PDF of the CM distribution resulting from removal of $h_T(x;m)$, those PDFs are given by:

$$g_X(x;m) = \frac{\beta}{\Gamma(\alpha/\beta+m)} \exp\left[-(\alpha+m\beta)(x-\mu) - \exp\{-\beta(x-\mu)\}\right] \quad (3)$$

$$h_T(t;m) = \frac{\beta\Gamma(\alpha/\beta+m)}{\Gamma(\alpha/\beta)(m-1)!}\{1-\exp(-\beta t)\}^{m-1}\exp(-\alpha t) \quad (4)$$

where α, β, and μ are three parameters of the CM distribution (Coale and McNeil, 1972).

Here $g_X(x;m)$ represents a distribution of age at entering a stage from which the process starts, and $h_T(t;m)$ is the distribution of the waiting time that is composed of m exponentially distributed waiting times. The mean and variance of the distribution $g_X(x;m)$ are respectively $\mu - \frac{1}{\beta}\psi\left(\frac{\alpha}{\beta}+m\right)$, and $\frac{1}{\beta^2}\psi'\left(\frac{\alpha}{\beta}+m\right)$.

Those of the distribution $h_T(x;m)$ are respectively $\sum_{j=1}^{m}\{\alpha+(m-1)\beta\}^{-1}$, and $\sum_{j=1}^{m}\{\alpha+(m-1)\beta\}^{-2}$.

As mentioned above, the exponential distribution with the largest mean convolved in distribution $h_T(t;m)$ has the parameter α. This is supposedly a distribution of the duration from entry into the marriage market to the meeting of the future husband. The exponential distributions of the second and third largest mean have the parameters, $\alpha+\beta$, and $\alpha+2\beta$. These are supposed to be distributions of the durations of dating and engagement. According to the parameter values of the CM standard age distribution of first marriage, which are derived from experiences of Swedish female cohorts, the mean duration from entry into marriage market to meeting of future husband is estimated as $(1/0.174)$ or 5.75 years. Similarly means of the second and third waiting durations are 2.16 years $(1/(0.174+0.2881))$ and 1.33 years $(1/(0.174+2\times0.2881))$ respectively (Coale and McNeil 1972).

As shown above, estimated parameter values for the CM distribution provides interesting behavioral measures as well. These interpretable features of the CM distribution are another advantage of the model in addition to its descriptive function.

Since the CM distribution is an alternative form of the generalized log-gamma distribution (GLG distribution) as identified by Kaneko(2003), the convolution formulations (5) and (6) are expressed in terms of the latter distribution. Namely:

$$g_X(x;m)=\frac{|\lambda|}{b\Gamma(\lambda^{-2}+m)}(\lambda^{-2})^{\lambda^{-2}+m}\exp\left[\left(\lambda^{-2}+m\right)\lambda\left(\frac{x-u}{b}\right)-\lambda^{-2}\exp\left\{\lambda\left(\frac{x-u}{b}\right)\right\}\right] \tag{5}$$

$$h_T(t;m)=\frac{|\lambda|\Gamma(\lambda^{-2}+m)}{b\Gamma(\lambda^{-2})(m-1)!}\left\{1-\exp\left(\frac{\lambda t}{b}\right)\right\}^{m-1}\exp\left(\frac{t}{b\lambda}\right) \tag{6}$$

where $\lambda\,(-\infty<\lambda<\infty,\neq 0), u\,(-\infty<u<\infty), b\,(>0)$ are three parameters, and Γ denotes the gamma function.

Some Other Convolution Models for First Marriage Process

Coale and McNeil also suggested a second type of convolution model for the first marriage process. In the course of deriving the first marriage model of the first type, they realized that repeatedly removing the exponential distribution from the CM distribution produced residual distributions that increasingly resemble the normal distribution. As a matter of fact, with random variable Z that follows the CM distribution, and T_j that follows the exponential distribution with mean $1/\{\alpha+(j-1)\beta\}$, they proved that $\sqrt{m}\left\{Z-\sum_{j=1}^{m}T_j-\left(a-\sum_{j=1}^{m}\frac{1}{\alpha+(j-1)\beta}\right)\right\}$ has a limiting distribution as $m\to\infty$ which is the normal distribution with mean zero and variance $1/\beta^2$ (Coale and McNeil 1972).

This implies that the CM distribution can be approximated by convolution of the normal distribution and some of the exponential distributions. After examining the cumulants of the distributions, they concluded that when m=3 the convolution is sufficiently close to their standard model derived from Swedish experience.

They did not derive any mathematical form of the convolution with a PDF or CDF in closed form. However, this formulation has certain technical advantages because of the abundance of techniques for working with the normal distribution. The basic formula for the convolution of the normal distribution and some distribution of non-negative waiting time with PDF $w(t)$ is given by:

$$g(z)=\int_{-\infty}^{z}\frac{1}{\sigma_e}\phi\left(\frac{t-\mu_e}{\sigma_e}\right)w(z-t)dt \tag{7}$$

where $\phi(x)$ is PDF of the standard normal distribution.

In the following, $N(\mu,\sigma)$ denotes the normal distribution with mean μ and standard deviation σ, and $E(\gamma)$ denotes the exponential distribution with mean $1/\gamma$. $A \oplus B$ stands for convolution of distributions, A and B. $\phi(x)$, $\Phi(x)$ are respectively the PDF and CDF of the standard normal distribution.

The convolution of the normal and one exponential distribution $N(\mu_e,\sigma_e) \oplus E(\gamma)$ is formulated by D'Souza (1974). Its PDF and CDF are given by:

$$g(z) = \gamma h(z) \quad (8)$$

$$G(z) = \Phi(\frac{z-\mu_e}{\sigma_e}) - h(z) \quad (9)$$

where $h(z) = \exp\left\{-\gamma(z-\mu_e) + \frac{1}{2}\sigma_e^2\gamma^2\right\}\Phi(z^*)$, in which $z^* = \{z-(\mu_e+\sigma_e^2\gamma)\}/\sigma_e$.

