Sankhyā : The Indian Journal of Statistics
2013, Volume 75-B, Part 2, pp. 374-408

Probability distributions of number of children and maternal age at various order births using age-specific fertility rates by birth order

Samba Siva Rao Pasupuleti
Indian Statistical Institute, Kolkata, India
Asis Kumar Chattopadhyay
Calcutta University, Kolkata, India

Abstract

In this study, a simple analytical framework to find the probability distributions of number of children and maternal age at various order births by making use of data on age-specific fertility rates by birth order was proposed. The proposed framework is applicable to both the period and cohort fertility schedules. The most appealing point of the proposed framework is that it does not require stringent assumptions. The proposed framework has been applied to the cohort birth order-specific fertility schedules of India and its different regions and period birth order-specific fertility schedules, including the United States of America, Russia, and the Netherlands, to demonstrate its usefulness.

AMS (2000) *subject classification.* Primary 62P99; Secondary 91D20.
Keywords and phrases. Wald model, Peristera Kostaki model, parity progression ratio, Total Fertility Rate, fertility models, maternal age at birth, variance of fertility distribution.

1 Introduction

Modelling reproductive behaviour of women has been of interest to demographers since the 1950s. In particular, many researchers have attempted modelling the distribution of the number of children per woman. Dandekar (1955) proposed a modified Poisson distribution to model the probability distribution of number of children. Later, researchers developed their own versions of stochastic models for the distribution of number of children by relaxing the strict assumptions made in the Dandekar (1955) model[1]. For instance, Singh (1963) derived an inflated binomial distribution for the number

[1]Strict assumptions in the Dandekar model (Dandekar, 1955) include: (1) each conception will result in a live birth, (2) probability of a conception in a trail (a menstrual

of children by taking into account that, for various biological reasons, not all women can give birth. Pathak (1970), on the other hand, derived the distribution of the number of children by assuming fecundity as a function of age. Later, Singh, Bhattacharya and Yadava (1974) proposed a parity-dependent model of number of births for a specific marital duration by considering fecundity as a function of parity. Considering the fact that not all conceptions will result in live births, Bhattacharya and Nath (1985) derived a bi-variate distribution of number of births and the pregnancy wastage. Perrins and Sheps (1964) introduced a class of stochastic models to approximate the reproductive process of women, including the number of children, by using the concept of renewal theory. Other notable efforts in modelling the distribution of number of children include Shalakani and Pandey (1989) and Pandey and Suchindran (1995). A detailed account of all these efforts in modelling the distribution of number of children can be found in the works of Pathak and Pandey (1993) and Pathak (1999). In this study, a different method to find the probability distribution of the number of children is introduced.

Unlike the distribution of number of children, which has been widely modelled, little effort has been made by demographers in modelling the distribution of maternal age at various order births[2]. Maternal age at various order births is specifically important in demography because 1) it has a strong influence on birth intervals, which in turn will influence the health of both mother and child (Cleland and Sathar, 1984; Koenig et al., 1990; Curtis, Diamond and McDonald, 1993; Nurual, 1995; Rousso et al., 2002; King, 2003), 2) it will dictate the overall timing of fertility, which controls population growth (Frejka, 1973; Bongaarts and Greenhalgh, 1985; Rajaretnam, 1990), and 3) it will affect period fertility indicators (Newell, 1988). In spite of this importance, there are very little efforts to study maternal age at various order births and their distributions. Although few

cycle) is constant and does not change from one trail to another, and (3) there will be no conception in the next 'm' months ('m' is a fixed constant) once there is conception in a month. We consider these to be strict assumptions because (1) not all conceptions will result in live births, (2) the scope of a non-pregnant woman to conceive in any month is not constant and depends on many factors, such as the occurrence of the menstrual cycle in that particular month, intercourse between couples and its frequency with regard to ovulation, usage of contraception, and the health condition of the couple, and (3) the rest period between conception and the next conception need not be a fixed constant and generally depends upon a number of other factors.

[2]Pandey and Suchindran (1997) have used vital rates to derive the distribution of maternal age at various order births, birth intervals, and the PPRs. Krishnamoorthy (1979) and Suchindran and Horne (1984) have derived the distribution of maternal age at first and last births by modifying the work of Hoem (1970) who developed models to study transition probabilities from a specific parity to the next higher parity.

demographers have attempted modelling birth intervals (which can help in the understanding of maternal age at various order births when maternal age at the first birth is known), they mostly confined themselves to the distribution of the first birth interval only (Singh, 1964; Singh and Singh, 1983; Bhattacharya et al., 1989; Nath et al., 1993; Nath, Land and Goswami, 1999).

In this study, a simple analytical framework to find the probability distributions of number of children and maternal age at various order births has been proposed. This framework makes use of data on age-specific fertility rates by birth order. The advantages of the proposed framework include: (1) the framework does not require stringent assumption(s), (2) validity of the models used (if used at all) in the framework can be easily assessed, (3) the primary infertility level in the population can be estimated, (4) model parameters throw light on various important characteristics of reproductive behaviour, such as mean maternal age at various order births and parity progression ratios (PPRs), (5) other important characteristics, such as Total Fertility Rate (TFR) can be easily found as a function of the parameters, and (6) the framework can be used with both period and cohort fertility data.

The proposed framework has been applied to the cohort birth order-specific fertility schedules (set of age specific fertility rates is commonly referred to as fertility schedule or fertility curve in the demographic literature. The same terminology has been followed in this study as well) of India and its different regions and the period (cross-sectional) birth order-specific fertility schedules of Austria, Canada, Czech Republic, Estonia, Finland, Hungary, the Netherlands, Russia, Slovakia, Slovenia, Sweden, and the United States of America to demonstrate its usage.

2 Framework to find the probability distributions of number of children and maternal age at various order births

Let B_i be the event of ever occurrence of the i^{th} order birth to the woman of our considered cohort[3] (i.e. B_i indicates whether the i^{th} order birth occurred during the entire reproductive period of the woman of our considered cohort).

[3] In the case of twin births, we have randomly taken one of the two births as the ith order birth and the other as the $(i+1)^{th}$ order birth. This way of considering birth orders may distort the timing of birth of various order births to which they correspond, for any considered cohort (real or synthetic). However, as twin births are rare (roughly one in 240 births as per our exploration of the Third National Family Health Survey data), the above issue is not a serious problem in reality.

Define

$$B_i = \begin{cases} 1 & if\ i^{th}\ order\ birth\ ever\ occurs \\ 0 & Otherwise. \end{cases}$$

By definition,

$$B_i = 1 \quad \Rightarrow \quad B_{i-1} = 1. \tag{2.1}$$

Also,

$$B_i = 0 \quad \Rightarrow \quad B_{i+1} = 0$$

(as the occurrence of the i^{th} order birth means occurrence of all previous order births and the non-occurrence of the i^{th} order birth means the non-occurrence of all the $(i+1)^{th}$ and higher order births) and $P(B_i = 1/B_{i-1} = 0) = 0$, for $i = 2,\ 3,\ 4,\ ....,\ z$. Here, z is the last observable birth order.

The following result is obvious from equation (2.1)

$$P(B_1 = 1) > P(B_2 = 1) > P(B_3 = 1) > ..... > P(B_z = 1).$$

Let $X_1,\ X_2,\ .....,\ X_z$ be the maternal age at the first, second, third, $...z^{th}$ order births, respectively.

Then, B_i can be written in terms of $X_i's$ as follows:

$$P(B_i = 1) = \sum_{x=\alpha}^{\beta} P(X_i = x),$$

where α and β are lower and upper age limits of reproductive period, respectively, and are assumed to be universally constant.

A common practise in modelling fertility schedules is

$$G(x) = cf(x), \tag{2.2}$$

where $f(x)$ is some probability density function that represents the frequency of occurrence of birth at age x, given that it ever occurred. $G(x)$ is an analytical fertility curve (for more details see Hoem et al., 1981; Pasupuleti and Pathak, 2010). Parameter c is interpreted as Cohort Total Fertility Rate (CTFR) in case of cohort fertility data and TFR in case of period (cross sectional) fertility data.

By extending the approach given in equation (2.2) to model various birth order-specific fertility schedules, we get

$$G_i(x) = c_i f_i(x/B_i = 1), \qquad i = 1,\ 2,\ 3,\ ...\ z, \tag{2.3}$$

and

$$ASFR_x = \sum_{i=\alpha}^{\beta} G_i(x),$$

where $G_i(x)$is the analytical fertility curve corresponding to the i^{th} order births and $f_i(x/B_i = 1)$ (or its discrete case equivalent $P(X_i = x/B_i = 1)$ from here on) is the conditional probability distribution of the occurrence of the i^{th}order birth at age x, given that it will ever occur.

