S. Guha Roy

Parameterizing Indian Fertility Experience by Gompertz Function

ORIGINALLY developed as a description of mortality curve, the Gompertz function (Gompertz, 1825) can be conveniently restated as

$$F(x) = FA^{B^x},$$

$F(x)$ being the cumulative fertility rates to age x, F the total fertility, A and B the positive constants less than unity. This was later found to be a good representation of fertility distributions (Brass, 1980, 1981; Martin, 1967; Murphy and Nagnur, 1972; Wunsch, 1966). Besides F, the other two parameters A and B of the curve have explicit demographic interpretations. Thus, A is the proportion of total fertility (F) completed by age x_0 (origin) and B is related to the extent of peakedness of the fertility distribution. An important property of the Gompertz curve is that the age specific fertility rates, obtained as

$$\frac{d}{dx} F(x) = F \log_e A \log_e B(A^{B^{x-x_0}} \cdot B^{x-x_0})$$

(x_0 being a suitable origin in the age axis), increase monotonically to a maximum and then decrease monotonically, and by the age at which the maximum fertility is attained, exactly $1/e$ or about 37 per cent of F is completed irrespective of the values of F, A, B and x_0 (Titus, 1972). In order to absorb some irregularities in the data, the model uses cumulative fertility rather than age specific fertility.

The problems of estimation of the parameters have been investigated in several studies. Wunsch (1966) used the method of partial totals and Martin (1967) applied selected points method to fit the Gompertz function

$$F(x) = FA^{B^{(x-x_0)}}$$

Thus, taking d as the interval between two consecutive selected age points (x_0, x_1 and x_2), and the observed cumulative fertility rates $F(0)$, $F(1)$ and $F(2)$ at the corresponding selected points, the parameters were estimated as follows :

$$\hat{B} = [(\log_e F(2) - \log_e F(1))/(\log_e F(1) - \log_e F(0))]^{1/d}$$

$$\hat{A} = \text{Exp}\,[(\log_e F(1) - \log_e F(0))/(B^d - 1)]$$

$$\hat{F} = \text{Exp}\,[\log_e F(0) - \log_e A].$$

The method is, however, considered unsatisfactory as different sets of selected points give rise to widely varying estimates. The problem is reduced, but not completely eliminated by fitting a cumulative curve. Murphy and Nagnur (1972) used least squares fits by taking recourse to an iterative procedure. As many as four iterations were required to reach the convergence. Although the method provides a much better fit, it requires the use of computer facilities.

In order to simplify the comparison of the Gompertz model with observations required for exploratory analysis, Brass made double logarithm $[-\log_e(-\log_e)]$ linearity transformation of the function

$$F(x)/F = A^{B^x}$$

Thus,

$$\log_e [F(x)/F] = B^x \log_e A$$

$$\log_e [-\log_e F(x)/F] = \log_e (-B^x \log_e A)$$

$$-\log_e [-\log_e F(x)/F] = -\log_e (-\log_e A) - x \log_e B$$

$$\text{Colog}\,[\text{Colog}\, F(x)/F] = \text{Colog}\,(\text{Colog}\, A) - x\, \text{Colog}\, B$$

that is, $Y(x) = a + bx$

where $Y(x) = \text{Colog}\,[\text{Colog}\, F(x)/F]$

$a = \text{Colog}\,(\text{Colog}\, A)$ and

$b = \text{Colog}\, B, -\alpha < a < \alpha$ and $0 < b < \alpha$.

This equation has been found to be a good representation of the fertility distribution in the central age range, but the fit to observations at the early and late ages is not good enough. As in the Brass's logit model life table system in which the assumption of a linear relationship of a given life table, $l(t)$, to a common standard, $l_s(t)$, has proved successful, the fit of the Gompertz model at the tails of the fertility distribution can be greatly improved by relating similarly $Y(x)$ to a suitable standard, $Y_s(x)$, rather than to x. Thus, the modified Gompertz function is written as

$$Y(x) = \alpha + \beta Y_s(x)$$

where α and β measure location and scale of the transformed variable and

$$Y_s(x) = \text{Colog}\,[\text{Colog}\,F_s(x)/F_s].$$

The standard schedule, $Y_s(x)$, can be derived from a set of fertility rates to represent an average pattern. Such a standard has been developed (Booth, 1979) from a series of reported fertility distributions and Coale-Trussell fertility models (Coale and Trussell, 1974). If the modified Gompertz model were exact, any one reported fertility schedule could be taken as $Y_s(x)$. In reality, however, the model does not hold exactly, and an appropriate standard for a specific country can be developed as an average of the available series of fertility schedules so as to minimize the deviations of individual schedule from linearity.

