Working Paper Series (E)

No.17

Transformations in First Marriage Behavior of Japanese Female Cohorts: Estimation and Projection of Lifetime Measures via Empirically Adjusted Coale-McNeil Model

Ryuichi Kaneko

December 2002

National Institute of Population and Social Security Research

2-2-3 Uchisaiwai-cho, Chiyoda-ku,
Tokyo, Japan

No.17

Transformations in First Marriage Behavior of Japanese Female Cohorts:Estimation and Projection of Lifetime Measures via Empirically Adjusted Coale-McNeil Model

Ryuichi Kaneko

December 2002

Transformations in First Marriage Behavior of Japanese Female Cohorts: Estimation and Projection of Lifetime Measures via Empirically Adjusted Coale-McNeil Model

Ryuichi Kaneko*

Abstract

In this study, historical development of first marriage behavior in Japan in terms of lifetime measures for female birth cohorts is reconstructed to examine current situation of the rapidly transforming marriage. The measures include timing indices (mean, mode, median, and sd of age at first marriage), and eventual occurrence levels (proportion ever-married and never-married at age 50). After two preliminary arrangements: empirical adjustment of the Coale-McNeil nuptiality model, and estimation of annual number of first marriage in postwar period, trends of the lifetime measures over cohorts born in 1935-1970 are estimated. It is found that the behavioral change relevant to the recent nuptiality and fertility decline is initiated with delaying marriage by cohort born in 1952, followed by diffusion of never marrying among cohorts born in and after 1959. Then the timing shift is gradually ending among cohorts born in and after 1965, while the proportion never marrying is still rising even at a seemingly accelerated pace. Mechanisms, implications, and prospects are briefly discussed.

Introduction

The present study aims at better understanding the current situation concerning the rapidly transforming issue of first marriages in Japan. For this purpose, we reconstruct the historical development of marriage behavior in the postwar period in terms of lifetime measures of first marriages for female birth cohorts. The measures include timing indices (the mean, mode, two kinds of median, and standard deviation of age at first marriage), and eventual occurrence levels (proportion married and never married at age 50).

Although lifetime measures are crucial for understanding behavioral changes and their causes in first marriages, they are not available for cohorts relevant to current or even one or two decades of marriage, since those cohorts will have yet to complete their first marriage processes for the distant future. Hence, some tool with a reliable predictive power for cohort life course for first marriages is needed. The Coale-McNeil standard model schedule for first marriages is the most widely used tool for this purpose (Bloom, 1982, Bloom and Bennett, 1990, Goldstein and Kenney, 2001, etc.). However, it has been pointed out that the model schedule shows substantial deviations from observed

* National Institute of Population and Social Security Research, Hibiya-kokusai Bldg. 6F, 2-2-3, Uchisaiwai-cho, Chiyoda-ku, Tokyo 100-0011, Japan. email: r-kaneko@ipss.go.jp

schedules even in a flexible free-shape form when it is applied to Japanese experiences (Takahashi, 1978, Kojima, 1985, Kaneko, 1991). This is a critical issue because prediction reliability heavily depends on the model's ability to trace reality, especially in application to young cohorts for which less information is available.

In this connection, we developed a new procedure to adjust the model through empirical correction so as to obtain a sufficient fit to Japanese first marriage experiences. The model with the adjustment indicates the high ability to describe the actual trajectory of first marriages. With this tool, the long-term trends of the lifetime measures of first marriages over cohorts born from 1935 to 1970 are constructed. Some interesting changing patterns of lifetime first marriage behavior are clarified in the results. According to a combination of timing and prevalence measures of first marriages, the history of first marriage behavior over cohorts born from 1935 to 1970 is divided into five phases, the latest three of which are relevant to the unprecedented nuptiality and fertility decline since the mid 1970s until today in Japan. In these three phases, the change of first marriage behavior was initiated with a delay in timing in the cohort born in 1952, followed by an onset of the steep increase in proportion never marrying after the cohort of 1959. Then the timing shift gradually ends and moves toward leveling off after the cohort of 1965, while the proportion never marrying is still rising even at a seemingly accelerated pace. These results of changing patterns in lifetime first marriage measures suggest some behavior hypothesis is relevant to the current marriage transformation, as well as prospects of its further development among younger cohorts now in their 20s.

Models: Coale-McNeil Model with Empirical Adjustment

Coale-McNeil Model

Following Coale's finding that the age distribution of first marriages for female cohorts from various countries shows virtually identical age patterns if the starting age and pace of process are adjusted (Coale, 1971), Coale and McNeil proposed a mathematical distribution representing the first marriage pattern with the closed form of the probability density function (PDF) given as:

$$g(x) = \frac{\beta}{\Gamma(\alpha/\beta)} \exp\left[-\alpha(x-\mu) - \exp\{-\beta(x-\mu)\}\right] \tag{1}$$

where Γ denotes the gamma function[1], $\alpha(>0)$, $\beta(>0)$, and $\mu(-\infty < \mu < \infty)$ are three parameters (Coale and McNeil, 1972).

They also proposed a fixed-shape version of this model derived from Swedish female cohort experiences to standardize the age distribution with two free dimensions, i.e. location and scale parameters. Rodriguez and Trussell revised the model so as to amend its mean 0 and variance 1^2(Rodriguez and Trussell, 1980). It is called the Coale-McNeil (CM) standard schedule, and has been

[1] $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$

[2] $\alpha = 1.145, \beta = 1.896, \mu = -0.805$

widely used for estimating underlying distribution from defected data, and projecting a halfway process to complete the entire schedule (Bloom, 1982, Bloom and Bennett, 1990, Goldstein and Kenney, 2001, etc.).

However, it has been pointed out that the model schedule shows substantial deviations from observed schedules when it is applied to Japanese experiences (Takahashi, 1978, Kojima, 1985), even in flexible free-shape (Kaneko, 1991). This is critical when it is used for projection purposes, because the reliability of the prediction heavily depends on the model's ability to trace the reality, especially in application to young cohorts for which less information is available.

In this connection, we developed a procedure to adjust the model with empirical correction so as to obtain a sufficient fit to Japanese first marriage experiences. Before describing its development, we introduce an alternative form of the CM model to make parameter estimation advantageous. Kaneko found that the CM distribution is mathematically identical to the generalized log gamma (GLG) distribution with somewhat different parameter space (Kaneko 1991, 2002). According to Prentice's parameterization (1974), the CM distribution is expressed by the GLG model with limited parameter space. The PDF of the GLG distribution is given by:

$$g(x) = \frac{|\lambda|}{b\Gamma(\lambda^{-2})}(\lambda^{-2})^{\lambda^{-2}} \exp\left[\lambda^{-1}\left(\frac{x-u}{b}\right) - \lambda^{-2}\exp\left\{\lambda\left(\frac{x-u}{b}\right)\right\}\right] \tag{2}$$

where $\lambda\ (-\infty < \lambda < \infty, \neq 0)$, $u\ (-\infty < u < \infty)$, $b\ (> 0)$ are three parameters, Γ denotes the gamma function defined above. We regard it as an equivalent of the CM distribution and use it for estimation and projection of the cohort first marriage process[3]. With the age distribution of first marriages $g(x)$, corresponding age schedule $f(x)$ (age specific first marriage rate at age x) is given as:

$$f(x; C, u, b, \lambda) = C\ g(x; u, b, \lambda). \tag{3}$$

where C denotes the proportion eventually marrying in the cohort.

Empirical Adjustment of the Coale-McNeil Model

No model fits actual data perfectly. Discrepancies consist of two types of errors; one is random error induced by exogenous factors such as measurement error, and the other is systematic error derived from insufficiency in specification of the model. The latter can be corrected by taking account of regularities perceived in the error pattern. Here we introduce empirical adjustment of the GLG model in seeking a better fit to actual experiences of Japanese female first marriages (Kaneko 2002).

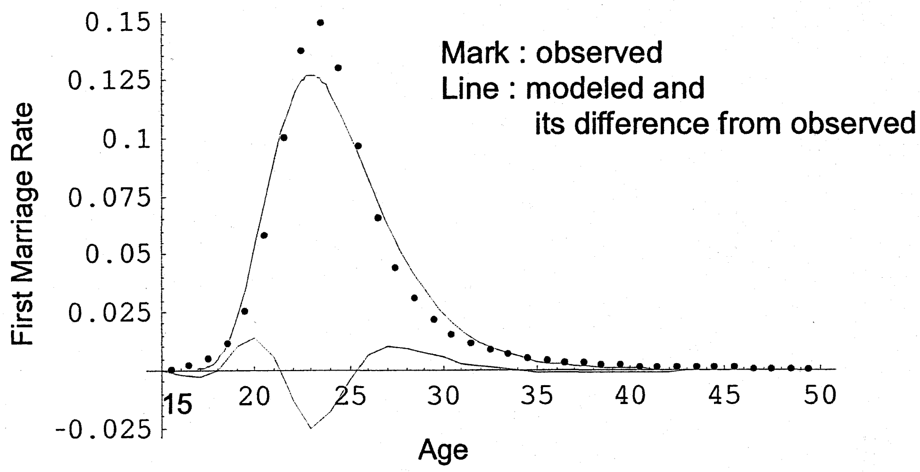
Performance of the GLG model even with free-shape (4-parameter model) is not satisfactory for first marriage experiences of Japanese female cohorts. In Figure 1, observed and modeled first marriage rates for Japanese female cohort born in 1950 are shown with the discrepancies between them. The model tends to be higher than the actual rate before mode, lower around mode, and higher again after then. Quite a similar pattern is found for every cohort that has completed its marriage

---

[3] For equivalence to the CM model, $\lambda < 0$.

process in our data set, and therefore is systematic. Here we develop an empirical adjustment procedure of the GLG model to Japanese female cohort.

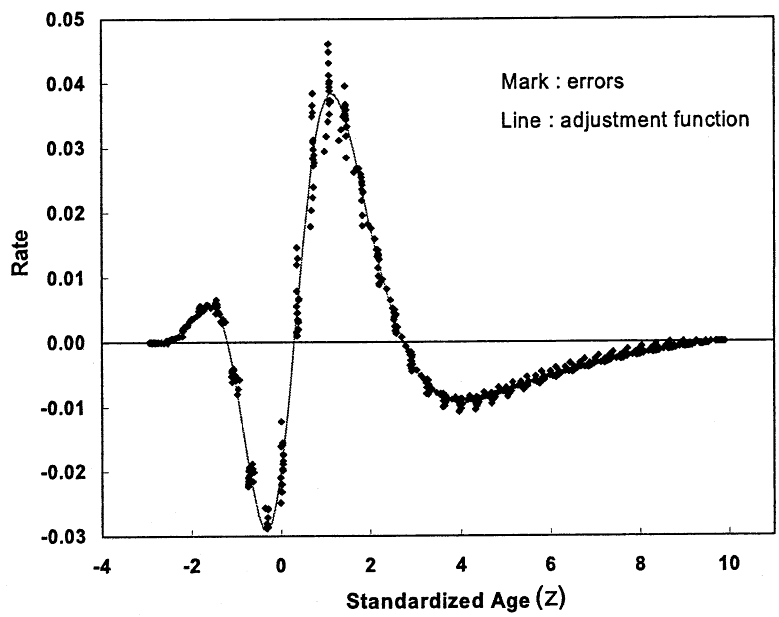
Figure 1 Observed Age Specific First Marriage Rates and Fitted GLG Model : Japanese Female Cohort born in 1950



First we try to identify regularity in the error pattern to be modeled. The difference in the cumulative first marriage rates by age between actual and fitted experiences for 16 cohorts (born in 1935 through 1950) who completed the marriage process are examined. We adjust the cumulative rate function (CDF) in our attempt because it is used in parameter estimation for aggregate representation of first marriage experiences (see **Method** for detail). Figure 2 shows the errors of the CDF for the cohorts. In the figure, horizontal coordinate is calibrated by standardized age z in terms of parameter u and b, i.e. for normal age x: $z=(x-u)/b$. The origin of the axis (0) indicates the location of mode, since parameter u designates mode of the GLG schedule. Let $\xi(z)$ denote the error as: $\xi(z)=F(u+bz)-\hat{F}(u+bz;C,u,b,\lambda)$, where $F(x)$ and $\hat{F}(x;\theta)$ are the cumulative function of first marriage rate of observed and model[4].

[4] Cumulative function of the model with standardized age is alternatively represented by $\hat{F}(z;C,\lambda,0,1)$.

