Special form of Gompertz curve for parametrising fertility experience of Indian female cohorts crossing childbearing age during 1993-2012

Samba Siva Rao Pasupuleti
Population Studies Unit, Indian Statistical Institute, 203 B T Road, Kolkata-700108. Tel: ++ 91-33-25753520
Email: srao113@gmail.com

Prasanta Pathak
Population studies unit, Indian Statistical Institute, 203 B T Road, Kolkata-700108. Tel: ++ 91-33-25753521 Email: prsnt@isical.ac.in

Abstract

The utility of mathematical models in understanding various demographic processes is very well known. In the present study the main aim is in providing a very satisfactory mathematical model for single year age fertility distribution. The basic objective of any modeling is to provide an alternative to data. Particularly we prefer those models where number of parameters are as less as possible and are interpretable in physical terms and are good enough to approximate all the relevant variations that are observable in the data. Keeping this in mind and drawing inspiration from Meyer et al. 1999, we have proposed here a special form of Gompertz curve which gives more insight into the problem and is helpful in providing a good platform for comparison of fertility experience of different cohorts across regions and over time horizon. Within the context of understanding fertility behavior, the current paper extends the basic literature by adding one more important dimension along with CTFR for better understanding of the problem. The proposed special form of Gompertz curve has been fitted to each female cohort who are crossing childbearing age in different calendar years during 1993-2012 in India to explain their fertility behavior. For the purpose of the current study we have used three National Family Health Survey data sets, NFHS-1(conducted during 1992-1993), NFHS-2 (conducted during 1998-1999), NFHS-3 (conducted during 2005-2006). The most important finding of the present study is that there is an estimated reduction of 1.4 children (i.e., from 5.55 to 4.15, meaning a 25.23% reduction) in completed fertility per female during the period 1993-2012. The other interesting findings are (1) the effective duration of childbearing age of a female i.e. the duration required to reach from 5% to 95% of saturation level is shrinking at the rate 2 months 16 days(i.e. 0.214 years) per year and (2) the age of giving birth to half of the children for a female has also decreased by 1.44 years(i.e., from 25.93 to 24.49) during the period 1993-2012.

Key words: CTFR, effective fertility period, paramerising fertility, fertility model, Special form of Gompertz curve.

Authors are thankful to Professor Anup Dewanji, Applied Statistics Unit, Indian Statistical Institute, Kolkata, Professor Asis Kumar Chattopadyay, Calcutta University, Professor Nityananda Sarkar, Economic Research Unit, Indian Statistical Institute, Kolkata, Dr. Arni Srinivasa Rao, Associate Professor, Population Studies Unit, Mr. Patho De, Scientist, Population Studies Unit, Dr. Sourabh Ghosh, Associate Professor, Human Genetics Unit, Indian Statistical Institute, Kolkata, for their comments and suggestions.

Introduction

For proper understanding of what changes have occurred in fertility behavior of female cohorts who are crossing childbearing age in the recent period(1993-2012) we need to compare various important characteristics of fertility behavior which can be derived from the fertility distributions of the corresponding cohorts. But we generally just compare cohort total fertility rates (CTFRs) of different cohorts for knowing the changes in this regard. This is just looking at the change in fertility behavior in one dimension only. Though this is the main dimension on which a policy maker is having an eye, there are also some other dimensions which need to be considered for better understanding of how the fertility behavior of female is changing. These new dimensions may also help the policy maker in providing some more clues that may help for better implementation of policies. The present work focuses on such other dimensions along with the main dimension CTFR. An easy way of finding out different characteristics of a fertility distribution is to build a suitable model for fertility distribution and derive various important characteristics of fertility distribution from the fitted model. While building any model we have to keep the following points in our mind. Models are alternative means for describing a given data. The basic objective of any modeling is to reduce confusing mass of numbers to a few intelligible basic parameters. Experts appreciate those models where number of parameters are as less as possible and are interpretable in physical terms and are also good enough to approximate all the relevant variations that are observable in the data. Over the years many researchers have tried and succeeded to model single year age fertility pattern. Following are the widely known functions for modeling single-year age specific fertility pattern.

I) *Gamma function (Hoem et al. ,1981) :-*

$$f(x)=R\left(\frac{1}{gamma(b)c^{b}}\right)(x-d)^{b-1}\exp(-(x-d)/c) \quad , \quad \text{for } x > d .$$

where f(x) is age specific fertility rate, d represents the lower age of childbearing, R indicates the level of fertility. Though b, c have no direct demographic interpretation Hoem et al. have substituted these by the mode m, the mean μ and the variance σ^2 in the following way .

$c = \mu - m$, $b = (\mu - d)/d = \sigma^2/c^2$

II) *Beta function (Hoem et al. ,1981):-*

$$f(x)= R\left(\frac{1}{beta(A,B)}\right)(\beta-\alpha)^{-(A+B-1)}(x-\alpha)^{A-1}(\beta-x)^{B-1} \qquad for \quad \alpha < x < \beta$$

The parameters are related to the mean μ and the variance σ^2 through the following relations

$$B = \left\{\frac{(\mu-\alpha)(\beta-\mu)}{\sigma^2}\right\}\left(\frac{\beta-\mu}{\beta-\alpha}\right) \quad \text{and} \quad A = B\left(\frac{\mu-\alpha}{\beta-\mu}\right)$$

where α is lower age limit of fertility, β is upper age limit of fertility . R determines the level of fertility.

III) *Hadweiger function (Hadwiger, 1940; Gilje, 1969):-*

$$f(x) = \frac{ab}{c}\left(\frac{c}{x}\right)^{\frac{3}{2}}\exp\left(-b^2\left(\frac{c}{x}+\frac{x}{c}-2\right)\right)$$

According to Chandola et al. (1999) the parameters may have the following demographic interpretation, the parameter a might be associated with the total fertility rate , the parameter c might be related to the mean age of motherhood , parameter b might be the height of the curve and the term $\frac{ab}{c}$ might be related to the maximum age specific fertility rate.

IV) *Normal curve (Peristera et al., 2007) :-*

$$f(x) = c\exp\left(-\left(\frac{x-\mu}{\sigma(x)}\right)^2\right)$$

. with $\sigma(x) = \begin{cases}\sigma_{11} & if \quad x \le \mu \\ \sigma_{12} & if \quad x > \mu\end{cases}$

Here c describes the basic level of the fertility curve and is associated with the total fertility rate. Additionally, μ gives modal age of age specific fertility, σ_{11} *and* σ_{12} are spreads of fertility distribution before and after its peak .

V) *Gompertz curve (Gompertz , 1825) :-*

$$f(t) = F(t+1) - F(t)$$

$$F(t) = Fa^{b^{(t-t_0)}} \quad , t = 15, 16, 17, \quad 49.$$

where $f(t)$ is the age specific fertility rate at age t and $F(t)$ is the average number of children born by exact age t. The parameter $F$is the saturation level (Cohort Total Fertility Rate), a is the proportion of total fertility attained by age t_0 , b is the intrinsic rate of growth of cumulative age specific fertility rate by age.

Wunch (1966), Martin (1967), Murphy and Nagnur (1972), Farid (1973), Brass, (1980, 1981); have suggested using Gompertz curve to model fertility distributions.

In the literature we also have many mixture models for modeling single year age specific fertility pattern some of such popular mixture models are given below.

VI) Mixer models:-

A) *Hadwiger mixer model(Chandola et al. ,1999) :-*

$$\text{f(x)} = m\frac{ab_1}{c_1}\left(\frac{c_1}{x}\right)^{\frac{3}{2}}\exp\left(-b_1{}^2\left(\frac{c_1}{x}+\frac{x}{c_1}-2\right) + (1-m)\frac{b_2}{c_2}\left(\frac{c_2}{x}\right)^{\frac{3}{2}}\exp\left(-b_2{}^2\left(\frac{c_2}{x}+\frac{x}{c_2}-2\right)\right.\right.$$

m is the mixture parameter that determines the relative sizes of two component fertility distributions .

According to the authors, parameter α is correlated with the total fertility rate , $c_1 and\ c_2$ are related respectively to the level and trend of the mean ages of births outside and inside marriage.

B) *Normal mixer model(peristera et al., 2007) :-*

$$f(x) = c_1 \exp\left(-\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) + c_2 \exp\left(-\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right)$$

The parameters c_1 & c_2 express the severity i.e. the total fertility rates of the first and the second hump respectively, μ_1 & μ_2 are related to the mean ages of two subpopulations the one with earlier fertility and the other with fertility at later ages. σ_1 , σ_2 reflect the variances of the two humps.

C) *Adjusted model of Peristera et al. model :-*

$$f(x) = c_1 \exp\left(-\left(\frac{x-\mu_1}{\sigma_1(x)}\right)^2\right) + c_2 \exp\left(-\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right)$$

$$\text{With} \quad \sigma_1(x) = f(x) = \begin{cases} \sigma_{11} & x \leq \mu_1 \\ \sigma_{12} & x \geq \mu_1 \end{cases}$$

The parameters have same meaning as above while σ_{11} & σ_{12} reflect respectively the spread of the first hump before and after its peak.

Further Scope and Proposal of a New Form of Gompertz Curve

Out of all the above models, the relative advantage of Gompertz model is that its parameter 'F' directly stands for completed fertility which is particularly a very important characteristic of the growth process(cumulative progression of births by age), in which particularly we are interested. Hence if we use Gompertz curve for modeling single year fertility distribution data of a cohort then from the value of F itself we can understand some part of the growth process like at what level average number births gets saturated. Though parameter b (called Gampertz growth rate) says how births progress by age but it is difficult to understand it instantaneously in physical terms. Hence, some modification is required at the level of parameter b in order to make Gompertz curve more useful. Drawing inspiration from Meyer et al. , (1999), we are proposing the following special form of Gompertz curve that helps in understanding the cumulative progression of births by age in a better manner.

The proposed model is

$$F(x) = Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{x-x_0}{b}}} \quad , \text{ with } F > 0 , 0 < a < 1 , b > 0$$

Here $F(x)$ is the cumulative fertility up to exact age x .

Let z_ξ be the exact age that is required to reach $100\xi\%$ of saturation level by the above growth process, where $\xi \epsilon (0,1)$.

