Volume Title: Aging Issues in the United States and Japan

Volume Author/Editor: Seiritsu Ogura, Toshiaki Tachibanaki and David A. Wise, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-62081-6

Volume URL: http://www.nber.org/books/ogur01-1

Publication Date: January 2001

Chapter Title: What Went Wrong with the 1991â€'92 Official Population Projection of Japan?

Chapter Author: Seiritsu Ogura

Chapter URL: http://www.nber.org/chapters/c10295

Chapter pages in book: (p. 361 - 402)

12

What Went Wrong with the 1991–92 Official Population Projection of Japan?

Seiritsu Ogura

12.1 Introduction

The track records of the last two official Japanese population projections have been very poor, judging from the accuracy of the figures for new births (fig. 12.1). Every five years, the National Institute of Population Research, which is now the National Institute of Social Security and Population Research, prepares three different sets of projections based on middle, high, and low fertility assumptions. The projection based on the middle fertility assumption is treated as the official population projection for Japan. This projection is linked almost automatically to many long-term plans the government prepares, from Public Pensions Projections to Energy Projections, and hence any sizeable error in the projection seriously affects resource allocation over time and income distribution across generations.

The 1986 projection, for instance, predicted a recovery of the Total Fertility Rate (TFR) to 1.86 but ended up overestimating 1991 births by almost 300,000. In fact, TFR dropped from 1.72 in 1986 to 1.51 in 1991. In spite of this experience, in the 1991–92 projection, demographers at the Institute again predicted a recovery in TFR beginning in 1995. At this time, however, there is no sign of the predicted recovery and, in fact, TFR continued to slip by another 0.1 during the five-year period. Thus, the projection ended up overestimating 1996 births by more than 120,000, and the
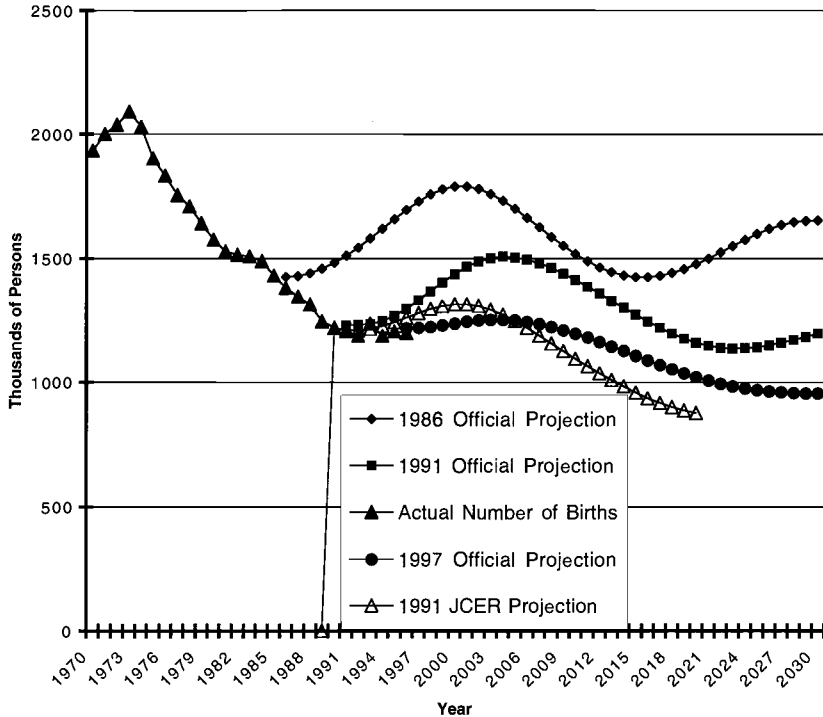
Fig. 12.1   Actual number of live births and official population projection in Japan

gap between projected and actual figures would have reached almost a quarter of a million early in the twenty-first century. Because all the Japanese social insurance programs are essentially income transfer mechanisms between the old and the young generations, these optimistic fertility assumptions allowed the government to understate the burdens on future generations very substantially. In a sense, these optimistic fertility projections also allowed the government to adopt a long series of ad hoc measures to solve only the immediate crises, without addressing the problems that occur in the long run.

From the purely technical point of view, to be sure, the 1991–92 projection was an innovative one. As we will see in detail shortly, it tried to estimate age-specific fertility rates of female cohorts by fitting their records to a class of special statistical distribution functions. This procedure allowed fertility behavior to be represented by a small number of parameters, which can be used to predict future fertility for this cohort as well as for future generations. Thus armed, the projection confidently declared that the observed decline in the fertility rate of Japanese women had been a temporary phenomenon resulting from delayed marriage and childbear-

ing, and that both would end by the mid-1990s, raising TFR to where it was in the middle of 1980s.

Unfortunately, this attempt to introduce rigorous science into the population projection rather than inventing fertility rates out of thin air was not very successful in terms of its outcome; it seems to have been abandoned in the latest projection. In fact, in the 1997 projection (which was made public in January 1997), the projection of fertility rates is based on marriage rates and the marital fertility rates of recent cohorts, as I had done in the last two Japan Center for Economic Research (JCER) population projections. Because marriages precede births by many years, and given substantial information on the marriage behavior of a cohort, one using the information should be able to predict the future course of the number of births by the cohort far more precisely than by, say, using their age alone.

This paper attempts to show the following: The government demographers failed in their 1991–92 projection because there was a fundamental flaw in their methodology. In fact, if we apply the same methodology in the 1997 projection, there is hardly any need to change their 1991–92 projection. Thus, their methodologically "correct" projection will continue to diverge from reality for another five years, when they are scheduled for another projection.

The rest of this paper is constructed as follows. In section 12.2, I present the formal demographic model used in the 1991–92 projection and report my own estimation results using new data made available since 1992. In section 12.3, I formulate a marriage/birth model and explore the possibility of misspecification as a source of the government model's sensitivity to truncation. In section 12.4, I formulate the age distribution of marital fertility rates and report my estimation results. In section 12.5, I look at significant changes in the marital behavior of Japanese women that took place in the last twenty years. In section 12.6, I analyze the causes of the drop in the fertility rates observed among three different cohorts, almost five years apart. In section 12.7, I provide concluding remarks.

12.2   Official 1992 Projection Model: Direct Estimation of Birth Rates

In the Japanese official population projection of 1992, the probability distribution of giving birth to a child of a particular order (first child, second child, etc.) in a female cohort is fitted to an incomplete log-gamma distribution function. Although it is not usually treated in introductory statistics, a complete log-gamma probability distribution that seems to be popular among demographers and sociologists (e.g., Coale 1971; Coale and McNeil 1972; Kaneko 1991) can be written as

$$(1) \qquad g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}|\mu|\exp\left\{\alpha\mu(x - u) - \frac{1}{\beta}\exp[\mu(x - u)]\right\},$$

with four distribution parameters (α, β, μ, u). Multiplying this function by a positive factor, C, which is less than 1, generates the incomplete log-gamma distribution function. Kaneko (1993), for instance, presents the distribution function used for the official population projection (Kenky-usho 1992) as

$$(2) \quad h(x) = \frac{C|\lambda|}{b\Gamma(\lambda^{-2})}\left(\frac{1}{\lambda^2}\right)^{\lambda^{-2}} \exp\left\{\frac{1}{\lambda}\left(\frac{x-u}{b}\right) - \frac{1}{\lambda^2}\exp\left[\lambda\left(\frac{x-u}{b}\right)\right]\right\}.$$

In equation (2), C is the incompleteness factor, which we shall ignore for the moment. By comparing equations (1) and (2), we can conclude initially that the parameters in these two expressions are related in the following manner:

$$(3) \qquad\qquad\qquad \alpha = \frac{1}{\lambda^2},$$

$$(4) \qquad\qquad\qquad \beta = \lambda^2,$$

$$(5) \qquad\qquad\qquad \mu = \frac{\lambda}{b}.$$

Second, I find that equation (2) imposes a restriction on the parameters of equation (1) such that

$$(6) \qquad\qquad\qquad \alpha\beta = 1,$$

where $\alpha\beta$ is equal to the mean of the gamma variable in the original gamma distribution. Thus equation (2) is equivalent to

$$(2') \quad g(x;a,\mu,u) = \frac{\alpha^\alpha}{\Gamma(\alpha)}|\mu|\exp\{\alpha\mu(x-u) - \alpha\exp[\mu(x-u)]\}$$

in my notation.

Finally, I define a new distribution function $f(x)$ obtained by multiplying equation (1) by a positive constant factor C that is smaller than 1, namely,

$$(7) \qquad\qquad h(x) = Cg(x;\alpha,\mu,u).$$

The cumulative distribution function for h is then defined by

$$(8) \qquad\qquad H(x) = \int_0^x h(\tau;a,\mu,u)dt = CG(x),$$

where $G(x)$ is the cumulative distribution function for g, or

$$(9) \qquad\qquad G(x) = \int_0^x g(\tau;a,\mu,u)dt.$$

Kaneko (1993) provides a fairly detailed account of how estimates of the parameters for different cohorts were obtained. He applied a standard

probability model in a multinomial distribution, where the probability of giving birth at age i, denoted by P_i, is defined as

$$(10) \qquad P(i) \; = \; H(i) \; - \; H(i \; - \; 1).$$

Now, suppose a cohort has accumulated complete data regarding the number of births at each age between fifteen and forty-nine, $(m_{16}, m_{16}, \ldots, m_{49})$. Because the probability of such a sequence of events taking place is

$$(11) \qquad L \; = \; P(15)^{m_{15}} \cdot P(16)^{m_{16}} \cdot \cdot P(i)^{m_i} \cdot \cdot P(49)^{m_{49}},$$

the parameters of an incomplete gamma distribution function for this cohort must be set in such a way as to maximize this joint probability, L.