Figure 2 Errors of GLG Model in First Marriage Rate for Japanese Female Cohort (1935-50) and Adjustment Function



In Figure 2, age pattern of the error is found to be highly systematic. Thus, it is expected that if the error pattern is modeled and built into the GLG model, the fit of the model is drastically improved. We simply add an average error pattern to the model. The resulting GLG model with the adjustment, $\overline{F}(x;\boldsymbol{\theta}), \boldsymbol{\theta}=(C,\lambda,u,b)$ is expressed as:

$$\overline{F}(x;C,\lambda,u,b)=\hat{F}(x;C,\lambda,u,b)+\hat{\xi}\left(\frac{x-u}{b}\right), \tag{4}$$

where $\hat{F}(x;\boldsymbol{\theta})$ is the GLG model, and $\hat{\xi}(z)$ is the average error function of the standardized age z.

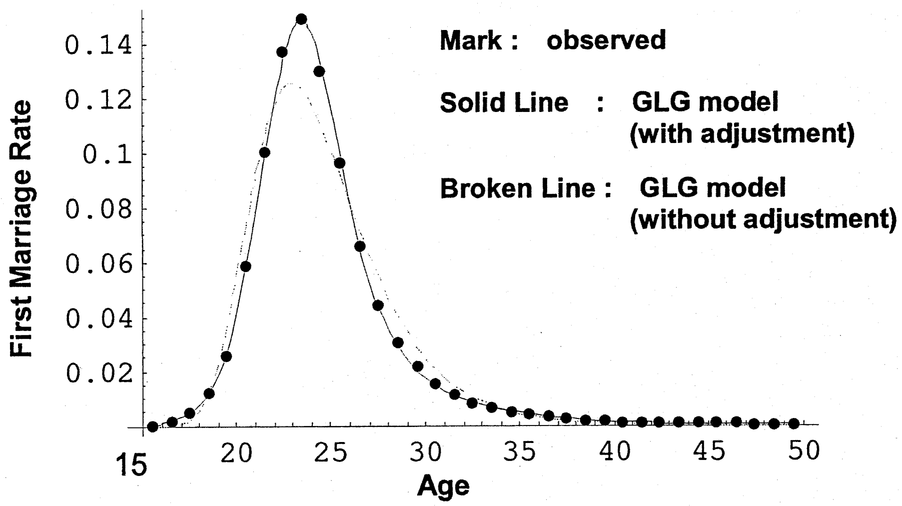
The function $\hat{\xi}(z)$ is shown in Figure 2 in a solid line along with the dots[5]. It is treated as continual function of age by means of the spline interpolation technique. In Figure 3, we see an improvement by the adjustment. The GLG model with the adjustment (solid line) traces the observed rates almost exactly, while the model with no adjustment (broken line) shows visible deviations. The

---

5 $\hat{\xi}(z)$ is to be zero as z goes to plus or minus infinity to keep parameter C intact in the adjustment. Definite integral of $\hat{\xi}(z)$ over the domain of z should be zero to keep the mean age of the schedule intact. We slightly adjust the average pattern to derive $\hat{\xi}(z)$ so that these properties of schedule are kept.

deviations are critical for the purpose of projecting the censored marriage process of young cohorts, for which the data of limited age range is available.

Figure 3 Observed Age Specific First Marriage Rates and Fitted GLG Model (with Adjustment): Japanese Female Cohort born in 1950



Data: Estimation of Annual Number of First Marriages

Since vital statistics are the only source that annually covers all of Japan, we rely on first marriage rates derived from vital statistics to make results represent over all the country. However, there is a major drawback with the statistics for this purpose. There is a substantial amount of delayed registrations, and the annually reported number of marriages is subject to omission if date of the delayed registrations are not corrected. Fortunately, the delayed registrations are reported with information on date and age at the time of marriage in later annual reports for a part of our target period. Therefore, it is possible to sum them up to obtain the eventual number of marriage occurrences. Nevertheless, information on delayed registrations is not available for the earlier period. Also, for recent years there is no following period long enough to cumulate the delay up to the eventual number. Hence, evaluation and estimation of the delayed registrations is required to obtain numbers of marriage occurrences on which we rely as a source of long-term observations with sufficient precision.

To accomplish this, we first examined a way to evaluate and estimate delayed registration numbers. Patterns of delayed registrations by age at marriage are observed over a period for which enough information are available. During this course, we found that a measure of moderation of delay ("average hazard of same year registration") is useful to observe the pattern and to project it into a period with no sufficient information. We modeled the age pattern of delayed registrations in terms of the measure, and applied it to estimate numbers of the delay in years whose information on delayed registrations is unavailable. In this course, intercensal numbers of first marriages by age are estimated from census results on population by marital status, and are used for guidance in determining levels of delayed registrations. As a result, annual numbers of first marriages by wives aged 15-49 in the year 1950 to 2000 (tentatively with 1948, 1949) are estimated. The details of the evaluation and estimation procedures are described in the Appendix.

After the numbers of first marriages are estimated by age and year, the age specific first marriage rates corresponding to cohort diamonds on the Lexis diagram are calculated with populations by age and year from censuses as denominators.

Method: Parametric Estimation of Cohort Marriage Schedules

Method of Parameter Estimation

In a standard situation where age at first marriage for those married and age at survey for the never married are available, the likelihood function is constructed as:

$$L = \prod_{i \in P} f(x_i;\mathbf{p})^{\delta_i} \left[1 - F(x_i;\mathbf{p})\right]^{1-\delta_i} \tag{5}$$

where $f(x;\mathbf{p})$ and $F(x;\mathbf{p})$ are respectively the density function and the cumulative function of first marriage model at age x with parameter set $\mathbf{P}$, x_i is age at marriage or age at survey of individual i depending on whether i is married or never married, δ_i is a indicator variable that takes value one if individual i is married and zero otherwise, and P denotes the sample set as a whole. We estimate a set of parameters $\mathbf{p}$ so as to maximize L, although the logarithm of L is to be maximized in practice for the sake of handiness in calculation.

In the situation above, age at marriage or at survey x_i is supposed to be exact. If only aggregated information, such as numbers or rates of marriage classified by age group, is available, the maximum likelihood method with the interval censoring is appropriate for parameter estimation. Even with data classified by single year of completed age, it applies[6], since grouping by completed age is a form of interval censoring. Most data at national level fall in these conditions.

Suppose that a female cohort of size N at exact age x had number of marriage m_a in completed age a, and with number of never married n_x, i.e. $N = \sum_{a=\alpha}^{x-1} m_a + n_x$. Then assuming marriages are independent of each other, the probability of having such a sample follows multinomial distribution with $x-\alpha+1$ parameters (m_a $(a = \alpha, \alpha+1, \cdots, x-1)$, n_x). Letting $F(x;\mathbf{p})$ denote the cumulative first marriage rate function, the probability (L) is given by:

[6] If date is classified by single year of completed age, estimation without taking account of interval censoring regarding completed age as exact age plus half year yields almost identical results with estimation of the interval censoring.

$$L = \frac{N!}{m_\alpha ! m_{\alpha+1} ! \cdots m_{x-1} ! n_x !} \left[\prod_{a=\alpha}^{x-1} \left(F(a+1;\mathbf{p}) - F(a;\mathbf{p}) \right)^{m_a} \right] \left(1 - F(x;\mathbf{p})\right)^{n_x}. \tag{6}$$

According to the maximum likelihood procedure, we estimate a set of parameters $\mathbf{p}$ so as to maximize L. Since it is equivalent to maximize log transformed of L eliminating constant factors, we use the following function to maximize:

$$\ln L' = \sum_{a=\alpha}^{x-1} m_a \ln\left(F(a+1;\mathbf{p}) - F(a;\mathbf{p})\right) + n_x \ln\left(1 - F(x;\mathbf{p})\right). \tag{7}$$

We use rates rather than numbers of people as m_a and n_x to focus on marriage behavior by eliminating influences from death and migration, since raw numbers are subject to those factors. Hence we replace m_a with observed first marriage rate at age a, and n_x with the complement of cumulative first marriage rate up to exact age x.

Censoring Effects on Parameter Estimation

If specification of the model to the data is not perfect, parameter estimation is affected by censoring. This happens in our research for cohorts who have not completed the marriage process (right censoring). This imposes a difficulty on us when we wish to perform a demographic projection of a future course of young cohort experiences. The extent of the censoring effects on parameter estimation depends both on exactness of model specification and data adequacy. Here we conduct some experiments in which censoring is artificially performed during parameter estimation using the data of non-censored cohorts to assess those effects on estimated value of parameters.

The estimated values of λ with artificial censoring show chaotic fluctuations under standardized age zero (age range before the mode). However, the values become stable and close to the estimates of no censor around and after standardized age 5.0, which approximately corresponds to normal age 36-40. Hence we may trust estimates of λ for cohorts who completed the marriage process at least up to 40. Similarly parameters C, u and b indicate that estimates with censor after standardized age 5.0 are mostly trustworthy.

Close examination of estimates of C (proportion eventually marrying) reveals that the differences between estimated and true value (estimates without censor) are within a range of −1.5% to 1.0% around and after standardized age 2.0, which corresponds to normal age 28-32 according to cohorts in Japan. Therefore we may expect that we can estimate proportion eventually marrying (consequently proportion never marrying) for the cohort who has completed the marriage process up to 32 with an error of $\pm 2\%$.

If true values of some parameters are given, other parameters are expected to be more reliably estimated. Parameter λ is supposed to be stable in value. In fact, the widely accepted CM standard schedule is nothing more than a λ-fixed version (= -1.287) of the GLG schedule according to Swedish experiences. Estimation experiments with simulated censoring giving the true value of

λ (estimated value without censor) are conducted. The result shows that the reliability of estimated values of C, u and b are improved, and the reliable range expands to younger ages as well.

As for C, differences between estimated and true value are within a range of −0.4% to 0.2% with censor at standardized age 2.0 and older. In this case we can reasonably expect that we are able to predict the proportion never married for cohorts who are above age 30 with an error of $\pm1\%$.

Not only C, but also parameter u and b are estimated more accurately all together if true λ is given. Parameter u, location parameter that appoints location of the mode, is estimated within a range of −0.015 to 0.01 of the target when censor at standardized age 2.0 and older. For the same condition, parameter b is estimated within a range of −0.05 to 0.01 around the target value. These are to be adequately accurate for most demographic applications. Parameter u and b are only determinants of the first two moments, i.e. the mean and variance of age at first marriage, if λ is fixed. Then similar stabilities are expected for the moments.

Hence identifying plausible values of λ for young cohorts is essential to predict demographic measures of their marriage behavior. In this connection, it seems imperative to inquire how values of λ are determined. A finding that a mixture of different marriage types with different timing, such as arranged marriages in particular, makes λ large (small in absolute value) is one important clue (Kaneko, 1991). However, other forces to alter λ are considered to exist. These are discussed later.

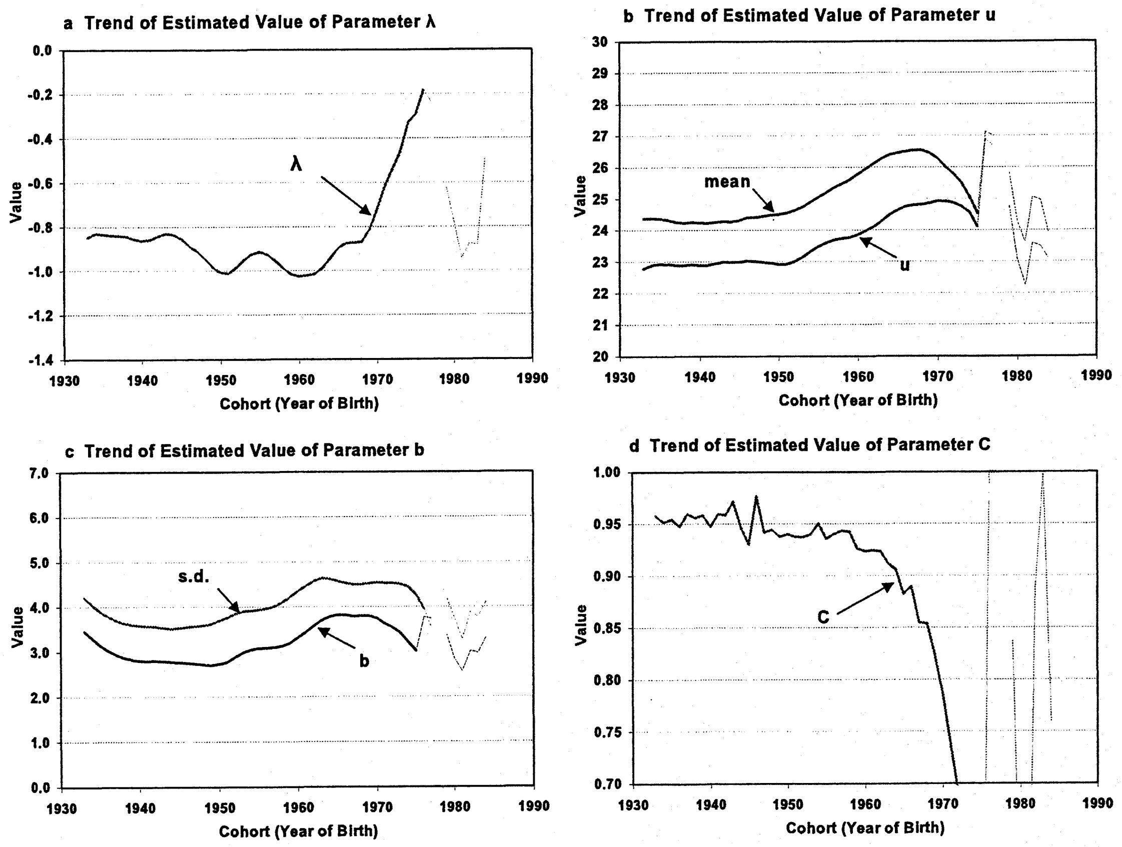
Parametric Estimation of Cohort Marriage Schedules

From the estimated first marriage rates by age and year from 1950-2000, the lifetime first marriage experiences of 16 single year cohorts from 1935-1950 are reconstructed over ages 15-49. However, the relevant cohorts to the unprecedented nuptiality and fertility decline in Japan since the mid 1970s are mostly those born after the 1950s. We make use of the adjusted GLG model on Japanese females described above for this purpose. It is fitted to cohort first marriage processes to estimate the lifetime behavioral measures.