Hence by definition z_ξ satisfies the following equation

$$Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{z_\xi - x_0}{b}}} = \xi F$$

This implies

$$a^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{z_\xi - x_0}{b}}} = \xi$$

Taking logarithm on both sides we get

$$\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{z_\xi - x_0}{b}} \log(a) = \log(\xi)$$

$$\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{z_\xi - x_0}{b}} = \left(\frac{\log(\xi)}{\log(a)}\right)$$

Once again taking logarithm on both sides we get

$$\frac{z_\xi - x_0}{b}\left(\log\left(\frac{\log(0.95)}{\log(0.05)}\right)\right) = \log\left(\frac{log(\xi)}{log(a)}\right)$$

This implies

$$z_\xi = x_0 + b\left(\frac{\log\left(\frac{\log(\xi)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)$$

The following results follow from the above result

i) the age of attaining half of the saturation level is $z_{0.5} = x_0 + b\left(\frac{\log\left(\frac{\log(0.5)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)$

ii) the age of attaining 5% of the saturation level is $z_{0.05} = x_0 + b\left(\frac{\log\left(\frac{\log(0.05)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)$

iii) the age of attaining 95% of the saturation level is $z_{0.95} = x_0 + b\left(\frac{\log\left(\frac{\log(0.95)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)$

from ii) and iii) it can be shown that $z_{0.95} - z_{0.05} = b$

Hence b can be safely interpreted as the length of age interval during which fertility level raises from 5% to 95% of saturation level. Here saturation level of a cohort, cohort total fertility rate and cohort completed fertility level they mean one and the same and are used interchangeably throughout this paper.

Here we have defined the *Period of effective fertility or effective fertility period* for a cohort as the age interval during which fertility level of that cohort reaches from 5% to 95% of saturation level.

i.e., *effective fertility period* is $(z_{0.05}, z_{0.95}) = \left(x_0 + b\left(\frac{\log\left(\frac{\log(0.05)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)\ ,\ x_0 + b\left(\frac{\log\left(\frac{\log(0.95)}{\log(a)}\right)}{\log\left(\frac{\log(0.95)}{\log(0.05)}\right)}\right)\right)$

Therefore the Gompertz model parameters are interpreted as follows,

F is saturation level(completed fertility level), a is proportion of total fertility attained by age x_0 ($where\ \ x_0$ is origin and generally we take it as 15 years) and b is the length of the *effective fertility period.*

As the model parameters are directly throwing light on the important characteristics of the fertility behavior and fit wide also this model is well comparable with other models which are used in this context and hence this form of Gompertz curve is very satisfactory model than other models to model single year age fertility distribution data.

It may be more appropriate to give some clarification about effective fertility period (EFP) and why we have paid some attention to it. We have defined effective fertility period for a cohort as the age interval in which most of the births (90%) occur, i.e. the age interval in which the fertility level of that cohort reaches from 5% to 95% of the saturation level (CTFR). For example, the uneducated females who are crossing childbearing age in the calendar year 2007 (these are the females who were born in the calendar year 1957 and who have entered into the childbearing age in the calendar year 1972) have the EFP (16.87,41.48), which means that, 5% (0.258 births) of the total births (5.16) were given by them by the age 16.87 years and 95% (4.9 births)of the total births were given by age 41.48 years and hence 90% of the total births i.e. 4.64 births are given in between 16.87 and 41.48 years of ages. EFP is not only an important characteristic of fertility behavior but it also gives some clues about the impact of some of the intermediate variables on fertility. For example lower limit of EFP gives some clue regarding the nuptial pattern and upper limit of EFP gives clue regarding the practice of birth control at the late ages (sterilization) of the females of that cohort. Furthermore, for a cohort with characteristic A and crossing childbearing age in the calendar year T if we know CTFR (F_A^T), EFP (which is denoted by($\eta_{0.05}^T(A)$, $\eta_{0.95}^T(A)$)) then b_A^T and a_A^T can be computed from the following relations.

$$b_A^T = \eta_{0.95}^T - \eta_{0.05}^T$$

$$a_A^T = exp\left(\frac{(\ \log(0.95)}{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{(\eta_{0.95}^T - 15)}{b}}}\right)$$

Rest of the characteristics of fertility behavior in which one is interested in can be easily found out from these values and using special form of Gompertz function given in the methodology.

Data

As a part of global demographic health surveys, India have started collecting data through it's nationally conducted survey so called National Family health Survey. NFHS-1 was conducted in India during 1992-93. A sample of 88,562 households and 89,777 ever-married women in the age group 13-49 was collected from 24 states and the then National Capital Territory of Delhi , which is now a separate state. The objective of the survey was to collect reliable and up-to-date information on fertility, family planning, mortality, and maternal and child health. The second National Family Health Survey (NFHS-2) was conducted in 1998-99. The survey covers a national wide representative sample of 90,303 ever-married women in the age group 15-49. The third National Family Health Survey (NFHS-3) was conducted during 2005-06. The survey covers a sample of more than 2,30,000 women in the age group 15-49 and men in the age group 15-54.

Formation of cohorts from NFHS data sets

In order to understand how the fertility behavior of females in India has been changing in the recent period we have made use of selected portions of NFHS-1, NFHS-2, NFHS-3 data sets. Let us assume that we have two females in our NFHS-1 survey who were born on 1st February 1945. We know that NFHS-1 was conducted during April 1992 – September 1993. Suppose the first female was interviewed on 1st May 1992 the second female was interviewed on 1st August 1993. Since the data on maternity history and other related information were collected from each surveyed female retrospectively up to the date of survey, we had that information on the first female till her age of 47 years and 3 months. on the other hand the corresponding information on the second female was relevant up to her age of 48 years and 6 months. Though both the females belong to the same birth cohort (birth cohort of 1945) we however are having information from them for different durations. Non-homogeneous information of this kind is rather undesirable when we compare fertility performance of one female cohort with another female cohort. To overcome this kind of problem, we consider in each survey the information on each female only up to a base-line time point of January 1 of the year of starting each survey. Thus, we consider for each surveyed woman the age (in completed years) at the base-line time point of the corresponding survey year. Simultaneously, we consider only the events that had taken place to each surveyed female by her completed years of age , i.e. by the corresponding base-line time point only.

So for all female respondents in any of the NFHS rounds, we calculate the age of each female by 1st January of the year of initiating the survey(which is considered as the base-line time point of a NFHS round) and ages of her at first , second, etc. births till the last birth. Thus we find the number of births to her by the exact ages 15,16,17 and so on up to her age (in completed years) at the base-line time point of a NFHS round.

All the respondents who were in NFHS-1 and whose age was 48 years (in completed years) at the base- line time point of the survey (i.e. 1st January 1992) were born in the calendar year exactly 48 years back from 1992 , i.e., these females belonged to the 1943 birth cohort. The females of this cohort would cross childbearing age in the calendar year 1993 and hence by 1st January 1994,these women were out of the childbearing age interval of 15 to 49 years. We call this cohort as a cohort crossing childbearing age during 1993 (symbolically, CC-1993). Similarly, the females who were in NFHS-1 and whose age was 43 years (in completed years) at the base-line time point of the survey (i.e. at 1st January 1992) were born in the calendar year exactly 43 years back from 1992, i.e. these females are in the 1948 birth cohort. This cohort would cross childbearing age in the calendar year 1998 and by 1st January 1999 they would be out of the childbearing age interval of 15 to 49 years. This cohort is called a cohort crossing childbearing age in 1998 (symbolically, CC-1998).

The females of 1949 birth cohort were covered in both NFHS-1(i.e. the respondents whose age had been 42 years at the base line time point of NFHS-1) and NFHS-2 (i.e. the respondents whose age had been 48 years at the base-line time point of NFHS-2). But the respondents of age 48 years at the base line time point of NFHS-2 survey provide more information on their fertility than the respondents of age is 42 years at the base-line time point of NFHS-1. So, the former group of respondents have been considered for our analysis while taking into account the females who have crossed their childbearing age in 1999. They are symbolically denoted as CC-1999. Similarly, we have labeled several birth cohorts as cohorts crossing childbearing age in the following way.

For respondents in the survey	For respondents having following age	The exact age up to which birth performance	Years of birth performance information that is lacking	Year of birth of corresponding respondents	Year of crossing childbearing age	Label used for the respondents

	(age at last birth day)at the base-line time point of corresponding NFHS round	history was considered for the present study K_T	in order to have Complete birth performance history			
NFHS-1	48	48	2	1943	1993	CC-1993
NFHS-1	47	47	3	1944	1994	CC-1994
NFHS-1	46	46	4	1945	1995	CC-1995
NFHS-1	45	45	5	1946	1996	CC-1996
NFHS-1	44	44	6	1947	1997	CC-1997
NFHS-1	43	43	7	1948	1998	CC-1998
NFHS-2	48	48	2	1949	1999	CC-1999
NFHS-2	47	47	3	1950	2000	CC-2000
NFHS-2	46	46	4	1951	2001	CC-2001
NFHS-2	45	45	5	1952	2002	CC-2002
NFHS-2	44	44	6	1953	2003	CC-2003
NFHS-2	43	43	7	1954	2004	CC-2004
NFHS-2	42	42	8	1955	2005	CC-2005
NFHS-3	48	48	2	1956	2006	CC-2006
NFHS-3	47	47	3	1957	2007	CC-2007
NFHS-3	46	46	4	1958	2008	CC-2008
NFHS-3	45	45	5	1959	2009	CC-2009
NFHS-3	44	44	6	1960	2010	CC-2010
NFHS-3	43	43	7	1961	2011	CC-2011
NFHS-3	42	42	8	1962	2012	CC-2012

For the women who were crossing childbearing age in the year 1993 the percentage of births between the exact ages 42-48 to the total number of births up to the exact age 48 is 1.47%. If births after the exact age of 48 years are negligible (the general convention is that births after the exact age of 50 years are negligible) then we may treat exact age of 42 years as 98.5th fertility percentile and over time age 42 (based on the information from the cohorts who are crossing childbearing age during 1999, 2006) raised more than 99th percentile with the above concept. Even if we restrict our analysis up to the exact age of 42 years, we can study 98.5% of fertility behavior of the women. However, we had used the full information that is available up to the reference period of each survey.