In equation (2.3), parameter c_i is the probability of a woman (proportion of women) of our considered cohort who have given i^{th} order birth, i.e., $P(B_i = 1) = c_i$.

Since each woman can give at most one i^{th} order birth, and not all women can give i^{th} order birth due to a variety of reasons, c_i can also be interpreted as the average number of children of i^{th} order births to the women of our considered cohort[4].

Hence, the PPR[5] of order i (denoted as P_i from here on) can be obtained as follows:

$$P_i = P(B_{i+1} = 1/B_i = 1) = (c_{i+1}/c_i).$$

Under the above mentioned framework, the probability of a woman of our considered cohort not to have child during her entire reproductive period is $(1 - c_1)$, the probability of a women to have at least one child in her life is c_1, and the probability she will have exactly one child is $c_1(1 - P_1)$ or $P_0(1 - P_1)$.

Similarly, the probability of a women to have at least i children is c_i or $P_1P_2P_3....P_{i-1}$ and to have exactly i children is $c_i(1-P_i)$ or $P_1P_2P_3....P_{i-1}(1-P_i)$, $i = 1, 2, 3, ..., z - 1$.

Furthermore, the probability of a woman to give i^{th} order birth at age x is P(i^{th} order birth at age x) $= P_1P_2P_3....P_{i-1} \times P(X_i = x/B_i = 1)$.

Once the PPRs $P_1, P_2, ..., P_z$ are estimated, then TFR can be estimated using the following formula:

$$TFR = P_1 + P_1P_2 + P_1P_2P_3 + ... + P_1P_2P_3....P_z.$$

However, in general, births after a particular birth order, for example k^{th} order, are small (this k varies from one population to another population

[4]Let there be N women in our considered cohort. As introduced earlier, let B_i^j indicate the event of ever occurrence of the i^{th} order birth to the j^{th} individual of our cohort. So, basically, B_i^j is a binary event taking values either zero or one. Hence, some ${B_i}^j$s are zeros and some ${B_i}^j$s are ones, depending upon the occurrence of the i^{th} order birth to the corresponding individuals. In this case, the total number of women who give an i^{th} order birth is $\sum_{j=1}^{N} B_i^j$. Hence, the proportion of women who give an i^{th} order birth is $\sum_{j=1}^{N} B_i^j/N$, which is the same as the average number of i^{th} order births to the women of our considered cohort.

[5]Parity is the number of children a woman has already had. PPR is the probability that a woman with a particular parity to progress to the next parity. For example, if a woman has already had three children, it is the probability of her having the fourth.

and even from one cohort to another cohort within the same population) and hence birth order-specific fertility schedules are zigzag after this k^{th} order birth. So, it is better not to model such a fertility curves (k^{th} and higher order fertility curves).

In this case, as higher order PPRs are not computed, TFR can be obtained using the following formula:

$$TFR = c_1 + c_2 + c_3 + ... + c_k, \tag{2.4}$$

where c_{k+} is the average number of children of order k and above.

3 Analytical choices for the conditional probability distributions

Even without any analytical choices (or functions) in place of $f_i(x/B_i = 1)$, the probability distributions of number of children and the maternal age at various order births can be obtained by replacing $f_i(x/B_i = 1)$ with $(m_x^i / \sum_{x=\alpha}^{\beta} m_x^i)$, where m_x^i is the observed age specific fertility rate at age x corresponding to the i^{th} order births.

The advantage of using some analytical models in place of the above conditional probability distribution is that (1) the models will provide an alternative to the data on age-specific fertility rates by birth order, (2) they will allow to quantify some of the characteristics related to fertility behaviour in the form of parameter estimates and (3) they will aid in analytically deriving various characteristics related to fertility behaviour. Therefore, the present study provides some analytical choices for the conditional probability distributions that are shown in equations (2.3).

It is important to remember that the shape of fertility schedules vary considerably from one population to another and also from one birth order-specific fertility schedule to another, within the same population. Hence, there is a need to choose conditional probability distributions according to the shape of the empirical birth order-specific fertility schedules.

The Gamma model (Hoem et al., 1981), Wald model (Hadwiger, 1940; Wald, 1947; Hoem and Erling, 1975), Beta model (Hoem et al., 1981) and Peristera and Kostaki model-1 (PK-1 model from here on; Peristera and Kostaki, 2007) are among the best possible choices for the conditional probability distributions as (1) all these models have fewer numbers of parameters and hence they are less complex and (2) they can fit well to the fertility schedules. However, the Beta model often suffers from non-convergence of its parameter estimates (Hoem et al., 1981), and therefore it can be safely excluded as a choice for conditional probability distributions. The Gamma and the Wald's model fits are often found to be very close to one another

(Hoem and Erling, 1975; Pasupuleti and Pathak, 2010), but unlike the Wald model, the Gamma model generally under estimates the level of fertility (Pasupuleti and Pathak, 2010). Apart from that, the parameters in the Gamma model do not have any demographic interpretation, while the parameters in the Wald model have clear demographic interpretation. Therefore, the Gamma model can also be omitted as a choice for the conditional probability distributions. The remaining ones namely the Wald model and the PK-1 models can be used as choices for conditional probability distributions, depending upon whether a particular birth order-specific fertility schedule is distorted or not (see Peristera and Kostaki, 2007; Chandola, Coleman and Horns, 1999, for details on distorted fertility patterns).

The Wald model form considered in this study is the following

$$\left(\frac{\lambda}{2\pi\, x^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda}{2\mu}\left(\frac{x}{\mu}+\frac{\mu}{x}-2\right)\right),$$

where μ is mean and μ^3/λ is variance. Hence, parameter λ can be thought of as a characteristic that is inversely associated with the variance of fertility distribution.

In the original form of the PK-1 model, none of the parameters have clear demographic interpretation. Pasupuleti and Pathak (2010) have proposed a re-parametrised PK-1 model of the following form:

$$\left(\frac{1}{\sigma(x)\sqrt{2\Pi}}\right)\exp\left(-0.5\left(\frac{x-m}{\sigma(x)}\right)^2\right),$$

with $\sigma(x) = \begin{cases} \sigma_1 & if \;\; x \leq m \\ \sigma_2 & if \;\; x > m. \end{cases}$

Here, σ_1 and σ_2 are the standard deviations of the fertility distribution before and after its peak. m is the mode of the age specific fertility distribution.

Unlike other models, such as the Beta model, Gamma model, and Wald model, the characteristics of the PK-1 model are not discussed either in the statistical or in the demographic literature. Therefore, the characteristics of the PK-1 model are presented in the form of the following lemma, and its proof is given in Appendix B.

Lemma: If X is a random variable with the following probability density function $f(x) = \left(1/\sigma(x)\sqrt{2\Pi}\right)\exp\left(-0.5\left(x-m/\sigma(x)\right)^2\right)$, where

$$\sigma(x) = \begin{cases} \sigma_1 & if \;\; x \leq m \\ \sigma_2 & if \;\; x > m, \end{cases}$$

then Moment Generating Function (MGF) of X is

$$M_X(t) = \quad e^{mt+(1/2)t^2\sigma_1^2}\Phi(-t\sigma_1) + e^{mt+(1/2)t^2\sigma_2^2}(1-\Phi(-t\sigma_2)),$$

the mean of X is

$$m + \left(\frac{\sigma_2 - \sigma_1}{\sqrt{2\Pi}}\right),$$

and the variance of X is

$$\left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right) - \frac{1}{2\pi}(\sigma_2 - \sigma_1)^2.$$

4 Application of the proposed framework to the cohort birth order-specific fertility schedules of India and its different regions

Data on maternity histories of women aged 43 years and above in the third National Family Health Survey (NFHS-3) conducted in India during 2005-2006 have been used in this section. In NFHS-3, information was collected from a sample of 124,456 women of age group 15-49 years. The same regional classification as in NFHS-3 has been followed in this study (IIPS and Macro International, 2007). However, for reader's convenience, different regions of India and their constituent states are shown in Figure 4 in Appendix A. The number of women from North, South, East, West, Central and North-east India that have been covered in the above mentioned survey and are above the age of 42 years are 1,735, 1,769, 1,282, 1,191, 1,594, and 1,335, respectively. The birth order-specific fertility schedules corresponding to these women cohorts have been used in this study.

Observed birth order-specific fertility schedules of India and its different regions can be seen from Figure 5 in Appendix A. It can be observable from this figure that there are no distortions in birth order-specific fertility schedules (see just observed birth order-specific fertility schedules in Figure 5 in Appendix A). Hence, the Wald model has been used as a choice for conditional probability distributions (the right side function in equation (2.3)) for India and its different regions. A zigzag pattern is evident in the age pattern of giving 7^{th} and higher order births for India and its different

Table 1: Fit Statistics of various birth order-specific fertility curves for India and its different regions.