Supposing that the normal distribution controls age at entry into marriage market, and the exponential distribution controls waiting time to marriage, I can derive the proportion who are in the marriage market at age *x*, $P_E(x)$, and the hazard rate of marrying for those in the market, $r_E(x)$, as:

$$P_E(x) = h(x)$$

$$r_E(x) = \gamma$$

Following the similar procedure, I formulate the convolution of the normal and two different exponential distributions (, $\gamma_1 \neq \gamma_2$). Its PDF and CDF are given by: given by:

$$g(z) = \frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\{h_1(z)-h_2(z)\} \quad (10)$$

$$G(z) = \Phi(\frac{z-\mu_e}{\sigma_e}) - \frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\left\{\frac{1}{\gamma_1}h_1(z) - \frac{1}{\gamma_2}h_2(z)\right\} \quad (11)$$

where $h_i(z) = \exp\left\{-\gamma_i(z-\mu_e) + \frac{1}{2}\sigma_e^2\gamma_i^2\right\}\Phi(z_i^*)$, in which $z_i^* = \{z-(\mu_e+\sigma_e^2\gamma_i)\}/\sigma_e$, $i=1,2$.

The proportion who is in the marriage market at age *x*, and the hazard rate of marrying for those in the market are given by:

$$P_E(x)=\frac{\gamma_1\gamma_2}{\gamma_2-\gamma_1}\left\{\frac{1}{\gamma_1}h_1(z)-\frac{1}{\gamma_2}h_2(z)\right\}$$
$$r_E(x)=\frac{h_1(x)-h_2(x)}{h_1(x)/\gamma_1-h_2(x)/\gamma_2}$$

To compute the convolution of the normal and two identical exponential distributions, observe that by the associativity of convolution the exponentials convolve to a gamma distribution. Thus I can view the total convolution as the convolution of the normal and gamma distributions. Its PDF and CDF are given by

$$g(z)=\sigma_e\gamma^2\left\{z^*h(z)+\phi(z^*)\right\} \tag{12}$$

$$G(z)=\Phi\left(\frac{z-\mu_e}{\sigma_e}\right)-\left\{h(z)+\frac{1}{\gamma}g(z)\right\} \tag{13}$$

where $h(z)=\exp\left\{-\gamma(z-\mu_e)+\frac{1}{2}\sigma_e^2\gamma^2\right\}\Phi(z^*)$.

The proportion who is in the marriage market at age x, and the hazard rate of marrying for those in the market are given by:

$$P_E(x)=h(x)+\frac{1}{\gamma}g(x)$$
$$r_E(x)=\frac{h(x)}{h(x)+g(x)/\gamma}$$

Non-independent "Convolution" Model

All those formulae discussed above depend critically on the assumption of independence between the distributions to be convolved. If this assumption does not hold for the actual target process, their utility as behavioral models is limited. This issue has not been discussed extensively in previous work due to a lack of empirical evidence.

Kaneko (1991a) pointed out the heavy dependence of waiting time (T2+T3) on age at first meeting (X1) based on survey data collected in Japan[9]. He reported, in particular, a strong negative correlation between X1 (age at first encounter with eventual spouse) and T2 (waiting time from encounter to engagement) as -0.449 for Japanese women with sample size N=4682. These correlations make variance of age at first marriage much smaller than that of age at first meeting, which should not happen in the convolution framework.

[9] The Ninth National Fertility Survey (NFS 9) conducted in 1987 in Japan by Institute of Population Problem (currently National Institute of Population and Social Security Research).

Hence the validity of convolution models of first marriage process as behavioral explanatory tools is questionable at least for some populations.

Therefore, in order to develop the behavioral multistage models I should consider models without convolution or with non-independent "convolution", by which I mean a procedure generating the distribution of a sum of random variables assuming dependencies among them. The general formula for the PDF, $g_z(z)$, of sum of two random variables X and T (>0) is given by:

$$g_z(z) = \int_{-\infty}^{z} g_x(x) w_t(z-x|x) dx \tag{14}$$

where $g_x(x)$ is PDF of X and $w_t(t|x)$ is conditional PDF of T given X=x.

With this integral formula, I would obtain $g_z(z)$ by providing functional form of $g_x(x)$ and $w_t(t|x)$.

One possible such model is one whose parameters of the waiting time distributions are functions of age at initiation of waiting, i.e. X. Here I attempt to develop a model of this type to conduct a numerical examination, in a search for appropriate models to describe the observed of multistage process.

I employ the generalized gamma (GG) distribution to represent the mathematical models for waiting time because it is one of the most flexible parametric models for survival times. Its PDF is given by:

$$w(t;\theta,c,v) = \frac{|\theta|}{c\Gamma(\theta^{-2})}(\theta^{-2})^{\theta^{-2}} \left(e^{v}t\right)^{\frac{1}{c\theta}-1} \exp\left[-v-\theta^{-2}\left(e^{v}t\right)^{\frac{\theta}{c}}\right] \tag{15}$$

where $\theta, c,$ and v are three parameters.

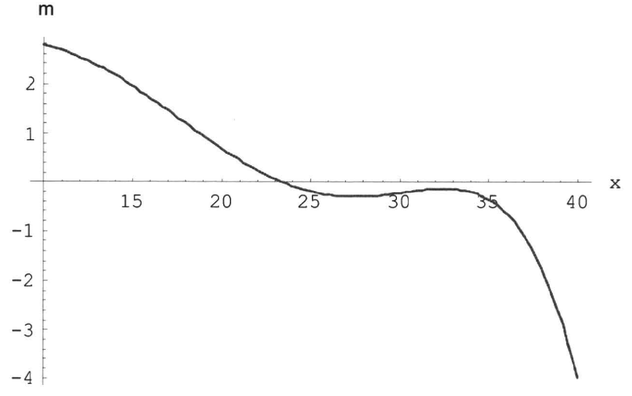
Let the scale parameter v be dependent on age at onset x, and expressed as a polynomial function of x, such as $v(x) = v_0 + v_1 x + v_2 x^2 \cdots$, where $v_0, v_1, v_2 \cdots$ are polynomial coefficients. Then PDF of the GG waiting time distribution is given by

$$w(t,x;\theta,c,v_0,v_1,v_2\mathrm{L}) = \frac{|\theta|}{c\Gamma(\theta^{-2})}(\theta^{-2})^{\theta^{-2}} \left(e^{v(x)}t\right)^{\frac{1}{c\theta}-1} \exp\left[-v(x)-\theta^{-2}\left(e^{v(x)}t\right)^{\frac{\theta}{c}}\right]. \tag{16}$$

It is possible to make the other parameters, θ and c dependent on X in the same way, though I do not attempt it here.