12.2.1   A Formulation Involving Right-Censored Data

If one insists on using only the cohort data with completed birth data, however, then the latest cohort available for obtaining distribution parameters for the last official projection (1991–92) would have been the cohort born in 1941. In estimating the degree of structural changes taking place, even the data from this cohort will have very limited value. As a practical matter, therefore, we have to find a way to extract information from those cohorts that are still in the middle of their child-bearing ages. One way to do so is to apply the same principle as above but allow for the possibility of right-censoring.[1]

Suppose that the age-specific birthrate data for a first child are available up to the age $x - 1$ for a particular cohort. Then a woman in this cohort either has already given birth to a child at some age between fifteen and $x - 1$, or she has not yet given birth by age $x - 1$. Suppose the proportion of women in the second group in the cohort is given by $S(x)$. The probability of observing a woman having a child between ages fifteen and $x - 1$ is given by

$$(12a) \qquad P(15)^{m_{15}} \cdot P(16)^{m_{16}} \cdot \cdot P(x \; - \; 1)^{m_{x-1}},$$

while the probability of observing $S(x)$ women not yet having a child is given by

$$(12b) \qquad \left[1 \; - \; \sum_{i=15}^{x-1} P(i) \right]^{S(x)}.$$

The product of these two probabilities gives the probability of the observed experience. Accordingly, the log-likelihood function is given by

$$(13) \qquad \ln L \; = \; \sum_{i=15}^{x-1} m_i \ln P(i) \; + \; S(x) \ln \left[1 \; - \; \sum_{i=15}^{x-1} P(i) \right].$$

1. I do not consider, however, using the left-censored data.

Since all the $P(i)$s are constructed by equation (10) from a log-gamma distribution function, they are functions of the parameters of the original log-gamma distribution function. Denoting a particular parameter of the log-gamma distribution function by θ, we can estimate its value by maximizing equation (13) with respect to θ. The first-order condition for maximization is given by

$$(14) \quad \frac{\partial \ln L}{\partial \theta} = \sum_{i=15}^{x-1} \frac{m_i}{P(i;\theta)} \frac{\partial P(i;\theta)}{\partial \theta} - \frac{S(x)}{1 - \sum_{i=15}^{x-1} P(i;\theta)} \sum_{i=15}^{x-1} \frac{\partial P(i;\theta)}{\partial \theta} = 0.$$

12.2.2   Updating the 1991–92 Estimation

Using the age-specific fertility rate of Japanese women born between 1950 and 1966, and considering their first, second, and third children, I estimate the five parameters of the incomplete log-gamma distribution functions. The fertility data used in the estimation are annual data from 1979 to 1994. The likelihood function is first maximized analytically with respect to Cs to yield a concentrated log-likelihood function, which is then optimized with respect to α, β, μ, and u using a Gauss-Newton method with a Berndt, Hall, Hall, and Hausman (BHHH) method for Hessian matrices. For computer software, I used GAUSS. The results are shown in tables 12.1 through 12.3 for selected years.

There are three possible sources of difference between my results and those obtained by the government demographers as a part of the previous official population projection. First, in my computation, I added data for three years (1992 to 1994) in the truncated estimation. Second, there are minor differences in the fertility data used, because I constructed my own data rather than using the cohort fertility data published by the National Institute of Population Research.[2] Third, the government's 1991–92 projection involves an error correction function on top of the estimated theoretical distribution, which I decided to skip because they perform, in my judgment, merely cosmetic adjustments to augment its goodness of fit.

The results of these updated estimations are summarized as follows:

1. According to table 12.1, the estimated completed fertility rate with respect to the first child seems to have come down from 0.90 of the cohort born in 1950 to around 0.81 of that born in 1960. The completed fertility rate seems to stay slightly below 0.8, but it seems to bounce back to 0.84 for those born in 1965. In any case, the official projection assumed the value of 0.821 as the long-run equilibrium value, and it seems to have been more or less appropriate.

---

2. This is done so that I can compare the results with another method, which I will explain later.

Table 12.1   Estimated Parameters for the Age-Specific First Child Fertility Rate (incomplete log-gamma function), by Birth Year

	1950	1955	1960	1961	1962	1963	1964	1965
α	1.4698	2.2380	3.4161	4.2767	3.6792	4.6745	4.5118	3.5082
β	0.2384	0.3474	0.2718	0.2646	0.3200	0.4038	0.4329	0.3504
μ	−0.2832	−0.2037	−0.1577	−0.1372	−0.1428	−0.1217	−0.1164	−0.1199
u	20.7266	23.8383	25.2534	26.7872	27.1037	31.4160	32.2879	28.7668
C	0.8991	0.8750	0.8143	0.7998	0.7928	0.7787	0.7844	0.8345
Likelihood	−2.6660	−2.7111	−2.5831	−2.5318	−2.4555	−2.3458	−2.1896	−1.9978
FOC								
α	−0.00382	0.00010	−0.00010	−0.00010	0.01073	−0.00009	−0.00006	−0.00003
β	−0.10808	−0.00006	0.00007	0.00010	0.12348	0.00010	0.00010	0.00010
μ	0.04461	−0.00006	−0.00002	−0.00007	0.05385	−0.00010	−0.00009	−0.00005
u	0.00191	0.00000	0.00000	0.00000	−0.00574	0.00000	−0.00001	0.00000

Source: All tables are based on author's estimation, as described in the text.

2. According to table 12.2, for the second child the story is similar to that for the first child. The estimated completed fertility rate with respect to the second child seems to have come down from 0.78 for the cohort born in 1950 to around 0.69 for that born in 1960. The official projection used the value of 0.685 as the long-run equilibrium. Again, a recovery seems to be taking place after the 1963 cohort, and for the 1965 cohort, fertility is projected to be 0.77.

3. The change in the fertility rate for a third child, on the other hand, seems to have followed a very different pattern (see table 12.3). The estimated completed fertility rate with respect to the third child was 0.26 for the 1950 cohort and reached 0.29 for the 1957 cohort. After a slight decline, it seems to be picking up again, reaching 0.32 for the 1965 cohort. The official projection used the value of 0.2510 as the long-run equilibrium, which seems to be rather modest.

Thus, if the government demographers had carried out the identical estimation of completed fertility rates in 1997 for the 1964 cohort, they should have obtained 0.8 for the first child, 0.69 for the second child, and 0.30 for the third child. Furthermore, they should have shown the fertility rates rising in more recent cohorts. There seems to be no reason for them to lower the long-run TFR of 1.8 assumed in the 1991–92 projection. In the 1997 population projection, however, the government did lower the long-run TFR level to 1.6, which is an admission of the general failure of the procedure used in 1991–92 projection.

12.3   The Implications of Conditioning Births to Marriages

In Japan, in contrast to most of the other developed countries, an overwhelming majority of births still take place within marriages, particularly in first marriages. Because marriages precede births by many years, it is natural to expect that one should be better able to forecast the number of future births by explicitly using information on marriages. One way to do this is to treat all the women in a given cohort who have married at a given age as a homogeneous group, because there seems to be a very stable (nonlinear) relationship between the number of years married and childbearing behavior.

Suppose that the probability is denoted by

(15) $\pi(s, k - s)$,

for a k-year-old woman who has been married for s years, to give a birth to a child of any given order. Here $k - s$ denotes the age at which she was married, which will be referred to as her marital age. (I ignore the possibility of divorce here.) Suppose also that there are $M(k - s)$ women at age k (or, in the same cohort) who were married at age $(k - s)$. In this birth order, the expected number of births from these women is given by

Table 12.2   Estimated Parameters for the Age-Specific Second Child Fertility Rate (incomplete log-gamma function), by Birth Year

	1950	1955	1960	1961	1962	1963	1964	1965
α	3.3370	3.9881	3.4645	3.8341	3.1418	3.1103	3.0436	2.9690
β	0.3010	0.2848	0.2894	0.2826	0.2954	0.2923	0.2813	0.2805
μ	−0.1721	−0.1488	−0.1523	−0.1393	−0.1447	−0.1362	−0.1267	−0.1191
u	27.4225	28.7419	28.2974	29.0294	28.1339	28.2849	28.3510	28.6731
C	0.7770	0.7605	0.6932	0.6819	0.6819	0.6894	0.7215	0.7658
Likelihood	−2.5745	−2.5616	−2.2618	−2.1423	−1.9759	−1.7758	−1.5314	−1.2692
FOC								
α	0.000705	−0.000268	−0.000326	−0.000206	−0.000170	−0.000077	0.000034	0.000026
β	0.000361	−0.000867	0.000437	0.000336	0.000084	−0.000282	0.000437	0.000096
μ	−0.000117	0.000600	−0.000983	−0.000965	−0.000997	−0.000630	0.000831	0.000936
u	−0.000019	0.000037	−0.000019	−0.000013	−0.000004	0.000011	−0.000016	−0.000003

Source: See table 12.1.

Table 12.3   **Estimated Parameters for the Age-Specific Third Child Fertility Rate (incomplete log-gamma function), by Birth Year**

	1950	1955	1960	1961	1962	1963	1964	1965
α	3.2266	3.8397	3.4810	3.4253	2.9392	3.0554	3.0323	2.5283
β	0.2672	0.2678	0.2839	0.2639	0.2759	0.2623	0.2477	0.2671
μ	−0.1609	−0.1448	−0.1365	−0.1302	−0.1303	−0.1211	−0.1124	−0.1155
u	28.9391	30.6122	30.9929	30.5911	30.1550	30.3962	30.3123	30.0219
C	0.2569	0.2881	0.2793	0.2797	0.2883	0.2943	0.3079	0.3221
Likelihood	−1.2605	−1.3376	−1.0233	−0.8961	−0.7559	−0.6111	−0.4689	−0.3444
FOC								
α	−0.0008	−0.0005	−0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
β	−0.0008	0.0009	−0.0001	0.0007	0.0000	−0.0002	−0.0001	0.0006
μ	−0.0004	−0.0001	−0.0010	0.0008	−0.0004	−0.0007	0.0005	0.0010
u	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Source: See table 12.1.

(16) $$\pi(s, k - s)M(k - s).$$

I assume that the lowest age of marriage is fifteen and the maximum age of reproduction is fifty. Assuming that it takes a year to give birth to a child after marriage, in equation (16), k goes from 16 to 50 and there will be $(k - 16)$ possible values of s. The expected number of births from women at age k, or $B(k)$, can be obtained as the sum of equation (16) over these $(k - 16)$ possible values of s; or

(17) $$B(k) = \sum_{s=1}^{k-16} \pi(s, k - s)M(k - s).$$

Equation (17) points to the source of the fundamental problem in the direct approach in the official 1991–92 projection. To see this more clearly, for the moment, we can take $M(k - s)$ as weights for $\pi(s, k - s)$. Let us denote the relevant characteristics of women who married at age $(k - s)$ by ω_{k-s}. We emphasize this by rewriting equation (17) as

(18) $$B(k) = \sum_{s=1}^{k-16} \pi(s, k - s, \omega_{k-s})M(k - s).$$

Thus, when we estimate an age-specific distribution function using the birthrate data up to age k, we implicitly take into account in the process all the relevant characteristics up to $k - 1$, or the information set at age k is given by

(19) $$I(k) = (\omega_{15}, \omega_{16}, \ldots, \omega_{k-1}; M_{15}, M_{16}, \ldots, M_{k-1}),$$

as far as this cohort is concerned.

In the next year, when data on $B(k + 1)$ become available,

(20) $$I(k + 1) = (\omega_{15}, \omega_{16}, \ldots, \omega_{k-1}, \omega_k; M_{15}, M_{16}, \ldots, M_{k-1}, M_k),$$

will be used for the estimation. However, it is not immediately clear how important this additional information can be in obtaining the correct parameters of the distribution. I have conducted experiments by controlling the last data used in the estimation, and have found that in the cohorts born after 1960, the estimated parameters, particularly C, are generally very sensitive to the addition of new data. In some cases, adding even a year's data completely changed the value of their s. A number of factors can contribute to this phenomenon; for instance,

1. *Misspecification problem.* If the $\pi(s, k - s)$s are not generated by an identical exponential distribution function, then the resulting age-specific distribution is no longer a log-gamma function as specified earlier. In such a case, the model is misspecified and the parameters are not consistent; hence, they may be unstable as more data are added (see appendix).