Region	**IGF of various order fertility curves**			
	1	**2**	**3**	**4**
North	99.0	98.9	98.2	98.2
South	99.1	98.9	97.6	97.8
East	98.3	97.4	98.3	98.7
West	97.6	97.8	98.9	97.3
Central	98.2	98.9	98.4	97.9
North-east	97.6	98.6	95.4	96.3
India	99.5	99.5	99.2	99.3

regions (data is not shown here). Hence, fertility curves of 7^{th} and higher order births are not modelled in this study. Instead, average number of births of order 7 and above has been calculated and is used in estimating the TFR with the formula given in equation (2.4). A procedure called 'Proc Nlin', available in the Statistical Analysis System package, has been used to fit various birth order-specific fertility curves to the corresponding empirical fertility schedules. Parameters have been estimated under the criterion of principle of least squares. Fit statistics corresponding to the fit of various birth order-specific fertility curves fitted to the corresponding empirical fertility schedules of India and its different regions are shown in Table 1.

Table 2: Parity progression ratios (up to fifth order birth) in India and its different regions.

Region	**Parity Progression Ratios of order**						**TFR**
	0	**1**	**2**	**3**	**4**	**5**	
North	0.967	0.969	0.819	0.671	0.582	0.576	4.628
South	0.948	0.938	0.692	0.567	0.535	0.551	4.040
East	0.936	0.909	0.830	0.736	0.720	0.649	4.568
West	0.950	0.932	0.753	0.597	0.514	0.444	4.148
Central	0.958	0.977	0.876	0.828	0.762	0.688	5.228
North-east	0.912	0.972	0.871	0.772	0.702	0.692	4.786
India	0.945	0.952	0.805	0.702	0.652	0.629	4.563

Table 3: Probability distribution of number of children for India and its different regions.

Region	P(no child)	P(Exactly one child)	P(Exactly two children)	P(Exactly three children)	P(Exactly four children)	P(Exactly five children)	P(Exactly six children)	P(Exactly seven or more children)
North	0.033	0.061	0.210	0.338	0.251	0.089	0.001	0.018
South	0.052	0.105	0.324	0.338	0.148	0.030	0.000	0.003
East	0.064	0.139	0.234	0.270	0.183	0.083	0.002	0.025
West	0.050	0.108	0.280	0.338	0.178	0.042	0.000	0.004
Central	0.042	0.061	0.161	0.236	0.241	0.167	0.004	0.089
North-east	0.088	0.104	0.184	0.252	0.216	0.111	0.004	0.041
India	0.055	0.095	0.234	0.303	0.209	0.082	0.001	0.021

In order to assess the fit of various birth order-specific fertility curves, we used a measure called the Index of Goodness of Fit (IGF), defined by Pasupuleti and Pathak (2010) as $\text{IGF} = 100(\sum_{t=\alpha}^{\beta}(y_x - Y)^2 - \sum_{t=\alpha}^{\beta}(y_x - z_x)^2) / \sum_{t=\alpha}^{\beta}(y_x - Y)^2$. Here, y_x is the observed Age Specific Fertility Rate (ASFR) at age x, Y is the mean of observed ASFRs, and z_x is the fitted ASFR at age x (all corresponding to a particular birth order). IGF values vary between 0 and 100. If a model perfectly fits to a given empirical fertility schedule (i.e., error in prediction is zero at all ages), then IGF takes a value of 100. In another extreme case of predicting all the ASFR values with Y(the mean of ASFRs), IGF takes a value of 0. If the IGF for a fitted fertility model is between 95 and 100, then we consider the model fit to be good. Otherwise, if the IGF value is between 95 and 90, then we consider the model fit to be fair. If the IGF value is less than 90, we consider the model fit to be poor.

Additionally, for better understanding the fits of various birth order-specific fertility curves, Figure 5 in Appendix A, which shows the observed and fitted birth order-specific fertility curves for India and its different regions, is provided. Parameter estimates of various birth order-specific fertility curves for India and its different regions are given in Table 9 in Appendix A, and the standard errors of the parameter estimates are listed in Table 10 in Appendix A. An excellent fit of the Wald model to all six observed birth order-specific fertility schedules can be seen in Figure 5 in Appendix A and from Table 1 (all IGFs are above 95). Characteristics of fertility behaviour, such as PPRs, probability distributions of number of children, and the mean maternal age at various order births for India and its different regions are presented in Tables 2, 3, and 4, respectively. Probability distributions of maternal age at various order births for India and its different regions are shown in Figure 1.

Table 4: Mean maternal age at various order births for India and its different regions.

Region	**Mean maternal age at various order births**					
	1	**2**	**3**	**4**	**5**	**6**
North	20.75	23.21	25.40	27.22	29.16	30.85
South	20.18	22.69	24.33	25.70	27.41	28.82
East	19.20	21.84	24.26	26.50	28.85	30.71
West	21.13	23.51	25.00	26.45	28.33	30.94
Central	19.57	22.12	24.52	26.97	29.42	31.37
North-east	21.14	24.01	26.41	28.31	30.21	32.11
India	20.28	22.84	24.96	26.92	29.08	31.07

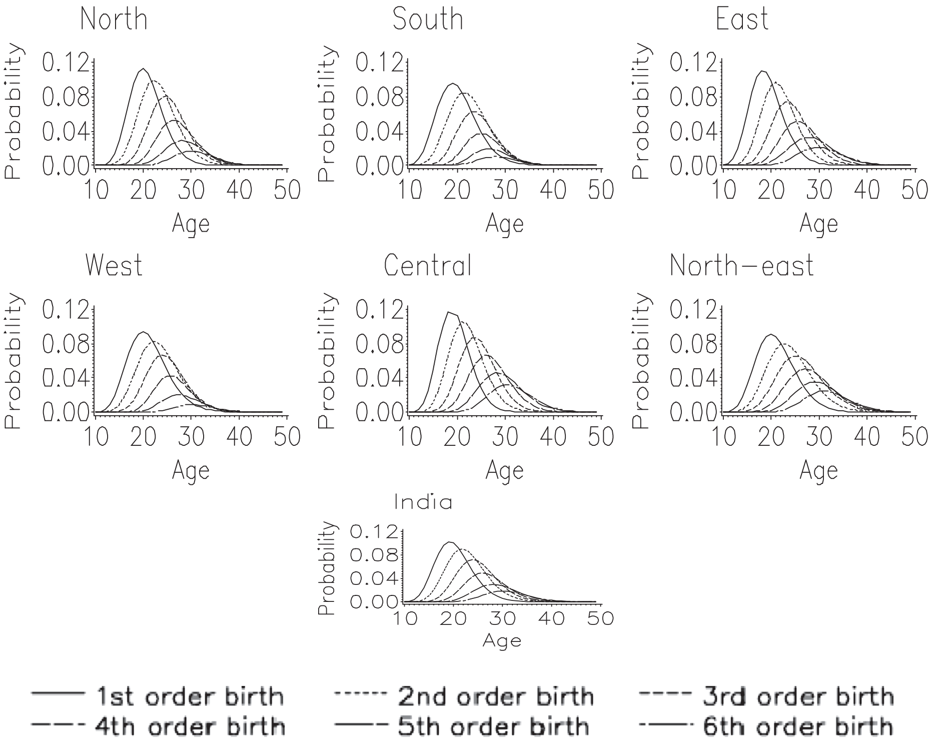

Figure 1: Probability distributions of maternal age at various order births in India and its different regions.

Table 3 shows that 5.47% of women in India, aged 43 years and above at the beginning of 2005, are childless. This means around one out of 18 women in India are childless. Similarly, in different regions of India, one in 11 in North-east India, one in 16 in East India, one in 19 in South India, one in 20 in West India, one in 24 in Central India, and one in every 31 women in North India, are childless. This indicates childlessness or primary infertility[6]

[6]Primary infertility level means the biological inability of couples to have at least one child in their life, despite their wish to have children. As the desire for children is universal in India (only 0.1% of married women reported their ideal number of children as zero, i.e., one in 1,000 women do not want any children (International Institute for Population Sciences and Macro International, 2007)), so, almost all couples who do not have children during their life can be treated as primary infertile (although most of these may have had children if proper treatment had been provided).