A 'Relational' Gompertz Fertility Model [this is a relational system as it relates an observed $Y(x)$ schedule to the standard $Y_s(x)$ by a simple function] can be fitted to the mean parities, P_i (P_1 for age group 15-19, P_2 for 20-24, etc.), as well so that

$$Y(i) = \text{Colog}\,(\text{Colog}\,P_i/F) = \alpha + \beta Y_s(i)$$

In the application of the relational model, the total fertility, F, must either be reliably known or estimated. If P_{45-49}, that is, the number of children ever born as reported by mothers of ages 45-49, is accurate, it is then roughly taken as F. But experience suggests that the total fertility reported by older women is low owing to omission of births and probably selection bias. In such a case, we have a three-parameter system in which F, α and β have to be estimated.

Zaba (1981) has developed a simple and elegant method of fitting the model by separating the fertility pattern measured by α and β from the estimation of fertility level (F). Instead of using $F(x)/F$ and P_i/F, which require reliable information on total fertility, the ratios $F(x)/F(x+5)$ and P_i/P_{i+1} are used and assumed to follow a Gompertz form. Thus defining,

$$\begin{aligned} Z(i) &= \text{Colog}\,(\text{Colog}\,P_i/P_{i+1}) \\ &= \alpha - \log_e\,[\exp\,(-\beta Y_s(i)) - \exp\,(-\beta Y_s\,(i+1))] \end{aligned}$$

and expanding it by a Taylor series in the vicinity of the standard value of $\beta = 1$, one gets,

$$Z(i) - e(i) = \alpha + 0.48\,(\beta - 1)^2 + \beta\,g(i),$$

where $g(i)$, put for the result of differentiating $Z(i)$ with respect to β, and $e(i) = Z_s(i) - g(i)$ are calculated from the standard distribution.

The Relational Gompertz Model (RGM) is illustrated (Guha Roy, 1983, India, Office of the Registrar General, 1983) with the application to the Indian fertility data collected in the 1972 fertility survey and 1979 infant and child mortality survey of the Registrar General. The model, however, is more appropriate for the cohort data rather than cross sectional data on fertility considered here. We thus implicitly assume that fertility remained constant in the past, which it really did upto 1970, at least approximately. Moreover, the decline of fertility thereafter has been slow, and may not possibly affect significantly the estimation of total fertility from the average parities in the age group 15-35. Besides the difficulty of application of the model to cross sectional data, the major limitation of the collected information on fertility has been the error of response. There has been a tendency to understate the number of children ever born due to memory lapse and to omit the children born alive but now dead. We have found that the ratios of cumulated current fertility and mean parity (P/F) show a gradual decline with age, reflecting ommission of births by older women. Further, the discrepency between the value of P/F and the expected value of 1 at young ages indicates most probably a period reference error in the current fertility reports. The mis-statement of age further complicates the matter.

The relational system discussed above is believed to be well suited to deal with such situations. We use the synthetic measure of the cumulative current marital fertility rather than using the direct information on children ever born to ever married women. This is because a fit of the model to the latter data has been found to be much poorer than that obtained in the case of parity of the synthetic cohort. This synthetic measure ($\hat{P}_i$) is derived from the cumulated fertility, F_i (upto ages 19.5, 24.5, etc.), as follows :

$$P_i = F_i + W_i f_i$$

$$= 5 \sum_{j=0}^{i=1} f_j + W_i f_i,$$

where f_i is the observed age (i) specific marital fertility rate and W_i is a set of adjustment factors derived by Brass (Brass and Coale, 1968, Table 3.1, p. 94) to obtain values of adjusted cumulated current fertility corresponding to mid-point of 5-year age groups, that is, to 17.5, 22.5 etc.

The fitted Gompertz function is obtained for different population groups as follows :

1971-72

Rural : $Y(i) = 0.019 + 0.873\ Y_s(i)$

Urban : $Y(i) = 0.169 + 0.949\ Y_s(i)$

All India : $Y(i) = 0.053 + 0.894\ Y_s(i)$

1978

All India : $Y(i) = 0.179 + 0.984\ Y_s(i)$, where $Y(i)$ is as defined above, and $Y_s(i)$ is taken from Booth (1979). The development of an independent standard schedule, $Y_s(i)$, from the historical series of fertility measures for India does not seem to be appropriate as they lack uniformity in their quality. Moreover, some of the earlier series were derived from the civil registration system, and are thus totally unacceptable.

