rate since the year 1951 specially in women above 30. It should be noted here that the data for the year 1962, represent only the first seven months, i.e., up to the month of July only. The estimates for 1962 refer to the seven-month figure inflated for twelve months. This does not take full account of the seasonal variations and hence inadequate.

ON THE VALIDITY OF FERTILITY DATA

The birth rates in each age group thus obtained are based on a rather small size of sample. As a result, some irregularities have been exhibited here and there, while for the oldest age group 41 years and above, not a single couple has been shown in the years 1951 and 1952. On the other hand, the number of couples observed were very few and that again with no child birth at all in the two years 1953 and 1955. The observed birth rates were therefore graduated by a second degree fit by the method of least squares, in the form :

$$y = a + bx + cx^2$$

where

y = birth rate

and

x = age levels.

The graduated values of birth rate have been shown in columns (2)-(7) of Table 2.

We have however to consider that the enquiry was confined to couples who were active in 1962. As we go back the number of active couples gradually fall out, many of them not having attained the active age at that time. On the other hand, couples who were still active in those early years, but have gone out of activity since, do not enter into our picture. As a result, the effective age-composition of the couples in all these years has undergone a serious change and does not represent the age pattern of 1962. The distribution of active couples according to their current age-groups as in the different years, have been given in Table 3. The percentage composition of the couples among the different age-groups has been worked out and shown in columns (2)-(7) of Table 4. It will be seen that the percentage composition of women in the younger ages have gone up, while those of an advanced age have gone down, as we proceed backwards from 1962 to the year 1951.

TABLE 1. OBSERVED AGE-SPECIFIC BIRTH RATE PER (000) OF ACTIVE COUPLES IN DIFFERENT YEARS ACCORDING TO THE LEVELS OF CURRENT AGE ATTAINED BY INDIVIDUAL WIVES (MOTHERS)

year of birth	observed birth-rates per (000) of active couples by mother's current age (years)						weighted overall age-composition of 1962
	15-20	21-25	26-30	31-35	36-40	41 and above	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1962	(226)	(297)	(135)	(098)	(034)	(907)	—
1961	298	265	147	102	048	000	119
1960	271	248	165	120	068	006	133
1959	235	247	184	122	093	013	140
1958	216	299	217	136	052	030	150
1957	280	236	238	123	107	066	164
1956	289	247	233	144	064	028	156
1955	218	294	142	100	174	—	107
1954	270	287	255	191	068	091	188
1953	291	255	183	105	107	—	143
1952	255	314	248	184	132	000	179
1951	200	247	208	204	111	000	170
number of couples interviewed	107	237	316	244	252	250	—

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

TABLE 2. GRADUATED AGE SPECIFIC BIRTH RATE PER (000) ACTIVE COUPLES IN DIFFERENT YEARS ACCORDING TO THE LEVELS OF CURRENT AGE ATTAINED BY INDIVIDUAL WIVES (MOTHERS)

year	graduated birth rates per (000) of active couples by mother's current age (years)						weighted overall ages	birth rate per (000) persons
	15-20	21-25	26-30	31-35	36-40	41 and above		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1961	292	216	149	93	47	12	119	15.6
1960	278	229	177	122	65	6	133	17.4
1959	246	223	188	141	83	13	140	18.3
1958	252	240	209	159	90	1	150	19.0
1957	284	243	200	155	108	69	164	21.4
1956	292	254	206	151	87	15	166	20.4
1955	226	240	226	187	120	28	167	21.8
1954	297	270	233	185	127	69	186	24.3
1953	295	242	188	131	73	12	143	18.7
1952	273	276	251	197	115	4	179	23.4
1951	291	259	220	176	123	65	179	23.4
number of couples (1962)	107	237	315	244	252	260	(1405)	