Parameter estimation of $w(t,x;\theta,c,v_0,v_1,v_2\cdots)$ is conducted by applying it to survey data of delay from age at first meeting to marriage in Japanese women through the maximum likelihood method[10]. The result indicates that the polynomial functions up to five produces statistically significant improvements in fitness, but little improvement is seen thereafter. Function $v(x)$, the scale parameter value of the GG waiting time for those whose age at first meeting is x, is shown in Figure-1 in the form of a fifth degree polynomial.

Figure 1 Scale Parameter of the GG Distribution for Waiting Time as a Function of Age at First Meeting with Eventual Husband: Japanese Women

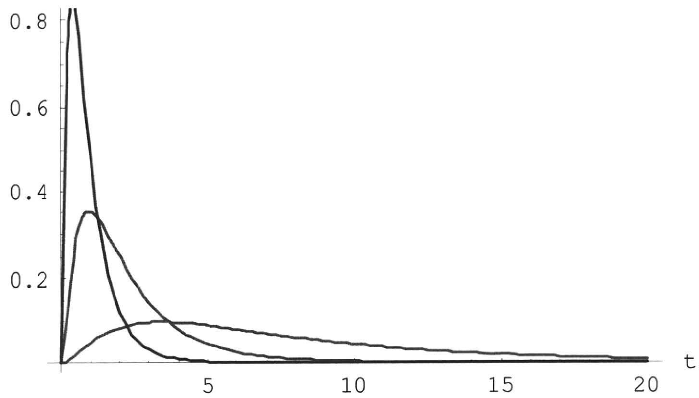


As age at meeting increases, the value of v goes down until mid twenties. Then it stabilizes at slightly below zero until the mid thirties followed by a steep fall. Note that the rapid decline after the late thirties is not reliable due to the small sample size.

Corresponding waiting time distributions for ages 15, 20, and 25 at first meeting are shown in Figure-2. The figure indicates that there are remarkable differences in waiting time distribution depending on age at first meeting, which implies again that convolution model with the independence assumption is simply not realistic.

[10] The Ninth through Eleventh National Fertility Survey conducted in 1987, 1992, and 1997 in Japan by National Institute of Population and Social Security Research.

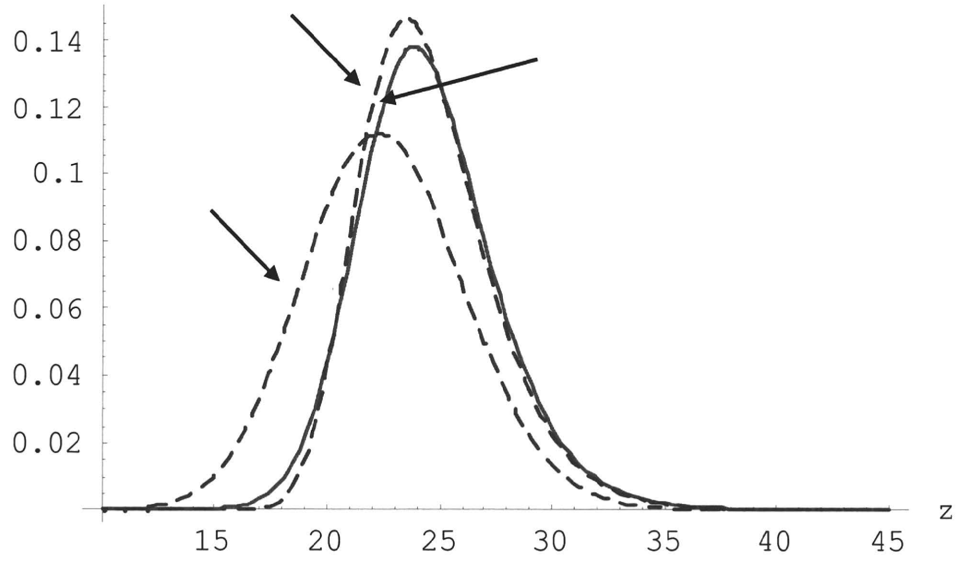
Figure 2 PDF of Estimated GG Distribution for Waiting Time by Age at First Meeting with Eventual Husband: Japanese Women



With this waiting time distribution combined with the estimated GLG distribution of age at first meeting on the same sample, age distribution of first marriage is reconstructed by numerical integration of the non-independent convolution formula (14) and is compared with directly estimated GLG distribution. The result is shown in Figure-3.

In the figure the reconstructed PDF of age at first marriage with the non-independent "convolution" of age at first meeting and dependent waiting time is matched up to the estimated PDF of the CM distribution of age at first marriage. The estimated PDF of age at first meeting with the eventual husband, which is used in the reconstruction, is also shown in the figure. The reconstructed distribution with non-independent convolution is fairly close to the directly estimated distribution, although differences are noticeable. Since the discrepancies remain even after introducing the dependence into the other parameters, it is due to inadequacy of the GG distribution for representing the waiting times, despite of the fact that the GG is one of the most flexible parametric models.

Figure 3 Observed and Estimated PDF of Age Distributions of First Marriage: Jap anese Women



Note: Estimated PDF of age at first marriage with non-independent "convolution" model is compared with estimated PDF of the GLG model. Estimated PDF of age at first meeting with eventual husband is also shown. Source data is from The Ninth through E leventh National Fer tility Survey conducted in 1987, 1992, and 1997 in Japan by National Institute of Population and Social Security Research.