Understanding fertility experience of Indian female cohorts crossing childbearing age during 1993-2012 :-

In order to understand the fertility experience of Indian female cohorts who are crossing childbearing age in different calendar years during 1993-2012, we have built separate models of the form $y_t^T = Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{b}}}$ for each female cohort crossing childbearing age in the said time span.

In the models above, y_t^T is average number of children born to the women cohort who are crossing childbearing age in the calendar year T by the time when each individual member of the cohort reaches an exact age of t years of their lives.

Criteria of estimating parameters in the model $y_t^T = Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{b}}}$ is finding such parameter estimates for F , a , b so that

$$\sum_{t=1}^{K_T}\left(y_t^T - Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{b}}} \right)^2 \quad \text{is minimum.}$$

After fitting above models to each cohort separately and on observing the trends in the parameters F , a and b what we have observed is that all the three parameter values are linearly decreasing over time. This has motivated us to build a more general model of the form

$$y_t^T = (F_1 + F_2\,(T-1993))(a_1 + a_2(T-1993))^{\left(\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{(b_1+b_2(T-1993))}}\right)}$$

Where $F_1$is CTFR of female cohort who has crossed childbearing age in the calendar year 1993, F_2 is the rate at which CTFR is falling over the cohorts who have crossed childbearing age during 1993-2012. a_1 is proportion of risky births (we considered births by age 15 is risky as infant mortality and maternal mortality is more associated with such births), a_2 is the rate at which proportion risky births has changed in subsequent cohorts. $b_1$is the length of effective fertility period for the females of CC-1993. b_2 is the rate at which EFP is shrinking over cohorts starting from CC-1993 to CC-2012. Just above six parameter model is enough to understand the fertility experience of various cohorts who are crossing childbearing age in different calendar years in the said time span.

Criteria of estimating parameters in the general model is

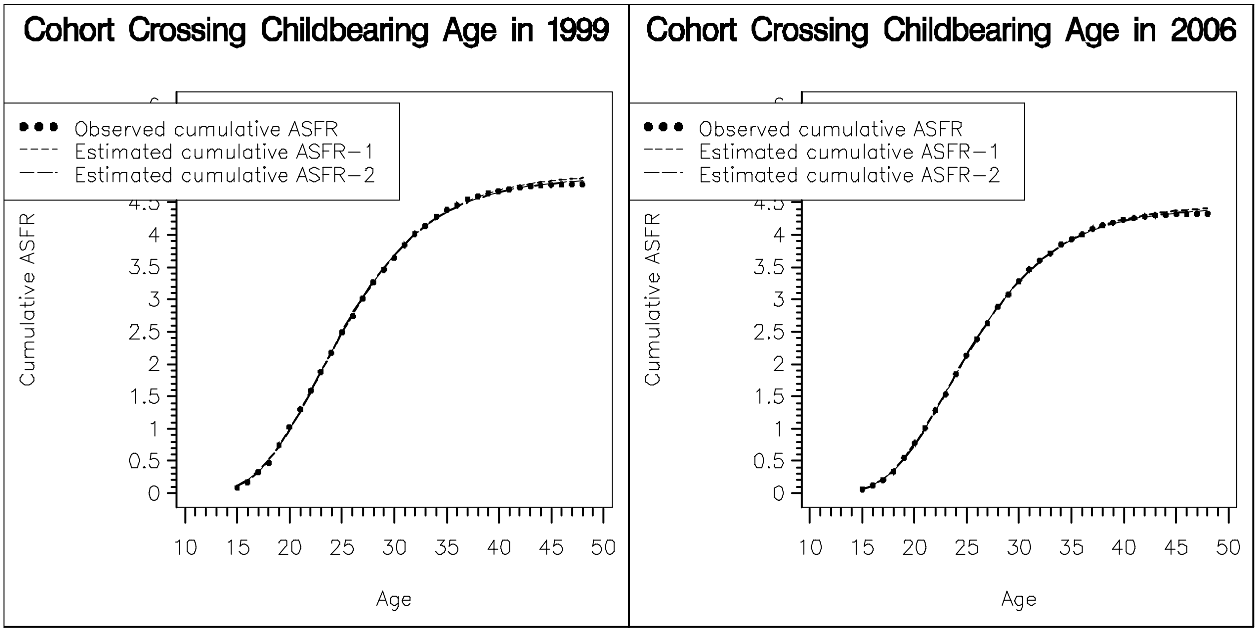
finding such parameter estimates for F_1, F_2 , a_1 , a_2 , b_1, b_2 so that

$$\sum_{T=1993}^{2012}\left(\sum_{t=1}^{K_T}\ (y_t^T - (F_1 + F_2(T-1993))\left((a_1 + a_2(T-1993))^{\left(\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{(b_1+b_2(T-1993))}}\right)} \right))^2 \right) \text{ is minimum.}$$

Here K_T is the exact age up to which the maternity history of the female cohort crossing childbearing age in the calendar year T was considered, T=1993,1994,……2012.

The pattern in which the average number of births progress by age for any cohort (meaning the progression of the cumulative ASFR) looks like a stretched 'S' curve and the Gompertz curve is a very good fit to it (see Figure-1 in Appendix). The important empirical observation that is the pivotal for the present study is that for a female cohort for whom maternity history is known up to the exact age of 48 years, if we estimate the characteristics of fertility behavior of that cohort separately by using the information (i) up to the exact age of 48 years (ii) up to the exact age of 42 years by fitting Gompertz curve to cumulative ASFR then the corresponding estimates of characteristics of fertility behavior are almost same. This fact can be observed from Figure-2.

Figure-2: *Observed and estimated cumulative ASFR for the female cohorts who are going to cross childbearing age in 1999, 2006 in India*



Estimated cumulative ASFR-1: Estimated cumulative age specific fertility rate based on the information up to the exact age of 42 years

Estimated cumulative ASFR-2: Estimated cumulative age specific fertility rate based on the information up to the exact age of 48 years

Here we had compared fits because if the fits are close enough then the corresponding estimates of characteristics of fertility behavior were also expected to be close. If we use the maternity history up to the exact ages of 43, 44, 45, 46, 47 years separately and derive characteristics of fertility behavior of the same cohort in a similar manner by fitting Gompertz curve, then the estimates of resulting characteristics of fertility behavior were also close to those of characteristics of fertility behavior derived based on the maternity history up to the exact ages of 42, 48 years. This is because for all cohorts most of the births are occurring by exact age of 42 years and cumulative ASFR curve which looks like a Gompertz curve (as cumulative ASFR is very closely following Gompertz law) is getting saturated at about age 42 years. So the lesson we have to take from this empirical observation is that for any female cohort for whom maternity history is known at least up to the exact age of 42 years, then we can estimate the characteristics of fertility behavior of that cohort with reasonable accuracy without bothering about the maternity history after age 42 years. For example, in order to estimate the characteristics of fertility behavior of the female cohort who are going to cross childbearing age in India in the calendar year 2011 (females whose age is 43 years at 1st January 2005 and who are in NFHS-3) we have used above result.

Analysis and Results

Understanding fertility experience of a particular cohort means understanding such aspects as the following. 1) How does the fertility level change over different ages for a cohort? 2) where does the average number of births converge ultimately to reach a saturation level with increasing ages for a cohort? 3) what is the age interval for a cohort within which most (say 90%) of the births occur ? etc.

For each cohort who are crossing childbearing age during calendar year T and whose age is A_T at the base-line time point of corresponding survey we have calculated average number of children born to them by exact ages 15, 16, 17, ….. up to A_T , T=1993,1993, 1994, ….2012.

From Figure-1 and Figure-2 it becomes very clear how the age pattern of fertility is changing and how average number of births progress by age for different cohorts who are crossing childbearing age during the calendar years 1993 ,1998, 2003, 2008, 2012 in India and hence gives us a glimpse of how the fertility behavior of the women is changing over cohorts who are crossing childbearing age during various calendar years in between 1993-2012 in India.

Figure-1: Change in age pattern of fertility over different female cohorts crossing childbearing age during 1993-2012 in India.