TABLE 3. DISTRIBUTION OF THE NUMBER OF ACTIVE COUPLES IN THE DIFFERENT YEARS ACCORDING TO DIFFERENT LEVELS OF WIFE'S (MOTHER'S) CURRENT AGE ATTAINED IN DIFFERENT YEARS

year	number of active couples according to mother's current age (years)						
	15-20	21-25	26-30	31-35	36-40	41 and above	total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1962	107	237	315	244	252	260	1405
1961	104	278	279	274	210	229	1374
1960	133	250	309	234	237	178	1341
1959	136	291	207	245	215	161	1305
1958	130	268	276	220	210	133	1246
1957	182	292	240	262	160	91	1216
1956	197	267	270	209	167	72	1172
1955	179	289	232	236	121	67	1114
1954	201	247	243	215	118	33	1057
1953	196	255	218	210	112	21	1012
1952	255	229	250	169	91	—	954
1951	224	259	207	157	72	—	919

ON THE VALIDITY OF FERTILITY DATA

TABLE 4. PERCENTAGE DISTRIBUTION OF ACTIVE COUPLES ACCORDING TO DIFFERENT LEVELS OF WIFE'S (MOTHER'S) CURRENT AGE ATTAINED IN DIFFERENT YEARS

year	percentage of births at different levels of mother's current age (years)						total
	15-20	21-25	26-30	31-35	36-40	41 and above	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1902	7.6	10.0	22.4	17.4	17.0	17.8	100.0
1901	7.6	20.2	20.3	19.9	15.3	16.7	100.0
1900	9.9	18.6	23.1	17.4	17.7	13.3	100.0
1959	10.4	22.3	20.4	18.8	16.5	11.6	100.0
1958	11.2	21.5	22.2	17.6	16.8	10.7	100.0
1957	15.0	24.0	19.7	20.7	13.1	7.5	100.0
1956	16.8	22.8	23.0	17.8	13.4	6.2	100.0
1955	16.1	25.9	20.8	21.2	10.9	5.1	100.0
1954	19.0	23.4	23.0	20.3	11.2	3.1	100.0
1953	19.4	25.2	21.5	20.7	11.1	2.1	100.0
1952	25.9	23.3	25.4	16.2	9.3	—	100.0
1951	24.4	26.2	22.5	17.1	7.8	—	100.0

In order to bring up the annual all-age birth rates on a comparable footing, the graduated age-specific birth rates in the different years were weighted with the age composition as in 1962. The weighted birth rates on this 1962 pattern have been worked out and shown in column (8) of Table 1. It will be seen, that leaving out the year 1962 which did not come under the full observation, the birth-rates tend to fall steadily from 1951 with advancing years. Thus, instead of the birthrate going down as we go backwards, which we may expect as a result of memory lapses, i.e., under-reporting of births, we are confronted with a trend of steady increase.

4. *Annual trend of birth (live) rate per (000) of active couples.* The situation revealed in Table 1 discussed above now has to be explained one way or the other. Apparently, the birth rate seems to have fallen from year to year in this decade. If there has been some recall lapse, the trend exhibited in Table 1 must have been damped, to a greater and greater extent as the year of reference recedes into distance. In other words, the true fall in birth rate must have been somewhat steeper than what is being apparently revealed here. It does not seem likely however that a lapse of recall could have been very high in as much as the birth rates in the early part of the decade are already high.

5. *Validity of the results.* Some sort of an external evidence is clearly needed to test the validity of the results obtained from this enquiry. That the birth rate is