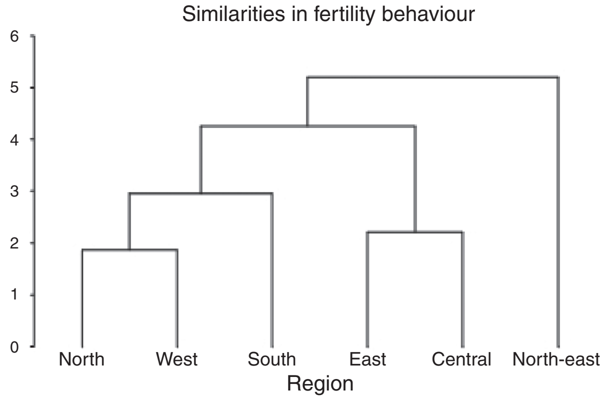

Figure 2: Similarities in fertility behaviour among different regions of India, obtained by using hierarchical clustering technique.

is a serious problem in North-east India compared to other regions of the country. In general, the tempo to progress to higher order births is relatively less in South India and West India, especially after the 3^{rd} order birth. This is not the case in other regions of India, particularly in Central India and East India. While just half of the women in the South and the West have progressed from 4^{th} order birth to 5^{th} order birth, more than three-fourth of the women did the same in case of Central India (Table 2). The average number of children born, i.e., CTFR, varies from a minimum of four children in the South to a maximum of 5.2 children in Central India. Irrespective of the region, with the exception of Central India, more than one-fourth of women have given exactly three births by the end of their reproductive period (Table 3). The average of birth intervals (average of the differences in mean maternal ages at successive births) is the smallest in South India (1.73 years), followed in increasing order by the West (1.96 years), North (2.02 years), North-east (2.19 years), East (2.30 years), and Central India (2.36 years), respectively. In general, the first birth interval (average gap between the first birth and the second birth) is larger than the remaining birth intervals (Table 4). Typically, the first order birth occurs around the age of 20 years in all the regions of India (Table 4). In East India it occurs one year earlier than the rest of India. Maternal age at various order births is typically the largest in the North-east India than in other regions of the country. The dendrogram in Figure 2 shows the disaggregated similarities

Table 5: Fit statistics of various birth order-specific fertility curves for various countries.

Country	**IGF of various order fertility curves**			
	1	**2**	**3**	**4**
Austria	98.7	99.7	99.6	98.2
Canada	98.9	99.8	99.8	98.4
Czech Republic	98.9	99.5	99.3	96.7
Estonia	97.8	99.3	99.3	95.1
Finland	98.8	99.5	99.6	96.8
Hungary	97.7	98.3	97.7	95.6
Netherlands	99.7	99.9	99.9	98.3
Russia	99.4	98.9	98.0	98.1
Slovakia	98.2	99.0	97.8	93.1
Slovenia	99.3	99.3	98.2	94.8
Sweden	97.4	97.8	97.3	98.1
USA	98.1	96.9	98.6	99.0

in fertility behaviour of women belonging to different regions of India, with respect to maternal age at various order births and the PPRs, calculated by using the single linkage hierarchical cluster analysis technique. It is clear from this Dendrogram (Figure 2) that Central India and East India are similar in terms of giving various order births, while the West is similar to the North and to the South. Fertility behaviour of women in the North-east is some what different from the rest of India.

5 Application of the proposed framework to the period birth order-specific fertility schedules of other countries

In order to show the application of the proposed framework to period birth order-specific fertility schedules, data from the Human Fertility Database (HFD) available online at http://www.humanfertility.org/cgi-bin/main.php has been used. At present, period birth order-specific fertility schedules are available from this database for Austria, Canada, Czech Republic, Estonia, Finland, Hungary, Netherlands, Russia, Slovakia, Slovenia, Sweden, and the United States of America (USA). The HFD (i.e. the fertility data of all the above mentioned countries) is entirely based on the officially registered birth counts by calendar year, mother's age (and/or cohort), and

Table 6: Parity progression ratios for various countries.

Region	**Parity Progression Ratios of order**					**TFR**
	0	**1**	**2**	**3**	**4**	
Austria	0.691	0.721	0.349	0.291	0.691	1.44
Canada	0.749	0.782	0.380	0.322	0.749	1.67
Czech Republic	0.726	0.735	0.271	0.260	0.726	1.46
Estonia	0.739	0.778	0.383	0.275	0.739	1.63
Finland	0.798	0.793	0.441	0.346	0.798	1.90
Hungary	0.626	0.649	0.397	0.362	0.626	1.30
Netherlands	0.830	0.774	0.348	0.255	0.830	1.78
Russia	0.799	0.673	0.275	0.260	0.799	1.55
Slovakia	0.697	0.620	0.372	0.362	0.697	1.42
Slovenia	0.758	0.720	0.293	0.245	0.758	1.52
Sweden	0.888	0.772	0.364	0.271	0.888	1.93
USA	0.879	0.803	0.527	0.405	0.879	2.20

biological birth order[7]. Except for the USA, Canada, and Austria, birth order-specific fertility schedules are available until the year 2009 for all the above countries. The same data are available until 2007 for the USA and Canada, and until 2008 for Austria. The most recent birth order-specific fertility schedules of these countries have been used in this study.

Since period fertility rates are sensitive tempo effects (i.e., affected by changes in the timing of childbearing), birth order-specific fertility schedules of all the above countries have been adjusted for tempo effects by using the Bongaarts and Feeney procedure (Bongaarts and Feeney, 1998) before using them in this study.

Some distortions in the birth order-specific fertility schedules of the above countries were noticed in this study. In particular, the fertility schedules corresponding to the first order births for Austria, Canada, Czech Republic, Estonia, Finland, and Sweden; the fertility schedules corresponding to the

[7]The registration of vital events is almost complete in all the countries that have been covered in HFD. Therefore, the sample size of women aged 'x' years is same as the corresponding population size of women aged 'x' years for all the countries that are covered in HFD.

Table 7: Probability distribution of number of children for various countries.

Country	P(no child)	P(Exactly one child)	P(Exactly two children)	P(Exactly three children)	P(Exactly four children)
Austria	0.309	0.193	0.324	0.124	0.051
Canada	0.251	0.163	0.363	0.151	0.071
Czech Republic	0.274	0.193	0.389	0.107	0.037
Estonia	0.261	0.164	0.355	0.160	0.060
Finland	0.202	0.165	0.353	0.182	0.097
Hungary	0.374	0.220	0.245	0.103	0.058
Netherlands	0.170	0.187	0.419	0.167	0.057
Russia	0.201	0.261	0.390	0.110	0.039
Slovakia	0.303	0.265	0.272	0.103	0.058
Slovenia	0.242	0.212	0.386	0.121	0.039
Sweden	0.112	0.203	0.435	0.182	0.068
USA	0.121	0.174	0.334	0.221	0.151

Table 8: Mean maternal age at various order births for the considered countries.

Country	**Mean age at first birth**	**Mean age at second birth**	**Mean age at third birth**	**Mean age at fourth birth**
Austria	26.94	30.23	32.29	33.63
Canada	27.29	31.00	31.84	32.31
Czech Republic	26.74	30.81	33.83	34.74
Estonia	24.92	30.14	33.80	35.57
Finland	26.64	30.62	32.48	33.20
Hungary	26.64	29.67	30.70	32.54
Netherlands	29.27	31.59	33.19	34.25
Russia	23.77	29.53	32.36	33.64
Slovakia	25.84	28.93	31.95	31.43
Slovenia	28.02	30.84	33.16	34.91
Sweden	28.10	31.44	33.51	34.70
USA	24.62	27.65	29.86	30.89

first and second order births for Slovakia and the USA; the fertility schedules corresponding to the first, second, and third order births for Hungary are distorted ones (see just observed birth order-specific fertility rates from Figure 6 in Appendix A). Therefore, in these cases we used the PK-1 model,