2. *Heterogeneity problem.* A cohort may not be very homogeneous.
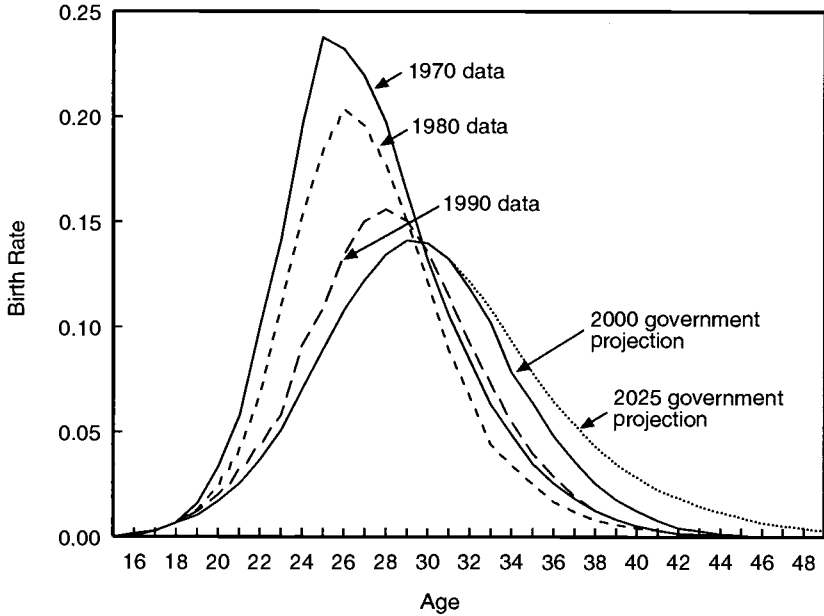
Fig. 12.2   Changes in the age-specific birthrates of Japanese women

Those married later may be very different from those married younger. Furthermore, those married later may be making choices under a very different set of circumstances from those of individuals who married younger. Heterogeneity thus may result in considerable and complex differences in the timing of having children after marriage that may not be captured by the very limited number of parameters in a log-gamma distribution function.

3. *Truncation and log-gamma variable.* Since the value of the log-gamma variable is defined from negative infinity to positive infinity, I may be getting distributions that technically fit well with the existing data but that have too much area in the truncated region to be consistent with human behavior. Such an example can be found in the 1992 official projection, in which women in their late forties are supposed to be giving birth almost as often as women in their late thirties today (fig. 12.2).

12.4   Estimation of Marital Fertility Functions

12.4.1   Beta Distribution Specification

In order to implement the fertility-rate projection on the basis of equation (17), we have to estimate both age distribution functions of marital

fertility rates conditioned on marital ages, or $(k - s)$, and age distribution functions of first marriages, or $m(k - s)$. For the former, I have selected incomplete beta distribution functions, and for the latter, incomplete log-gamma functions. A beta distribution is defined as

$$(21) \qquad f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}, \qquad 0 < x < 1.$$

The mean and the variance of this distribution are given by

$$(22) \qquad\qquad\qquad \mu = \frac{\alpha}{\alpha + \beta};$$

$$(23) \qquad\qquad\qquad \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}.$$

Thus, given β, a larger α moves the mode of the distribution to the right, and makes the mountain taller. Proportional increases in both α and β do not change the location of the mode in terms of the horizontal axis, but make the mountain taller and steeper as they reduce the variance. Although a beta function is not nearly as flexible as a log-gamma function given by equation (1), its domain is between 0 and 1, and it will not give us a phantom baby boom in women in their forties or fifties (as the government's projection does).

We must account for the possibility of women's not having children of that particular order or higher, and hence, I have added an incompleteness factor, C_M, to the beta distribution function. Thus, for each age of marriage in a cohort, the distribution of marital fertility rate is assumed to be given by

$$(24) \qquad f(x) = C_M \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1}, \qquad 0 < x < 1,$$

where the unit interval starts at the age of marriage and ends at age fifty. I again have accounted for the right-censoring problem by constructing a log-likelihood function similar to equation (15); I then optimize the likelihood function with respect to C_M and transform it into a concentrated likelihood function.

12.4.2   Data Problems

Underreporting

Marital fertility-rate data for each cohort have been computed from the vital statistics table on the number of first marriages for each woman's age and the cross-tabulation of the mother's age and the number of years married for each order of birth. The latter figures (seem to) include the births

from couples who had started living together without formally reporting their marriages, but who decided to report just before the birth of a child. As a result, in some marital age groups, there are more first children born than the number of first marriages reported. Another possible explanation for this phenomenon could be the ignored second and third marriages, because I focused only on first marriages to economize the estimation. It is clear that second marriages will soon become an important factor that should not be ignored in the analysis of birth rates.

In order to deal with this problem in the simplest possible way, I multiplied all the first marriages by a factor of 1.25 in the denominator of marital fertility rate. This factor was chosen because, prior to this adjustment, the highest cumulative marital fertility rate was recorded at 1.21 by those born in 1950 and married at age thirty. Thus, the marital fertility rates I actually used are about 25 percent smaller than the ratio of the number of children born and the number of first marriages in all marital age groups.

Similarly, I noticed some women giving birth to second or third children within a year of their marriages; most of these women are presumably married for the second or the third time. In figure 12.3, I show the age distribution of the marital fertility rate for the third child among those married at age thirty-three for three different cohorts. Clearly, this seems to be constructed as a composite of two different curves: one a downward-sloping curve and the other a single-peaked curve. In order to capture such reproductive behavior, I have assumed that the distribution function is a sum of a beta distribution function and an exponential distribution function, given by

$$(25) \quad f(x) = C_M \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1 - x)^{\beta-1} + C_K \delta \exp(-\delta x), \quad 0 < x < 1.$$

Censoring

Because Vital Statistics did not begin to publish the mother's age and the years married for each order of birth until 1979, my computation is limited to cohorts and their subgroups for which there is no left-censoring. Thus I have limited the exercise to groups of women on whom I have complete birth records starting from the first year of marriage and up to year 1994. If I were to estimate equation (25) for each cohort, because I have five parameters (namely, α, β, δ, C_M, C_K) to estimate, I would need at least six or seven observations for each cohort. The marital fertility data of cohorts that fall short of this requirement would not be used in the estimation at all. In order to avoid throwing out these data, I have assumed the values of α, β, and δ are common for all cohorts, but that the values of C_M and C_k vary across cohorts, and have used all the available data.
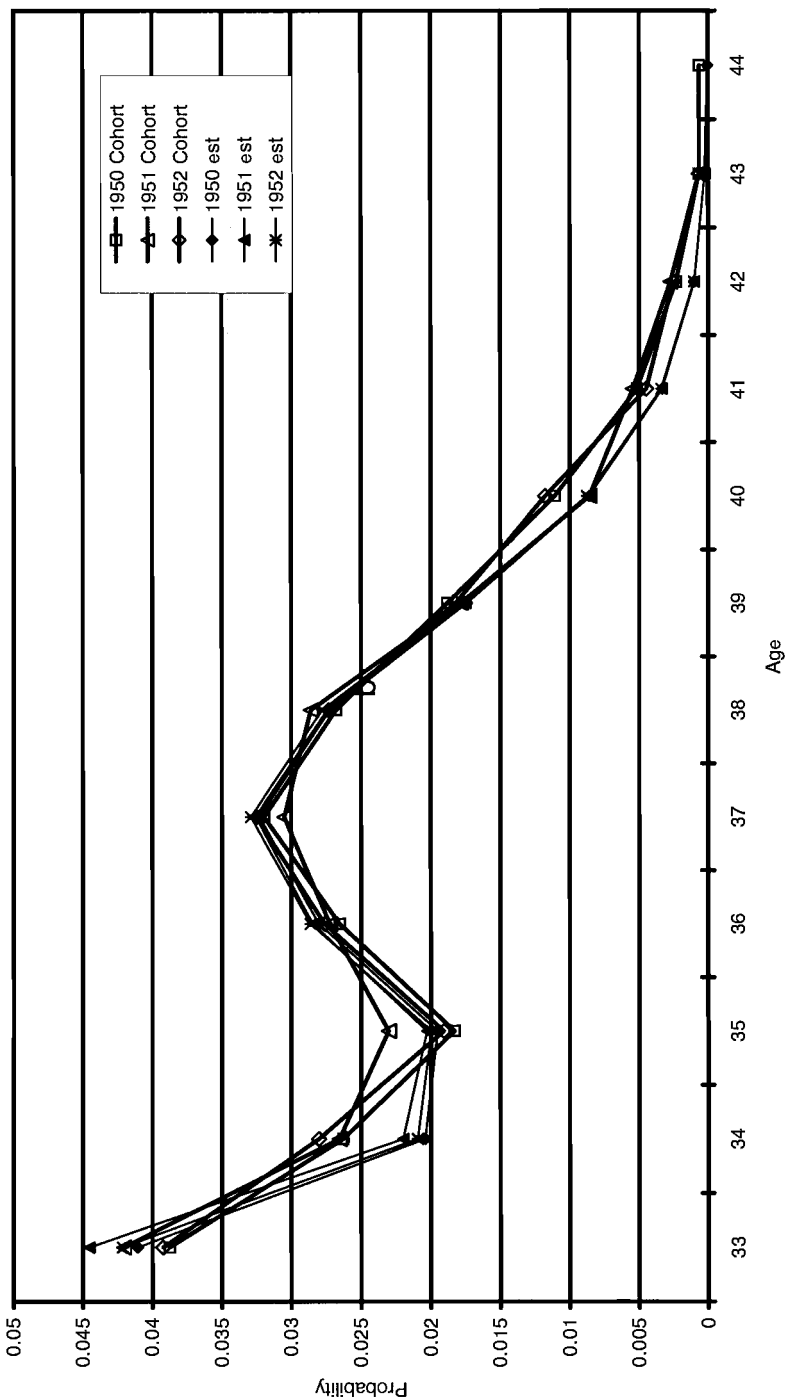
Fig. 12.3 Age distribution of third child fertility of women who married at age thirty-three (1950, 1951, and 1952 cohorts)

The actual estimation was carried out in two steps. In the first step, I estimated one set ($\hat{\alpha}$, $\hat{\beta}$, $\hat{\delta}$, $\hat{C}_M$, $\hat{C}_k$) by the maximum-likelihood method using all the available data. In the second step, by keeping the same values of ($\hat{\alpha}$, $\hat{\beta}$, $\hat{\delta}$), I estimated (C_M, C_k) for each cohort using only the data belonging to the cohort.

Cross-Cohort Comparisons of Marital Fertility Rates

Estimation result. The result of my estimation is summarized in tables 12.4 through 12.6, where only the sums of C_M and C_k are reported. The blank spaces in the tables represent the absence of estimates. There are two large blank areas in these tables: one in the upper-left region and the other in lower-right region. The upper-left blank region reflects the left censoring problem; that is, prior to 1979, the necessary data for estimation are not available as far as these groups are concerned. If earlier data on the marriage and childbearing behaviors are made available by the government, then the blank space in this region can be eliminated. Most of the second region, however, is attributable to the combination of the birth year and marital age that are yet to come. As such, they can never be filled in completely.