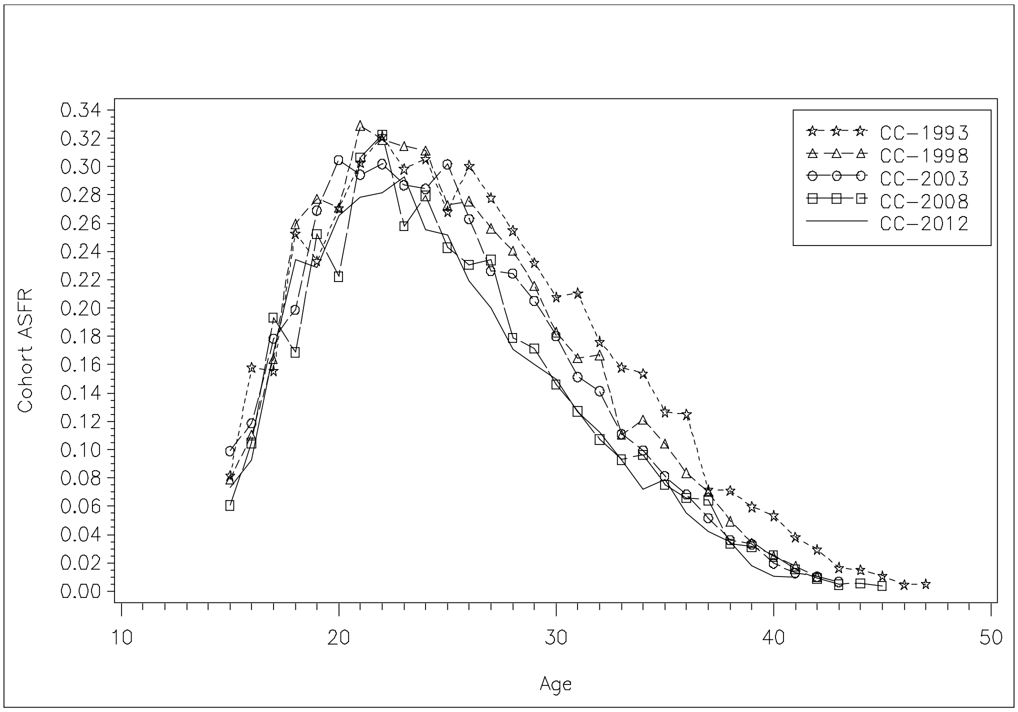
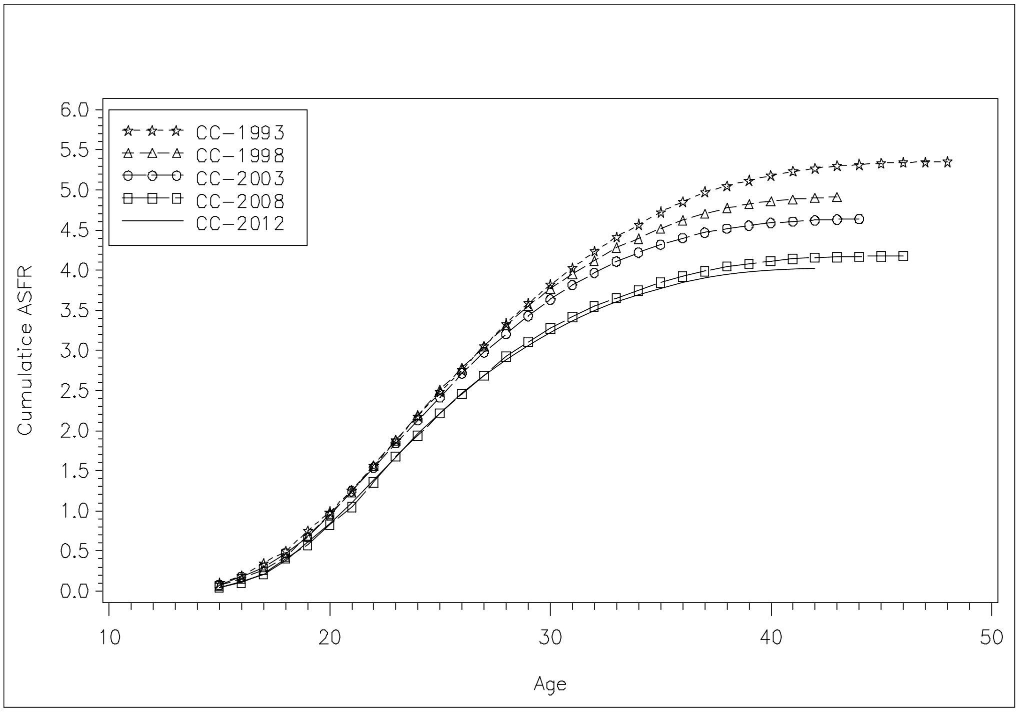


Figure-2: Change in cumulative age pattern of fertility over different female cohorts crossing childbearing age during 1993-2012 in India



The nature in which average number of births progresses by age for each cohort suggests that the Gompertz curve can be used to model progression of average number of births by age. We have fitted Gompertz curve of the form $y_t = Fa^{\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{b}}}$ to each female cohort crossing childbearing age during the calendar years 1993-2012 and derived some of the important characteristics of fertility behavior from the fitted model. The results are as as shown in the following Table-1.

Table-1: Some characteristics fertility behavior of Indian female cohorts who are crossing childbearing age in different calendar years during 1993-2012 as derived by using special form of the Gompertz curve.

Calendar year when a Cohort crossing childbearing age	Sample size used to estimate parameters	Estimate of average no. of children born to a cohort by the time it crosses childbearing age (F)	Proportion of total fertility attained by age 15 (a)	Estimate of age of attaining half of the completed fertility	Length of *effective fertility period* (b)	*Period of effective fertility/ Effective fertility period*
1993	1168	5.55	0.02177	25.93	26.01	(16.57,42.58)
1994	1561	5.35	0.01677	25.71	24.54	(16.88,41.42)
1995	1565	5.43	0.02303	25.82	25.98	(16.47,42.45)
1996	1635	5.33	0.02226	25.50	25.08	(16.47,41.55)

1997	2270	5.06	0.02099	25.22	24.19	(16.51,40.70)
1998	1767	5.12	0.01786	25.19	23.57	(16.71,40.28)
1999	812	4.90	0.02188	24.84	23.45	(16.40,39.85)
2000	1640	4.79	0.01666	24.90	22.67	(16.74,39.41)
2001	1523	4.88	0.02094	24.63	22.79	(16.43,39.21)
2002	1708	4.79	0.01895	24.73	22.70	(16.57,39.26)
2003	2214	4.81	0.01863	24.85	22.90	(16.60,39.51)
2004	1991	4.62	0.02071	24.43	22.28	(16.41,38.69)
2005	1876	4.65	0.02105	24.20	21.80	(16.36,38.16)
2006	834	4.43	0.01293	25.27	22.75	(17.08,39.84)
2007	1677	4.32	0.01010	25.17	21.88	(17.30,39.18)
2008	1768	4.27	0.01534	24.73	22.04	(16.80,38.85)
2009	1802	4.19	0.01395	24.65	21.59	(16.88,38.47)
2010	2825	4.22	0.01792	24.66	22.35	(16.62,38.97)
2011	2247	4.14	0.01630	24.44	21.56	(16.68,38.25)
2012	2157	4.15	0.01660	24.49	21.73	(16.67,38.40)

From Table-1 it is very clear that the estimate of completed fertility for a cohort who have crossed childbearing age in the calendar year 1993(CC-1993) is 5.55 whereas the same estimate for the females of CC-2012 is 4.15. So, the completed fertility has decreased by 1.4 number of children per woman in between 1993-2012. Other interesting findings are (i) the age of attaining half of the children has decreased by 1.44 years(i.e., it has decreased from 25.93 to 24.49) and (ii) *effective fertility period* has been shrinking at the rate of 2 months 16 days(i.e. 0.214 years) per year during the period 1993-2012.

As all the three parameters namely, F, a and b are changing linearly over cohorts(as evident from graphs drawn for parameters F, a ,b over cohorts based on Table-1), we have build a six parameter Gompertz model of the form $y_t^T = (F_1 + F_2\,(T-1993))(a_1 + a_2(T-1993))^{\left(\left(\frac{\log(0.95)}{\log(0.05)}\right)^{\frac{t-t_0}{(b_1+b_2(T-1993))}}\right)}$ to explain fertility behavior of all the female cohorts crossing childbearing age during 1993-2012.

The parameter estimates of the above model are

Table-2: Parameter estimates of the six parameter special form of Gompertz model.

Parameter	Estimate
F1	5.467005
F2	-0.07709
a1	0.021688
a2	-0.00041
b1	25.16112
b2	-0.23353

From the parameter estimates of above table we can understand that CCFR of cohort that have crossed childbearing age in 1993 is 5.467005 and CCFR has fallen over subsequent cohorts at a rate of 0.07709 per year during 1993-2012. Proportion of risky births has decreased with a rate of 0.00041 per year in between 1993-2012 starting from 0.021688 at 1993. Length of EFP has shrinked at a rate of 2 months 24 days over subsequent cohorts during 1993-2012 starting from an EFP of (16.52,41.68) at 1993. See the figures in Appendix under the side heading of Individual model fits (individual Gompertz model fits) and General model fit (six parameter Gompertz fit) to visualize how good the model fit.

Parameters of Table-2 are approximately similar to the same story that was being told with the help of Table-1.

Comparison of the special form of Gompertz model, six parameter Gompertz model with other models

As far as the degree of fit is concerned there is absolutely no difference between the special form of Gompertz curve and the original form of Gompertz curve. The difference lies only in parameter interpretation and that too only in respect of interpretation of parameter b.

We have also fitted other prominent models which are used in this context, namely, Hadwiger model ,Gamma model, Normal curve model, Logistic model and so on to the data on fertility distribution and a comparison of the fits in terms of 1000*(Error sum of squares) are shown in below Table-3.

Table-3 Comparison of special form of Gompertz curve as well as the six parameter Gompertz curve with other models used in this context.

Cohort crossing childbearing age during the calendar year	Hadweiger model	Hadweiger mixture model	Peristera model	Peristera mixture model	Gamma model	Special form of Gompertz model	Special form of six parameter Gompertz model	Logistic model
1993	7.7568	5.622	13.2596	5.9074	7.3111	9.6007	11.293	35.6082
1994	7.5855	8.1427	10.6358	5.6318	6.7365	10.8176	10.5745	32.3969
1995	10.1892	8.1585	14.3677	8.3378	9.705	13.0349	14.9615	41.4611
1996	6.7293	5.7679	11.7292	5.6466	6.4356	8.5892	9.235	34.4071
1997	5.2403	7.1477	11.0111	3.1303	4.9462	6.5454	6.6033	30.781
1998	6.7475	4.4347	18.0293	5.0571	5.4011	6.8516	7.231	38.5632
1999	12.0335	15.8236	19.6686	11.3762	11.5299	13.2929	14.8107	40.5045
2000	4.8904	6.1323	9.3189	4.5713	4.6088	7.1392	6.2321	23.8086
2001	3.9928	4.9068	9.9195	3.7018	3.8496	5.5138	7.7665	28.2862
2002	3.3717	4.8942	7.773	3.136	3.1768	5.9412	7.0596	29.1755
2003	3.7158	6.9817	9.8245	2.8207	3.5775	4.8481	5.3536	26.7949
2004	3.7836	5.2676	11.0936	4.0228	3.5242	4.2003	6.4857	25.7853
2005	4.5435	7.0989	9.7668	5.321	4.3384	6.1549	11.9499	27.0983
2006	7.7556	7.6131	15.8801	8.2175	7.4047	8.2585	12.9682	27.7307
2007	7.1151	6.9687	18.3333	6.2408	6.1488	6.2921	11.0894	29.3047
2008	10.0217	11.81	20.026	9.003	9.1566	9.5379	9.7824	33.9668
2009	9.3706	10.0445	20.5896	7.0762	8.1233	8.0686	8.1076	33.8672
2010	3.7651	4.823	11.2675	3.1838	3.549	3.6711	4.962	22.6507
2011	7.2992	8.5685	16.6366	5.3128	6.2751	6.6341	6.8714	30.2067
2012	4.3558	5.7243	13.7388	3.0326	3.0423	3.8719	4.2722	28.3139

Comparison of the special form of Gompertz model as well as the six parameter Gompertz model with other models for cohorts crossing childbearing age in 1993, 1999, 2006 are shown graphically in the Appendix through Figure-4, Figure-5, Figure-6 (For these cohorts fertility information is available up to exact age 48 years i.e., age specific fertility rate data is available up to age 47 years).