gradually falling in this city is the generally accepted popular view. As has already been mentioned, the population covered in this enquiry were all persons usually residing in these two centres arbitrarily chosen within the Calcutta Corporation in its northern section. For a comparison, reference was made to the official figures for birth rates in the Calcutta city area. The vital statistics on monthly births and deaths published by the Public Health Department, Govt. of West Bengal are based on these data compiled by the Corporation of Calcutta. But unfortunately, the published statistics refer to the city area as a whole, and smaller breakdowns by Wards or Constituencies are not available. As these two centres represent only small portions of three different Wards, the City as a whole does not provide us with a reasonable basis for comparison. Besides, these centres are situated in old and settled quarters of the city, which suffered the least of disturbing influences from migrating population compared to the other quarters of the city. An attempt was therefore made to compile the Ward-wise figures from the original weekly returns maintained by the Statistical Section of the Calcutta Corporation, which has very kindly made all their records accessible to us for purposes of our compilation and studies. It may be noted here that 32 Wards of the city had been split up in the year 1953 into 80 smaller Constituencies. Although, the individual 'Constituencies' containing these two centres would have constituted a population with the nearest correspondence, we had to fall back on the old Wards, much larger areas, so that a correspondence in the total population could be maintained between 1951 and 1961. Total births for these Wards within which the two centres fall, were therefore compiled week by week for the years 1951 to 1961.

5.1. *Source of official statistics on births and deaths.* The number of births maintained by the Statistical Section of the Calcutta Corporation are compiled from

- (a) births reported by the residents individually,
- (b) births reported from hospitals and clinics.

Up to the year 1952, (b), i.e., hospital reports were collected from the city as a whole and it is not possible to have them distributed among the individual Wards, where the births have actually taken place. But from the year 1953 onwards, the hospital reports are being analysed and each birth in the hospitals has been posted into their respective Wards on the basis of addresses furnished by the patients. For our present study the over-all contribution from hospitals in the two years 1951 and 1952 has been allocated among the different Wards in proportion to their total populations in 1951. This contribution from the hospitals formed however a very insignificant proportion of total births in the early years of 1951 or 1952. It is in the recent years only that hospitals and clinics report the major proportion of total births. This has been discussed in an earlier note based on the present enquiry.

ON THE VALIDITY OF FERTILITY DATA

5.2. *Comparison of birth-rates based on the present enquiry against the same based on official records.* Table 5 in column (3) gives the computed birth-rates based on the total number of births and total population in Wards 1, 4 and 6, taken together (within which the two centres are contained). The population figures for the intermediate years were obtained by a linear interpolation between the census figures for 1951 and 1961, and the corresponding number of total live births were compiled from the municipal records.

TABLE 5. COMPARISON OF BIRTH RATES BASED ON THE
PRESENT ENQUIRY (GFS) AGAINST THE SAME BASED
ON OFFICIAL RECORDS

year	Calcutta Corporation (Wards 1, 4, 6) as per municipal records		birth rate per (000) persons as per present enquiry (GFS)
	number of live births	birth rate per (000) person (c.a.)	
(1)	(2)	(3)	(4)
1901	6551	19.1	15.5
1960	8914	20.0	17.4
1959	8632	19.0	18.3
1958	7577	21.6	19.6
1957	7461	21.1	21.4
1956	8264	23.3	20.4
1955	8050	24.2	21.8
1954	7968	22.1	24.3
1953	8169	22.6	18.7
1952	7929	21.8	23.4
1961	8619	23.6	23.4

col. (3) is based on Calcutta Corporation data (C.C.)

col. (4) is based on the present "Growth of Family Size" enquiry (GFS)

It should be noted that the birth-rates given in column (3) of Table 5 based on the present enquiry refer to the number of births per (000) of active couples, while the municipal birth-rates given in column (4) of Table 5 refer to the number of births per (000) of gross population. The rates based on the present enquiry has therefore been expressed in terms of births per (000) persons in column (4) of Table 5, using a conversion factor of 0.1305, the proportion of active couples to the number of total persons covered by this survey. It will be seen, that the two parallel sets of results are remarkably close, considering that the population under reference is not strictly identical. In fact only a small portion of the three Wards are being contrasted against the three in full. Besides, the proportion of active couples in the total population is based on a small material as in 1962, which is assumed to have remained unchanged during the entire period.