Estimation of the completed marital fertility rates. In order to construct estimates of completed fertility rates for the cohorts, we need a complete set of estimated marital fertility functions for all marital ages, and we must somehow fill the blank spaces in tables 12.4 through 12.6. To fill these spaces, I have taken a simple shortcut: For each order of birth, I take the log value of the estimated Cs, regress them on age-married, age-married squared, and age-married cubed, and cross terms with age-married (raised to the third power) and with birth year, age, and birth year dummies. More specifically, for a woman born in year j who was married at age k, the completed marital fertility rate for her sth-order child $C_{k,j}(s)$ is

$$(26) \qquad \ln C_{k,j}(s) = \alpha + \sum_{r=1}^{3} \beta_r(s) \times k^r + \sum_{r=1}^{3} \lambda_r(s) \times j \cdot k^r$$

$$+ \sum_{j=1950}^{1964} \delta_j(s) \text{Dummy}_j + \varepsilon_{k,j}(s),$$

where Dummy$_j$ is a dummy variable for the cohort born in year j.

As far as the first- and second-child fertility rates are concerned, this specification has shown a good fit. For the first child, as the marital age approaches forty, the completed fertility drops sharply. As for the second child, a sharp drop occurs at the marital age of thirty-five. For the third child, however, the coefficient of the cubed age term turns out to be positive, and the fertility rate does not fall to zero as the marital age increases. Thus, for the third child I have dropped the third-power terms altogether and estimated parameters of the following function:

Table 12.4   First Child Marital Fertility Rate

	18	19	20	21	22	23	24	25	26	27
1950										
1951										
1952										0.922
1953									0.927	0.946
1954								0.855	0.943	0.910
1955							0.789	0.859	0.911	0.892
1956						0.719	0.809	0.847	0.899	0.913
1957					0.682	0.727	0.764	0.827	0.884	0.887
1958				0.693	0.634	0.655	0.656	0.702	0.743	0.793
1959			0.627	0.667	0.678	0.701	0.762	0.829	0.812	0.918
1960		0.699	0.639	0.660	0.667	0.702	0.759	0.793	0.834	0.869
1961	0.710	0.705	0.630	0.656	0.682	0.711	0.728	0.805	0.784	0.829
1962	0.708	0.701	0.624	0.652	0.669	0.684	0.738	0.771	0.790	0.831
1963	0.710	0.688	0.645	0.665	0.676	0.695	0.711	0.738	0.793	0.772
1964	0.687	0.697	0.663	0.636	0.674	0.671	0.692	0.728	0.749	0.741

(*continued*)

Table 12.4   (continued)

1950	28	29	30	31	32	33	34	35	36	37	38
1950			0.968	0.919	0.863	0.862	0.845	0.785	0.759	0.602	0.540
1951	0.962	0.915	0.954	0.907	0.908	0.871	0.839	0.792	0.757	0.588	0.597
1952	0.976	0.878	0.961	0.943	0.910	0.833	0.828	0.786	0.693	0.628	0.594
1953	0.950	0.874	0.952	0.952	0.868	0.856	0.804	0.755	0.762	0.632	0.576
1954	0.937	0.884	0.965	0.883	0.880	0.845	0.790	0.764	0.619	0.620	
1955	0.956	0.873	0.922	0.915	0.859	0.796	0.782	0.757	0.712		
1956	0.946	0.852	0.953	0.886	0.820	0.798	0.778	0.774			
1957	0.914	0.851	0.897	0.847	0.845	0.794	0.796				
1958	0.878	0.869	0.833	0.854	0.872	0.856					
1959	0.882	0.828	0.883	0.860	0.863						
1960	0.831	0.842	0.880	0.902							
1961	0.841	0.788	0.895								
1962	0.786	0.754									
1963	0.757										

Source: See table 12.1.

Table 12.5 **Second Child Marital Fertility Rate**

	18	19	20	21	22	23	24	25	26	27	28
1950											
1951											0.694
1952										0.767	0.720
1953									0.791	0.779	0.704
1954								0.759	0.801	0.757	0.694
1955							0.707	0.769	0.775	0.751	0.714
1956						0.646	0.722	0.737	0.769	0.755	0.696
1957					0.611	0.650	0.677	0.721	0.762	0.736	0.666
1958				0.577	0.617	0.630	0.663	0.715	0.738	0.691	0.653
1959			0.518	0.585	0.610	0.620	0.672	0.710	0.697	0.675	0.614
1960		0.547	0.508	0.571	0.595	0.615	0.661	0.667	0.691	0.635	0.594
1961	0.529	0.542	0.526	0.562	0.602	0.619	0.630	0.658	0.640	0.603	0.607
1962	0.521	0.549	0.514	0.561	0.601	0.582	0.615	0.614	0.609	0.584	0.550
1963	0.509	0.531	0.528	0.571	0.584	0.584	0.580	0.590	0.590	0.520	0.540
1964	0.503	0.546	0.537	0.559	0.575	0.547	0.555	0.579	0.541	0.457	0.460

(*continued*)

Table 12.5   (continued)

	29	30	31	32	33	34	35	36	37
1950	0.656	0.629	0.552	0.497	0.435	0.365	0.279	0.221	0.135
1951	0.652	0.623	0.536	0.508	0.415	0.371	0.311	0.230	0.100
1952	0.646	0.620	0.558	0.491	0.427	0.351	0.284	0.193	0.110
1953	0.640	0.632	0.540	0.482	0.433	0.341	0.269	0.206	0.098
1954	0.656	0.626	0.519	0.478	0.395	0.321	0.290	0.198	0.082
1955	0.642	0.596	0.539	0.458	0.402	0.348	0.281	0.176	0.073
1956	0.621	0.608	0.487	0.426	0.406	0.334	0.278	0.182	
1957	0.615	0.561	0.493	0.461	0.375	0.314	0.234		
1958	0.573	0.550	0.501	0.419	0.359	0.283			
1959	0.553	0.557	0.458	0.390	0.344				
1960	0.557	0.523	0.445	0.382					
1961	0.511	0.499	0.425						
1962	0.494	0.474							
1963	0.454								

Source: See table 12.1.

Table 12.6   Third Child Marital Fertility Rate

	18	19	20	21	22	23	24	25	26	27
1950										
1951										
1952										0.269
1953									0.280	0.282
1954								0.281	0.292	0.280
1955							0.264	0.287	0.282	0.276
1956						0.263	0.271	0.272	0.280	0.278
1957					0.249	0.270	0.258	0.268	0.273	0.270
1958				0.242	0.253	0.259	0.250	0.260	0.263	0.253
1959			0.226	0.250	0.248	0.256	0.251	0.257	0.245	0.251
1960		0.234	0.225	0.241	0.239	0.256	0.243	0.232	0.239	0.242
1961	0.238	0.224	0.230	0.238	0.246	0.228	0.226	0.220	0.229	0.229
1962	0.221	0.217	0.228	0.230	0.232	0.217	0.215	0.211	0.219	0.179
1963	0.215	0.217	0.218	0.216	0.231	0.204	0.210	0.207	0.177	0.148
1964	0.213	0.220	0.210	0.201	0.224	0.202	0.201	0.167	0.132	0.163

(*continued*)

Table 12.6 (continued)

	28	29	30	31	32	33	34	35	36	37	38
1950		0.234	0.234	0.207	0.209	0.199	0.188	0.173	0.156	0.158	0.154
1951	0.254	0.244	0.235	0.228	0.223	0.206	0.203	0.176	0.186	0.148	0.166
1952	0.263	0.239	0.243	0.224	0.222	0.204	0.203	0.188	0.164	0.155	0.166
1953	0.260	0.243	0.246	0.227	0.220	0.200	0.190	0.167	0.176	0.166	0.155
1954	0.261	0.251	0.249	0.219	0.217	0.202	0.180	0.167	0.160	0.159	0.178
1955	0.265	0.241	0.227	0.220	0.202	0.183	0.163	0.158	0.138	0.132	
1956	0.268	0.232	0.230	0.207	0.192	0.186	0.156	0.155	0.113		
1957	0.251	0.228	0.209	0.194	0.181	0.163	0.142	0.136			
1958	0.246	0.210	0.204	0.187	0.174	0.148	0.142				
1959	0.246	0.190	0.185	0.160	0.141	0.126					
1960	0.239	0.179	0.160	0.130	0.157						
1961	0.135	0.112	0.123								
1962	0.104	0.138									
1963	0.095										

Source: See table 12.1.

(27) $\ln C_{k,j}(3) = \alpha + \sum\limits_{r=1}^{2} \beta_r(3) \times k^r + \sum\limits_{r=1}^{2} \lambda_r(3) \times j \cdot k^r$

$$+ \sum\limits_{j=1950}^{1964} \delta_j(3) \text{Dummy}_j + \varepsilon_{k,j}(3),$$

where the number 3 in the parenthesis indicates that the relevant parameters are for the third child.

The results of this estimation for each birth-order/cohort/marital age are shown in tables 12.7 through 12.9. They indicate that the sharpest drop in fertility occurred (or is occurring) with the third child, and in particular, in the 1964 cohort; women marrying at age thirty-five or above are essentially not expected to produce a third child. For three cohorts (1955, 1960, and 1964), the shapes of the estimated C_{kj} function are shown in figures 12.4 through 12.6. In the same figures, I have added the point estimates of the completed fertility rates of each cohort. There are substantial differences in the goodness of fit among these three cohorts, because the cohort dummies alone have absorbed the differences across cohorts.

12.5   Age Distribution of First Marriages

In equation (17), the number of births from a given cohort in a given year is obtained as the sum of the products of the marital fertility rate and the number of women in all the groups, which consist of women who married at the same age for the first time. Thus, I need to obtain estimates of the age distribution of marriages for respective cohorts. For this purpose, I have followed the demographers' tradition and fitted incomplete log-gamma functions using maximum likelihood estimation.

The first marriage rates have been computed by combining the census data with the Vital Statistics data. Here I have experienced relatively little difficulty in obtaining reasonably good estimates of the actual age distributions by maximum likelihood method. The estimated parameters are shown in table 12.10. By comparing the age distribution data of first marriages with the estimated distributions, even for cohorts born after 1960, the estimated log-gamma distributions seem to fit the data relatively well, in spite of a substantial right-censoring problem (fig. 12.7).