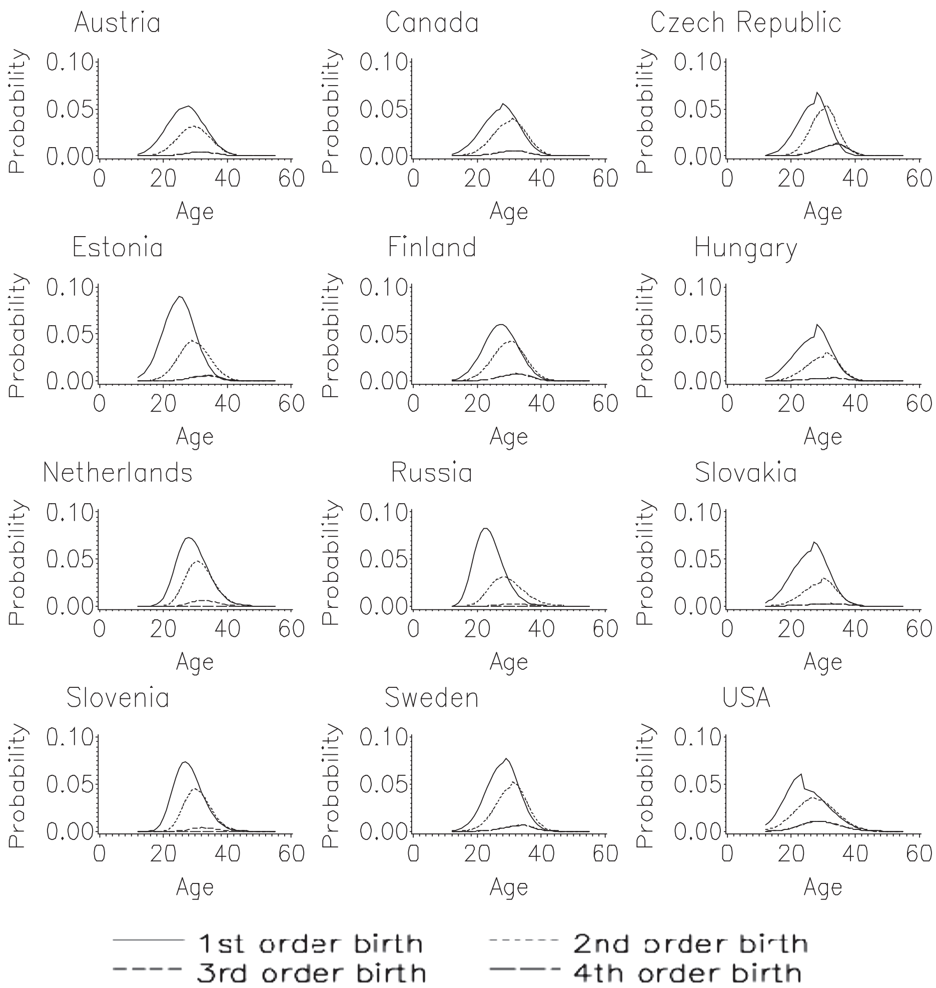

Figure 3: Probability distribution of maternal age at various order births for the considered countries.

while in the remaining cases we used the Wald model for the conditional probability distributions.

Fit statistics related to the fit of various order fertility curves, fitted to the corresponding empirical fertility schedules of various countries are provided in Table 5. Additionally, for better clarity of the fits, we provided the observed and fitted birth order-specific fertility schedules for all the considered countries in the form of Figure 6 in Appendix A. An excellent fit of various birth order-specific fertility curves to the corresponding empirical fertility schedules can be seen from this figure and Table 5 (all IGFs are above 95 except for 4^{th} order births of Slovakia and Slovenia). Parameter estimates of various birth order-specific fertility curves for the considered countries are given in Table 11 in Appendix A. The standard errors of the parameter estimates are given in Table 12 in Appendix A. The dynamics related to PPRs and the probability distributions of number of children are shown in Tables 6 and 7, respectively. Mean maternal age at various order births are shown in Table 8. Probability distribution of maternal age at various order births for all the considered countries are depicted in Figure 3.

A clear variation in the pattern of giving various order births among the considered countries can be seen from Tables 6, 7, and 8. With the future continuation of these age-specific fertility rates by birth order, it is expected that roughly 37% of women in Hungary, 31% of women in Austria, and 11% of women in Sweden will remain childless during their life. The TFR for the considered countries varies from a minimum of 1.3 in Hungary to a maximum of 2.2 in the USA (Table 6). The likelihood of a woman to have just one child is more in Slovakia and Russia than the remaining countries (Table 7). The probability of a woman to give 3^{rd} and higher order births is higher in the USA than the remaining countries. There is considerable variation in the maternal age at various order births from one country to another country. The mean maternal age at the first order birth is the lowest in Russia (23.8 years) and the highest in Netherlands (29.3 years), of all the considered countries (Table 8).

6 Summary

In this study, a simple analytical framework (using a set of analytical models) to find the probability distributions of number of children and the maternal age at various order births has been proposed. The proposed framework makes use of data on age-specific fertility rates by birth order.

Strengths of the proposed framework include (1) it does not require stringent assumptions, (2) validity of the models used (if used at all) in the framework can be easily assessed, (3) the framework will help to estimate the primary infertility level in our population, (4) model parameters shed light on various important characteristics of reproductive behaviour, (5) characteristics, such as TFR and PPRs, can be easily obtained based on the parameter estimates, and (6) the framework can be used with both the period and the cohort fertility schedules. However, care must be taken while using the proposed framework with the period fertility data as period fertility data often suffers from tempo effects.

The proposed framework has been applied to the cohort birth order-specific fertility schedules of India and its different regions and the period birth order-specific fertility schedules of other countries, including the USA, Russia, and the Netherlands. Some of the interesting findings that we obtained include: (1) the primary infertility level and the maternal age at various order births are relatively higher in North-east India compared to the remaining regions in the country and (2) more than 30% of women are expected to remain childless in the near future in Hungary and Austria, given the persistence of current age-specific fertility rates by birth order of these countries.

Given the increasing interest in PPRs, maternal age at various order births, and various birth order dynamics, as a result of increased focus of government policies on birth order dynamics, we strongly recommend considering the proposed framework in future efforts to study various birth order dynamics of women.

7 Further scope and future work

Efforts are underway to develop useful statistical tools that will help in the inferential aspects related to PPRs, by assuming an appropriate distributional form for errors.

Acknowledgement. We gratefully acknowledge the anonymous referees for their constructive comments on an earlier version of this paper.

References

BHATTACHARYA, B.N. and NATH, D.C. (1985). On some bivariate distributions of number of births. *Sankhya*, **47**, 372–384.

BHATTACHARYA, B.N., SINGH, K.K., SINGH, U. and PANDEY, C.M. (1989). An extension of a model for first birth interval and some social factors. *Sankhya*, **51**, 115–124.

BONGAARTS, J. and FEENEY, G. (1998). On the quantum and tempo of fertility. *Popul. Dev. Rev.*, **24**, 271–291.

BONGAARTS, J. and GREENHALGH, S. (1985). An alternative to the one-child policy in China. *Popul. Dev. Rev.*, **11**, 585–617.

CHANDOLA, T., COLEMAN, D.A. and HORNS, R.W. (1999). Recent European fertility patterns: fitting curves to 'distorted' distributions. *Pop. Stud.-J. Demog.*, **53**, 317–329.

CLELAND, J.G. and SATHAR, Z.A. (1984). The effect of birth spacing on childhood mortality in Pakistan. *Pop. Stud.-J. Demog.*, **38**, 401–418.

CURTIS, S., DIAMOND, I. and MCDONALD, J. (1993). Birth interval and family effects on postneonatal mortality in Brazil. *Demography*, **30**, 33–43.

DANDEKAR, V.M. (1955). Certain modified forms of binomial and Poison distributions. *Sankhya*, **15**, 237–250.

FREJKA, T. (1973). *The future of population growth: alternative paths to equilibrium.* John Wiley and Sons, New York.

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

HOEM, J.M. (1970). Probabilistic fertility models of life table type. *Theor. Popul. Biol.*, **1**, 12–38.

HOEM, J.M. and ERLING B. (1975). Some problems in Hadwiger fertility graduation. *Scand. Actuar. J.*, 129–144.

HOEM, J.M., MADSEN, D., NIELSEN, J.L., OHLSEN, E.-M., HANSEN, H.O. and RENNERMALM, B. (1981). Experiments in modeling recent Danish fertility curves. *Demography*, **18**, 231–244.

INTERNATIONAL INSTITUTE FOR POPULATION SCIENCES AND MACRO INTERNATIONAL (2007). National Family Health Survey (NFHS-3), 2005-06: India: Volume I, Mumbai: IIPS.

KING, J.C. (2003). The risk of maternal nutritional depletion and poor outcomes increases in early or closely spaced pregnancies. *J. Nutr.*, **133**, 1732S–1736S.

KOENIG, M.A., PHILLIPS, J.F., CAMPBELL, O.M. and D'SOUZA, S. (1990). Birth intervals and childhood mortality in rural Bangladesh. *Demography*, **27**, 251–265. doi:10.2307/2061452.

KRISHNAMOORTHY, S. (1979). Family formation and the life cycle. *Demography*, **16**, 121–129.

NATH, D.C., LAND, K.C. and GOSWAMI, G. (1999). Effects of the status of women on the first birth interval in Indian urban society. *J. Biosoc. Sci.*, **31**(1), 55–69.