However, the examination indicates the possibility of employing the non-independent "convolution" model as a behavioral explanatory tool. It is shown that the model is able to reproduce distribution of age at first marriage fairly accurately, if the dependency of waiting time on age at meeting is properly represented.

Estimation of waiting time distributions with survey data

First, I estimate waiting time distributions. Second, using the model of waiting time, I estimate preceding event distribution from that of the resulting event. For example, I estimate age distribution of first marriage from distribution of age at first birth. Then I proceed to application to the vital statistics data which has no information on waiting time distributions. The following is the data source and variables.

Data (survey data)

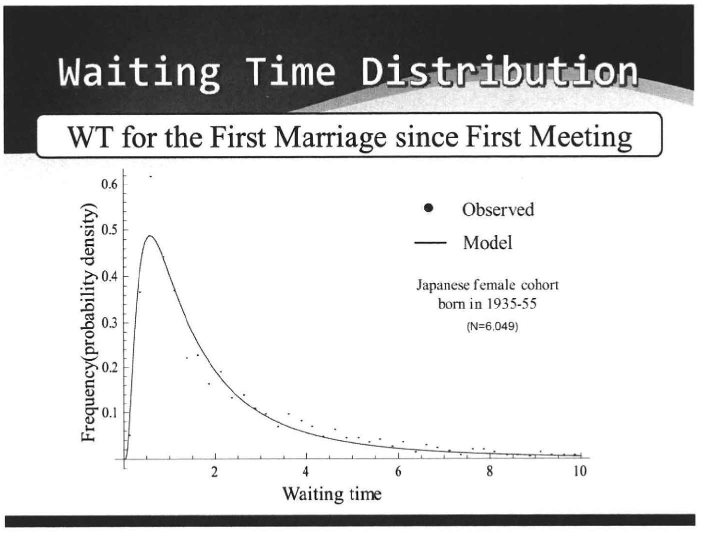
Japanese National Fertility Survey (NFS): 9th-13th (1987-2005) --- nationally representative samples of married Japanese female cohort born in 1940-62

Variables

Age at first meeting with present spouse (N=6,049)
Age at first marriage (N=5,200)
Age at first and second birth (N=4,699)
and waiting times between life events above

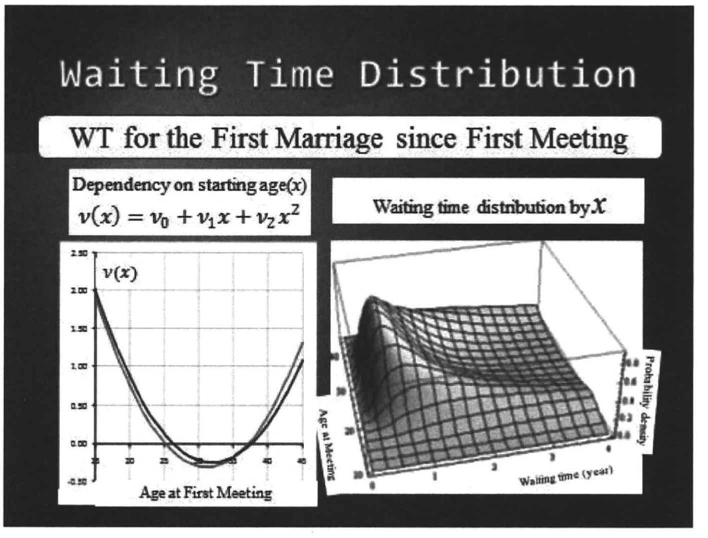
Figure 4 shows the waiting time distribution from the time at first meeting with present spouse to the time at first marriage for Japanese female cohort in our data set. The actual observed distribution is indicated by dots, while the thin line denotes a model of the waiting time, i.e. the Generalized Gamma distribution with the estimated parameters. The model indicates moderately good fit to the data.

Figure 4 Observed and Modeled PDF of Waiting Time from First Meeting to Marriage



The waiting time is dependent on age at first meeting. This dependency is expressed by parameter form which is dependent on starting age. In this case, the model fits the data when the dependency is expressed as a quadratic form. This implies that when you are young, the later you meet your spouse, sooner you get married, while when you get older, waiting times tend to prolong. The resulting waiting time model is shown right in 3D graph in Figure 5.

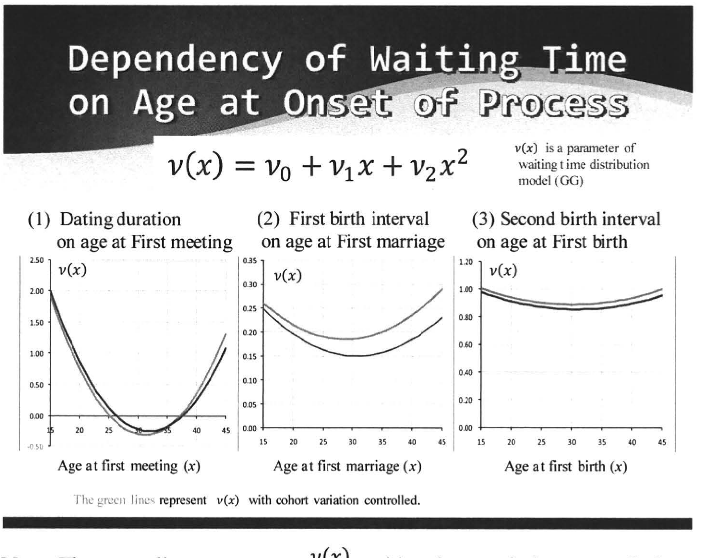
Figure 5 Dependency of the Waiting Time on Age at First Meeting



Note: The green lines represent $v(x)$ with cohort variation controlled.

In the similar way, I examined two other life events with this survey data, i.e. the first and second birth. While the dependency of waiting time for first marriage on age at first meeting is heavy, this is not true for waiting time for first birth, and for second birth as shown in Figure 6 (2) and (3).