From Table-3, it is clear that the Gompertz model fits well and is well comparable with the other models used in this context.

Conclusion :-

We have built a special form of the Gompertz model for each female cohort who are crossing childbearing age during 1993-2012 to understand their fertility experience. The most important finding of the present study is that there is an estimated reduction of 1.4 children (i.e., a fall from 5.55 to 4.15 or a 25.23% fall) in the completed fertility per female during the period 1993-2012. The other interesting findings are (1) the effective duration of childbearing age(i.e. the duration required to reach from 5% to 95% of the saturation level or the completed fertility level)has been shrinking at the rate of 2 months 16 days (i.e. 0.214 years) per year and the age of attaining half of the children has also decreased by 1.44 years(i.e., the age has decreased from 25.93 to 24.49 years)during the period 1993-2012. In the context of understanding the fertility experience of the Indian female cohorts, the proposed special form of Gompertz model fits well and is very comparable with other existing models (like peristera et.al. model and Hadwiger model, Gamma model etc.). The best part of the proposed model is that all the parameters are interpretable and throws light on the characteristics of growth process in which we are generally interested in. Thus, this model helps us in having a good platform for comparing the fertility performances of different cohorts across regions and over time horizon.

References :-

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.

Brass, W., 1981, The use of the Gompertz relational model to estimate fertility, *International Population Conference : Manila*, Vol. 3, pp. 345-362, IUSSP, Liege.

Chandola. T., Coleman, D. A., Horns, R. W. (1999). Recent European fertility patterns: fitting curves to ‘distorted’ distributions. *Population Studies*, 53, 3: 317-329.

Farid, S. M. (1973). “On the pattern of cohort fertility.”, *Population Studies*, 27: 159-168.

Gilje, E. (1969). Fitting curves to age-specific fertility rates: some examples. Statistical Review of the Swedish National Central Bureau of Statistics III, 7:118-134.

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.

Hadwiger, H. (1940). “Eine analytische reprodutions-funktion fur biologische Gesamtheiten.”, *Skandinavisk Aktuarietidskrift*, 23, 101-113.

Hoem, J. M., Madsen, D., Nielsen, J. L., Ohlsen, E., Hansen, H. O., Rennermalm, B. (1981). “Experiments in modelling recent Danish fertility curves.”, *Demography*, 18: 231-244.

Martin, M. D. P., 1967, Uapplication des fonctions de Gompertz a l`etude de la fecondite d`une cohort, population, 22, 1085-1096.

Meyer, Perrin S; Jason W. Yung and Jesse H. Ausbel(1999). ‘ A primer on Logistic Growth and Substitution: The Mathematics of the Loglet Lab Software. *Technical forecasting and social change, Vol.61, No.3, pp. 247-271.*

Murphy. E, M. and Nagnur, D. N., 1972, A Gompertz fit that fits : Application to
Canadian fertility patterns, *Demography*, 9 (1), 35-48.

Paraskevi Peristera and Anastasia Kostaki(2007). ‘Modeling fertility in modern populations’. *Demographic Research. Volume 16, article 6, pages 141-194.*

Wunsch, G., 1966, Courbes de Gompertz et perspectives de fecondite, Rech. Econ. Louvain, Sl, 457-468.

Appendix:-

Comparison of the special form Gompertz model fit with several other models which are commonly used in this contest are shown in the next few pages.

Figure-3: Comparing the fit of special form of Gompertz curve with other models for cohort crossing childbearing age during 1993

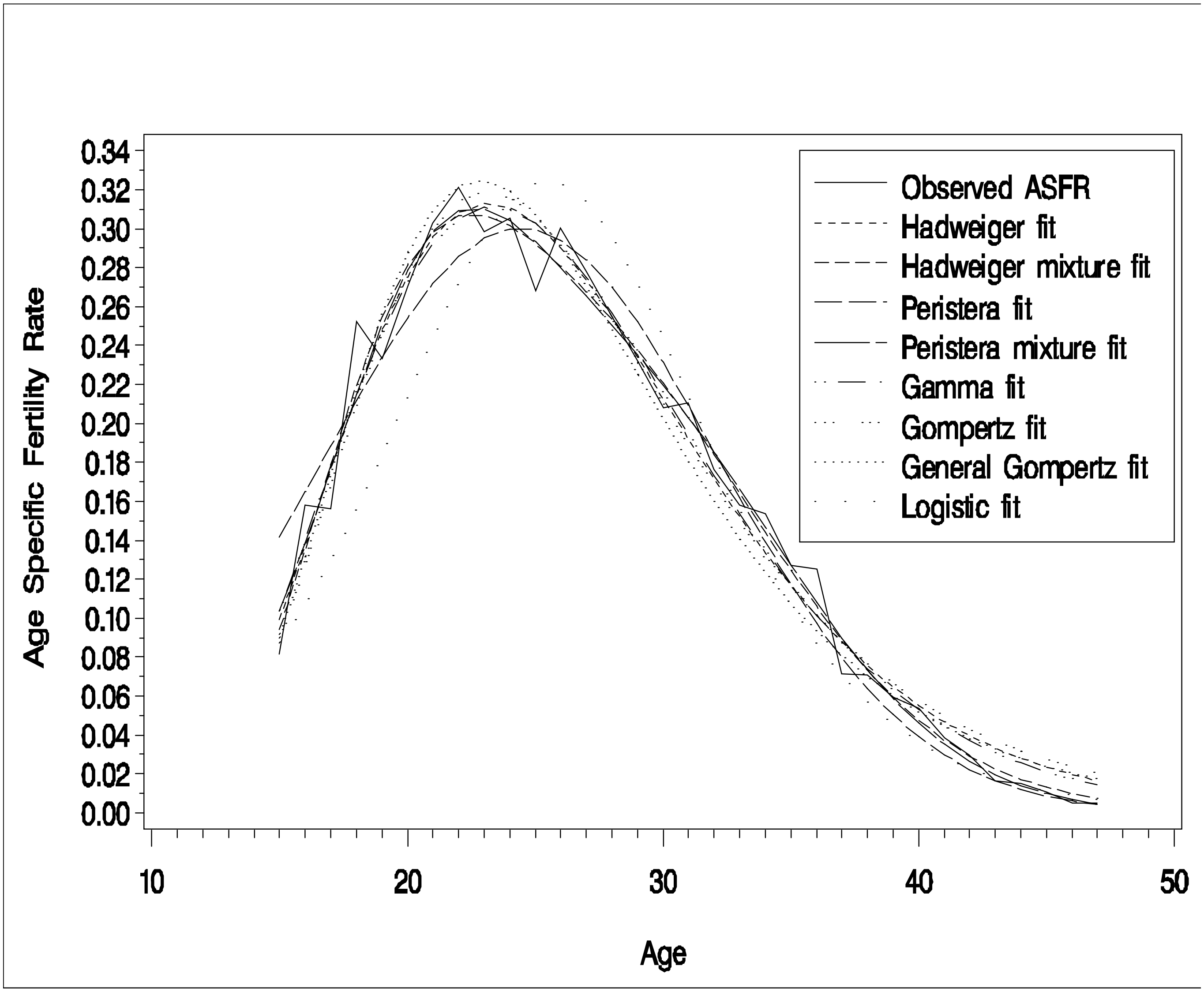


Figure-4: Comparing the fit of special form of Gompertz curve with other models for cohort crossing childbearing age during 1999