NATH, D.C., SINGH, K.K., LAND, K.C. and TALUKDAR, P.K. (1993). Age at marriage and length of first-birth interval in traditional Indian society: Life table and hazard model analysis. *Hum. Biol.*, **65**, 783–797.

NEWELL, C. (1988). *Methods and models in demography.* Wiley, Chichester.

NURUL, A. (1995). Birth spacing and infant and early childhood mortality in a high fertility area of Bangladesh: age-dependent and interactive effects. *J. Biosoc. Sci.*, **27**, 393–404. doi:10.1017/S0021932000023002.

PANDEY, A. and SUCHINDRAN, C.M. (1995). Some analytical models to estimate ma ternal age at birth using age-specific fertility rates. *Sankhya*, **57**, 142–150.

PANDEY, A. and SUCHINDRAN, C.M. (1997). Estimation of birth intervals and parity progression rations from vital rates. *Sankhya*, **59**, 108–122.

PASUPULETI, S.S.R. and PATHAK, P. (2010). Special form of Gompertz model and its application. *Genus*, **66**, 95–125.

PATHAK, K.B. (1970). A time dependent distribution for number of conceptions. *Artha Vijnana*, **12**, 429–435.

PATHAK, K.B. (1999). Stochastic models of family building: an analytical review and prospective issues. *Sankhya*, **61**, 237–260.

PATHAK, K.B. and PANDEY, A. (1993). *Stochastic models of human reproduction.* Himalaya Publishing House, Bombay.

PERISTERA, P. and KOSTAKI, A. (2007). Modeling fertility in modern populations. *Demogr. Res.*, **16**, 141–194.

PERRIN, E.B. and SHEPS, M.C. (1964). Human reproduction: a stochastic process. *Biometrics*, **20**, 28–45.

SHALAKANI, M.E. and PANDEY, A. (1989). Distribution of births in an abrupt sequence: a stochastic model. *Math. Biosci.*, **95**, 1–11.

SINGH, S.N. (1963). Probability models for the variation in the number of births per couple. *J. Amer. Statist. Assoc.*, **58**, 721–727.

SINGH, S.N. (1964). On the time of first birth. *Sankhya*, **26**, 95–102.

SINGH, S.N., BHATTACHARYA, B.N. and YADAVA, R.C. (1974). A parity dependent model for number of births and its application. *Sankhya*, **36**, 93–102.

SINGH, V.K. and SINGH, U.N. (1983). Period of temporary separation and its effect on first birth interval. *Demography India*, **12**, 282–88.

SUCHINDRAN, C.M. and HORNE, A.D. (1984). Some Statistical Approaches to the Modelling of Selected Fertility Events. In *Proceedings of American Statistical Association (Social Statis tics Section).* Washington DC, pp. 629–634.

RAJARETNAM, T. (1990). How delaying marriage and spacing births contributes to population control? An explanation with illustrations. *J. Fam. Welfare*, **36**(4), 3–13.

ROUSSO, D., PANIDIS, D., GKOUTZIOULIS, F., KOURTIS, A., MAVROMATIDIS, G. and KALAHANIS, I. (2002). Effect of the interval between pregnancies on the health of mother and child. *Eur. J. Obstet. Gyn. R. B.*, **105**, 4–6.

WALD, A. (1947). *Sequential analysis.* John Wiley & Sons, Inc., New York.

Appendix A

Figure 4: Different regions of India and their constituent states.

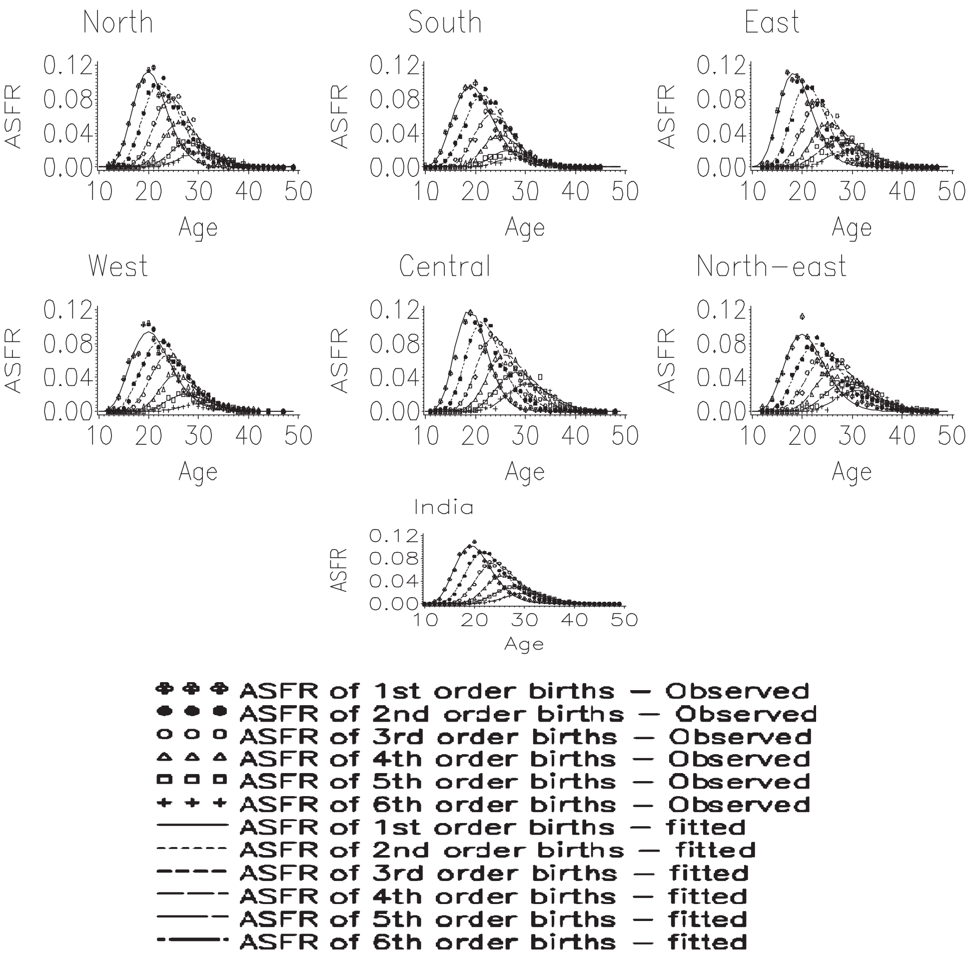

Figure 5: Fit of various birth order-specific fertility curves for the considered countries.