Figure 6 Dependencies of Three Waiting Times of Life Events on Age at Onset

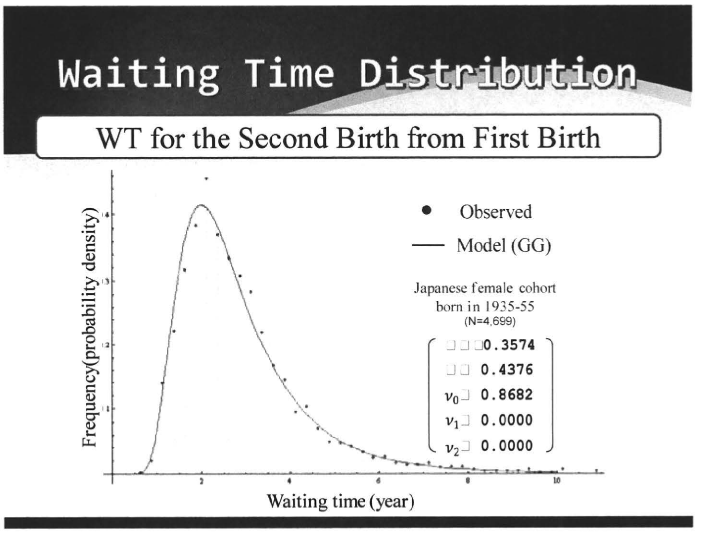


Note: The green lines represent $v(x)$ with cohort variation controlled.

According to these observations, the original independent model could be applied to the first birth and second birth with reasonable goodness of fit. Avoidance of the complexities introduced by the dependency scheme has great advantage in practical applications.

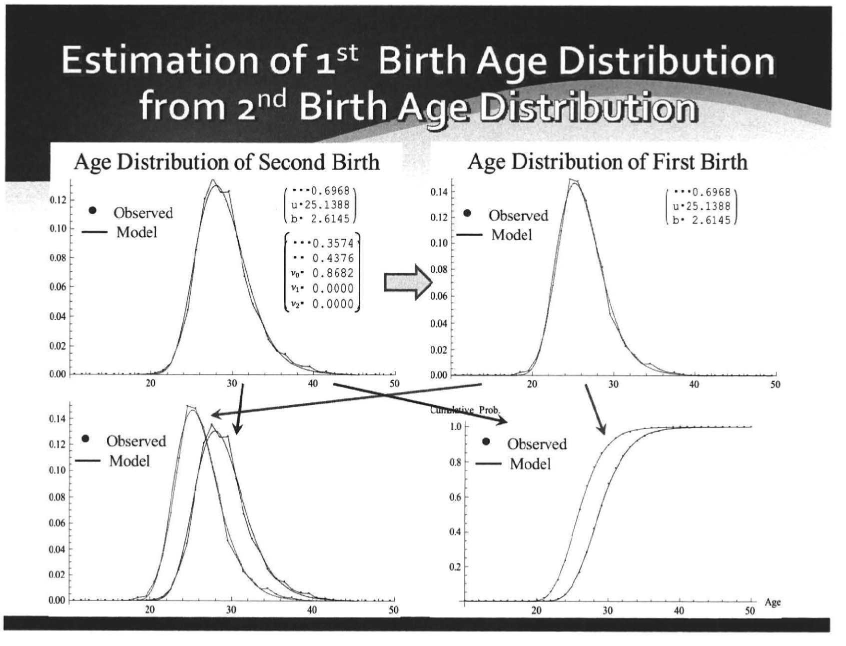
Figure 7 shows the observed and modeled distribution of waiting time from the first birth to the second.

Figure 7 The Observed and Modeled Distribution of Waiting Time from the First Birth to the Second



The model (the Generalized Gamma distribution) well replicates the observed. The age distribution of the first birth is to be estimated by applying the model with this waiting time model to data of age at second birth. The estimated first birth distribution from the second birth data is compared with the actual observed distribution in Figure 8.

Figure 8 Estimation of Age Distribution of the First Birth from the Distribution of the Second Birth: Application to the Survey Data



The estimation of age distribution of age at first birth replicates the actual distribution quite well. In other words, the estimation of age distribution of the first birth is conducted by applying the multistage model to the second birth. In a sense this is an inevitable outcome, because the model utilized the well fitted waiting time distribution to the actual data. Then I attempted estimation of age distribution of the first birth of the vital statistics.

Figure 9 Estimation of Age Distribution of the First Birth from the Distribution of the Second Birth: Application to the Vital Statistics Data