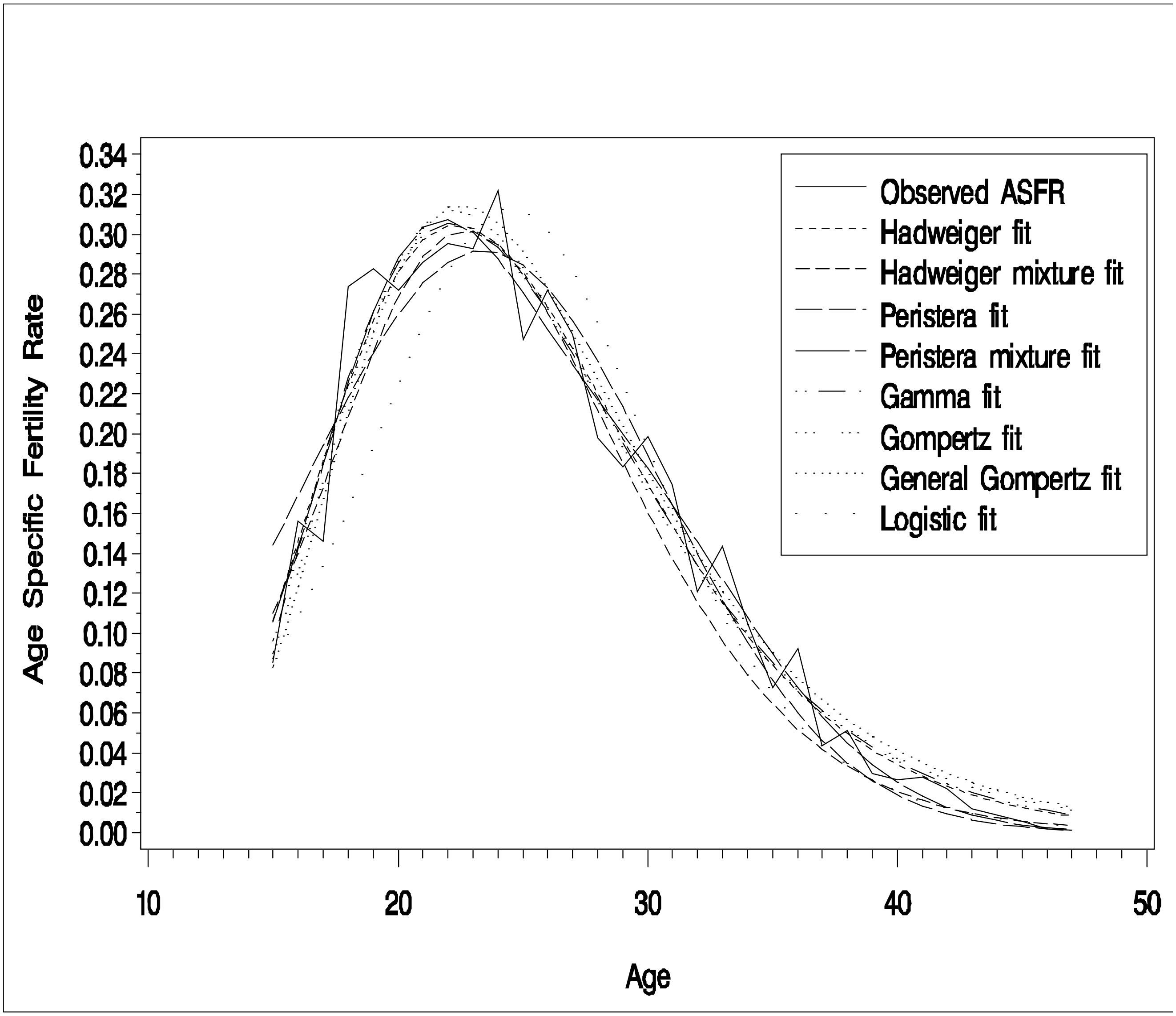
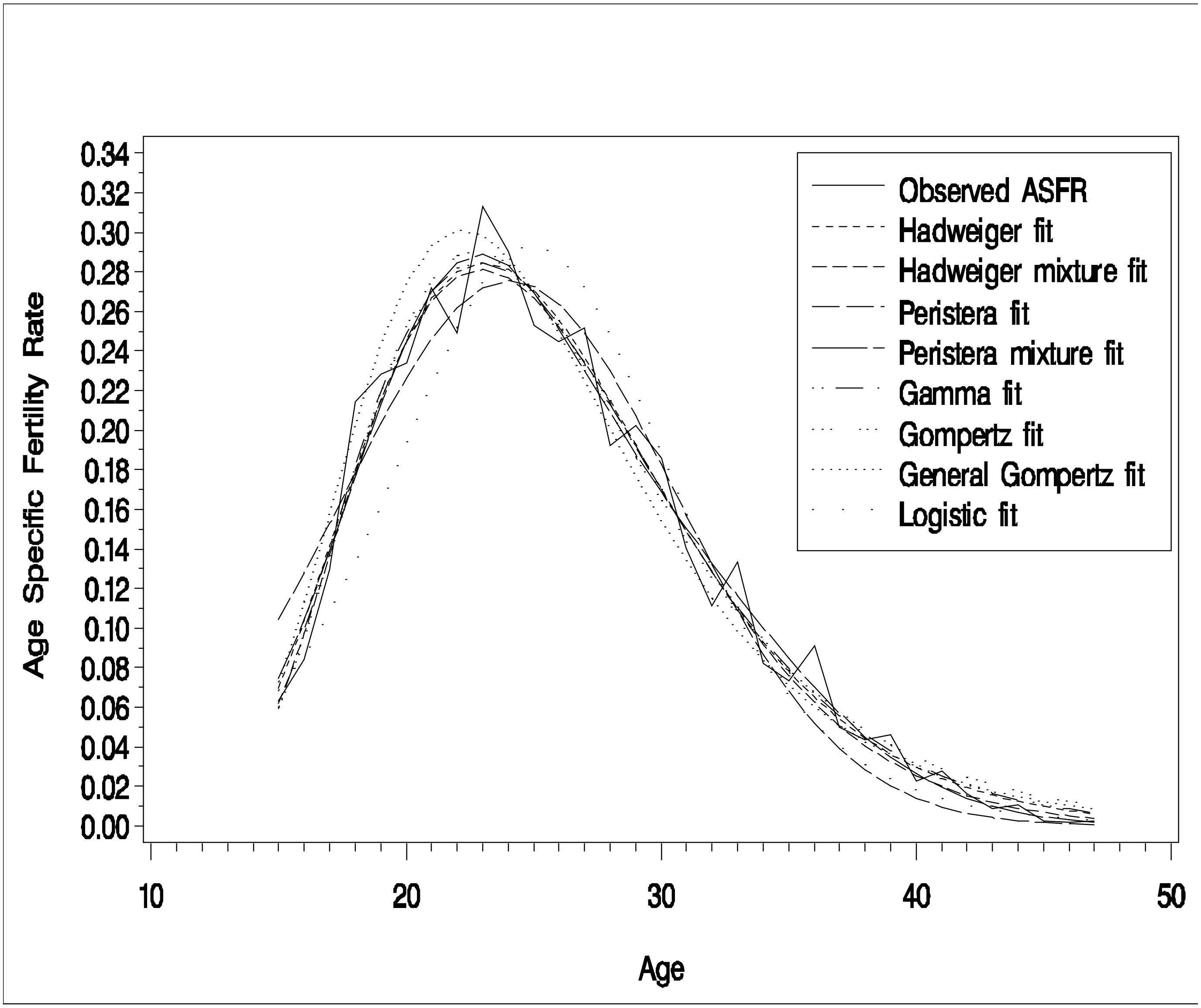
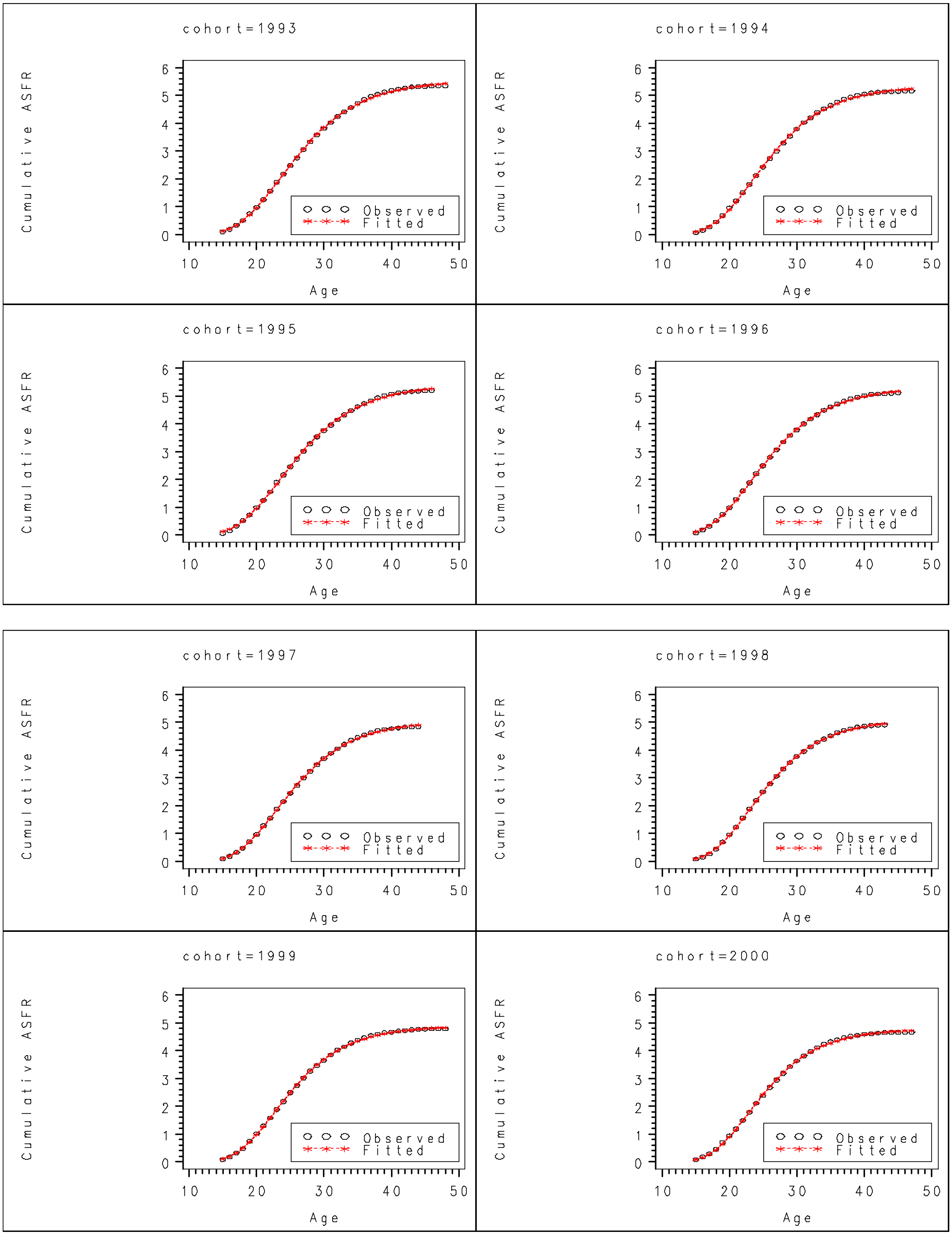
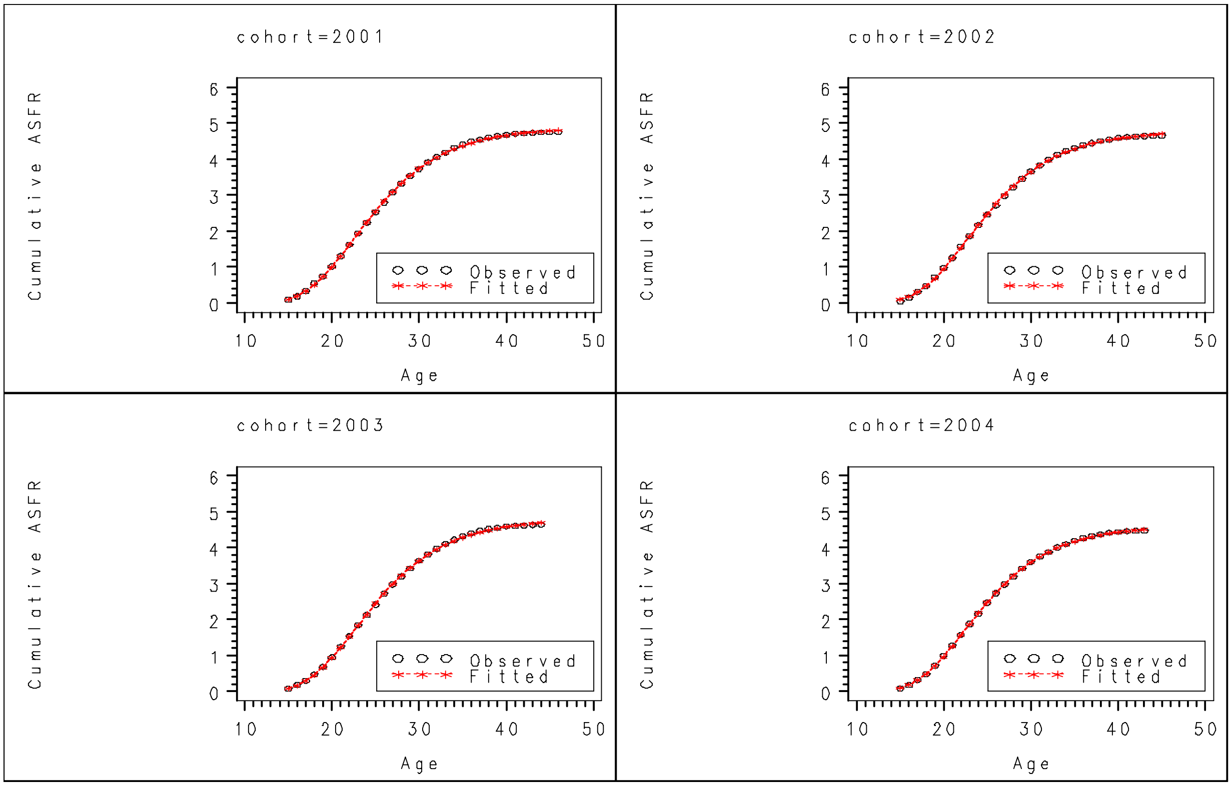