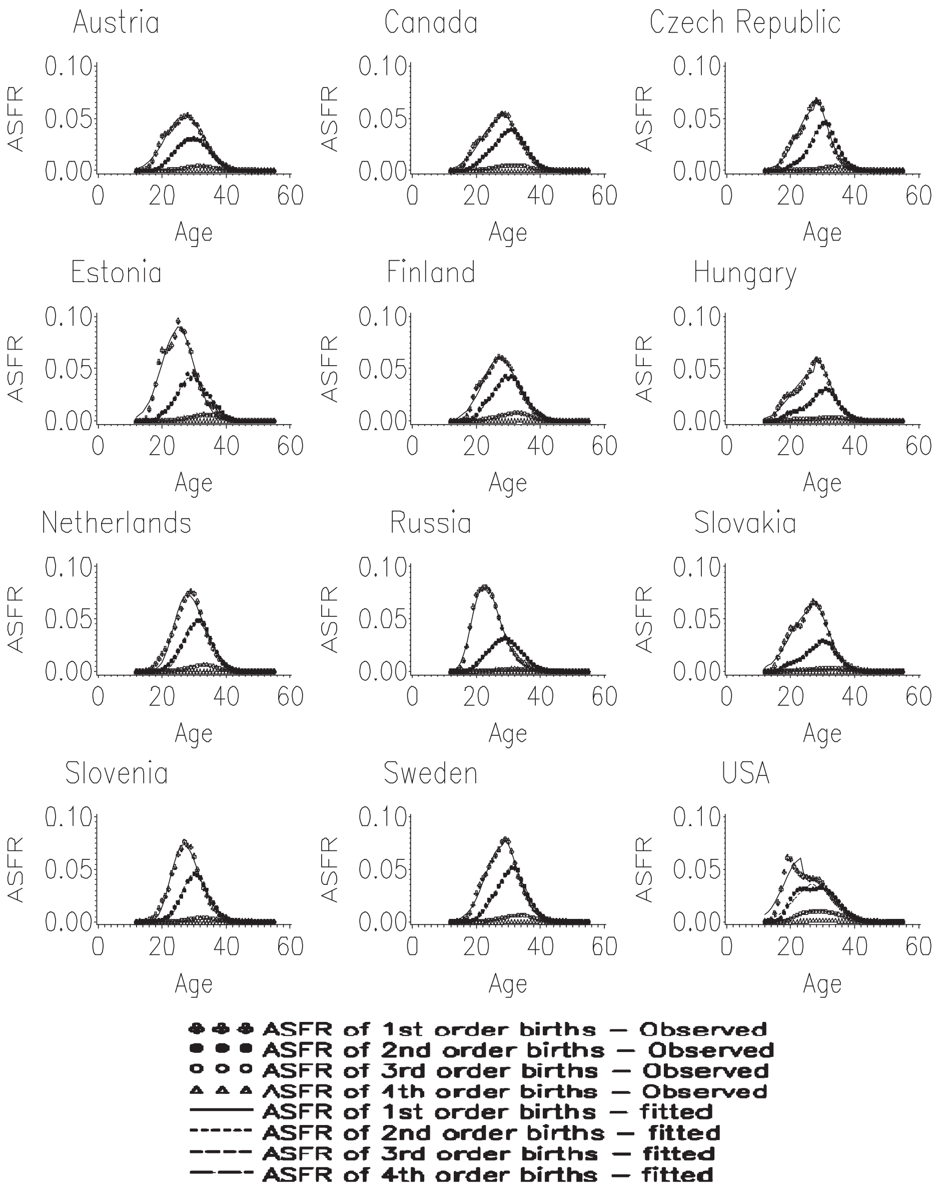

Figure 6: Fit of various birth order-specific fertility curves for the considered countries.

Table 9: Parameter estimates of various birth order-specific fertility curves fitted to the cohort birth order specific fertility schedules of India and its different regions (variance of conditional probability distributions is also provided).

Birth order	C	λ	μ	Variance
North				
1	0.967	708.1	20.75	12.62
2	0.937	827.9	23.21	15.1
3	0.768	1091.7	25.4	15.01
4	0.515	1280.5	27.22	15.75
5	0.3	1392.5	29.16	17.81
6	0.173	1645.7	30.85	17.84
South				
1	0.948	480.9	20.18	17.08
2	0.889	621.6	22.69	18.8
3	0.616	882.3	24.33	16.32
4	0.349	1179.1	25.7	14.4
5	0.187	1357.3	27.41	15.17
6	0.103	1394.5	28.82	17.16
East				
1	0.936	580.7	19.2	12.2
2	0.851	786.6	21.84	13.24
3	0.706	939	24.26	15.21
4	0.52	1083.7	26.5	17.18
5	0.375	1090.9	28.85	22.02
6	0.243	1289.4	30.71	22.47
West				
1	0.95	524.8	21.13	17.96
2	0.886	652.2	23.51	19.92
3	0.668	925.5	25	16.89
4	0.398	1275	26.45	14.52
5	0.205	1320.7	28.33	17.21
6	0.091	1636.2	30.94	18.1
Central				
1	0.958	709.8	19.57	10.57
2	0.937	810.8	22.12	13.35
3	0.821	993	24.52	14.85
4	0.68	1112.5	26.97	17.63

Table 9: (continued)

Birth order	C	λ	μ	Variance
5	0.518	1189.1	29.42	21.42
6	0.356	1492.8	31.37	20.68
North-east				
1	0.912	544.6	21.14	17.36
2	0.887	651.4	24.01	21.26
3	0.772	769.6	26.41	23.94
4	0.596	960.3	28.31	23.63
5	0.419	1167.8	30.21	23.61
6	0.29	1471.7	32.11	22.49
India				
1	0.945	572.1	20.28	14.59
2	0.9	706.4	22.84	16.87
3	0.724	916.1	24.96	16.97
4	0.508	1105.4	26.92	17.65
5	0.332	1198.1	29.08	20.53
6	0.209	1410.9	31.07	21.25

Table 10: Standard errors of the parameter estimates* of various birth order-specific fertility curves fitted to the cohort birth order-specific fertility schedules of India and its different regions.

Birth order	C	λ	μ
North			
1	0.13	5.26	0.273
2	0.129	5.759	0.289
3	0.133	7.546	0.327
4	0.109	8.162	0.33
5	0.11	11.248	0.45
6	0.082	12.069	0.443
South			
1	0.124	4.172	0.288
2	0.127	5.02	0.31
3	0.129	7.332	0.363
4	0.096	8.393	0.346
5	0.098	12.527	0.487
6	0.073	12.832	0.508

Table 10: (continued)

Birth order	C	λ	μ
East			
1	0.15	5.587	0.319
2	0.157	7.139	0.355
3	0.128	6.988	0.329
4	0.101	6.907	0.311
5	0.115	9.306	0.446
6	0.108	11.765	0.52
West			
1	0.163	5.708	0.381
2	0.152	6.158	0.377
3	0.112	6.242	0.304
4	0.11	9.345	0.371
5	0.092	11.084	0.451
6	0.09	18.157	0.669
Central			
1	0.153	6.24	0.309
2	0.131	5.78	0.284
3	0.134	6.999	0.318
4	0.128	7.785	0.349
5	0.121	8.692	0.395
6	0.095	9.236	0.37
North-east			
1	0.158	5.768	0.374
2	0.133	5.377	0.335
3	0.164	7.749	0.456
4	0.137	8.246	0.43
5	0.117	9.249	0.436
6	0.112	12.008	0.495
India			
1	0.107	3.934	0.237
2	0.104	4.346	0.244
3	0.102	5.429	0.266
4	0.083	5.816	0.261
5	0.065	5.85	0.262
6	0.067	8.279	0.344

Table 11: Parameter estimates of various birth order-specific fertility curves fitted to the period birth order-specific fertility schedules of the considered countries (mean (for Wald's model is mean) and variance* of conditional probability distributions are also given).

Birth order	c	λ	μ	Variance	c	m	σ_1	σ_2	Mean	Variance
Austria										
1					0.692	27	5.92	5.75	26.94	34.05
2	0.498	953.6	30.23	28.96						
3	0.174	1254.8	32.29	26.83						
4	0.051	1478	33.63	25.73						
Canada										
1					0.749	27.55	6.13	5.47	27.29	33.69
2	0.585	1048.9	31	28.41						
3	0.222	1151.6	31.84	28.04						
4	0.072	1145.2	32.31	29.46						
Czech Republic										
1					0.726	27.12	5.21	4.26	26.74	22.51
2	0.533	1722.2	30.81	16.99						
3	0.144	1805.2	33.83	21.44						
4	0.038	1309.5	34.74	32.01						
Estonia										
1					0.7391	24.988	5.1239	4.9418	24.9154	25.3326
2	0.5751	1088.1	30.1412	25.17						
3	0.2201	1690.9	33.7985	22.83						
4	0.0605	2065.1	35.5698	21.79						

Table 11: (continued)

Birth order	c	λ	μ	Variance	c	m	σ_1	σ_2	Mean	Variance
Finland										
1					0.798	27.35	5.48	5.5	27.36	30.09
2	0.632	1034.8	30.62	27.74						
3	0.279	1330.6	32.48	25.74						
4	0.097	1282.9	33.2	28.53						
Hungary										
1					0.626	27.23	6.44	4.95	26.64	32.63
2					0.406	30.12	5.64	4.51	29.67	25.84
3					0.161	31.42	7.3	5.48	30.7	41.13
4	0.059	579	32.54	59.5						
Netherlands										
1	0.83	1069.8	29.27	23.44						
2	0.643	1613.6	31.59	19.53						
3	0.224	2025.3	33.19	18.06						
4	0.057	2079	34.25	19.33						
Russia										
1	0.799	654.5	23.77	20.53						
2	0.538	903.8	29.53	28.49						
3	0.148	1115	32.36	30.41						
4	0.039	1263.7	33.64	30.12						
Slovakia										
1					0.697	26.22	6.08	5.1	25.84	31.35
2					0.433	29.35	5.63	4.57	28.93	26.1
3	0.161	753.6	31.95	43.3						
4	0.058	500.7	31.43	62						

Table 11: (continued)

Birth order	c	λ	μ	Variance	c	m	σ_1	σ_2	Mean	Variance
Slovenia										
1	0.758	1062.3	28.02	20.71						
2	0.546	1707	30.84	17.19						
3	0.16	2077	33.16	17.55						
4	0.039	2157.3	34.91	19.73						
Sweden										
1					0.888	28.1	5.34	4.86	27.91	26.03
2	0.685	1326.9	31.44	23.43						
3	0.25	1716	33.51	21.93						
4	0.068	1745.9	34.7	23.94						
USA										
1					0.88	23.8	5.58	7.66	24.62	44.22
2					0.706	27.51	6.2	6.55	27.65	40.65
3	0.372	687.8	29.86	38.71						
4	0.151	844.3	30.89	34.9						

*Variance of the fertility distribution is obtained by multiplying variance of the conditional probability distribution with the reciprocal of the squared value of c^2. Mean of the fertility distribution is same as the mean of the conditional probability distribution.