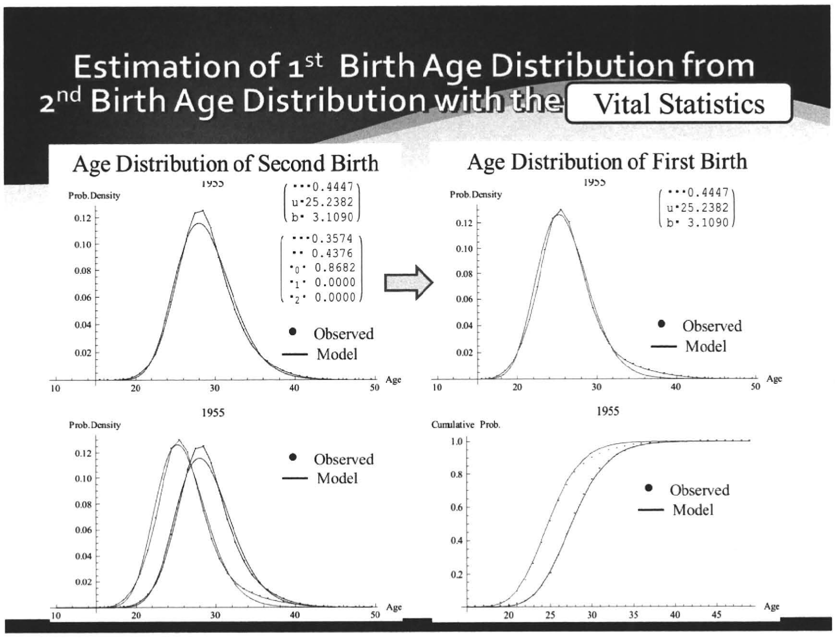
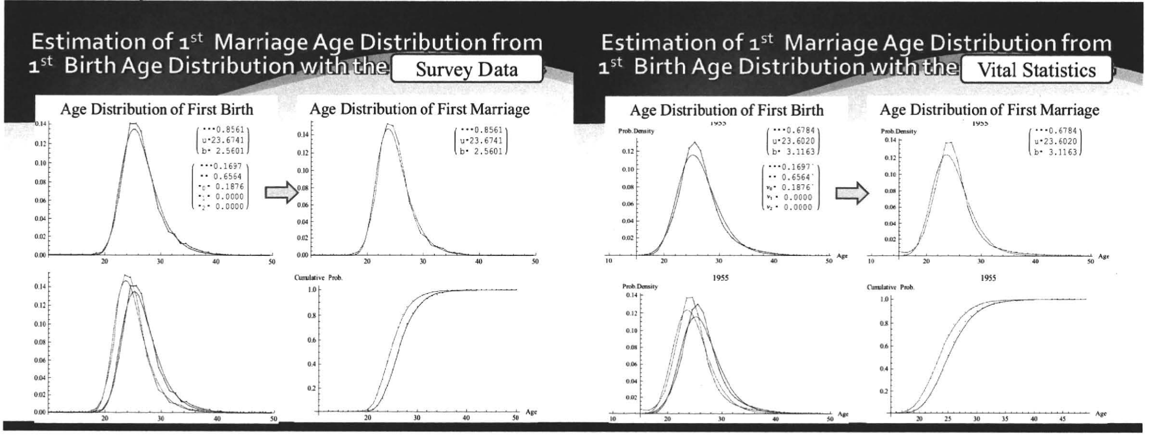


Figure 9 indicates the results of estimation of age distribution of the first birth from the second with the vital statistics data by applying the same procedure as case for the survey data assuming that the waiting time distribution observed in the survey sample does not vary a lot in population of the vital statistics. The fit is fairly good though it is not as good as the case applied to the survey data from which the information on waiting time is derived.

I applied the same procedure to estimate age distribution of the first marriage from that of the first birth. Figure 10A shows results of the application for the survey date, which indicate again that the estimation represents the actual data fairly well. Figure 10B shows results of the application for the Vital Statistics. It indicates that the estimation is moderately accurate though it does not represent the actual data as well as the case for the survey data. The central tendency and dispersion may be well approximated by the model.

Figure 10 Estimation of Age Distribution of the First Marriage from the Distribution of the First Birth

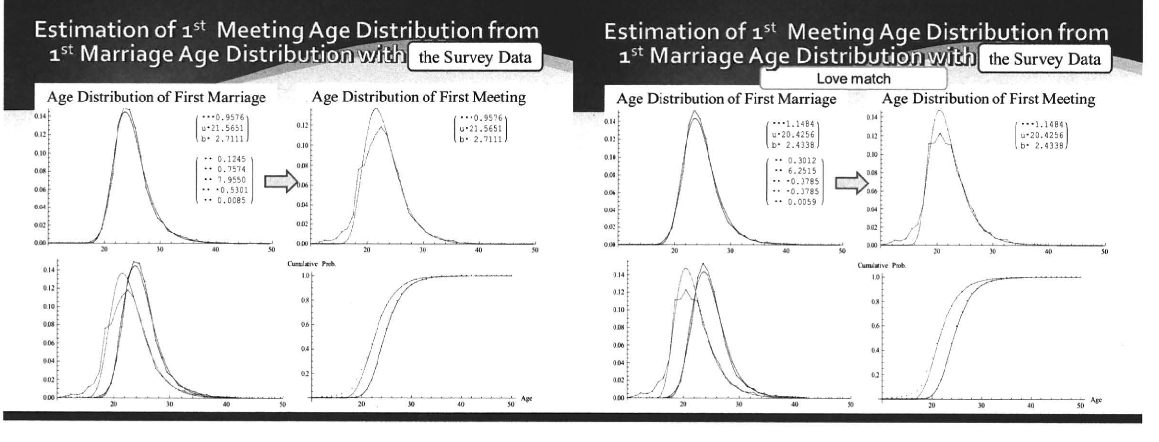
A. Survey data B. Vital statistics



Finally, I conducted an estimation for age distribution of the first meeting with present spouse from data on age distribution of the first marriage. In this case, remember that the waiting time is heavily dependent on age at meeting with spouse. Then I used the dependent convolution framework explained before for this case. The result is shown in Figure 11A. Though the distribution of age at meeting with present spouse itself shows somewhat irregular shape, it is observed that the estimated model catches the main properties of the distribution.

Figure 11 Estimation of Age Distribution of the First Meeting from the Distribution of the First Marriage

A. Survey data (Total) B. Survey data (Love-match)



One of causes of irregularity in the observed distribution of age at meeting is derived from the mixture of different type of marriage, i.e. arranged marriage and love match. Then I applied same procedure to only marriage from love match. The result in Figure 11B shows modest improvement, though distribution itself is still irregular.

A summary for model fits in terms of the mean and the standard deviation are presented in Table 1. The results indicate that the models represent the central tendency and dispersion of age distribution of those latent events fairly well.

Table 1 Actual and Estimated Mean and S.D. of Age at Onset of Process

Target event	Mean	S.D.
Survey data		
Age at first meeting (est. from 1st marriage)		
Actual	22.95	3.73
Model	23.05	3.41
Δ	**0.10**	**- 0.32**
Age at first marriage (est. from 1st birth)		
Actual	24.87	3.04
Model	24.90	3.08
Δ	**0.03**	**0.04**
Age at first birth (est. from 2nd birth)		
Actual	26.12	3.01
Model	26.12	2.96
Δ	**0.00**	**- 0.05**
Vital statistics		
Age at first marriage (est. from 1st birth)		
Actual	24.37	3.73
Model	24.16	3.25
Δ	- 0.21	- 0.48
Δ*	**- 0.08**	**- 0.24**
Age at first birth (est. from 2nd birth)		
Actual	26.34	3.77
Model	25.95	3.27
Δ	- 0.39	- 0.50
Δ*	**- 0.01**	**- 0.15**

A possible application of these procedures is an estimation of age distribution of the union formation from the distribution of the first birth in a society where cohabitation is common and relevant statistics is limited.