This tends to show that data collected in the present enquiry through personal interviews, does not materially differ from the official statistics collected through routine reports. The incidence of memory lapse on the part of the informant which should increase as the length of recall period increases, is also not noticeable. On the contrary it is only in the very recent years, where the computed birth rates are at a lower level compared to the official rates, while they remain in excellent agreement during the earlier years. It is quite possible however that both of them suffer equally from inaccuracies resulting from an under-reporting of births, and thus both are giving underestimates of the true birth rate.

As has already been discussed, official statistics is compiled from two different sources : (a) from private residents, (b) through institutions like hospitals and clinics etc. It is only in the former that an under-reporting of birth incidents may be suspected which in fact is considered to be a basic weakness in all our vital statistics in India. But in a city like Calcutta this factor should be of much less importance compared to the rural areas in general.

Finally, the proportion of births reported from private residences in the city of Calcutta is going down rapidly as has been discussed in an earlier note based on the present enquiry. The under-reporting factor is thus becoming less and less important. It is in the earlier years only that the official rates may have been underestimated. But since the earlier years are showing higher birth rates in spite of this, it seems that this factor had never been serious in a city like Calcutta, unlike the rural areas where the situation may be quite different.

6. *Seasonal trend in birth rate.* It will be interesting to study the effect of recall lapse within a relatively short period of time, say within a span of one year. The size of the present sample is unfortunately too small to show up the monthly features in the rate of birth. The monthly rates of birth computed from the present survey have been given in column (3) of Table 6 against which the corresponding rates based on the Corporation material are shown in column (2). For smoothing, three month moving averages were calculated and duly entered. The official vital statistics clearly indicate a seasonal trend with a maximum of births during the months September-December. In fact, this boom in births during the winter-months is a known feature in the seasonal pattern of birth rates everywhere in India. The corresponding birth rates based on the present sample shows a general rise towards the autumn months i.e. somewhat staggered with spurious ups and downs here and there. The seasonal pattern is more clearly brought out in columns (4) and (5) of Table 6 where the birth rates have been expressed as indices to the respective annual averages. The size of sample, namely, the number of couples under observation was obviously too small to indicate firm results.

ON THE VALIDITY OF FERTILITY DATA

TABLE 6. BIRTH RATES IN DIFFERENT CALENDAR MONTHS OF 1961 BASED ON THE CALCUTTA CORPORATION RECORDS AND THE PRESENT ENQUIRY "GROWTH OF FAMILY SIZE" (GFS)

month	birth rate per (000) person		indices to the all months average rate as base (= 100)	
	Calcutta Corporation	GFS* (3 months moving average)	of col. (2)	of col. (3)
(1)	(2)	(3)	(4)	(5)
January 1961	1.54	0.95	97.5	71.4
February "	1.27	1.30	80.4	97.7
March "	1.36	1.30	86.1	97.7
April "	1.29	1.30	81.0	97.7
May "	1.29	1.30	81.6	97.7
June "	1.32	1.40	83.5	105.3
July "	1.52	1.68	96.2	126.3
August "	1.88	1.40	119.0	105.3
September "	1.08	1.52	125.3	114.3
October "	2.14	1.43	135.4	107.6
November "	1.88	1.36	119.0	102.3
December "	1.40	1.01	94.3	75.9

*Birth rate per (000) couples has been expressed in terms of birth rate per (000) persons using a conversion factor 0.1203 being the proportion of native couples in 1962.

It may be recalled here that the present enquiry was completed in about two months time. A recall period ranging from one month to twelve months counted from this short period of enquiry will thus practically coincide with the calendar months of the year. The months of recall will be concentrated over one or two months only and the seasonal effect is not likely to be confounded.

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : Series B

But in surveys sampling over 'time', i.e., with sub-samples spread over the different months of the years, one has to ensure that the monthly sub-samples are even in size and that all the months in the calendar are represented equally. Since the monthly birth-rates are patterned, certain months having considerably larger birth rates than others, any such imbalance may produce appreciable bias in the estimated monthly birth rates and thus lead to wrong conclusions.

Paper issued : February, 1961.

Paper received : June, 1961.