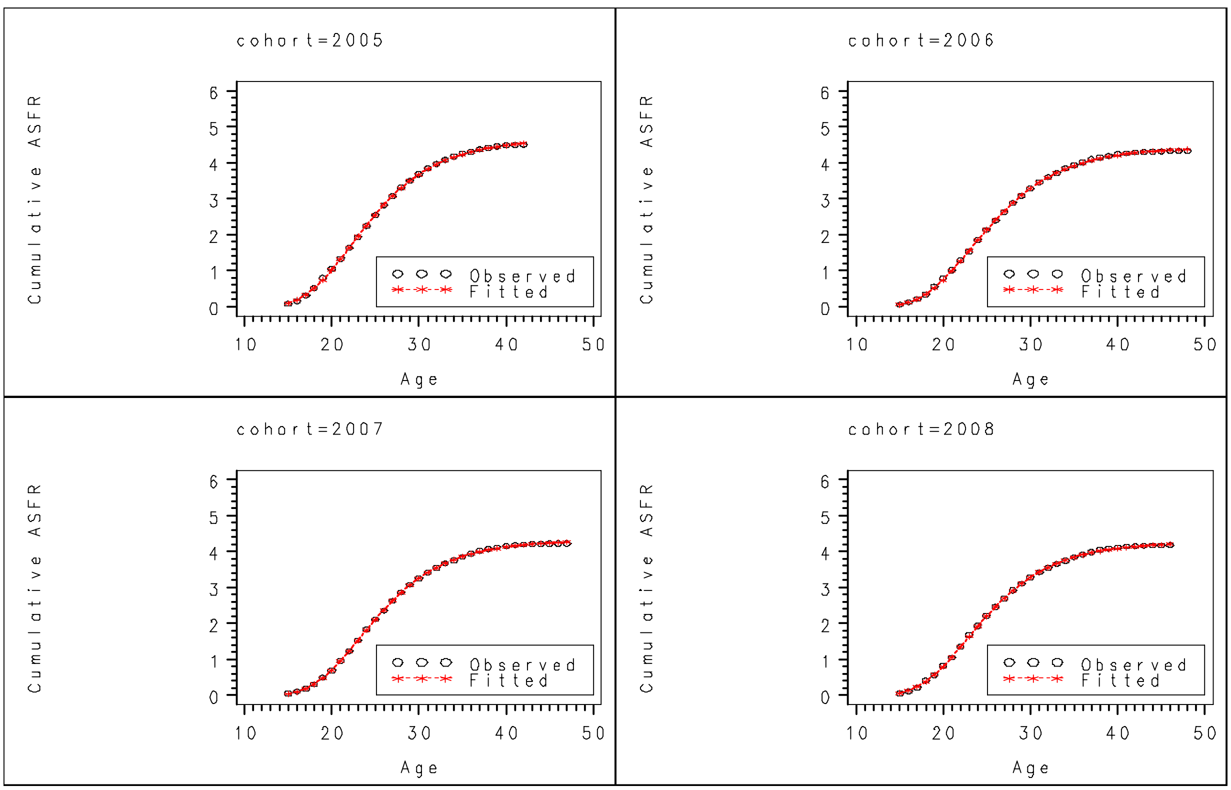
Figure-5: Comparing the fit of special form of Gompertz curve with other models for cohort crossing childbearing age during 2006