The model schedule is fitted to each cohort experience by estimating model parameter values specific to the cohort through the maximum likelihood method, which is applicable to censored data generated by young cohorts. Firstly, parameter estimations are performed without any constraint on parameter values in order to obtain estimated and projected marriage trajectories for cohorts that both have fully and substantially completed first marriage schedules (cohorts 1935-60). Then we try to extend estimation to younger cohorts who are at relatively early stages in the process by fixing parameter λ at feasible values.

Figure 4 illustrates trends of estimated parameter values with no constraints. For cohorts who have completed the marriage process, i.e. those born in years up to 1950, predicted measures by the model agree almost exactly to the observed, since model schedules fit the actual experiences quite well. However, censoring effects on estimates are apparent in younger cohorts born after the mid 1960s, making estimation results increasingly implausible afterward. According to criterion of reliability in estimated value of C assessed above, we employ free estimation up to cohorts with a censor at standardized age 5.0 (approximate normal age 36-40). The boundary corresponds to cohorts born in 1960 in our data set.

Figure 4 Trends of Estimated Parameter Values for Female Birth Cohorts: Without Constraint



Note: Convergence was not attained for cohort 1978.

As for cohorts born after 1960, the value of the shape parameter λ is to be fixed while the other parameters are freely estimated. According to all free estimation, values of λ go apparently anomalous starting from the cohort of 1969 after a short plateau during 1965-68 (Figure 4-a), and should be discarded. The criteria for reliable estimation with fixed λ described above also suggests that the border of feasible estimation is around the cohort of 1970. Hence we limit our observation up to the cohort born in 1970.

Which value should we fix λ to for cohorts born from 1961 to 1970? According to the free estimation, the value of λ shows upward development during 1961 to 1970. It is not certain if the trend is actually happening or is a just pretense due to the censoring effect. Previously, we found that the shape value becomes larger (smaller in absolute value) when marriages are a mixture of different types of marriages with distinct time schedules, in particular with a coexistence of non-arranged and arranged marriages (Kaneko, 1991). Since arranged marriages have been diminishing throughout the postwar period, the value of λ is expected to decrease instead of to increase as the results of free estimation indicate. Thus, first we fix λ at the level of 1960 so as not to let λ increase.

However, since an upward turn is also observed in the trend of λ for the previous cohorts with supposedly reliable estimates (i.e. cohorts of 1952-55), we cannot fully exclude the possibility of λ to rise for the younger cohorts of 1961-68. In particular, if there exist upper barrier to marriage propensity in late ages, delay in marriage could result in more symmetric shape of marriage schedule, and in rise of λ. Hence, we provide an alternative prediction in which a free estimation is employed for cohorts of 1961-68, and then λ is fixed at the level of 1968 for cohorts of 1969 and 1970.

The results of parameter estimation are presented in Table 1 with five other measures of marriage behavior derived from predicted schedules generated by estimated parameters. These include a proportion never married at age 50 (γ), the mean, two types of median, and standard deviation (SD) of age at first marriage. We provide two types for median. Those are median 1: the median of age at first marriage (age under which a half of those eventually marrying have got married), and median 2: the median of first marriage schedule (age at which the proportion married attains 50%).

Table 1 Estimated Parameter Values and Some Measures of First Marriage Schedules of Japanese Female Cohorts

a. Free Estimation

Cohort	Estimated Parameter Values				Measures of Schedule				
(Birth Year)	λ	u (mode)	b	C	γ (%)	mean	SD	median1	median2
1933	-0.848	22.78	3.46	0.958	4.6	24.38	4.21	23.83	24.00
1934	-0.832	22.89	3.31	0.952	5.1	24.39	4.04	23.87	24.07
1935	-0.835	22.93	3.16	0.955	4.7	24.37	3.90	23.87	24.04
1936	-0.838	22.92	3.04	0.948	5.3	24.33	3.80	23.83	24.03
1937	-0.839	22.90	2.95	0.960	4.1	24.26	3.71	23.78	23.92
1938	-0.843	22.89	2.88	0.956	4.4	24.24	3.65	23.76	23.91
1939	-0.855	22.91	2.84	0.958	4.2	24.27	3.61	23.78	23.92
1940	-0.863	22.90	2.80	0.947	5.3	24.25	3.59	23.76	23.94
1941	-0.859	22.90	2.80	0.960	4.0	24.24	3.58	23.75	23.89
1942	-0.843	22.95	2.81	0.959	4.1	24.27	3.57	23.79	23.93
1943	-0.831	23.00	2.80	0.971	2.9	24.29	3.54	23.83	23.92
1944	-0.835	22.99	2.78	0.946	5.4	24.28	3.52	23.82	24.00
1945	-0.854	23.00	2.77	0.930	7.0	24.32	3.54	23.84	24.08
1946	-0.888	23.03	2.76	0.977	2.3	24.41	3.56	23.90	23.98
1947	-0.911	23.01	2.74	0.942	5.8	24.42	3.57	23.90	24.10
1948	-0.945	22.99	2.72	0.945	5.5	24.45	3.59	23.90	24.08
1949	-0.982	22.96	2.70	0.938	6.2	24.49	3.63	23.90	24.11
1950	-1.006	22.93	2.73	0.940	6.0	24.51	3.69	23.90	24.10
1951	-1.012	22.93	2.77	0.938	6.2	24.55	3.76	23.92	24.14
1952	-0.984	23.01	2.88	0.937	6.3	24.63	3.85	24.01	24.24
1953	-0.949	23.14	2.98	0.940	6.1	24.74	3.90	24.15	24.37
1954	-0.923	23.32	3.05	0.950	5.1	24.89	3.92	24.32	24.51
1955	-0.915	23.48	3.08	0.936	6.5	25.05	3.94	24.48	24.73
1956	-0.927	23.60	3.09	0.940	6.1	25.20	3.96	24.62	24.85
1957	-0.956	23.70	3.10	0.943	5.8	25.36	4.02	24.75	24.97
1958	-0.987	23.75	3.14	0.943	5.9	25.49	4.10	24.84	25.07
1959	-1.015	23.78	3.20	0.926	7.5	25.61	4.21	24.93	25.23
1960	-1.024	23.86	3.31	0.924	7.8	25.77	4.34	25.06	25.39

Table 2 Estimated Parameter Values and Some Measures of First Marriage Schedules of Japanese Female Cohorts (continued)

b. λ is fixed after 1961

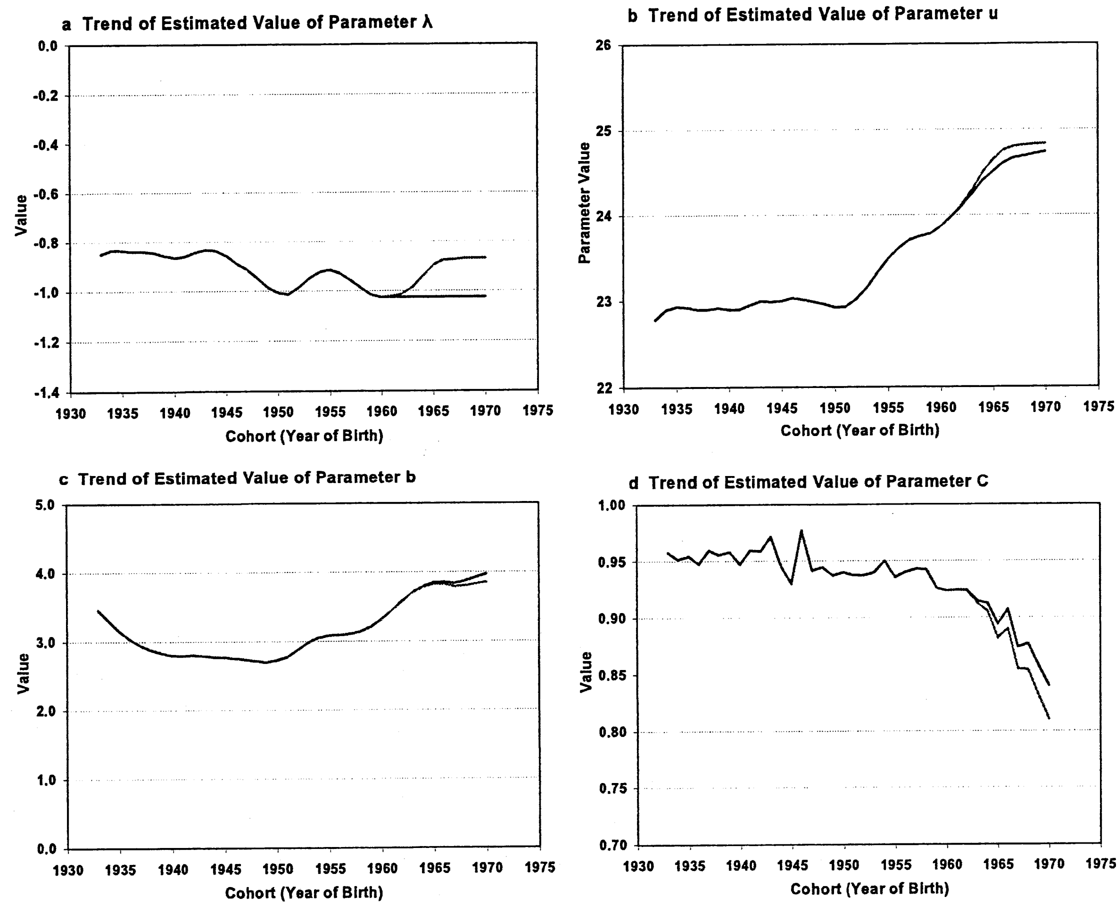
Cohort	Estimated Parameter Values				Indices of Schedule				
(Birth Year)	λ	u (mode)	b	C	γ (%)	mean	s.d.	median1	median2
1961	-1.024	23.98	3.44	0.925	7.8	25.95	4.46	25.23	25.56
1962	-1.024	24.10	3.57	0.924	7.9	26.13	4.57	25.40	25.75
1963	-1.024	24.25	3.70	0.915	8.9	26.33	4.64	25.59	26.00
1964	-1.024	24.40	3.79	0.912	9.2	26.52	4.62	25.77	26.21
1965	-1.024	24.51	3.84	0.894	11.1	26.65	4.58	25.90	26.44
1966	-1.024	24.61	3.85	0.907	9.8	26.75	4.53	26.00	26.47
1967	-1.024	24.66	3.84	0.874	13.1	26.80	4.50	26.05	26.71
1968	-1.024	24.69	3.87	0.877	12.8	26.83	4.50	26.09	26.73
1969	-1.024	24.71	3.92	0.858	14.8	26.88	4.52	26.13	26.91
1970	-1.024	24.74	3.97	0.840	16.6	26.92	4.54	26.17	27.08

c. λ is fixed after 1969

Cohort	Estimated Parameter Values				Indices of Schedule				
(Birth Year)	λ	u (mode)	b	C	γ (%)	mean	s.d.	median1	median2
1961	-1.021	23.98	3.44	0.925	7.8	25.95	4.46	25.23	25.56
1962	-1.013	24.12	3.58	0.924	8.0	26.12	4.57	25.40	25.75
1963	-0.985	24.30	3.70	0.913	9.2	26.28	4.64	25.59	26.01
1964	-0.941	24.50	3.78	0.906	9.9	26.40	4.62	25.76	26.22
1965	-0.897	24.64	3.82	0.882	12.2	26.45	4.58	25.86	26.46
1966	-0.875	24.75	3.82	0.890	11.4	26.50	4.53	25.94	26.49
1967	-0.874	24.80	3.79	0.855	14.9	26.53	4.50	25.98	26.72
1968	-0.868	24.82	3.80	0.854	15.0	26.54	4.50	25.99	26.74
1969	-0.868	24.83	3.83	0.832	17.2	26.56	4.52	26.01	26.90
1970	-0.868	24.83	3.85	0.810	19.4	26.57	4.54	26.02	27.07

In Table 2, alternative results with different settings to λ for cohorts after 1961 are presented in separate tables. The trends of estimated parameters are also portrayed in Figure 5. Branches of the graphs after 1960 indicate the results of alternative estimations. The trends show smooth continuous transition cohort to cohort except relatively large fluctuation in C for cohorts born at the end of World War II, probably caused by some inconsistency in original statistics.