Tentative Conclusion

In this paper, the CM model is enhanced as a behavioral model so as to describe latent processes behind occurrences of first marriage such as entering into marriage market and searching for a mate.

The CM model is viewed as a multistage model, by which we mean a process that consists of multiple processes required for a target event (*i.e.* marriage in this case) to occur. Its convolution structure can be regarded as an expression of a multistage process consisting of attainment of marriageable age, waiting time for meeting the eventual spouse, dating stage, and engagement period. This interpretable nature of the CM model is carefully examined and the corresponding equivalent convolution structure of the GLG model is presented. Various other convolution models for first marriage process are also examined. The critical problem associated with the multistage view of the CM model is that the assumption of the model over independence between the sub-processes is not satisfied in the reality. The analysis on a national representative survey of the first marriage process in Japan detected the substantial dependency between age at first meeting of a couple and subsequent waiting time for the first marriage (Kaneko 1991a, Kaneko 1999). This is likely the case for other societies. Hence we proposed a multistage model in which the independence among the sub-processes is not assumed (here, we call it non-independent "convolution" model) by introducing liner relationships among parameters so that realistic measurement of the sub-processes can be obtained. Though it may spoil the elegance to some extent, the model acquires additional applicability to the actual process. Our verification of the new model indicates reasonable agreements with the empirical data. Actually applications to the vital statistics date in Japan with waiting time distribution from survey indicates good estimates to latent preceding events like first marriage, or first meeting with present spouse. It demonstrates the outstanding potential of the CM model to represent the behavioral process of first marriage and conceivably other demographic events as well such as birth.

The new model is to be applicable in many ways to behavioral studies of such vital events as marriage and birth separated by birth order. For example, it can be applied to estimation of size of marriage market, i.e. population size of marriageable people in a society. The size is possible to measure by estimating age distribution of people's becoming marriageable, which is major part of the model. The model should be effective to detect latent processes of births, since they are viewed as multistage processes. For application, it should be useful for estimating the distribution of "real" age at onset of sexual union by applying the model to the first birth for the society in which marriage is no longer viewed as the onset of sexual relationships among couples. Some of these applications should be demonstrated in the complete version of the present paper.

Selected References

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

D'Souza, S., 1974, *Closed Birth Intervals: A Data Analytic Study*, New Delhi: Sterling Publishers.

Hernes, G., 1972, "The Process of Entry into First Marriage," *American Sociological Review*, Vol. 37: 173-182.

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

—, 1999, "Below-replacement fertility in Japan: Trends, Determinants and Prospects," 1999, 6, *Population Bulletin of the United Nations*, Nos. 40/41, pp.266-291

--, 2003, "Elaboration of the Coale-McNeil Nuptiality Model as The Generalized Log Gamma Distribution: A New Identity and Empirical Enhancements," *Demographic Research*, Vol. 9-10:223-262. http://www.demographic-research.org/Volumes/Vol9/10/9-10.pdf

Liang, Z. 2000. *The Coale-McNeil model. Theory, generalisation and application.* Amsterdam, The Netherlands: Thela Thesis.

Stacy, E. W. 1962. "A Generalization of the Gamma Distribution." *The Annals of Mathematical Statistics* 33: 1187-1192.

4　2005年以降の合計出生率反転の要因：都道府県別データを用いた空間分析の応用

岩澤　美帆
金子　隆一

和文抄訳

本研究は、2005年以降上昇傾向にある日本の合計出生率（Total fertility rate: 年齢標準化した出生力指標で、仮に女性が当該年次の年齢別出生率に従って生涯にわたって子どもを生んだ場合に実現する子ども数に相当する）が、どのような要因で上昇したかを理解するために、都道府県別の出生率変化を説明する回帰モデルを用いた検証を試みたものである。その際、都道府県別データのような空間データでは、人口規模の違いによる分散不均一性の問題や、近隣県の指標間で高い相関がみられる点に配慮し、誤差項における空間自己相関を明示的にモデル化した重みつき空間誤差モデル weighted spatial error model を利用した。

日本は南欧諸国や東欧諸国、東アジアの一部の地域と同様、1.3を下回る極めて低い合計出生率を経験している。これらの国はしばしば超低出生力国 lowest-low fertility countries と呼ばれ、そうした地域では、出産の先送り傾向（晩産化）が著しいこと、出生力の高い移民などが少ないこと、経済の低成長やグローバル経済によって若者の雇用情勢が悪いこと、高学歴化によって出産・子育ての機会費用が高まっているにもかかわらず、十分な両立支援体制が整っていないこと、家族のネットワークが強力で、子育てに女性以外が参加することを期待しない文化的風潮をもつこと（家族主義）、などが指摘されている(Kohler, Billari, and Ortega 2002 ,Frejka and Westoff 2008, Perelli-Harris 2005, Zuanna and Micheli 2004, Reher 2007, McDonald 2006)。こうしたことから、超低出生力地域の出生率は長期にわたって低迷するとの見通しが主流であったが、2000年前後から、一部の地域で出生率の反転上昇が見られ、日本でも2005年に1.26の最低値を記録して以降、2009年の1.37まで回復した。南欧など欧州についての先行研究によれば、晩産化の進展がとまり（人口学的にはテンポ効果の消滅とみなせる）、移民の増加や景気の回復、子育て支援、とくに両立支援政策の充実などが要因として指摘できるという(Castiglioni and Dalla Zuanna 2008, Billari 2008, Goldstein et al. 2009, Caltabiano et al. 2009)。また南欧では、本来、再生産に有利であった家族主義的な南部地域よりも、西欧諸国に特徴的な家族ライフスタイルが急激に進展している北部地域で出生率の回復が著しいことが指摘されている。これらの要因が日本における出生率の上昇の説明としても有効かどうかを検証した。