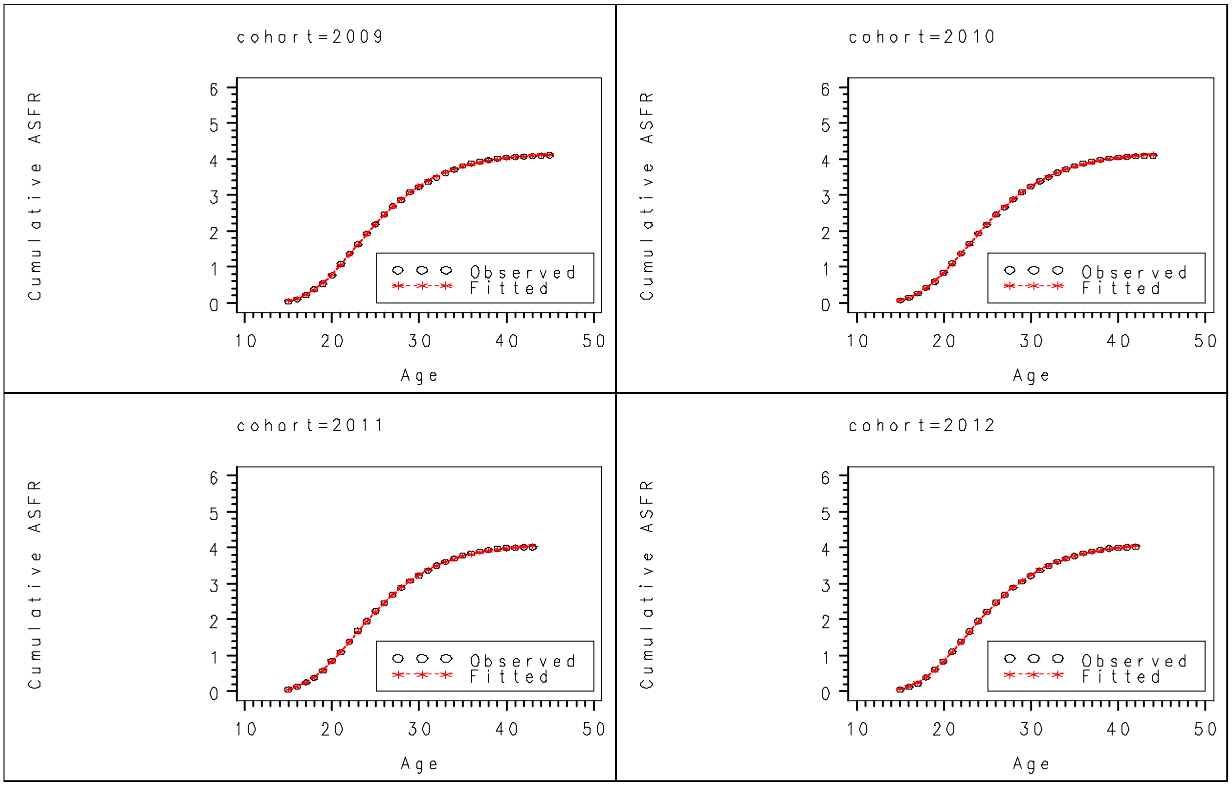


Individual Model fits

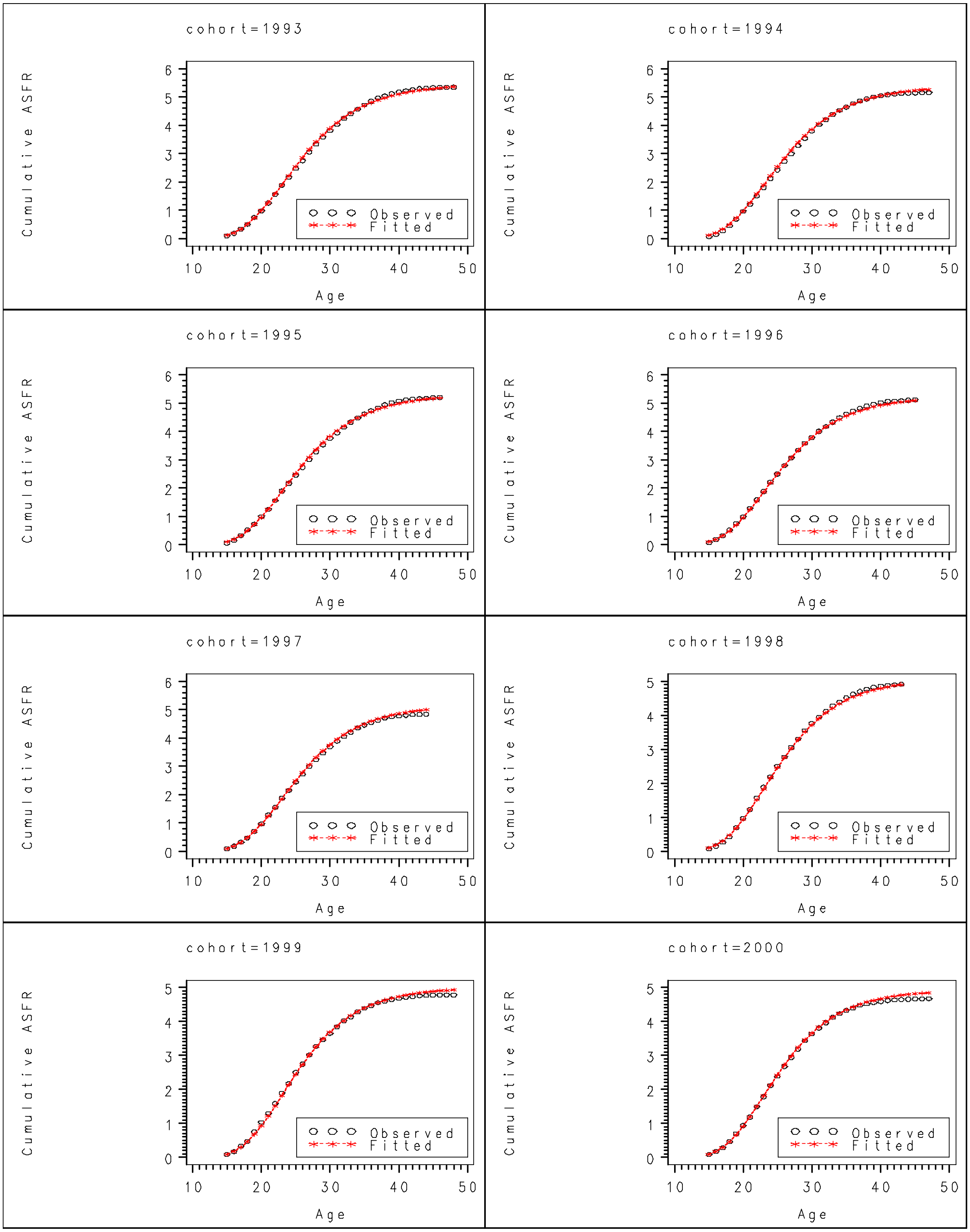


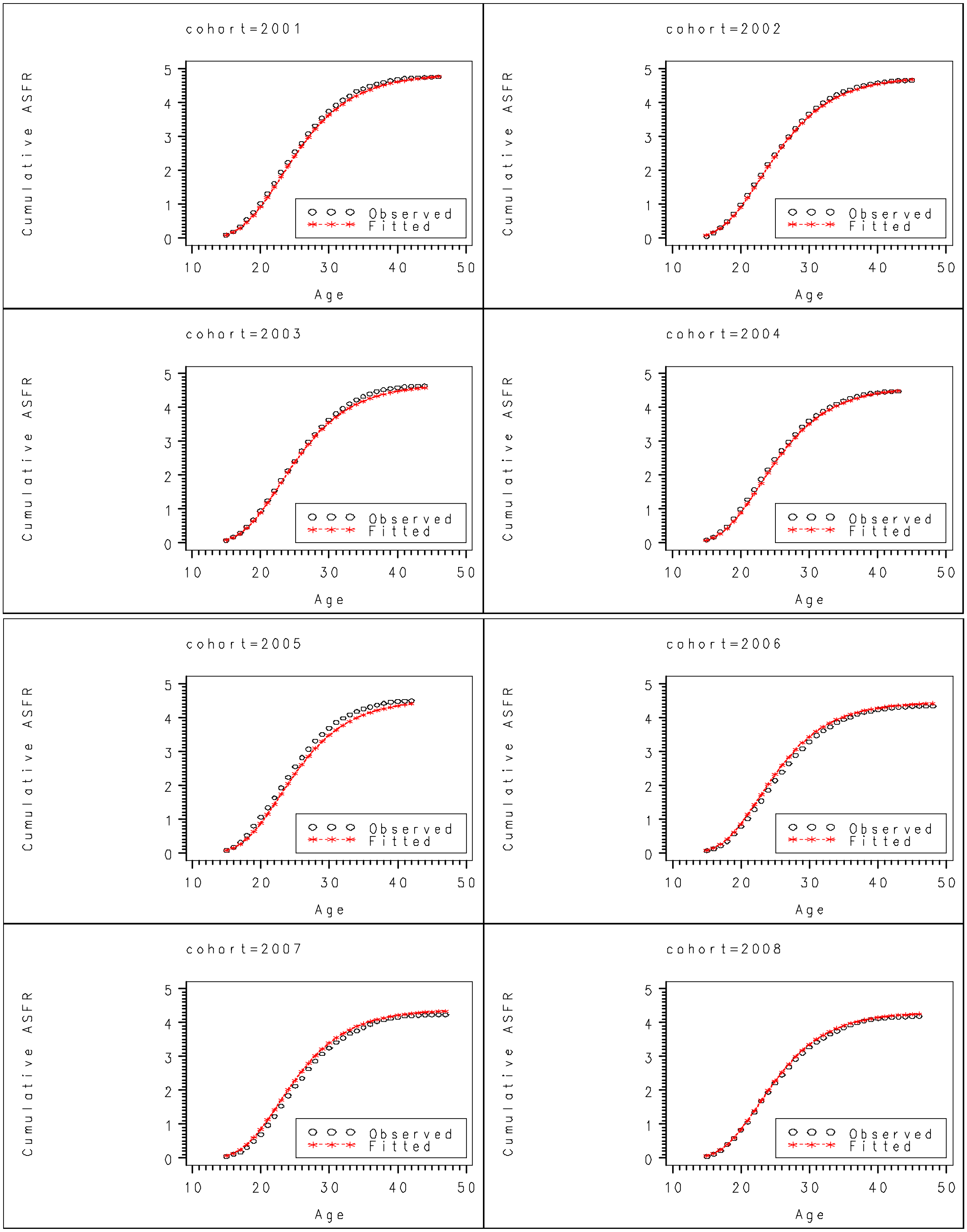

cohort=2001
cohort=2002
cohort=2003
cohort=2004
Cumulative ASFR
Age
Observed
Fitted
0
1
2
3
4
5
6
10
20
30
40
50

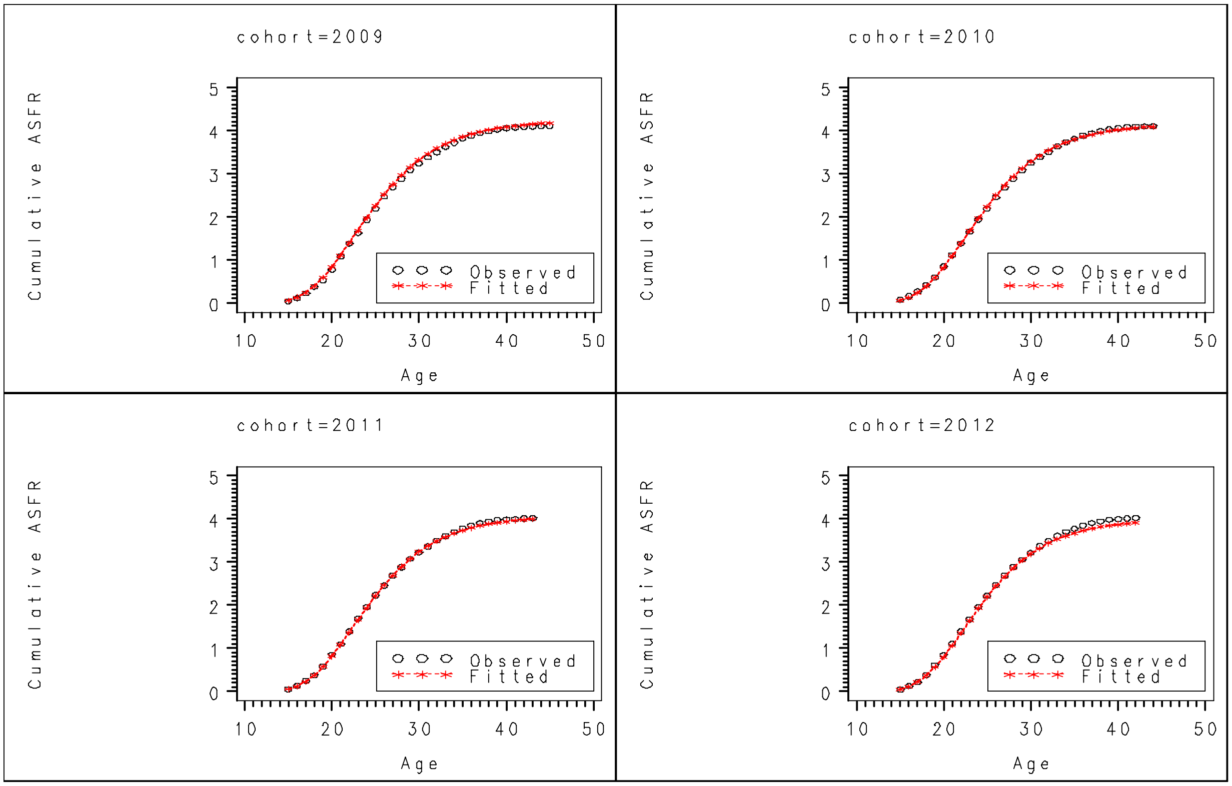

cohort=2005
cohort=2006
cohort=2007
cohort=2008
Cumulative ASFR
Age
Observed
Fitted
0
1
2
3
4
5
6
10
20
30
40
50


cohort=2009
cohort=2010
cohort=2011
cohort=2012
Cumulative ASFR
Age
Observed
Fitted

General Model Fit




cohort=2001
cohort=2002
cohort=2003
cohort=2004
cohort=2005
cohort=2006
cohort=2007
cohort=2008
Cumulative ASFR
Age
0
1
2
3
4
5
10
20
30
40
50
Observed
Fitted


cohort=2009
cohort=2010
cohort=2011
cohort=2012
Cumulative ASFR
Age
Observed
Fitted
0
1
2
3
4
5
10
20
30
40
50