Table 12: Standard errors of the parameter estimates* of various birth order-specific fertility curves fitted to the period birth order-specific fertility schedules of the considered countries.

Birth order	c	λ	μ	c	m	σ_1	σ_2
Austria							
1				0.112	0.33	0.374	0.38
2	0.085	5.57	0.308				
3	0.054	6.91	0.324				
4	0.032	8.16	0.348				
Canada							
1				0.103	0.305	0.364	0.333
2	0.121	7.64	0.4				
3	0.069	7.43	0.369				
4	0.037	6.93	0.35				
Czech Republic							
1				0.121	0.298	0.359	0.323
2	0.11	9.45	0.334				
3	0.067	11.24	0.412				
4	0.042	11.67	0.56				
Estonia							
1				0.158	0.36	0.426	0.396
2	0.097	6.3	0.313				
3	0.076	10.03	0.387				
4	0.045	12.6	0.433				
Finland							
1				0.116	0.32	0.362	0.367
2	0.115	6.94	0.363				
3	0.086	8.88	0.4				
4	0.051	8.77	0.413				
Hungary							
1				0.126	0.365	0.456	0.385
2				0.088	0.35	0.425	0.377
3				0.063	0.6	0.713	0.661
4	0.052	7.7	0.671				

Table 12: (continued)

Birth order	c	λ	μ	c	m	σ_1	σ_2
Netherlands							
1	0.127	6.86	0.337				
2	0.108	8.12	0.308				
3	0.069	9.96	0.329				
4	0.034	9.80	0.325				
Russia							
1	0.097	4.15	0.256				
2	0.091	5.61	0.317				
3	0.055	7.21	0.372				
4	0.028	7.61	0.367				
Slovakia							
1				0.128	0.342	0.414	0.371
2				0.076	0.307	0.375	0.328
3	0.085	8.65	0.602				
4	0.055	7.53	0.718				
Russia							
1	0.101	5.67	0.271				
2	0.084	7.06	0.252				
3	0.058	9.93	0.322				
4	0.037	13.14	0.430				
Slovakia							
1				0.113	0.279	0.323	0.313
2	0.124	8.2	0.361				
3	0.079	9.88	0.374				
4	0.037	9.02	0.346				
Slovakia							
1				0.19	0.548	0.581	0.695
2				0.125	0.45	0.504	0.521
3	0.083	5.32	0.377				
4	0.047	5.25	0.325				

*Variance of the fertility distribution is obtained by multiplying variance of the conditional probability distribution with the reciprocal of the squared value of c^2. Mean of the fertility distribution is same as the mean of the conditional probability distribution.

Appendix B

Proof of Lemma.

Moment Generating Function (MGF)

$$\begin{aligned}MGF &= E(e^{tx})\\ &= \int_{-\infty}^{\infty} e^{tx} f(x)dx\\ &= \int_{-\infty}^{m} e^{tx}\left(\frac{1}{\sigma_1\sqrt{2\Pi}}\right)\exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma_1}\right)^2\right)dx+\\ &\int_{m}^{\infty} e^{tx}\left(\frac{1}{\sigma_2\sqrt{2\Pi}}\right)\exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma_2}\right)^2\right)dx.\end{aligned}$$

Taking $x-m/\sigma_1 = z_1$and $x-m/\sigma_2 = z_2$, the above expression can be rewritten as

$$\begin{aligned}MGF = &\int_{-\infty}^{0} e^{t(\sigma_1 z_1+m)}\left(\frac{1}{\sqrt{2\Pi}}\right)\exp\left(-\frac{1}{2}z_1^2\right)dz_1\\ &+\int_{0}^{\infty} e^{t(\sigma_2 z_2+m)}\left(\frac{1}{\sqrt{2\Pi}}\right)\exp\left(-\frac{1}{2}z_2^2\right)dz_2\\ &= e^{tm}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{-\infty}^{0} e^{t\sigma_1 z_1-\frac{1}{2}z_1^2+\frac{1}{2}t^2\sigma_1^2-\frac{1}{2}t^2\sigma_1^2}dz_1\\ &+ e^{tm}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{0}^{\infty} e^{t\sigma_2 z_2-\frac{1}{2}z_2^2+\frac{1}{2}t^2\sigma_2^2-\frac{1}{2}t^2\sigma_2^2}dz_2\\ &= e^{tm+\frac{1}{2}t^2\sigma_1^2}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{-\infty}^{0} e^{-\frac{1}{2}(z_1-t\sigma_1)^2}dz_1\\ &+ e^{tm+\frac{1}{2}t^2\sigma_2^2}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{0}^{\infty} e^{-\frac{1}{2}(z_2-t\sigma_2)^2}dz_2.\end{aligned}$$

By taking $z_1 - t\sigma_1 = y_1$and $z_2 - t\sigma_2 = y_2$it can be shown that the above expression is equal to

$$MGF = e^{tm+\frac{1}{2}t^2\sigma_1^2}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{-\infty}^{-\sigma_1 t} e^{-\frac{1}{2}y_1^2}dy_1 + e^{tm+\frac{1}{2}t^2\sigma_2^2}\left(\frac{1}{\sqrt{2\Pi}}\right)\int_{-t\sigma_2}^{\infty} e^{-\frac{1}{2}y_2^2}dy_2.$$

$$\therefore M_X(t) = e^{mt+(1/2)t^2\sigma_1^2}\,\Phi\left(-t\sigma_1\right) + e^{mt+(1/2)t^2\sigma_2^2}(1-\Phi\left(-t\sigma_2\right)),$$

where $\Phi(x)$ is the cumulative probability distribution function of standard normal distribution.

Mean

$$Mean = \frac{d}{dt}M_X(t)|_{t=0}.$$

Using the result

$$\frac{d}{dt}\left(\int_{\xi(t)}^{\psi(t)} f(x)dx\right) = f(\psi(t))\,\frac{d}{dt}\,(\psi(t)) \;-\; f(\xi(t))\,\frac{d}{dt}\,(\xi(t))\,,$$

it is possible to show that

$$\begin{aligned}\frac{d}{dt}M_X(t)|_{t=0} &= \left[e^{mt+(1/2)t^2\sigma_1^2}\Phi'\left(-t\sigma_1\right) + \Phi(-t\sigma_1)e^{mt+(1/2)t^2\sigma_1^2}(m+t\sigma_1^2)\right] + \\ &\quad \left[e^{mt+(1/2)t^2\sigma_2^2}(-\Phi'(-t\sigma_2)) + (1-\Phi(-t\sigma_2))e^{mt+(1/2)t^2\sigma_2^2}(m+t\sigma_2^2)\right] \\ &= \left[\frac{-\sigma_1}{\sqrt{2\Pi}} + \frac{m}{2}\right] + \left[\frac{\sigma_2}{\sqrt{2\Pi}} + \frac{m}{2}\right] \\ &= m + \frac{1}{\sqrt{2\Pi}}\left(\sigma_2 - \sigma_1\right).\end{aligned}$$

$$\therefore Mean = m + \left(\frac{\sigma_2 - \sigma_1}{\sqrt{2\Pi}}\right)$$

<u>*Variance*</u>

$$E(X^2) = \frac{d}{dt^2}M_X(t)|_{t=0}$$

Proceeding in the same way as was done previously, it can be shown that

$$E(X^2) = \left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right) + m^2 + \frac{2m}{\sqrt{2\pi}}\left(\sigma_2 - \sigma_1\right),$$

and

$$\begin{aligned}variance(X) &= E(X^2) - [E(X)]^2 \\ &= \left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right) - \frac{1}{2\pi}\left(\sigma_2 - \sigma_1\right)^2.\end{aligned}$$

$$\therefore Variance = \left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right) - \frac{1}{2\pi}\left(\sigma_2 - \sigma_1\right)^2.$$

It can also be shown that

$$Median = m,$$

$$1^{st} quartile = m + \sigma_1(-0.67449),$$

and

$$3^{rd} quartile = m + \sigma_2(0.67449).$$

Samba Siva Rao Pasupuleti
Population Studies Unit
Indian Statistical Institute
Kolkata, India
E-mail: srao113@gmail.com

Asis Kumar Chattopadhyay
Department of Statistics
Calcutta University
Kolkata, India
E-mail: akcstat@caluniv.ac.in

Paper received: 10 January 2012; revised: 27 September 2012.