A very substantial drop has been occurring in the lifetime probability of a Japanese woman's experiencing first marriage. Between the two cohorts born in 1950 and in 1965, I observe a 10 percentage point decline. The decline in the rate between the 1950 and 1957 cohorts, however, is hardly noticeable, and the decline between the 1957 and 1960 cohorts is relatively mild. The rapid decline seems to have taken place in the 1960 cohort, whose probability of getting married in the first half of their twenties is only half that of their predecessors. Although in subsequent cohorts the probability of marriage in the few years after age twenty-seven has

Table 12.7   Estimated First Child Completed Marital Fertility Rate, by Birth Year and Marital Age

	18	19	20	21	22	23	24	25	26
1950	0.807	0.785	0.775	0.775	0.783	0.796	0.815	0.836	0.858
1951	0.817	0.796	0.787	0.787	0.796	0.810	0.830	0.852	0.875
1952	0.810	0.790	0.782	0.783	0.792	0.808	0.827	0.850	0.873
1953	0.803	0.784	0.777	0.779	0.789	0.805	0.825	0.848	0.872
1954	0.778	0.760	0.754	0.757	0.768	0.784	0.804	0.827	0.850
1955	0.767	0.751	0.746	0.750	0.761	0.777	0.798	0.821	0.845
1956	0.752	0.737	0.733	0.738	0.749	0.766	0.787	0.810	0.834
1957	0.728	0.715	0.711	0.716	0.728	0.746	0.767	0.790	0.813
1958	0.675	0.663	0.661	0.666	0.678	0.694	0.715	0.737	0.759
1959	0.701	0.689	0.687	0.694	0.707	0.725	0.746	0.770	0.793
1960	0.689	0.679	0.678	0.685	0.698	0.717	0.738	0.762	0.786
1961	0.676	0.667	0.666	0.674	0.688	0.707	0.729	0.753	0.776
1962	0.666	0.658	0.659	0.667	0.681	0.700	0.723	0.747	0.771
1963	0.661	0.653	0.654	0.663	0.678	0.698	0.721	0.745	0.769
1964	0.648	0.642	0.644	0.653	0.668	0.688	0.711	0.736	0.760

	27	28	29	30	31	32	33	34	35	36	37
1950	0.879	0.898	0.910	0.915	0.910	0.893	0.862	0.818	0.760	0.691	0.612
1951	0.897	0.915	0.928	0.933	0.927	0.909	0.877	0.831	0.771	0.699	0.618
1952	0.895	0.914	0.927	0.932	0.925	0.907	0.874	0.827	0.766	0.693	0.612
1953	0.895	0.913	0.926	0.930	0.924	0.904	0.871	0.823	0.761	0.688	0.605
1954	0.873	0.891	0.903	0.907	0.900	0.881	0.847	0.800	0.739	0.666	0.585
1955	0.867	0.886	0.898	0.902	0.894	0.874	0.840	0.792	0.730	0.657	0.576
1956	0.857	0.875	0.887	0.890	0.882	0.862	0.827	0.779	0.717	0.644	0.564
1957	0.835	0.853	0.865	0.868	0.860	0.839	0.805	0.757	0.696	0.624	0.545
1958	0.780	0.796	0.807	0.810	0.802	0.782	0.749	0.704	0.646	0.578	0.503
1959	0.815	0.833	0.844	0.847	0.838	0.817	0.782	0.733	0.672	0.600	0.522
1960	0.808	0.825	0.837	0.839	0.830	0.808	0.773	0.724	0.662	0.591	0.512
1961	0.798	0.816	0.827	0.828	0.819	0.797	0.761	0.712	0.651	0.579	0.501
1962	0.793	0.810	0.821	0.822	0.813	0.790	0.754	0.705	0.643	0.571	0.493
1963	0.792	0.809	0.820	0.821	0.811	0.788	0.751	0.701	0.638	0.566	0.488
1964	0.782	0.800	0.810	0.811	0.801	0.787	0.741	0.690	0.628	0.556	0.478

(*continued*)

Table 12.7 (continued)

	38	39	40	41	42	43	44	45	46	47	48	49	50
1950	0.528	0.442	0.358	0.280	0.211	0.153	0.107	0.071	0.045	0.028	0.016	0.009	0.004
1951	0.532	0.444	0.358	0.280	0.210	0.152	0.105	0.070	0.044	0.027	0.015	0.008	0.004
1952	0.525	0.437	0.352	0.273	0.205	0.147	0.102	0.067	0.042	0.025	0.014	0.008	0.004
1953	0.518	0.430	0.345	0.267	0.199	0.143	0.098	0.064	0.040	0.024	0.014	0.007	0.004
1954	0.499	0.413	0.331	0.255	0.190	0.135	0.092	0.060	0.038	0.022	0.012	0.007	0.003
1955	0.491	0.405	0.323	0.248	0.184	0.131	0.089	0.058	0.036	0.021	0.012	0.006	0.003
1956	0.479	0.394	0.313	0.240	0.177	0.125	0.085	0.055	0.034	0.020	0.011	0.006	0.003
1957	0.461	0.378	0.300	0.229	0.168	0.118	0.080	0.051	0.031	0.018	0.010	0.005	0.003
1958	0.426	0.348	0.275	0.209	0.153	0.107	0.072	0.046	0.028	0.016	0.009	0.005	0.002
1959	0.440	0.359	0.283	0.214	0.156	0.109	0.073	0.046	0.028	0.016	0.009	0.004	0.002
1960	0.431	0.350	0.275	0.208	0.151	0.105	0.070	0.044	0.027	0.015	0.008	0.004	0.002
1961	0.420	0.341	0.267	0.201	0.145	0.101	0.066	0.042	0.025	0.014	0.008	0.004	0.002
1962	0.413	0.334	0.260	0.195	0.141	0.097	0.064	0.040	0.024	0.013	0.007	0.004	0.002
1963	0.407	0.328	0.255	0.191	0.137	0.094	0.061	0.038	0.023	0.013	0.007	0.003	0.002
1964	0.398	0.320	0.248	0.185	0.132	0.090	0.059	0.036	0.021	0.012	0.006	0.003	0.001

Source: See table 12.1.

Table 12.8   Estimated Second Child Completed Marital Fertility Rate, by Birth Year and Marital Age

	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1950	0.366	0.404	0.446	0.492	0.540	0.587	0.632	0.670	0.698	0.713	0.710	0.689	0.648	0.590	0.517	0.436
1951	0.381	0.418	0.460	0.505	0.551	0.596	0.639	0.674	0.699	0.710	0.705	0.681	0.638	0.579	0.506	0.425
1952	0.402	0.439	0.480	0.524	0.569	0.613	0.654	0.686	0.709	0.717	0.708	0.681	0.636	0.575	0.501	0.420
1953	0.425	0.462	0.503	0.546	0.590	0.632	0.670	0.701	0.720	0.725	0.713	0.683	0.636	0.572	0.497	0.415
1954	0.457	0.494	0.534	0.577	0.620	0.662	0.698	0.726	0.743	0.745	0.729	0.696	0.645	0.578	0.501	0.417
1955	0.479	0.515	0.555	0.596	0.638	0.677	0.710	0.735	0.749	0.747	0.729	0.692	0.639	0.571	0.493	0.409
1956	0.485	0.519	0.556	0.595	0.633	0.668	0.698	0.719	0.729	0.724	0.703	0.665	0.612	0.545	0.469	0.388
1957	0.493	0.525	0.559	0.595	0.630	0.662	0.688	0.705	0.711	0.704	0.680	0.641	0.587	0.521	0.447	0.369
1958	0.498	0.527	0.559	0.591	0.623	0.651	0.674	0.687	0.690	0.679	0.654	0.614	0.560	0.495	0.423	0.348
1959	0.502	0.529	0.558	0.588	0.616	0.641	0.659	0.670	0.669	0.656	0.629	0.588	0.534	0.471	0.401	0.329
1960	0.508	0.532	0.558	0.585	0.610	0.631	0.646	0.653	0.649	0.634	0.605	0.563	0.510	0.448	0.380	0.311
1961	0.513	0.534	0.558	0.581	0.603	0.621	0.633	0.637	0.630	0.612	0.582	0.539	0.486	0.426	0.360	0.294
1962	0.511	0.530	0.550	0.571	0.589	0.603	0.612	0.612	0.603	0.583	0.552	0.510	0.458	0.399	0.337	0.274
1963	0.508	0.524	0.542	0.559	0.574	0.585	0.590	0.588	0.576	0.555	0.523	0.481	0.430	0.374	0.314	0.255
1964	0.509	0.523	0.537	0.551	0.563	0.571	0.574	0.569	0.555	0.532	0.499	0.457	0.408	0.353	0.296	0.239

(*continued*)

Table 12.8 (continued)

	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1950	0.352	0.272	0.199	0.139	0.092	0.057	0.033	0.018	0.009	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1951	0.343	0.264	0.193	0.134	0.088	0.055	0.032	0.018	0.009	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1952	0.337	0.259	0.189	0.132	0.087	0.054	0.031	0.017	0.009	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1953	0.333	0.255	0.186	0.129	0.085	0.053	0.031	0.017	0.009	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1954	0.333	0.255	0.186	0.129	0.085	0.052	0.031	0.017	0.009	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1955	0.326	0.249	0.181	0.125	0.082	0.051	0.030	0.016	0.008	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1956	0.309	0.235	0.171	0.118	0.077	0.048	0.028	0.015	0.008	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1957	0.293	0.222	0.161	0.111	0.073	0.045	0.026	0.014	0.007	0.004	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1958	0.276	0.209	0.151	0.104	0.068	0.042	0.025	0.014	0.007	0.003	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1959	0.260	0.196	0.142	0.098	0.064	0.040	0.023	0.013	0.007	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000
1960	0.245	0.185	0.133	0.092	0.060	0.037	0.022	0.012	0.006	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000
1961	0.231	0.174	0.125	0.086	0.056	0.035	0.020	0.011	0.006	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000
1962	0.215	0.161	0.116	0.079	0.052	0.032	0.019	0.010	0.005	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000
1963	0.199	0.149	0.107	0.073	0.048	0.029	0.017	0.010	0.005	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000
1964	0.186	0.139	0.100	0.068	0.044	0.027	0.016	0.009	0.005	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000

Source: See table 12.1.