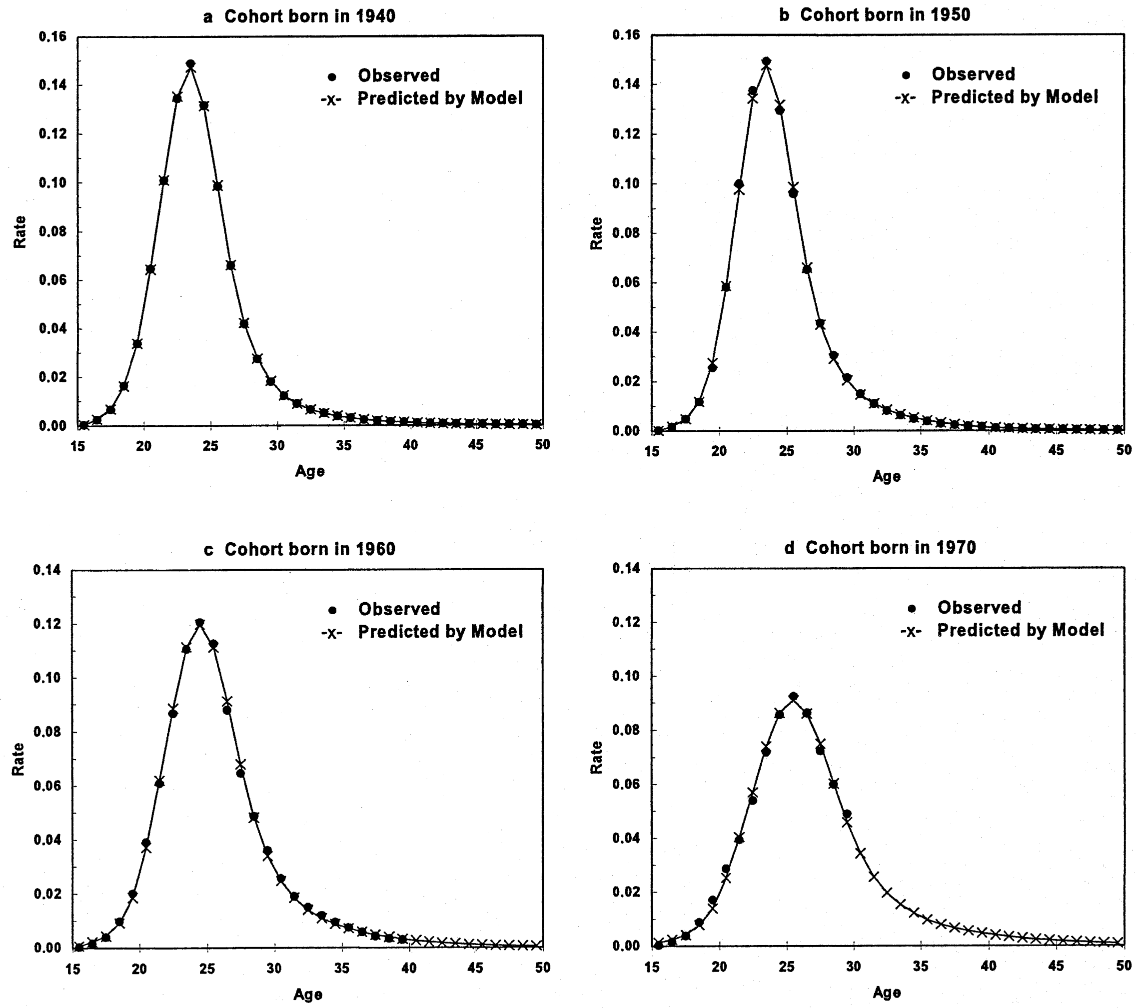
Figure 5 Trends of Estimated Parameter Values for Female Birth Cohorts: With λ Fixed for Cohorts Born after 1961



Predicted marriage schedules from the results of parameter estimation are contrasted with those observed in Figure 6. The model follows the actual experiences quite well, though exactitude of fit becomes slightly weak in younger cohorts[7].

[7] Predicted schedule for the cohort of 1970 in Figure 5-4 is one from estimation with λ fixed at level of 1960. Alternate schedule with λ fixed at level 1968 fits only slightly better.

Figure 6 Observed and Predicted Age Specific First Marriage Rate for Selected Cohorts



Note: Predicted Schedule for cohort born in 1970 in this figure is from estimation with fixed λ at level of 1960.

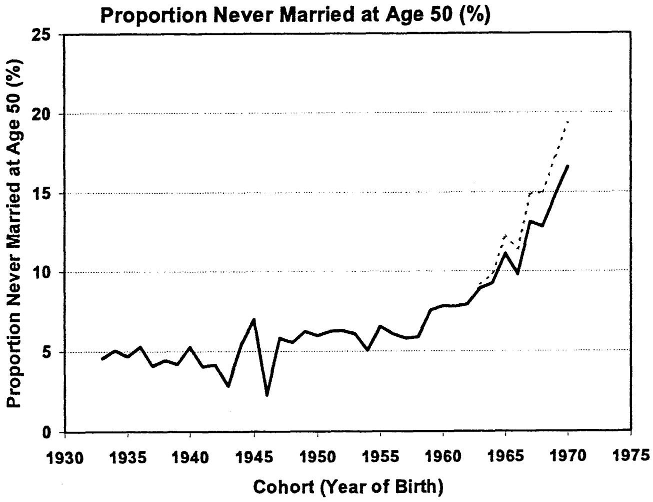
Results

For the cohorts of 1935-1970[8], trends of the proportion never married at age 50 (γ), the mean, two types of median and the standard deviation (SD) of age at first marriage are shown in Table 1 and Figure 7. The mode of age at first marriage is also shown as a trend of parameter u in Table 1and Figure 5.

[8] In the figures and tables tentative estimates for cohorts born in 1933 and 1934 are included.

Figure 7 Trends of Estimated and Projected Measures of First Marriage Schedule

a. Trend of Estimated and Projected Value of Proportion Never Married at Age 50 (%)



b. Trend of Estimated and Projected Value of Mean Age at First Marriage (Year)

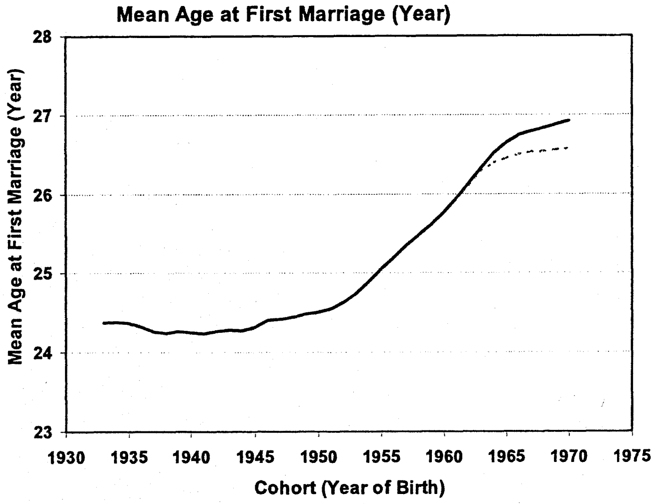
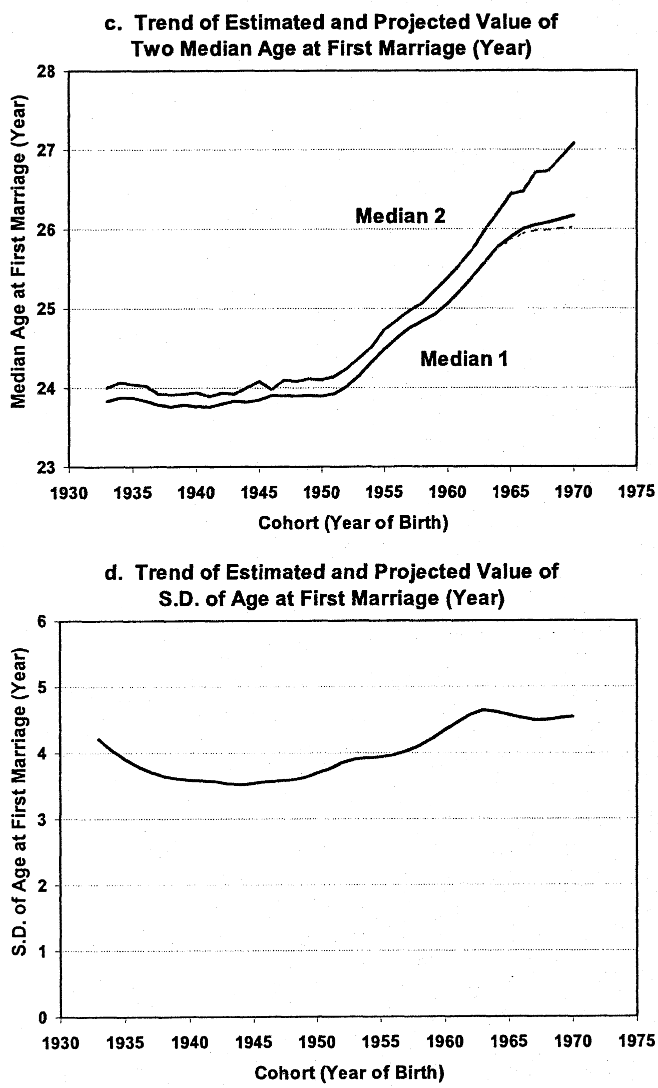


Figure 7 Trends of Estimated and Projected Measures of First Marriage Schedule (continued)



Note that observation of trends over cohorts born in from 1951 up to 1970 with certain reliability is made possible only through application of the GLG model with the adjustment we devised. As noted above, relying solely on observed statistics, the lifetime measures are available only for cohorts of 1935-1950, which have little relevance to the recent drastic decline in fertility and nuptiality in Japan.

As for the proportion never married, the trend for cohorts born in the mid 1930s through 1970 are divided into three phases. First, for the cohort born before the end of the World War II, the proportion had stayed at a level slightly below five percent. From the oldest in our data set, at least 10 cohorts are included in this phase until the measure is disturbed in cohorts born in the years at the end of the war. In the second phase, the proportion rose across the five percent line to about six percent and stayed there until the cohort born in 1958, involving 12 single year cohorts. In the third, for the following cohorts starting from 1959, the proportion started to increase very steeply until the cohort of

1970, which is the last cohort we safely estimate. The alternative projection from estimation with a relaxed shape suggests even a steeper proportional rise. This change has sustained for at least 12 cohorts so far and seems to continue into the following cohorts judging from a trend at the end.

As for the mean age at first marriage for those who are eventually marrying in each cohort, the trend is characterized also by a sharp increase in the latter half of the target span. However timing of the onset for the change differs from that of the proportion never married. The mean had been at halfway of 24 years of age until the cohort born in the early 1950s. Then it showed a remarkable increase for some thirteen cohorts until it decelerated in the cohort born in the late 1960s. It seems that the mean is about leveling off afterward. This new trend is even clearer in the alternative projection.