Table 1 presents the results of fitting the modified Gompertz function to period data for the Indian currently married women. The present estimate of total marital fertility (TMF) of 7.00 for 1971-72 is of the same order of magnitude as the estimates of 6.68 and 6.95 derived elsewhere by us by Brass's *P/F* correction and Coale's model of marital fertility respectively.

The efficiency of the fit of the Gompertz curve may be tested by estimating standardized errors. Choosing the age pattern of Indian females in 1971 Census, adjusted by a transitional age structure model (Guha Roy, 1984), as the standard population (S_i), we calculate the errors of fit as

$$\text{net error} = \sum_i (f_i - \hat{f}_i)\, S_i / \Sigma f_i S_i$$

$$\text{gross error} = \sum_i |f_i - \hat{f}_i|\, S_i / \Sigma f_i S_i$$

where f_i is the reported, and $\hat{f}_i$ the estimated marital fertility rates by age i. Thus, the errors for current period rates are

	Net error	*Gross error*
Rural India, 1971-72	0.0408	0.0691
Urban India, 1971-72	0.0314	0.1038
All India, 1971-72	0.0291	0.0782
All India, 1978	0.0094	0.0638

Given the criterion that the net error is most appropriate measure of fit for period fertility distribution (Murphy and Nagnur, 1972), the errors of less than one per cent to four per cent are not overwhelming in relation to the convenience of estimating age specific fertility by only three parameters of the Gompertz function. Though the estimate of the fertility rate for age group 15-19 is unstable when derived by the conventional *P/F* method, the Gompertz fit is good enough even at this age range.

The per cent change in the estimated marital fertility between 1971-72 and 1978 is as follows :

15-19	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*
−15.8	−13.7	−21.4	−27.7	−34.1	−42.8

TABLE 1—MODIFIED GOMPERTZ FERTILITY ESTIMATES (MODEL FITTED TO MEAN PARITIES OF MARRIED WOMEN 15-35)

	Age group of married women							*TMF*
	15-19	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*	*45-49*	
(a) Rural India : 1971-72								
Mean parity :								
synthetic	0.59	1.98	3.55	4.92	5.94	6 57	6 85	
model	0.59	2 01	3.55	4.94	6.11	6.95	7.26	
Marital fertility :								
reported	0.212	0.313	0.303	0.249	0.170	0.094	0 032	6.9
model	0 213	0.321	0.287	0.267	0.206	0.136	0.025	7.3
(b) Urban India : 1971-72								
Mean parity :								
synthetic	0.62	2.03	3.54	4.72	5.51	5.91	5.87	
model	0.61	2.06	3.50	4.70	5.61	6.19	6.38	
Marital fertility								
reported	0.221	0.313	0.284	0.201	0.124	0.052	0.016	5.9
model	0.216	0.333	0.246	0.230	0.147	0.089	0.016	6.4
(c) All India : 1971-72								
Mean parity :								
synthetic	0.59	1.99	3 55	4.88	5.85	6.44	6.70	
model	0.58	2.01	3.53	4.88	5.97	6.74	7.02	
Marital fertility :								
reportad	0.213	0 313	0.303	0.240	0.161	0.089	0.029	6.7
model	0.209	0.328	0.276	0.260	0.185	0.126	0.022	7.0
(d) All India : 1978								
Mean parity :								
synthetic	0.48	1.69	2.97	3.97	4.67	5.07	5.26	
model	0.48	1.71	2.96	3.97	4.72	5.19	5.33	
Marital fertility :								
reported	0.178	0.272	0.236	0.174	0.114	0.056	0.024	5.3
model	0.176	0.283	0.217	0.188	0.122	0.072	0 008	5.3

Consistent with the age pattern of practice of fertility control methods in India, the fertility declined most among married women of over 30 years of age, and least in the peak fertility age group 20-24. The lifetime fertility rate (TMFR) declined by 24 per cent from 7.0 to 5.3 during 1972-78. This estimate of TMFR agrees well with that obtained earlier by Coale's procedure for all India, 1971-72, but the agreement is not so good with that derived by the parity/fertility ratio method of Brass.

Perhaps it would be useful to fit a relational Gompertz model to mean parities of all women rather than married women only*. The results of fitting are shown in Table 2. Consideration of married women only, as done above, may create problems about interpreting cumulated age specific marital fertility rates since it does not properly correspond to any cohort. The total fertility rate for Taiwan in 1981, for example, was only 2.46, whereas the cumulated age specific marital fertility rates stood at 7.23. The reason put forward for this large difference is the multiple counting of first and low order births in cumulated age specific marital fertility rates.