Table 12.9   Estimated Third Child Completed Marital Fertility Rate, by Birth Year and Marital Age

	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1950	0.115	0.128	0.141	0.154	0.167	0.178	0.188	0.197	0.204	0.209	0.212	0.212	0.211	0.207	0.201	0.193
1951	0.140	0.155	0.169	0.183	0.196	0.208	0.218	0.226	0.232	0.235	0.236	0.235	0.230	0.224	0.215	0.204
1952	0.160	0.176	0.192	0.206	0.219	0.230	0.239	0.246	0.250	0.251	0.249	0.245	0.238	0.228	0.217	0.204
1953	0.181	0.198	0.214	0.228	0.241	0.251	0.258	0.263	0.265	0.264	0.259	0.252	0.242	0.230	0.216	0.200
1954	0.198	0.215	0.231	0.244	0.256	0.264	0.270	0.273	0.272	0.268	0.261	0.251	0.239	0.224	0.208	0.191
1955	0.206	0.222	0.236	0.249	0.258	0.265	0.268	0.268	0.265	0.259	0.249	0.237	0.223	0.207	0.190	0.172
1956	0.221	0.237	0.251	0.261	0.269	0.274	0.275	0.273	0.267	0.258	0.246	0.232	0.216	0.198	0.180	0.161
1957	0.225	0.240	0.252	0.260	0.266	0.269	0.267	0.263	0.255	0.244	0.230	0.215	0.198	0.180	0.161	0.142
1958	0.232	0.245	0.255	0.262	0.266	0.266	0.263	0.256	0.246	0.233	0.218	0.201	0.183	0.164	0.146	0.127
1959	0.232	0.243	0.251	0.257	0.258	0.256	0.251	0.242	0.230	0.216	0.200	0.183	0.165	0.146	0.128	0.111
1960	0.231	0.240	0.247	0.250	0.250	0.246	0.239	0.228	0.215	0.200	0.183	0.166	0.148	0.130	0.112	0.096
1961	0.225	0.232	0.237	0.238	0.236	0.230	0.222	0.210	0.196	0.181	0.164	0.147	0.129	0.112	0.096	0.081
1962	0.220	0.227	0.230	0.229	0.225	0.218	0.208	0.195	0.181	0.165	0.148	0.131	0.114	0.098	0.083	0.069
1963	0.222	0.226	0.228	0.226	0.220	0.211	0.200	0.186	0.171	0.154	0.137	0.120	0.103	0.088	0.073	0.061
1964	0.212	0.216	0.215	0.212	0.205	0.195	0.183	0.169	0.153	0.137	0.121	0.105	0.089	0.075	0.062	0.050

(*continued*)

Table 12.9   (continued)

|	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1950	0.184	0.173	0.161	0.149	0.136	0.122	0.109	0.097	0.085	0.073	0.063	0.053	0.044	0.037	0.030	0.025	0.020
1951	0.192	0.178	0.164	0.149	0.134	0.119	0.105	0.092	0.079	0.067	0.057	0.047	0.039	0.032	0.026	0.020	0.016
1952	0.189	0.173	0.157	0.141	0.125	0.110	0.095	0.082	0.069	0.058	0.048	0.040	0.032	0.026	0.020	0.016	0.012
1953	0.184	0.166	0.149	0.132	0.116	0.100	0.085	0.072	0.060	0.050	0.041	0.033	0.026	0.021	0.016	0.012	0.009
1954	0.173	0.154	0.137	0.119	0.103	0.088	0.074	0.062	0.051	0.041	0.033	0.026	0.021	0.016	0.012	0.009	0.007
1955	0.154	0.136	0.119	0.102	0.087	0.073	0.061	0.050	0.040	0.032	0.026	0.020	0.015	0.012	0.009	0.007	0.005
1956	0.142	0.124	0.107	0.091	0.076	0.063	0.052	0.042	0.033	0.026	0.020	0.016	0.012	0.009	0.007	0.005	0.003
1957	0.124	0.107	0.091	0.076	0.063	0.052	0.042	0.033	0.026	0.020	0.016	0.012	0.009	0.006	0.005	0.003	0.002
1958	0.110	0.093	0.078	0.065	0.053	0.043	0.034	0.027	0.021	0.016	0.012	0.009	0.006	0.005	0.003	0.002	0.002
1959	0.094	0.079	0.066	0.054	0.043	0.034	0.027	0.021	0.016	0.012	0.009	0.006	0.005	0.003	0.002	0.002	0.001
1960	0.081	0.067	0.055	0.044	0.035	0.027	0.021	0.016	0.012	0.009	0.007	0.005	0.003	0.002	0.002	0.001	0.001
1961	0.067	0.055	0.045	0.035	0.028	0.021	0.016	0.012	0.009	0.007	0.005	0.003	0.002	0.002	0.001	0.001	0.001
1962	0.057	0.046	0.037	0.029	0.022	0.017	0.013	0.009	0.007	0.005	0.003	0.002	0.002	0.001	0.001	0.000	0.000
1963	0.049	0.039	0.031	0.024	0.018	0.014	0.010	0.007	0.005	0.004	0.003	0.002	0.001	0.001	0.001	0.000	0.000
1964	0.040	0.032	0.025	0.019	0.014	0.011	0.008	0.005	0.004	0.003	0.002	0.001	0.001	0.001	0.000	0.000	0.000

Source: See table 12.1.
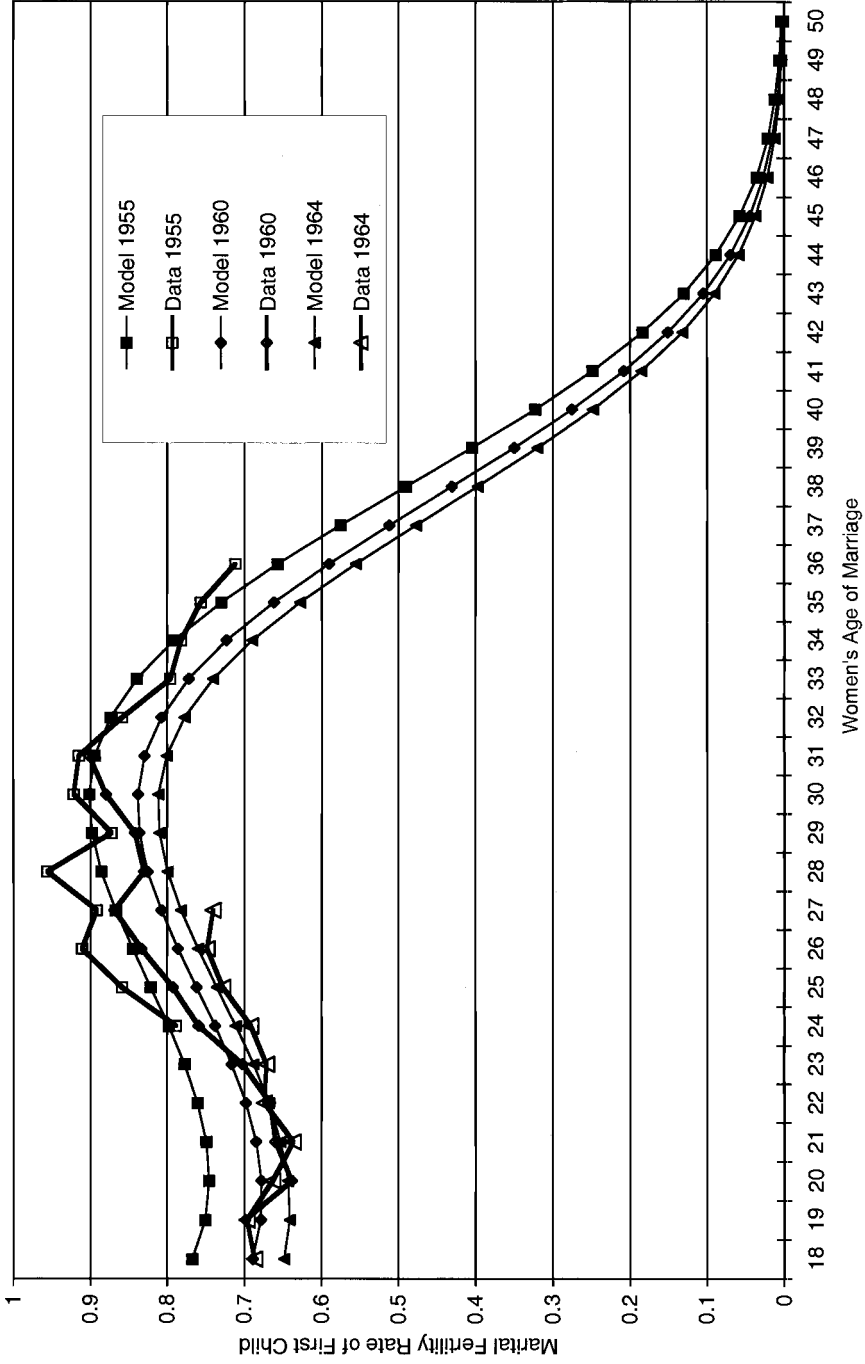
Fig. 12.4 Estimated completed marital-fertility rates of first child of a female for three cohorts
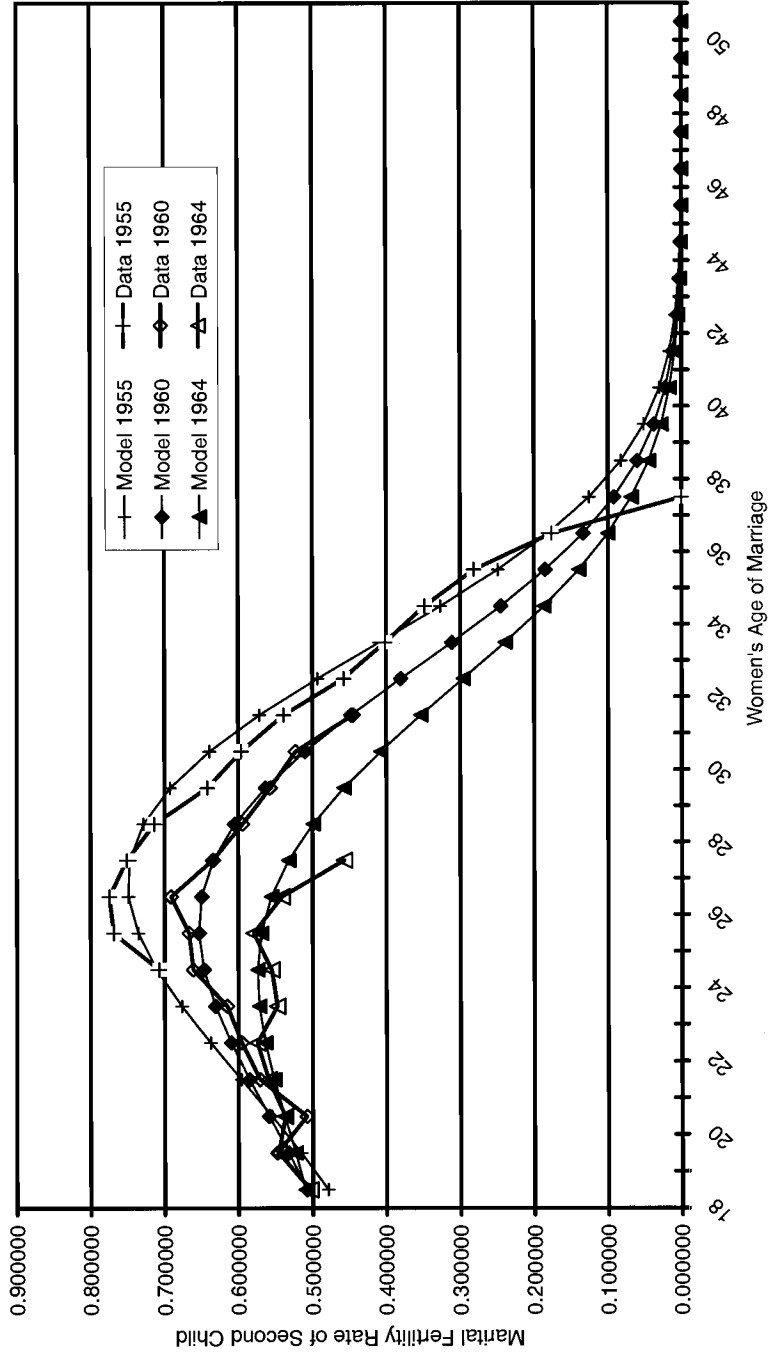
Fig. 12.5 Estimated completed marital-fertility rates of second child for three cohorts
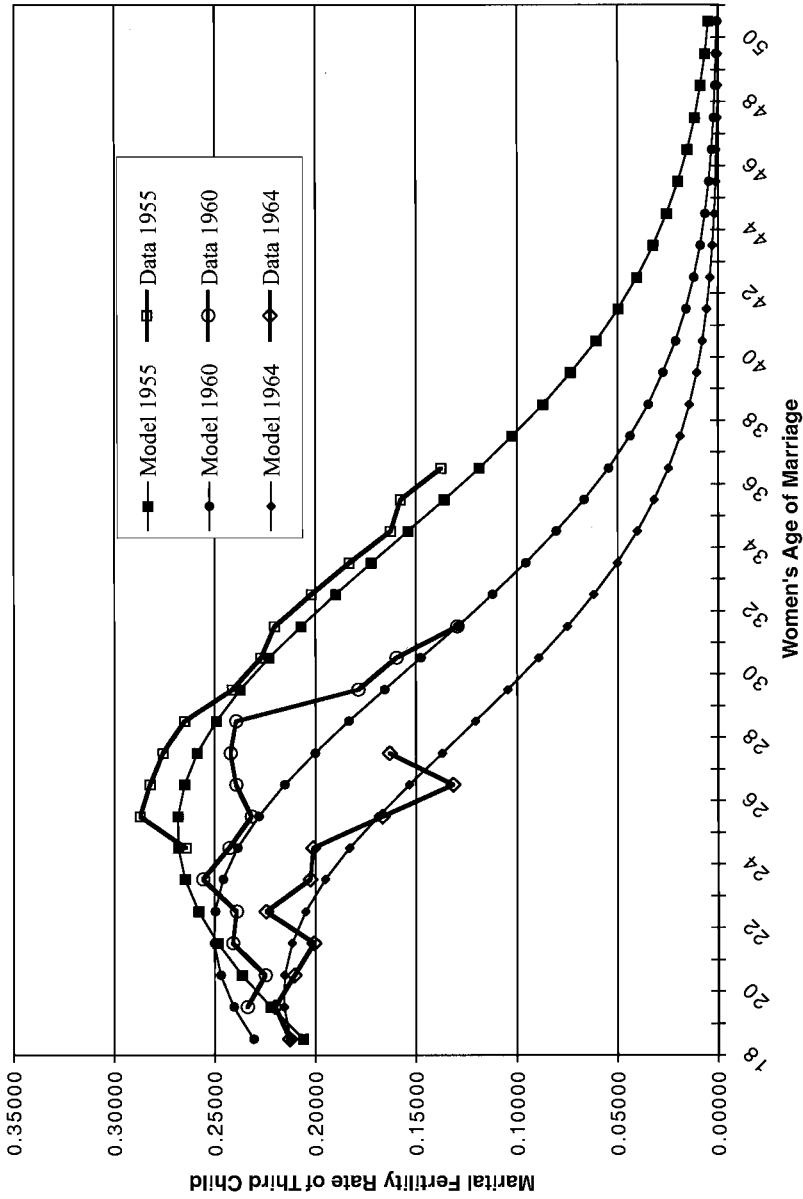
Fig. 12.6 Estimated completed marital-fertility rates of third child for three cohorts