The steep increase in the mean indicates that first marriage schedules have rapidly shifted toward older on age axis. The shift is also represented by a change in the mode, whose values are carried by parameter u. The trend of u graphed in Figure 5 illustrates similar development to that of the mean with even clearer turning points. It started to increase with the cohort of 1952, and started to decelerate with the cohort of 1965, and almost ceased to increase after the cohort of 1968. In the alternative projection, the level off is clearly attained after the cohort of 1968. Median 1 and 2 shows almost parallel trends with the mean and mode until the mid 1960's, but afterward, median 1 still follows the parallel path with the mean and mode to level off, while median 2 alone continues to rise. All these indices are measures of marriage timing, and indicate all together that a rapid timing shift started from the cohort of 1952, and slowly ended after the mid 1960s. Only median 2, which conveys quantum factor of cohort nuptiality as well as timing, takes a different course after the mid 1960s.

The trend of the standard deviation (SD) of age at first marriage demonstrates a somewhat different course from the other measures in cohorts born before World War II. It decreased at first until cohort 1944, then turned to increase moderately throughout postwar cohorts until 1963, being followed by a leveling off or even minor decrease afterward. The recent leveling off of SD also suggests even more clearly that rapid timing transformation of first marriage schedules from the viewpoint of dispersion is about to end.

In summary, the demographic history of lifetime first marriage behavior among Japanese women is divided into five phases represented by the following groups of cohorts:

Group A (cohort born in 1933–1944 (age 56-67 as of 2000), 10 cohorts);

First marriage behavior is stable except minor reduction of variance of marriage timing.

Group B (cohort born in 1947–1951 (49-53 year-olds), 4 cohorts);

The proportion never married shifted to a slightly higher level by 1.5 %.

Group I (cohort born in 1952–1958 (42-48 year-olds), 7 cohorts);

The mean age at first marriage (and the mode) started to rise, while the proportion never married is unchanged.

Group II (cohort born in 1959–1964 (36-41 year-olds), 5 cohorts);

The proportion never married started to increase, while the mean, mode and SD continued to rise in similar rates as before.

Group III (cohort born in 1965–1970 (30-35 year-olds), 6 cohorts);

The mean, mode and SD decelerated and are about to level off, while the proportion never married continues to increase even at an accelerated pace.

Cohorts born in 1945 and 1946 are regarded as a transition generation between Group A and B, and are to be included in either of the groups, which is uncertain due to their radical cohort size changes leading to a disturbance in their statistical features.

Among the phases, behavioral changes relevant to the prolonged period of fertility decline since the mid 1970s are carried by Group I and after. These changes of female marriage behavior in Japan were initiated with a timing delay by Group I (born in 1952-58). It was followed by a new tendency of gradual diffusion of those who stay never married throughout reproductive ages, in addition to continuing marriage postponement, in Group II (born in 1959-64). Then the timing shift decelerates and is about to end in Group III[9], while the diffusion of lifetime never married continues at even a faster pace.

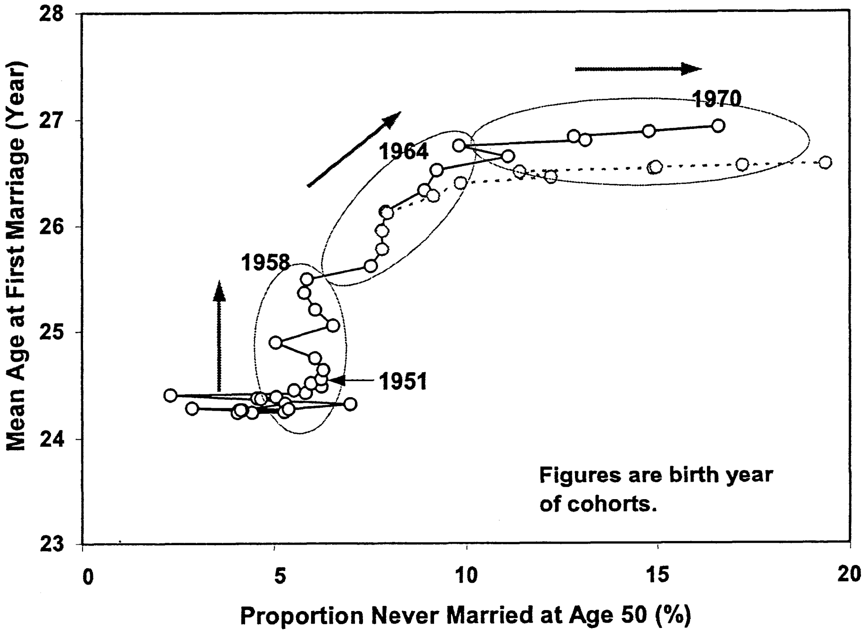
Despite the surmises by many, female baby boomers did not change their marriage behavior in terms of aggregate features as compared to those of preceded generations. Instead, the onset of all changes started from the cohort of 1952, which are a few years younger than the baby boomers, and many of which are supposed to be partners of male baby-boomers.

A new finding reveals that there was a time lag of seven years for cohorts between the onset of timing shift in marriage and that of the diffusion of never married. It has also been found in very recent cohorts that the timing shift is about to end while retreat from marriage has accelerated. Although this delay and retreat are expected to be closely related, it has been confirmed that either may solely take place in a cohort. The relationship between the mean and proportion seen in Japanese female cohorts is illustrated in Figure 8 as a sequential scattergram, where the change started with a vertical rise (phase I) followed by a diagonal rise (II) and then by a horizontal shift (III). The delay and retreat went together only in the part of the diagonal rise.

---

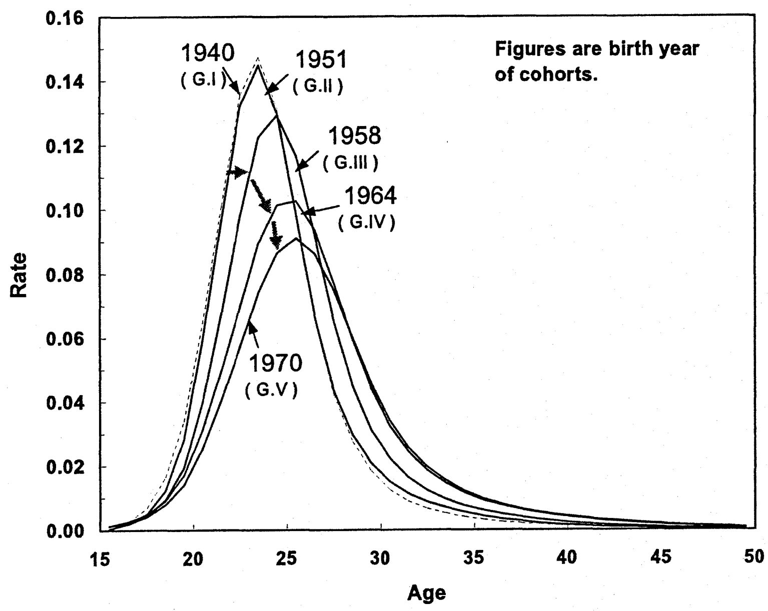
[9] Deceleration of increase in SD of age at first marriage already stated from late cohorts in Group II. The boundary of the trend in timing shift is somewhat ambiguous.

Figure 8 Relationships between Proportion Never Married and Mean age at First Marriage



Implication of this combined changing pattern on marriage schedules is illustrated in Figure 9. Before cohort of 1951, no visible change is found in marriage schedules. From cohort 1951 (Group B) to 1958 (I) only a horizontal shift took place (with some horizontal dispersion). Then toward cohort 1964 (II) a combined change of horizontal shift and reduction in area under the curves occurred, followed by a change from reduction alone to cohort 1970 (III).

Figure 9 Changes in Cohort First Marriage Schedule of Japanese Female



As a result of a timing shift, the mean age at first marriage of Japanese female cohorts has increased from 24.6 in cohort 1951 (end of Group II) to 26.9 in cohort 1970 (most recent in Group V) by 2.4 years in 20 years (2.0 years in the alternative estimation). The corresponding figures in the mode are 1.8 years between cohort 1951 and 1970 (1.9 years in the alternative). Two years of shift appear to be small. But suppose that the proportion married of a cohort doubles in a year around the peak age of marriage, which is typically the case. Then a single year shift of schedule toward older makes the proportion married half at the same age of previous cohort around the peak, which is drastic enough to impact society if it happens in a short period of time. In fact, it has exercised its impact on fertility by removing births that would be given from lost marriages in youth. However according to our results, these are expected to end in the cohorts after 1970.

The proportion never married was 5.9% (or equivalently one out of 17 women) in the 1958 cohort, the last cohort before the rise stated (in Group II). Since then it has increased up to 16.6% (or one out of 6 women) in the cohort 1970 (19.4 or one out of five women in alternative estimation). This triple fold increase had been taking place only in 12 years of cohort. Furthermore, judging from the slope at the end of the observation, an even higher proportion of women are anticipated never marrying for life in the following generations. It is feasible to assert that lives in which one woman out of five remains unmarried throughout her life will be lived by generations now in their 20s. These estimated proportions are extremely high as measures at national level even in international and historical context. On the grounds that marriage has been the very basis of family formation in society, the impact of this unprecedented situation on society and people's lives would be immeasurable.

Discussion and Conclusion

The present study aims at better understanding the current development concerning the rapidly transforming first marriage situation in Japan. For this purpose, we reconstructed the historical development of marriage behavior in the postwar period in terms of lifetime measures of first marriages for female birth cohorts born after 1935. However, by relying only on observed statistics, the lifetime measures are available only for cohorts born in 1935-1950. Hence the generalized log gamma (GLG) model, as an alternative parameterization of the Coale–McNeil (CM) model with some statistical advantages, is applied to estimate and project the cohort first marriage schedule so as to provide long-term trends of the lifetime measures for the cohorts. However, since the model performs relatively poorly for Japanese female schedules, which is critical against a reliable prediction of the measures for young cohorts that have yet to complete the marriage process, we devised a procedure to combine the GLG model and the empirical adjustment specific to Japanese females. Then we applied the resulting model to the time series data set of age-specific first marriage rates during the postwar period to construct trends of lifetime measures of first marriage behavior. With these techniques, estimates of lifetime marriage measures for cohorts born in 1951-70, which are relevant to massive decline in fertility and nuptiality since mid 1970's in Japan, are obtained.

As a result, we found that the history of Japanese cohort first marriage behavior starting from the cohort born in 1935 is divided into five phases. Firstly, among cohorts born in 1933-1944 (Group A), the measures are stable except for a minor reduction in the variance of marriage timing. Secondly, among the cohort born in 1947-1951 (Group B), which are the baby boomers and a few successive cohorts, the proportion never married shifted to a slightly higher level by 1.5 %, but otherwise little changes were found. Then in the cohort born in 1952-1958 (Group I), the mean age at first marriage (with the mode and median) started to rise, while the proportion never married is unchanged. In the cohort born in 1959-1964 (Group II), the proportion never married started to rise, along with the delay in marriage timing continuing at the same pace as before. Finally, in the cohort born in 1965-1970 (Group III), the mean, mode, median (of the first kind) and SD decelerated and are about to level off, while the proportion never married continues to increase even at an accelerated pace. It seems that the tendency of Group III still continues into the following cohorts.

The important findings about change in the cohort marriage behavior are summarized as follows;

(1) The cohort that started the delay and therefore started the whole process of the historical marriage and fertility transfiguration is the one who was born in 1952. This is younger than the baby boomers (1947-49) by three to five years, and many in it had partnerships with male baby boomers.

(2) There is a time lag between the onsets of delaying marriage timing and retreating from marriage. The delay in marriage timing started by the cohort of 1952 and the retreating from marriage started to move toward diffusion by the cohort of 1959. So the time lag is seven years in the cohort.

(3) The delay decelerated after the cohort of 1965, while the retreat continues to diffuse at even a faster pace. According to the trend up to the cohort of 1970, the timing shift appeared to cease in the following cohorts that are in there 20s today. Instead, the proportion never marrying continues to increase at an unprecedented level. It is possible for them that the

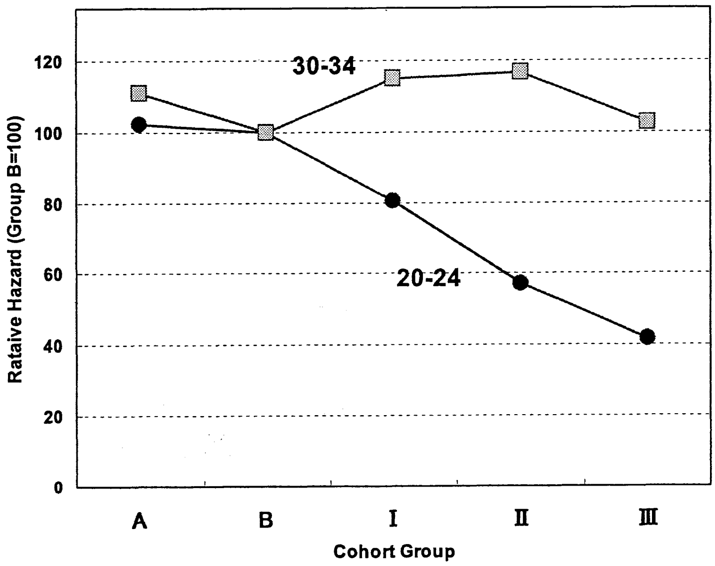
proportion exceeds 20 percent.