都道府県別出生率の2005年～2008年の変化分を、テンポ効果の縮減(Bongaarts-Feeney(1998、2005)の period didtortion index の変化を使用)、外国人母による出生割合の変化、経済の回復（就業率の変化）、両立支援策の充実を示す母親の就業率の変化、家族主義を示す三世代家族割合の各効果で説明する回帰モデルを推定する。その際、都道府県による人口規模の違いを考慮し、再生産年齢女性の人口を重みにした重み付け最小二乗回帰モデルを用いるとともに、近隣県間で誤差項に相関がある場合は、重み付き空間誤差モデルを用いることを検討した。空間誤差モ

デルは以下のように定式化される。

$$y = X\beta + u,$$
$$u = \lambda Wu + \varepsilon,$$
$$\varepsilon \sim N(0, \sigma^2 I)$$

ここでyは従属変数を示す(n x 1)ベクトルである。Xはk-1の独立変数をしめす(n x k)行列である。βは推定されるべき(k x 1)ベクトルである。uは(n x 1)の誤差項を示すベクトルであるが、次行の式で示された構造をもつ(Anselin 1988, Ward and Gleditsch 2008)。そこでは空間自己回帰係数であるλと、地域間の関係を示す(n x n)の加重行列Wによって空間自己相関が表現され、εが独立に分布した（地域間で相関しない）誤差項ベクトル(i.i.d.)を示す。加重行列のための近隣構造の定義に際しては、当該都道府県に全方向で隣接する場合を近隣と見なす一重クイーン方式 first order queen convention を採用した。モデルは以下のように表される。Δは差分を示す。推定は出生順位別に行った。推定にはRのspdepパッケージを利用した。

ΔTFR (2005-2008) = 切片
+Δテンポ効果指標 (2004-2007)
+Δ外国人母による出生割合 (2005-2008)
+Δ就業率 (2002-2007)
+Δ核家族世帯に暮らす未就学児を持つ母親の就業率 (2002-2007)
+ 未就学児を含む世帯に占める三世代家族割合（固定効果）(2005)
(+ 空間自己相関誤差項 λWu)

いずれのモデルでも、誤差項に空間相関が認められたので、通常の最小二乗法ではなく空間誤差モデルによる推定が望ましいことが分かった。テンポ効果指標の変化、就業率の変化、外国人母による出生割合の変化が出生率の変化と正の関係を示し、晩産化の停止、景気の回復、外国人の増加が近年の出生率上昇の一翼を担っていることがわかった。家族主義を示す三世代家族世帯割合は、第1子や第2子など、低い出生順位で、南欧における結果と同様、負の関係を示していた。また分散全体の15%程度を説明していた。子どもをもつ母親の就業率の変化は明確な正の関係を示さなかった。ただし、説明変数で説明できる部分は3割程度であり、残りは全国に共通する要因が存在したと解釈できる。

両立支援策の効果については、3歳未満の子どもをもつ母親の就業率変化を代理変数として関係性を検証したが、有意な関係は確認することはできなかった。出生率が上昇した都市部では保育園の待機児童が増加するなど、両立のニーズはありながらも体制が追いついていない可能性がある。また2006年以降国際結婚が減少し、2008年以降は失業率も再上昇していることから、今後上昇効果が一時的に薄れる可能性も考えられる。しかしその効果は0.02程度と限定的であろう。空間相関の高さは、その地域で、説明変数では説明できない出生率上昇の要因が存在していたことを意味する。誤差項の空間相関の高い地域を地図上で示すと九州地方で顕著であった。こうした地域に特有な文化や施策と出生率との関係を検証していくことが有効であろう。

Explanations for the fertility reversal after 2005 in Japan

Miho Iwasawa
Ryuichi Kaneko

Abstract

The goal of this paper is to evaluate explanations for the total fertility rate (TFR) upturn in Japan after 2005. Drawing on recent research on the retreat from lowest-low fertility in European countries, we focus on diminished tempo effects, increasing numbers of foreign mothers, improving economic conditions, and policy efforts to support work-family balance. We also examine the role of familistic culture. Decomposition analyses based on the results of weighted spatial error models indicate that TFR upturn at the national level is partially explained by diminished tempo effects, increases in foreign mothers, and economic improvement. However, change in maternal labor force participation – an indirect measure of policy efforts to improve work-family balance – was not significantly associated with the TFR reversal. Our measure of familistic culture – the proportion of extended family households – was negatively associated with fertility increase for lower-order births. Our results also suggest that over half of the TFR increase is not explained by the factors in our models. Subsequent decline in the number of international marriages and the recent economic downturn may contribute to a slow-down or reversal in the upward trend in TFR, but the impact of these changes should be relatively small.

This research was contributed by Kenji Kamata (National Institute of Population and Social Security Research), James M. Raymo (University of Wisconsin), and Kimiko Tanaka (Rochester Institute of Technology).

1. Introduction

Contrary to the expectations of classic demographic transition theory, post-transitional total fertility rates (TFR) range from nearly replacement level to below 1.2 (Morgan 2003).[1] The future of below-replacement fertility is an open question and efforts to identify the causes and characteristics of extremely low fertility have been a central part of demographic research since the late 1990s. Some scholars have proposed comprehensive explanations for lowest-low fertility – defined as TFR under 1.3 – and suggested that such levels may persist for several decades (Kohler, Billari, and Ortega 2002, Lutz, Skirbekk, and Testa 2006, McDonald 2006, Reher 2007). However, fertility rates have recovered steadily since the end of the 1990s in Italy and Spain, two of the first countries to reach lowest-low fertility. Since 2000, an increase in TFR has also been observed in other lowest-low fertility nations in Central

[1] TFR is an age-standardized period measure that can be interpreted as the number of births a woman would have if she experienced current age-specific rates throughout her lifetime (and did not die prior to the end of reproduction). Replacement-level fertility is slightly over two births per women (Morgan and Taylor 2006).