Table 12.10   **Estimated Parameters of Age Distribution Function of First Marriage (incomplete log-gamma function), by Birth Year**

	1950	1955	1960	1961	1962	1963	1964	1965
α	1.0740	1.6386	2.5796	3.1566	3.8334	4.8363	5.5960	5.8053
β	0.4990	0.4648	0.2774	0.2966	0.3416	0.4719	0.3452	0.4461
μ	−0.3746	−0.2687	−0.2085	−0.1842	−0.1644	−0.1432	−0.1293	−0.1233
u	21.4220	22.6289	22.6800	24.0757	26.2199	30.5007	30.0440	32.8875
C	0.8771	0.8623	0.8352	0.8242	0.8114	0.7909	0.7822	0.7796
Likelihood	−2.5812	−2.6236	−2.5429	−2.5120	−2.4601	−2.3767	−2.2639	−2.1043
FOC								
α	0.000003	−0.000062	−0.000094	−0.000099	−0.000100	−0.000096	−0.000097	−0.000098
β	−0.000100	0.000099	0.000094	0.000047	0.000096	0.000099	0.000043	0.000061
μ	0.000052	−0.000074	−0.000090	−0.000093	−0.000036	−0.000069	−0.000444	−0.000048
u	0.000019	−0.000012	−0.000005	−0.000003	−0.000005	−0.000007	−0.000002	−0.000003

Source: See table 12.1.
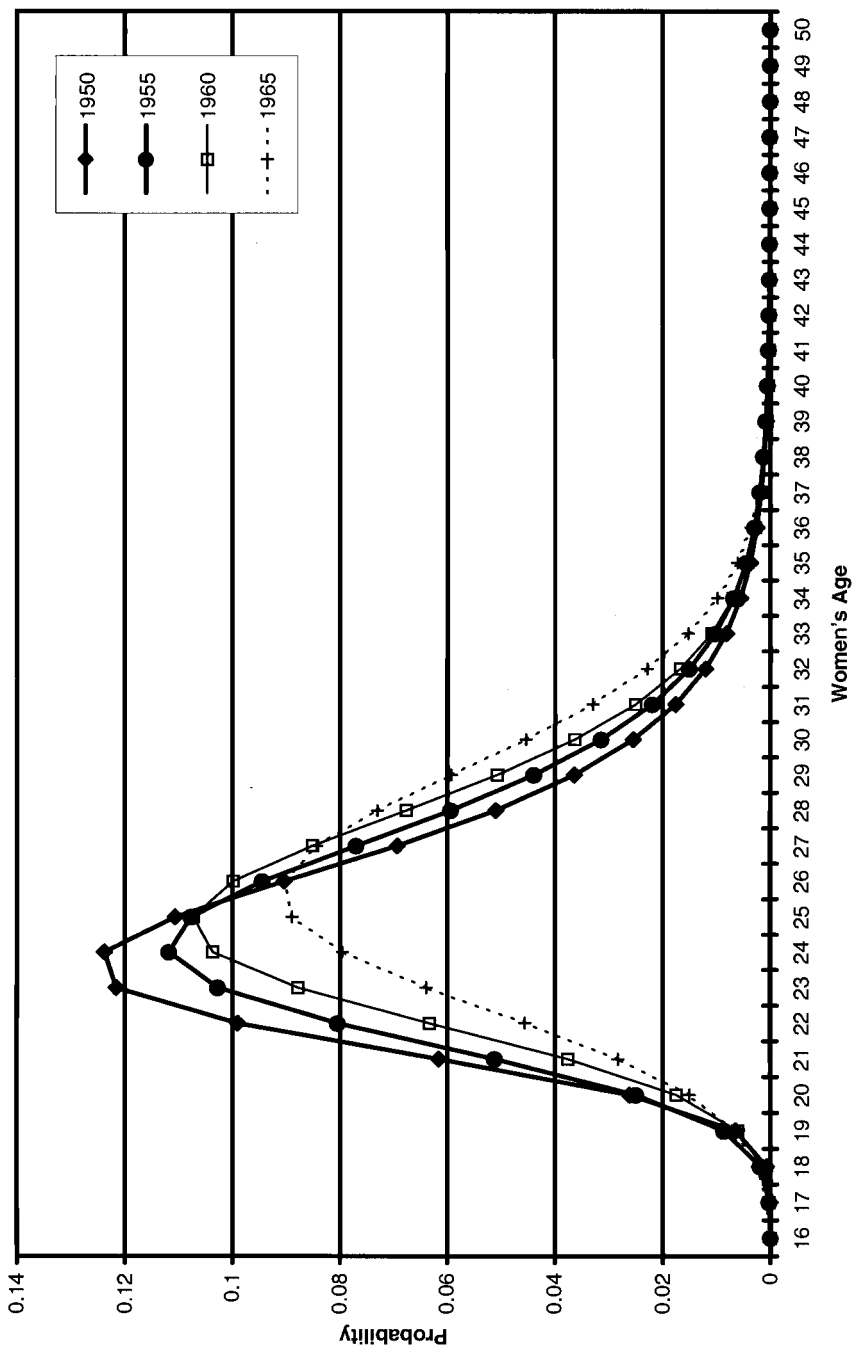
Fig. 12.7 Estimated age distribution functions of marriage probability of Japanese females by birth years

increased by a few percentage points, this increase does not seem to extend beyond the first half of the thirties. In the 1991–92 projection, it was assumed implicitly that the 1960 cohorts would complete the structural adjustment in childbearing behavior. My estimation suggests that in the next five cohorts the lifetime probability of first marriage declined by 7 percentage points.

12.6   Analyzing the Sources of the Decline in Fertility Rates

In my notation, for a woman born in year j married at age k, $C_{kj}(s)$ gives the expected number of children of the order s in her lifetime. Thus, as far as the children of order s are concerned, the completed fertility of a woman born in year j is obtained as the weighted sum of $C_{kj}(s)$'s, or

$$(28) \qquad \hat{C}(s) = \sum_{k=15}^{49} C_{k,j}(s) \cdot m(k),$$

where k stands for the age at first marriage. Using equation (28), in table 12.11 I have computed the fertility rates of each cohort between 1950 and 1964 for their first three children.

According to the table, the completed fertility rate shows a consistent decline for the first child, with the birth year of the woman. For the second and the third child, however, the fertility rates seem to rise until 1954 and then decline. The latter is not an actual phenomenon, but is a problem caused by insufficient information in my "extrapolation" method[3] when cohorts are considerably heterogeneous. For women born in 1950, for example, I began my observation of their reproductive behavior only for age twenty-nine, which is probably too late to start observations even for the second child in those days. In contrast, for cohorts born after 1955, my extrapolation should be far more reliable, since for the 1955 cohort, the observation started at age twenty-four. Thus, 1955 should serve as a good starting point for comparative analysis. If I concentrate on cohorts born after 1955, I notice a general decline in fertility rates between the 1955 and 1960 cohorts. The estimated fertility rates have dropped by almost 0.1 for each order of birth, a drop of at least 0.3 in the TFR.

Has the drop in TFR been brought about by changes in marriage rates or by changes in marital fertility rates? In order to answer this question, I computed the hypothetical completed fertility rates using the estimated marriage-rate function of the 1955 cohort. The differences between these hypothetical fertility rates and those in table 12.12 are attributable solely to changes in marital behavior between the 1955 cohort and the previous cohort.