The finding (1) raises the question whether radical changes in cohort size affected the marriage behavior of the cohort of 1952 as a trigger of the following prolonged shifts through favorable condition of marriage market for them. It is suggested that a further pinpointed investigation at or around the cohort of 1952 is effective to be conducted on this issue.

As for the findings (2) and (3), some hypotheses about the mechanism of behavioral change in first marriages are set up. The presence of time lag between onsets of postponement and retreat in marriage behavior suggests that when adaptation in marriage behavior to new situations is needed, cohorts tend to start adjusting the timing at first because of its lower impact on their life. Then they start to resort a substantive alteration of life by retreating from marriage itself when further adaptation is required. There are two mechanisms hypothesized to make it take place. First, There may be a certain limit for marriage propensity to rise in late ages so that some of postponed marriage should be lost for life, when the limit is attained by prolonged postponement. Second, postponement and retreat may be somewhat independent behavioral adaptation to resolve somewhat different difficulties so that they could emerge in different timings.

In order to test these hypotheses, we derived the hazard rate for each of cohort group by age group. Figure 10 indicates relative hazards of two age groups, young (20-24) and late (30-34) for the cohort groups. The relative hazards are chosen so as to be 100 for Group B, which is those before the behavioral changes started. According to examination on the hazard pattern, the first hypothesis seems applicable to Phase II. In cohorts of Phase I, the hazard to marry in young ages decreased, while it increased in late twenties through early thirties. This pattern of hazard change implies that those who postponed their marriage in young ages among cohort group I got married later. On the contrary, in cohorts of Phase II, the hazard rates in late ages sit around the same level as Phase I, although it continued to decrease in young ages. It implies that some of those who postpone their marriage in this cohort group gave up to marry even later ages so as not to raise hazard in these ages. These patterns are predicted by the hypothesis that there is a certain limit in propensity to marry in these late ages, and postponement brings about rise in proportion never marrying, which we actually observed in Phase II.

Figure 10 Changes in Relative Hazard to Marry in Young and Late Ages: Female Cohort Average



However, in cohorts in Phase III, the hazard in early thirties are found to decline slightly, while it continued decline in early twenties. These patterns are predicted by the second hypothesis that retreat from marriage could take place somewhat independently of postponement, because decline in hazard in late ages could not be explained either by simple postponement or by the first hypothesis.

Consequently, our simplified conclusions concerning results (2) and (3) are as follows: the rise in the proportion never marrying in Phase II was mainly result from prolonged marriage postponement that had started earlier in the previous cohorts, because of presence of upper limit in marriage propensity in late ages: however the further rise in the proportion never marrying in Phase III was from somewhat independent cause of postponement, and was genuine decline in marriage propensity through lifetime as a new behavioral adaptation especially in later ages, although the behavioral shift from Phase II and III should be taking place continuously.

The emergence of the non-postponement-related rise in proportion never marrying in the recent female cohorts (Group III) in Japan leads to a prospect of further substantive decline in nuptiality and fertility, which would be unbearable to the society that is already among the world lowest fertility society. However, it should be noted that what we found is life course transformation that is *now* taking place in the marriage processes of the young generations, although the measures used are lifetime indices predicted as of the distant future. It indicates that a totally different setting is emerging for family formation among generations now in their 20s and early 30s in Japan.

Another significant question about the result of the present study is whether the Japanese changing pattern of first marriage behavior epitomizes changes of marriage in other countries. Although rises in age at marriage have been universally witnessed in the developed countries during late 20th century, Japanese experience is quite unique in that the transformation of marriage behavior has been taking place without substantial increase in cohabitation and in birth out of wedlock. This is suggesting that there are different paths in transformation of family formation behavior. International comparisons among those processes from the viewpoint of here hypothesized sequence of behavioral change, i.e. pure marriage postponement, postponement with postpone-related retreat, and non-postpone-related retreat, may provide a platform to clarify the differences in casual interactions with social settings and in future courses of development.

Some precautions are needed about the lifetime measures estimated in this study. They should be accurate only if the model reproduces precisely the actual cohort marriage schedules. We employed the GLG model with the empirical adjustment for the estimation, and it is confirmed that the model reproduces the schedule almost exactly for cohorts that completed their marriage processes. However, whether it fits equally well or not for the young cohorts that have not completed the marriage processes is uncertain, since there is no way to confirm.

The key issue is represented by course of the shape value (λ) of the GLG model. Although the shape value is constant for widely used Coale-McNeil model, it is allowed to vary in our study to express distinctiveness of Japanese female cohorts and its shifts in time. If the shape value is predicted correctly, it is believed that the estimated lifetime measures are fairly accurate. The free estimation of λ indicated steep rise from cohorts born early 1960's. Since we do not know if it is genuine shift of the shape or merely artificial effect from censoring, two separate courses of λ for the young cohorts born in 1961and after are employed.

Two antithetic speculations about course of the shape may be made: (1) As presence of arranged marriage make the shape value large (toward symmetry) and proportion of arrange marriage has decreased in recent Japan, the shape should become more skewed in younger cohorts, and (2) Under the hypothesis that there is a certain upper limit in marriage propensity in later ages, which seems valid in cohort Group II, the shape should become more symmetric due to the deficit of would-be-marriage in those ages. As the shape become more symmetric than that of the model used here, estimate of the proportion never marrying should be higher, and the mean age at marriage is lower. Since our primary estimates presented here follow (1), and the alternative estimates mostly follow (2), it is reasonable to consider the true values fall between them.

References

Bloom, David E., 1982, What's happening to the age at first birth in the United States? A study of recent cohorts, *Demography*, Vol. 19(3): 351-370.

Bloom, David E., and Bennett, Neil G., 1990, Modeling American marriage patterns, *Journal of The American Statistical Association*, Vol. 85, No. 412:1009-1017.

Coale, A.J., McNeil, D. R., 1972, "The Distribution by Age of the Frequency of First Marriage in Female Cohorts", *Journal of American Statistical Association*, Vol. 67: 743-749.

Coale, A. J., 1971, "Age Patterns of Marriage", *Population Studies*, Vol. 25(2): 193-214.

Kaneko, R., 1991, "Demographic Analysis of the First Marriage Process," *The Journal of Population Problems*, Vol.47, No. 3: 3-27 (in Japanese with English summary)

Kaneko, R., 1993, "A Projection System for Future Age-specific Fertility Rates," *The Journal of Population Problems*, Vol.49, No. 1: 17-38 (in Japanese with English summary).

Kaneko, R., 2002, "Elaboration of the Coale-McNeil Nuptiality Model as The Generalized Log Gamma Distribution: New Identity and Empirical Enhancement," National Institute of Population and Social Security Research, Working Paper Series (E), No.16: Tokyo.

Kojima, H., 1985, "Coale-McNeil no Kekkon Model ni Tuite (About the Coale-McNeil Nuptiality Model)," *Jinkou Mondai Kenkyu*, No. 175: 36-47.

Rodriguez, G., Trussell, J., 1980, *Maximum Likelihood Estimation of the Parameters of Coale's Model Nuptiality Schedule from Survey Data*, World Fertility Survey, Technical Bulletins, No.7.

Takahashi, S., 1978, "Kekkon model schedule no Nihon eno Tekiyouni Tuite (Application of the Nuptiality Model Schedule to Japan)," *Kokumin Keizai Zasshi*, Vol. 138, No. 6: 80-95.

Appendix

Estimation of the Number of First Marriages by Age and Year in Postwar Japan: Correction for Delayed Registration in Vital Statistics

The annual number of marriages reported in the vital statistics does not agree with the true number of marriages that occurred in the same year because a substantial amount of delayed registrations exist. Here, we develop a new procedure of estimation for the number of marriages with the presence of delayed registration, and apply it to Japanese female experiences to estimate the numbers and rates of first marriages by year and age. First, we derive a feasible measure for the delayed registration, and observe it for the period in which it is available. Then, we estimate the number of marriages in intercensal periods as a reference for the estimation of the number by single year from the vital statistics. Finally, we estimate the number and rates of first marriages for cohorts by combining it with the model of delayed registrations.

Measure of Delayed Registration; Average Hazard of Same Year Registration

It is advantageous to have a measure with an interpretative nature for the purpose of prediction. We here derive such a measure of delayed registration of first marriages for the Japanese vital statistics.

Let N_1 and N denote number of first marriages registered in the same year of their occurrence and eventual number of occurrence, and $n(t)$ be density of marriage occurrence in a certain year at time t originated from the beginning of the year. In this case, $N = \int_0^1 n(t)dt$ follows. With $G(y)$ as the distribution function of delayed registration at delay y, N_1 is given as;

$$N_1 = \int_0^1 n(t)G(1-t)dt. \tag{A.1}$$

Assuming occurrences of marriage distribute evenly over the year,

$$N_1 = \int_0^1 n(t)dt \int_0^1 G(1-t)dt = N\int_0^1 G(y)dy. \tag{A.2}$$

Then true number of marriages in that year N is given by;

$$N = \frac{N_1}{\int_0^1 G(y)dy}. \tag{A.3}$$

If the hazard rate of delayed registration within the year is assumed even, its distribution is exponential in this time span that is expressed as $G(y) = 1 - e^{-ry}$ with parameter r. Therefore,

$$N = \left\{1 - \frac{1}{r}(1-e^{-r})\right\}N_1. \tag{A.4}$$

Solving the equation for r,

$$r = W(-qe^{-q}) + q \tag{A.5}$$

where W denotes the Lambert's W-function[10], and $q=N/(N-N_1)$. r is regarded as average hazard rate of delayed registration within the same year of marriage occurrence under assumptions above.

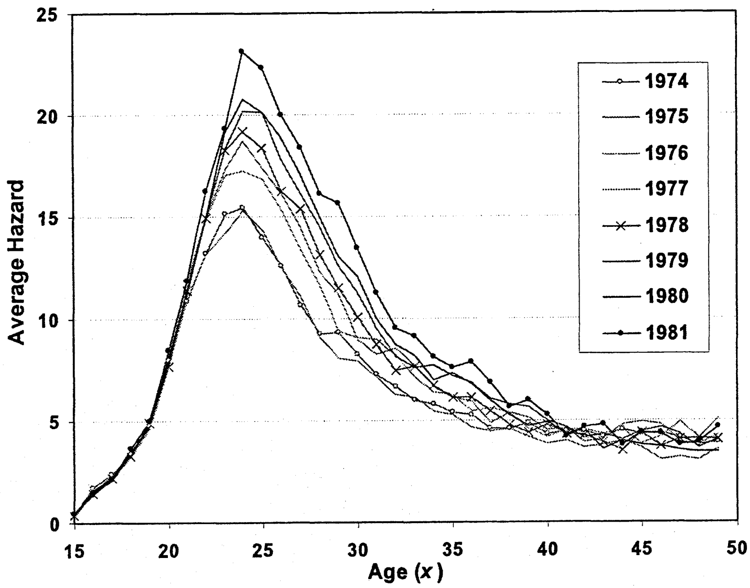
Inversely, if r is given, according to equation (A.4) the following coefficient relates number of first year registration, N_1, to ultimate number of registration, N, as $N = R N_1$:

$$R = \left\{1 - \frac{1}{r}(1 - e^{-r})\right\}. \tag{A.6}$$

The r and R heavily depend upon age, and ought to be functions of age x, as $r(x)$, $R(x)$.

Regarding N_{20}, cumulative number of registrations up to 20 years from marriage incidence, as N, the ultimate number of registrations, $r(x)$ can be observed for the years from 1974 to 1981, since later reports on delayed registration are available for these years. Let $r_{20}(x)$ denote this, since it stands for average hazard of same year registration among cumulative registrations up to 20 years. Plots of the observed $r_{20}(x)$ are presented in Figure A-1.