TABLE 2—MODIFIED GOMPERTZ FERTILITY ESTIMATES (MODEL FITTED TO MEAN PARITIES OF ALL WOMEN AGED 15-35)

	Age group of women							*TFR*
	15-19	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*	*45-49*	
				(a) 1971-72				
Mean parity :								
synthetic	0.17	1.18	2.58	3.81	4.68	5.19	5.43	
model	0.17	1.19	2.58	3.83	4.75	5.28	5.43	
Fertility rates :								
reported	0.087	0.262	0.276	0.217	0.143	0.076	0.029	5.4
model	0.087	0.265	0.272	0.227	0.152	0.075	0.009	5.4
				(b) 1978				
Mean parity :								
synthetic	0.12	0.98	2.16	3.13	3.79	4.15	4.31	
model	0.12	0.98	2.15	3.12	3.78	4.12	4.20	
Fertility rates :								
reported	0.068	0.228	0.228	0.166	0.106	0.049	0.019	4.3
model	0.065	0.232	0.221	0.170	0.105	0.044	0.003	4.2

Indeed for India, the fit of the model to fertility of all women is better (except perhaps the age group 45-49) than that to marital fertility, but the disparity is nowhere near Taiwan's. Thus, the estimation of total fertility rates, being the

*This was suggested by Prof. K. S. Srikantan, Gokhale Institute of Politics and Economics, Pune in a personal communication to the author.

main focus of interest, is broadly consistent between married and all women. The level of this estimate is however somewhat lower than that derived elsewhere by us by Brass's procedure and also by Bongaart's model of proximate determinants of fertility (Jain and Adlakha, 1982), the implementation of the latter model in the case of India requiring several assumptions.

References

1. Booth, H., 1979, The estimation of fertility from incomplete Cohort data by means of the transformed Gompertz Model. *Ph.D. Thesis*, University of London.
2. Brass, W. and Coale, A. J., 1968, Methods of analysis and estimation. *In* : W. Brass, *et al.* (eds.), *The Demography of Tropical Africa*. Princeton : Princeton University Press, pp. 88-150.
3. Brass, W., 1980, Assessment of the validity of fertility trend estimates from maternity histories, *Regional Workshop on Techniques of Analysis of WFS Data*, ESCAP, Bangkok, pp. 50-57.
4. Brass, W., 1981, The use of the Gompertz relational model to estimate fertility, *International Population Conference : Manila*, Vol. 3, pp. 345-362, IUSSP, Liege.
5. Coale, A. J. and Trussell, T. J., 1974, Model fertility schedules : Variations in the age structure of childbearing in human populations, *Population Index*. 40, 185-258.
6. Gompertz, B., 1825, On the nature of the function expressive of the law of human mortality, *Philosophical Transactions of the Royal Society*, 115, 513-585.
7. Guha Roy, S., 1980, Population estimates for India (short version), *Population Index*, 46 (2).
8. Guha Roy, S., 1983, Parameterizing Indian fertility experience by Gompertz function. *Presented at the Ninth Annual Conference of the IASP*, Jaipur, December 22-26, 1983.
9. Guha Roy, S., 1984, Population estimates for India, *Demography India*, 13 (1 & 2), 168-192.
10. India, Office of the Registrar General, 1983, *Census of India 1981*. Occassional paper No. 1 of 1983.
11. Jain, A. K. and Adlakha, A. L., 1982, Preliminary estimates of fertility decline in India during the 1970s, *Population and Development Review*, 8 (3), 589-606.
12. Martin, M. D. P., 1967, Une application des fonctions de Gompertz a l'étude de la fécondité d'une cohorte, *Population*, 22, 1085-1096.
13. Murphy, E. M. and Nagnur, D. N., 1972, A Gompertz fit that fits : Application to Canadian fertility patterns, *Demography*, 9 (1), 35-48.
14. Titus, A. L., 1972, Some properties of the Gompertz function, *Demography*, 9(1), 48-50.
15. Wunsch, G., 1966, Courbes de Gompertz et perspectives de fécondite, *Rech. Econ. Louvain*, 32, 457-468.
16. Zaba, B., 1981, Use of the relational Gompertz Model in analysing fertility data collected in retrospective surveys, *CPS Working Paper*, Vol. 81, No. 2, London School of Hygiene and Tropical Medicine.