For instance, for the 1955 cohort, the sum of the completed fertilities of

---

3. Needless to say, this problem will be solved once the relevant data prior to 1979 are made public.

Table 12.11 **Estimated Fertility Rates of First Three Children of a Female Born after 1950**

	First Three Children	First Child	Second Child	Third Child
1950	1.785	0.909	0.670	0.206
1951	1.823	0.920	0.669	0.234
1952	1.840	0.912	0.676	0.251
1953	1.876	0.914	0.693	0.269
1954	1.916	0.904	0.730	0.282
1955	1.875	0.881	0.724	0.270
1956	1.859	0.873	0.711	0.275
1957	1.832	0.861	0.705	0.266
1958	1.697	0.781	0.667	0.250
1959	1.702	0.817	0.649	0.236
1960	1.637	0.797	0.622	0.218
1961	1.570	0.778	0.596	0.196
1962	1.501	0.761	0.562	0.178
1963	1.427	0.741	0.523	0.164
1964	1.363	0.725	0.495	0.144

Source: See table 12.1.

Table 12.12 **Hypothetical Fertility Rates of Women Born after 1950 with Marital Behavior of 1955 Cohort**

	First Three Children	First Child	Second Child	Third Child
1950	1.766	0.899	0.663	0.204
1951	1.815	0.916	0.666	0.233
1952	1.842	0.913	0.677	0.251
1953	1.869	0.911	0.691	0.268
1954	1.879	0.887	0.716	0.276
1955	1.875	0.881	0.724	0.270
1956	1.851	0.869	0.708	0.274
1957	1.804	0.846	0.694	0.264
1958	1.721	0.788	0.676	0.257
1959	1.725	0.823	0.659	0.243
1960	1.687	0.815	0.643	0.229
1961	1.642	0.804	0.627	0.211
1962	1.598	0.797	0.604	0.197
1963	1.563	0.795	0.580	0.188
1964	1.517	0.785	0.562	0.171

Source: See table 12.1.

the first three birth orders that I have estimated is 1.875, whereas for the 1954 cohort it was 1.916. Thus, the total change is −0.041. If I had applied the estimated marriage-rate function of the 1954 cohort to the 1955 cohort, the sum of the three fertility rates would have been 1.913 instead of 1.875. In other words, if the marriage behavior remained unchanged, the sum of

the fertilities would have changed only very slightly (or $1.913 - 1.916 = -0.003$, to be exact). Thus the remainder of the change (namely $1.875 - 1.913 = -0.038$) is attributable to changes in marital fertility rates.

The same analysis is repeated for cohorts born between 1955 and 1964, and the results are shown in table 12.12. The results suggest that, of a 0.55 drop in the sum of these three fertility rates, changes in marital behavior explain -0.19, while changes in marital fertility rates account for -0.36.

12.7   Concluding Remarks

It should be clear by now why the last official population projection failed. In an attempt to estimate directly the age-specific fertility rate distribution, the projection assumed a priori the stability of parameters beyond the 1960 cohort, probably the last cohort for which demographers were able to obtain stable distribution parameters. This is equivalent to assuming that this cohort would complete the behavioral changes in childbearing for Japanese women. Instead, it took another five cohorts for this happen and, in the meantime, lifetime marital rates dropped by 15 percentage points.

In the new population projection announced in January 1997, the government forecasters again have predicted that TFR will recover to 1.6 by the end of the second decade of the next century. On this point, I have only two things to say. First, the TFR for the latest cohort (those born in 1964) that I was able to estimate stands at slightly above 1.4, or about the current TFR of all Japanese women, because fertility rates attributable to a birth order higher than three add very little. Thus, the present Japanese women's TFR is at its equilibrium level, not at a transitory or disequilibrium low level as government demographers have argued.

Furthermore, there is no inherent demographic mechanism for a rebound in fertility rates to occur. In fact, the lifetime rates that I used to derive the expected completed fertility rates are based on the estimated dynamic behavioral equations of childbearing. Hence, I have already taken dynamic adjustment behaviors into account. As far as I know, there are no demographic grounds for a fertility rebound to occur, but there are very strong economic grounds for expecting another drop to take place. Since the labor force is expected to decline at the rate of almost one percent a year starting early in the next century, female labor is expected to fill part of the shortage. This will mean better employment opportunities and better pay for women, raising the opportunity costs of having children even further. In Tokyo, TFR stands at about 1.1, and we should not be surprised to see the national TFR drop to that level.

What does this all mean in terms of our (or our children's) lives in the twenty-first century? First of all, if one limits the analysis to the first quarter of the century, differences in the TFR matter very little. This is true of

pay-as-you-go public pension program costs or of public health insurance costs. The difference will show up in the second quarter of the century. Second, the difference in these costs is substantial. The new population projection is equivalent to the government's officially admitting that the costs of these programs have been underestimated by at least 20 percent. As I suspect, even their new projection turns out to be too optimistic. If no rebound occurs, the per capita costs of these programs will increase by at least another 10 percent. It is clear that the time has come to begin reengineering the fundamental frameworks of our social insurance programs, particularly their financing mechanisms, rather than putting more money into the terminal care of these programs.

Appendix

Source of Instability in the Estimated Parameters

Let us explore the source of the instability in the estimated parameters as more data are added. Collecting the terms in equation (14), and noting

$$S(x) = 1 - \sum_{\tau=15}^{x-1} m_\tau,$$

we obtain the first-order condition

(14′) $$\sum_{\tau=15}^{x-1} \left[\frac{m_\tau}{P(\tau;\theta)} - \frac{1 - \sum_{\tau=15}^{x-1} m_\tau}{1 - \sum_{\tau=15}^{x-1} P(\tau;\theta)} \right] \frac{\partial P(\tau;\theta)}{\partial \theta} = 0.$$

Let us define the theoretical probability of not giving birth of the child of that order by $P(51;\theta)$, or the probability of giving birth at age fifty-one. Similarly, if the proportion of women not giving birth to a child is given by m_{51}, we have

(15′) $$1 - \sum_{\tau=15}^{x-1} P(\tau;\theta) = \sum_{\tau=x}^{51} P(\tau;\theta)$$

$$1 - \sum_{\tau=15}^{x-1} m_\tau = \sum_{\tau=x}^{51} m_\tau.$$

Furthermore, the optimality condition (14′) with respect to C is particularly simple and is given by

(16′) $$\sum_{\tau=x}^{51} m_\tau = \sum_{\tau=x}^{51} P(\tau;\theta).$$

For the rest of the parameters of log-gamma distribution function, (14′) can be written as

$$(17') \quad \Psi_\theta(x-1;\hat\theta) = \sum_{\tau=15}^{x-1} \frac{m_\tau}{P(\tau;\hat\theta)} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} - \left[\frac{\sum_x^{51} m_\tau}{\sum_x^{51} P(\tau;\hat\theta)}\right] \sum_{\tau=15}^{x-1} \frac{\partial P(\tau;\hat\theta)}{\partial\theta}$$

$$= \sum_{\tau=15}^{x-1} \frac{m_\tau}{P(\tau;\hat\theta)} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} - \sum_{\tau=15}^{x-1} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} = 0.$$

Moreover, the sum of all the possibilities including $P(51;\theta)$ is always equal to one, regardless of the value of θ; hence we have

$$(18') \qquad\qquad \sum_{\tau=15}^{51} \frac{\partial P(\tau;\theta)}{\partial\theta} = 0.$$

Therefore, equation (17′) implies θ

$$(19') \quad \Psi_\theta(x-1;\hat\theta) = \sum_{\tau=15}^{x-1} \frac{m_\tau}{P(\tau;\hat\theta)} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} + \sum_{\tau=x}^{51} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} = 0.$$

Adding the fertility data at age x, the optimality condition is given by

$$(20') \quad \Psi_\theta(x;\tilde\theta) = \sum_{\tau=15}^{x} \frac{m_\tau}{P(\tau;\tilde\theta)} \frac{\partial P(\tau;\tilde\theta)}{\partial\theta} + \sum_{\tau=x+1}^{51} \frac{\partial P(\tau;\tilde\theta)}{\partial\theta} = 0.$$

If the right-hand side of the equation is evaluated at the previous optimal value $\hat\theta$,

$$(21') \quad \Psi_\theta(x;\hat\theta) = \left[\sum_{\tau=15}^{x-1} \frac{m_\tau}{P(\tau;\hat\theta)} \frac{\partial P(\tau;\hat\theta)}{\partial\theta} + \sum_{\tau=x}^{51} \frac{\partial P(\tau;\hat\theta)}{\partial\theta}\right]$$

$$+ \left[\frac{m_x}{P(x;\hat\theta)} - 1\right]\frac{\partial P(x;\hat\theta)}{\partial\theta},$$

and we obtain the dynamic equation

$$(22') \quad \Psi_\theta(x;\hat\theta) = \Psi_\theta(x-1;\hat\theta) + \left[\frac{m_x}{P(x;\hat\theta)} - 1\right]\frac{\partial P(x;\hat\theta)}{\partial\theta}.$$

It is clear from equation (22′) that the estimated parameters are consistent even when one truncates samples provided that the fertility data at every age have been generated by the theoretical distribution. If the sample size gets arbitrarily large for every t, we have

$$m_t \rightarrow P(t;\theta),$$

and, from equation (22′), at every age x, we should have

$$\lim_{n\to\infty} \Psi(x - 1;\theta) = \lim_{n\to\infty} \Psi(x;\theta) = 0.$$

Ordinarily, if we have one million observations, the law of large numbers should work to give us this condition. If, on the other hand, the estimated parameters remain unstable, it is most likely that the fertility data have not been generated by that particular theoretical distribution.

The source of the instability is the second term of equation (22′). The larger the divergence of the new fertility data at age x, or m_x, and the estimated fertility rate $P(x;\hat{\theta})$ using the data up to age $(x - 1)$, the absolute value of the coefficient in the second term,

$$\frac{m_x}{P(x;\hat{\theta})} - 1,$$

will be larger, and the revision of the estimated parameters will be larger.

References

Coale, A. J. 1971. Age patterns of marriage. *Population Studies* 25 (2): 193–214.

Coale, A. J., and D. R. McNeil. 1972. The distribution of age of the frequency of first marriage in a female cohort. *Journal of the American Statistical Association* 67 (340): 743–49.

Kaneko, Ryuichi. 1991. "Shokon kateino jinkoga kuteki bunseki" (A demographic analysis of first marriage process). *Jinkogaku Kenkyu* 47:3–27. Koseisho Jinko Mondai Kenkyusho (Institute of Population Problems, Ministry of Health and Welfare).

———. 1993. "Nenrei betsu shusseiritsu no shourai suikei sisutemu" (A projection system for future age-specific fertility rates). *Jinkogaku Kenkyu* 49:17–38. Koseisho Jinko Mondai Kenkyusho (Institute of Population Problems, Ministry of Health and Welfare).

Kenkyusho, Koseisho Jinkomondai, ed. 1992. *Nihonno shorai suikei jinko: Heisei 4 nen 9 gatu suikei* (Population projections for Japan: 1991–2090). Kosei Tokei Kyokai.