Figure A-1 Average Hazard Function of Same Year Registration, *r*(*x*): 1974-1981



Observation of Average Hazard of Same Year Registration, $r(t, x)$

To confirm the regularity in the average hazard of same year registration, $r(t, x)$, over the years, we conducted the principal component analysis. In this context $r(t, x)$ is modeled as

$$r(t,x) = \bar{r}_0(x) + k_1(t)\,\bar{r}_1(x) + k_2(t)\,\bar{r}_2(x) + \cdots \tag{A.7}$$

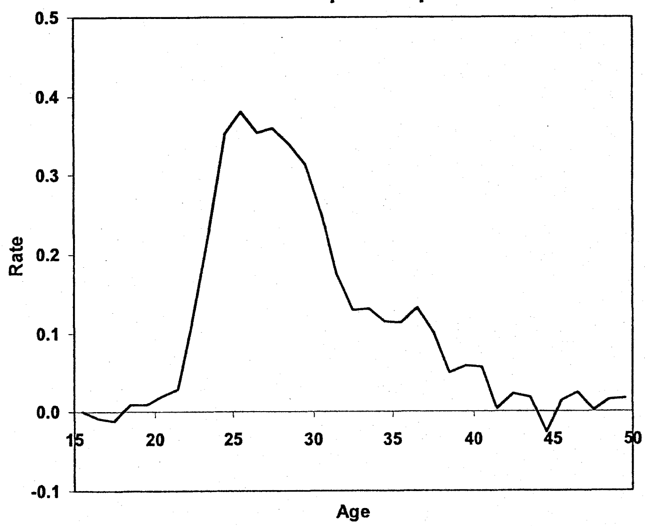
where $\bar{r}_0(x), \bar{r}_1(x), \bar{r}_2(x), \cdots$ are intercept, and first, second, ... principal components, and $k_1(x), k_2(x), \cdots$ are corresponding principal scores. $\bar{r}_0(x)$ is interpreted as average age pattern of $r(t, x)$ over the observation period, and is presented in Figure A-2. Similarly $\bar{r}_1(x), \bar{r}_2(x), \cdots$ are

[10] The Lambert's W-function, $w=W(x)$ is function that make $x = we^w$ happen.

average age pattern of change in $r(t, x)$ in form of linear combination with $k_1(x), k_2(x), \cdots$. Only $\overline{r_1}(x)$ is shown in Figure A-2.

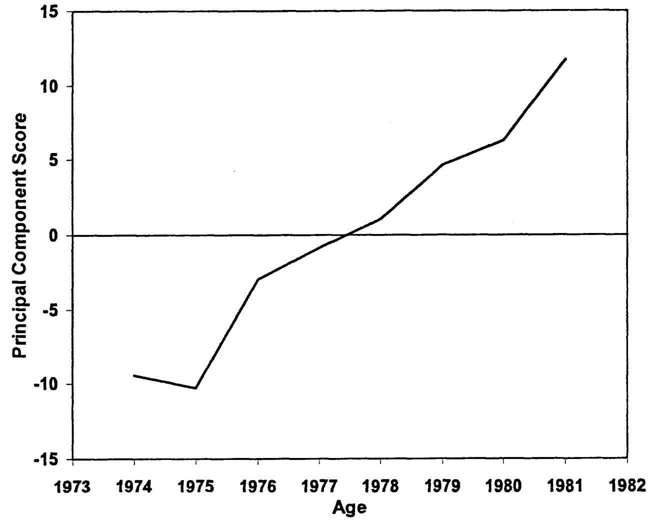
Figure A-2 Average Age Pattern of $r(x)$: Intercept of Principal Component Analysis

Figure A-3 Average Age Pattern of $r(x)$ Change: The First Principal Component



For our observation period, the first component explains 93.4% of the annual variation of the pattern with coefficient $k_1(x)$ in somewhat liner manner over the period (Figure A-4).

Figure A-4 Deviation of $r(t, x)$ from Average by Year: Score of First Principal Component



These results indicate that a high regularity is maintained in the delayed registrations, especially in the age pattern. It looks like the principal component model would be directly applicable for extrapolating $r(t, x)$ into years for which it is to be estimated. However, preliminary examination revealed it inadequate, probably because trends of the principal score are not linear outside of the observed period, and the regularity itself is shifting especially in recent years. Tentative estimates did not agree with the number of married women enumerated by censuses.

Consequently, we cannot rely on a mechanical procedure for the principal component model to extrapolate $r(t, x)$ into years concerned. We need another source of information on the level of delayed registrations or number of marriages itself even in crude form. The most feasible source seems a census that provides information on the population by marital status. Hence, we attempt to estimate the number of marriages in an intercensal period from population ever married enumerated by successive censuses to be referred to when estimation from vital statistics is conducted.

Estimation of Number of Marriages in a Five Year Interval from Census Results

Let M_1 and M_2 denote the number of ever-married women of the same cohort at two consecutive censuses. If there is no migration or death in the cohort during the intercensal period, the number of marriages in this period is simply given by M_2 - M_1. If migration and mortality cannot be ignored, discount for them is necessary. In Japanese cases, number of marriage M_2 - M_1 become mostly minus after age late 30s, which implies that the correction for migration and death is necessary. Here we employ the cohort change rate method to adjust the number to accommodate mobility and mortality.

Let p the cohort change rate, i.e. $p=P_2/P_1$ where P_1 and P_2 are the number of all women in the cohort in each census. Then average hazard of cohort change μ is given by $\mu = -(1/n)\ln p$, where n is the interval of two censuses. Suppose the census cohort change rate is equally applicable to ever-married and never-married women, the number of surviving stayers of the newly married in this period is given as $\int_0^n m(t)e^{-\mu(n-t)}dt$, where $m(t)$ is a density function of marriage occurrence at time

t elapsed since the first census. Assuming intersensal marriage rate $m(t)$ is constant by value m, [11] the integral become $\frac{1-e^{-\mu n}}{\mu} m$. Since M_2 consists of two groups of people, the surviving stayers out of M_1, which amounts $p\,M_1$, and those who stayed alive out of newly married during the period, M_2 is given as:

$$M_2 = pM_1 + \frac{1-e^{-\mu n}}{\mu} m. \tag{A.8}$$

Then the number of first marriages occurred for the cohort during intercesal period N_c is given by;

$$N_c = n\,m = \frac{\mu n}{1-e^{-\mu n}}(M_2 - pM_1) = -\frac{\ln p}{1-p^{-\frac{1}{n}}}(M_2 - pM_1). \tag{A.9}$$

Note that if $\frac{M_2}{M_2} < \frac{P_2}{P_2}$, the estimated number of marriages becomes minus. This would happen if ever married people are more likely to out migrate or die than the total population, or merely from a measurement error mainly taking place where the number of marriages is very small. In the former we should know how much more likely the married is to exit the population than the rest. In the latter, it is probably feasible to set the number of marriage at zero. Even in the former case, the number should be small, and therefore it may be feasible to set it at zero.

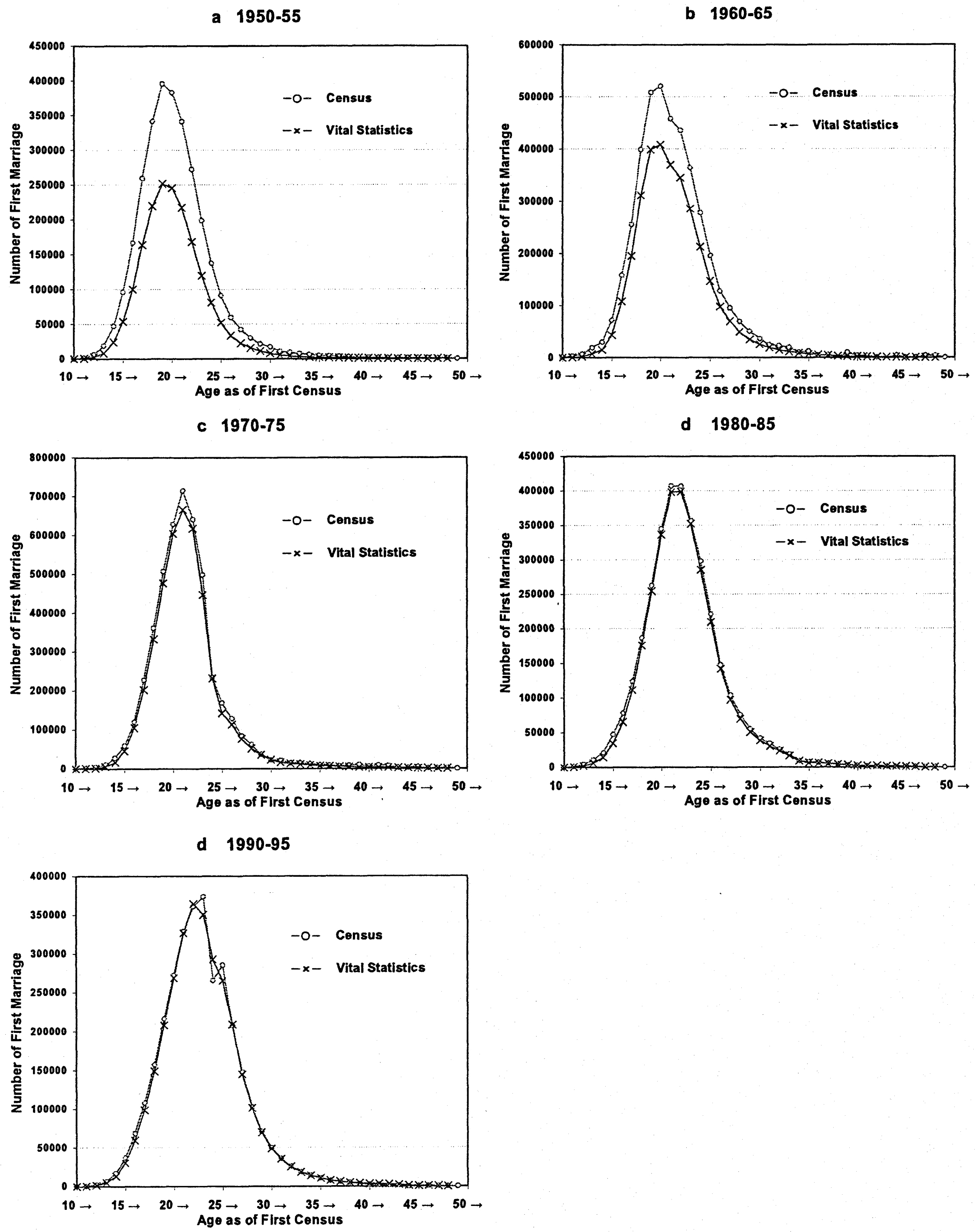
The estimated numbers of marriages for female age cohorts for 10 intercensal periods in Japan are presented in Table A-1. The comparisons with those numbers from the vital statistics without correction of delayed registration for selected period are presented in Figure A-4.

[11] This is a crude approximation for ages in which marriage accelerates or decelerates. However this is only relevant to the adjustment factor, and effect on results is generally small.

Table A-1 Estimated Number of First Marriage of Age Cohorts for Intercensal Periods

Age		1950-55	1955-60	1960-65	1965-70	1970-75	1975-80	1980-85	1985-90	1990-95	1995-2000
at starting census	at ending census										
10	15	224	190	1,399	7,671	989	315	113	61	46	189
11	16	1,304	1,152	2,545	8,424	1,908	1,430	1,381	747	569	891
12	17	6,298	4,792	6,803	12,157	4,596	4,079	4,138	3,532	2,254	3,244
13	18	18,891	15,949	18,931	22,437	11,639	10,135	11,056	9,771	6,562	8,336
14	19	47,402	39,543	29,972	45,378	28,465	22,375	21,149	21,719	17,131	18,330
15	20	96,039	79,817	71,355	98,700	58,775	44,443	47,441	42,209	37,188	33,877
16	21	167,099	133,415	158,136	195,938	120,136	84,030	78,115	70,187	68,419	54,795
17	22	259,438	244,316	255,420	328,181	227,759	138,298	124,450	107,812	108,311	81,905
18	23	341,588	360,065	398,845	452,681	361,625	216,127	187,147	159,290	157,516	117,232
19	24	395,708	444,345	507,908	372,229	507,022	327,648	262,993	184,047	216,464	168,847
20	25	382,905	472,786	520,570	461,217	628,552	418,350	344,902	299,307	272,488	230,105
21	26	341,662	446,292	458,349	583,829	715,271	450,706	407,034	337,627	328,798	297,269
22	27	271,955	395,870	436,347	514,449	641,933	450,076	407,014	360,057	361,112	343,906
23	28	198,818	312,608	363,803	420,369	497,985	394,439	356,917	340,684	373,537	353,168
24	29	137,781	226,558	278,372	298,363	234,190	311,804	297,991	292,915	265,527	333,030
25	30	91,527	153,842	196,012	192,923	169,007	233,220	221,400	229,860	285,357	288,294
26	31	59,931	106,501	128,114	112,824	129,024	165,198	147,641	167,598	207,601	234,787
27	32	42,072	72,923	95,309	85,551	83,681	111,875	104,291	113,945	148,018	180,811
28	33	30,012	49,758	69,128	59,114	62,231	76,384	75,778	76,067	102,313	137,706
29	34	21,375	38,552	50,495	47,990	37,713	31,109	55,316	54,335	70,995	78,912
30	35	17,341	25,995	36,103	29,911	23,497	23,945	41,832	41,330	49,985	78,110
31	36	10,768	19,746	24,721	25,935	21,634	24,913	34,425	29,624	36,460	55,189
32	37	9,303	14,358	22,343	23,458	15,145	18,636	25,824	23,918	25,818	40,949
33	38	7,565	14,509	19,581	17,210	15,428	11,849	18,803	18,921	18,556	29,873
34	39	6,185	11,398	10,264	9,430	13,167	13,951	9,360	15,875	14,359	22,243
35-39	40-44	17,856	21,250	31,909	30,340	45,794	23,538	28,185	45,659	33,500	52,778
40-44	45-49	7,494	10,624	11,131	6,226	29,012	21,975	7,202	14,388	7,553	19,166
Total		2,988,540	3,717,152	4,203,865	4,462,935	4,686,177	3,630,849	3,321,898	3,061,484	3,216,434	3,263,942

Figure A-4 Estimated Number of First Marriage for Intercensal Period from Censuses and Corresponding Number Registered in Vital Statistics



According to the comparison, the reported numbers of marriages by vital statistics for the years before 1970 are substantially smaller than those estimated from the census, suggesting that the delayed registrations beyond year of occurrence are massive for those years. However, for years after then, those estimates from two sources of statistics match well. Therefore, we conclude that some correction for the number of marriages against the reduction by delayed registrations for the years before 1970 is indispensable to have proper views of first marriage behavior, although the correction is not as relevant as those years then after.

It is possible to rely fully on the estimates from censuses to analyze marriage behavior. It is especially plausible for the period before the end of World War II, when the settings of vital statistics do not match with those of the postwar period. However, when relying only on censuses, the estimates are available only for five-year intervals, which is too broad to analyze age patterns of marriage in detail.

Estimation of "True" Number of Marriages of Vital Statistics with Census Estimation

Our aim here is to estimate the ultimate number of marriage registrations that would be observed in many years after marriage actually took place, i.e. N_{∞}. Instead of infinity we think N_{20}. For marriages that occurred in the years from 1974 to 1981, observed numbers of N_{20} are available. For years before this period in our data set, namely 1950-73, no information on delayed registrations is available. For the years of 1982-96, the number of registrations during the first five years after marriage occurs, N_5 is available instead of N_{20}.

For the years of 1950-73, the number of marriages during the intersensal period estimated from population by marital status at relevant censuses (N_c) are related to the number of marriage registrations within the same year of occurrence in the vital statistics (N_1) as:

$$\hat{N}_c = \sum_{y=1}^{n} c\, R_{20,y}\, N_{1,y} \qquad \text{(A.10)}$$

where $N_{1,y}$ is the same year registration in year y of intercensal period, $R_{20,y}$ is the ratio that connects N_1 to N_{20} given in (A.4) with N_{20} as N, specific to year y, and c is the adjustment between estimates from census and the vital statistics obtained from observation for the years of 1974-81 in which all the other quantities are known. $R_{20,y}$ is directly calculated from $r_{20,y}$, the average registration hazard in first year estimated from N_{20}, with formula (A.6).

So far in this section we have omitted the dimension of age for simplicity. But some evidence suggests that $r_{20,y}$ and $R_{20,y}$ shift on age axis along with marriage frequency. Therefore, we need more precise notation for them as functions of time and age, $r_{20}(t, x)$ and $R_{20}(t, x)$, where t and x are year and age at marriage respectively. Then we consider the changes of $r_{20}(t, x)$ over time as:

$$r_{20}(t,x) = B(t)\, \overline{r}_{20}(x - A(t)) \qquad \text{(A.11)}$$

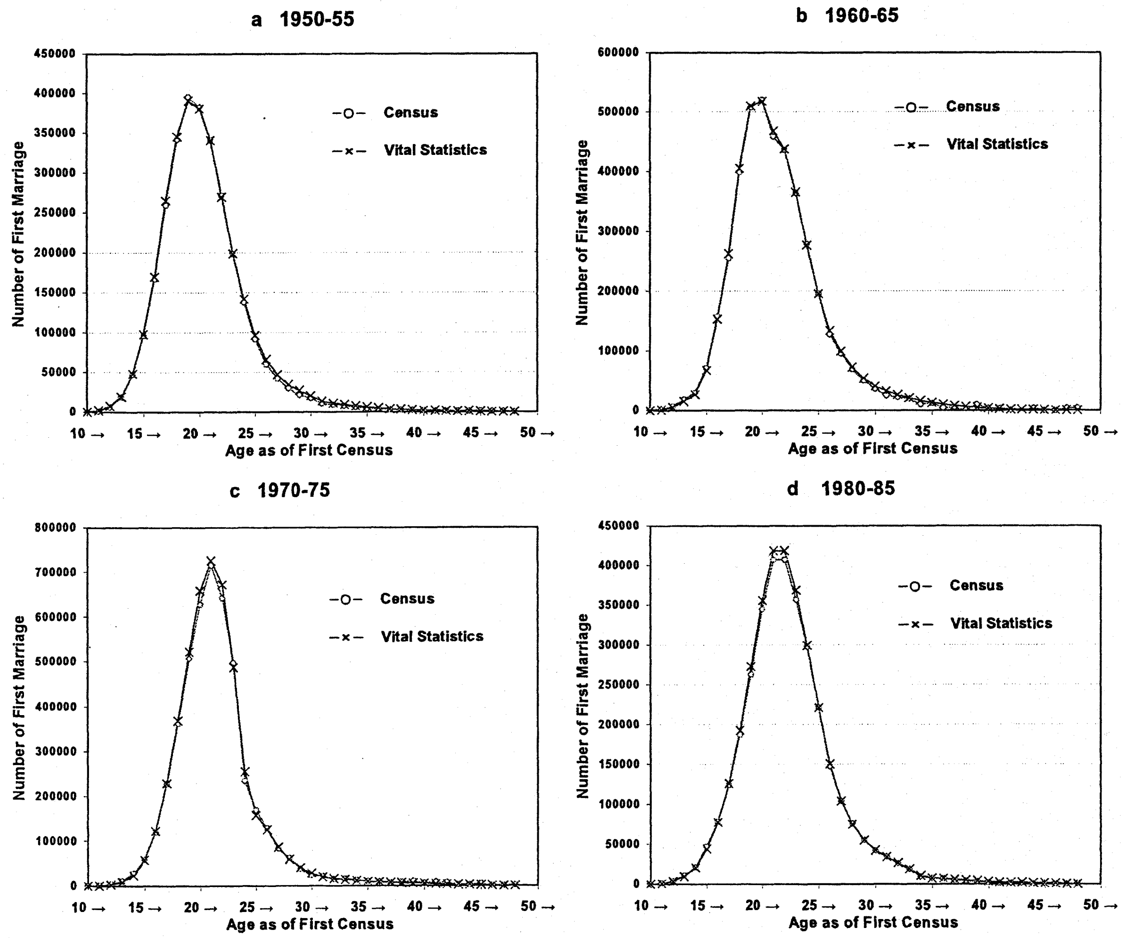
where $\overline{r}_{20}(x)$ is a model pattern obtained from principal component analysis on delayed registration for the years of 1974-81 adjusting to the pattern as of 1974, A(t) and B(t) are location and level adjustments to the model pattern specific to year t. We constrained A(t) and B(t) as the logistic function of time connecting to 0 and 1 at 1974. Then parameters of the logistic function are estimated so as to minimize the square difference of estimated marriage numbers by census and the vital statistics, $(\hat{N}_c - N_c)^2$, where $\hat{N}_c$ and N_c are given above as (A.10) and (A.9) respectively.

For the years of 1981-1996, R_5, the ratio of five year cumulated number of registrations to the number of same year registrations, is available instead of R_{20}. Therefore, we can estimate N_5 for these years. To correct N_5 to N_{20}, we compared them for the period of 1974-81, and obtained the average ratio, $I_{5\to20}$. Then we estimated N_{20} as $I_{5\to20}\ R_5\ N_1$.

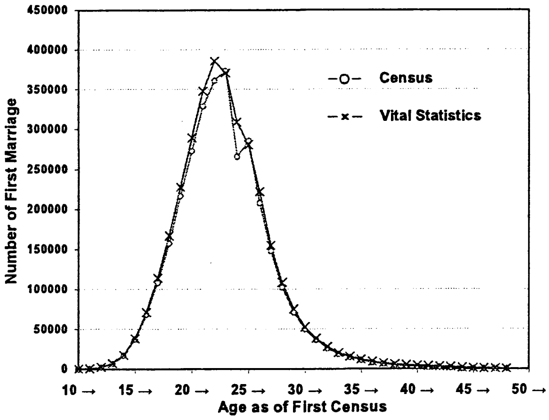
For the years of 1997-2000, we employed the same procedure as applied to the years 1950-73, using $\overline{r}_5(x)$ adjusted to the pattern as of 1996, and the relation; $\hat{N}_c = cI_{5\to20}\ R_5\ N_1$.

Figure A-5 shows comparison between estimates from censuses described in the previous section and estimates from the vital statistics described here, for intercensal periods. Estimated N_{20}'s are somewhat greater than the census estimate.

Figure A-5 Comparison of Estimated Numbers of First Marriage for Intercensal Period from Censuses and from Vital Statistics



d 1990-95



450000
400000
350000
300000
250000
200000
150000
100000
50000
0
Number of First Marriage
Census
Vital Statistics
10 → 15 → 20 → 25 → 30 → 35 → 40 → 45 → 50 →
Age as of First Census

Working Paper Series(E)

＊ *Working Paper Series are provided for discussions and critical comments, and their circulation is limited.*
Do not quote and cite without a written consent from the authors.

1 The Pattern of Intergenerational Mobility Among Men and Woman
(Sawako Shirahase)
July 1997

2 Pigou' s Welfare Economics and Ethical Belief
(Yuichi Shionoya)
February 1998

3 Health expenditure of deceased elderly in Japan
(Tetsuo Fukawa)
February 1998

4 Population Development and Social Problems in the Inner City and Suburbs of the Tokyo Metropolitan Area
(Satoshi Nakagawa)
April 1998

5 A Social Procedure for Choosing Fair Allocation Rules
－Formalizing the Rawlsian Principles of Justice－
(Reiko Gotoh and Naoki Yoshihara)
April 1998

6 Economy and Morality:A Conceptual Framework
(Yuichi Shionoya)
May 1998

7 Schumpeter on the Relationship between Economics and Sociology from the Perspective of Doctrinal History
(Yuichi Shionoya)
May 1998

8 Women and the Class Structure in Japan and Great Britain
(Sawako Shirahase)
March 1999

9 The Transformation of Partnerships of Japanese Women in the 1990s:
(Miho Iwasawa)
March 2000

10 Effects of Ethnic Concentration on migration in peninsular Malaysia
(Yoshimi Chitose)
June 2000

11 Japanese Income Inequality by Household Types in Comparative Persepective
(Sawako Shirahase)
September 2001

12 Effects of Government Policy on Internal Migration in Peninsular Malaysia: A Comparison between Malays and Non-Malays
(Yoshimi Chitose)
March 2002

13 What Determines Earning Differentials in Urban China?
(Yanfei Zhou)
April 2002

14 Women's Working Pattern and the Support to Working Mothers in Contemporary Japan
(Sawako Shirahase)
September 2002

15 Leaving Home in Japan: Its Trends, Gender Differences, and Determinants
(Toru Suzuki)
November 2002

16 Elaboration of the Coale-McNeil Nuptiality Model as
The Generalized Log Gamma Distribution:
New Identity and Empirical Enhancement

(Ryuichi Kaneko)

December 2002

17 Transformations in First Marriage Behavior of Japanese Female Cohorts:
Estimation and Projection of Lifetime Measures via
Empirically Adjusted Coale-McNeil Model

(Ryuichi Kaneko)

December 2002