J. Biosoc. Sci., (2016) **48**, 577–592, © Cambridge University Press, 2015
doi:10.1017/S0021932015000437 First published online 28 Dec 2015

COHORT MARRIAGE KINETICS IN THE CONTEXT OF MIGRATION, WITH A CASE STUDY OF JAPAN, 1920–1940

JIANGHUA LIU*†1

**Max Planck Institute for Demographic Research, Rostock, Germany and †Institute for Population & Development Studies, Xi'an Jiaotong University, Xi'an, China*

Summary. The concept of marriage squeeze expects a positive association between marriage formation and the availability of preferred mates. Previous research to test the hypothesis has had mixed results owing to inconsistent marriage measures, inconsistent age focuses and insufficient attention to migration. This study derives kinetics equations of marriage formation to link cohort age-specific mate availability to migration-adjusted marriage rate/incidence, a measure in contrast to nominal marriage rate. On testing the equations with Japanese census data for 1920–1940, it is found that, in female cohorts, mate availability impacts first marriage rate at the life-course stage from 15–19 to 20–24 years, but not at later stages. Among young females, the decline in mate availability accounted for about 21% of the decline in first marriage rate over the period 1920–1940, when there was a trend towards later but not less marriage in Japan. The study suggests that the flexibility of mate/spouse choice in females varies along the marriageable life course and is more manifest at older ages. At young ages, however, the marriage squeeze hypothesis could hold, presumably because young women are evolutionarily shaped to be choosier, perhaps postponing their marriages when preferred mates are in short supply.

Introduction

Referring to a shortage of mates of a certain sex, 'marriage squeeze' has been a major concept in demography since it was introduced 50 years ago (Glick *et al.*, 1963). It implicitly assumes a rigid age preference in mate choice (Ní Bhrolcháin, 2001) and expects a positive association between marriage formation and the availability of mates with a preferred age difference or simply preferred mates, and a substantial proportion of members of the sex in surplus ending in permanent celibacy under a monogamous marriage system (Akers, 1967). The concept has stimulated major interest in the

[1] Corresponding author. Email: liujianghua@tsinghua.org.cn

imbalance between numbers of men and women in the prime marriage ages and its consequences for marriage formation in historical and contemporary populations.

Many studies have supported the marriage squeeze hypothesis in identifying factors responsible for a gender imbalance: cohort growth (Akers, 1967; Mari Bhat & Halli, 1999), sex-specific migration (Goodkind, 1997), differential mortality by sex (Saxena *et al*., 2004) and an imbalanced sex ratio at birth (Kim, 2004; Guilmoto, 2012). However, there are also examples of studies contradicting it. Dixon (1971) found that across 57 countries in 1960, the sex ratios of men aged 20–44 to women aged 15–39 and of men aged 25–29 to women aged 20–24 correlated weakly with the proportions ever-married in men and women aged 40–44 and 20–24, respectively. By checking 382 US labour-market areas in 1980, Lichter *et al*. (1991) found that the sex ratio of men to women aged 20–29 was not significantly correlated with the proportion of ever-married women in this age group. Using a marriage squeeze measure based on all marriageable ages, Schoen and Baj (1985) indicated that marriage squeeze was not a major factor influencing age at marriage and proportion ever-married, although it cannot be ignored.

There were several reasons for the mixed results from previous studies of marriage squeeze. Firstly, different marriage measures were used. Proportion ever-married or single (e.g. Dixon, 1971; Lichter *et al*., 1991) was used more often than marriage rate, i.e. marriages per person-year (e.g. Akers, 1967); in other words, prevalence was a more popular measure of marriage formation than incidence (South & Lloyd, 1992; Ní Bhrolcháin & Sigle-Rushton, 2005). The problem with the prevalence measure is that lower marriage prevalence at a given age may be caused by lower marriage incidence at ages before the age; thus, analysing the relationship between mate availability and marriage prevalence at that age might lead to a misleading conclusion. Secondly, different studies focused on different ages or age groups. In some studies, the focus was on a single age group (e.g. Akers, 1967; Dixon, 1971; Lichter *et al*., 1991); in others, the focus was on all men and women at marriage ages (e.g. Schoen & Baj, 1985; Mari Bhat & Halli, 1999). Few studies analysed marriage squeeze and its consequences for marriage behaviour in each age group along the marriageable life course (Fossett & Kiecolt, 1991), but the situation has changed somewhat in recent decades (e.g. Lampard, 1993; Raymo, 1998). It is possible that the association between mate choice and mate availability changes with age, as suggested by Lampard (1993). Consequently, the pattern observed in one age or age group might not be observed in others or the whole marriageable life course. Thirdly, these studies generally focused on populations that were not close to migration, but few of them made adjustment for migration. As Lichter *et al*. (1991) pointed out, migration could be a serious confounding factor of the relationship between marriage formation and mate availability (see also Lampard, 1993).

This study suggests a new framework for estimating the effect of mate availability on marriage formation in a cohort, which leads naturally to a test of the marriage squeeze hypothesis. In this theoretical framework, the linkage between the availability of mates with a preferred age difference and first marriage rate/incidence in a cohort at a given age is formulated, with both mortality and migration considered. The framework is thus robust to age effect and situation, e.g. contemporary China (Zhang & Song, 2003) and early 20th century Japan (Nojiri, 1949), where migration is not negligible. Here, the preferred age difference could be different from the observed or acceptable age difference (Buss, 1989; Lampard, 1993); it is the former that is directly relevant to the marriage squeeze hypothesis (Ní Bhrolcháin, 2001; Matthews & Garenne, 2013). Then, the formulation is tested at each stage (e.g. from 15–19 to 20–24 years) along the marriageable life course of female cohorts using Japan's 1920–1940 census data. As preferred by the test (see Methods), one merit of the Japanese census data in contrast to data from most other countries is a 5-year rather than a longer intercensal period.

If the estimated effect of mate availability is significant at some life-course stage, the marriage squeeze hypothesis will not be rejected at the stage. Relevant estimations can then be used to infer to what an extent change in mate availability contributed to change in intercensal marriage experience in terms of first marriage rate at specific life-course stages, and thus change in Japan's marriage pattern from 1920 to 1940, when the female singulate mean age at marriage (SMAM) increased from about 21 years to more than 23 years, i.e. a trend towards later marriage (Hajnal, 1953; Retherford *et al.*, 2001). Although an effect of marriage squeeze on the change in marriage pattern has been noticed (e.g. Hajnal, 1965; Mari Bhat & Halli, 1999), the current study raises the issue of an age-specific effect, i.e. an effect along the marriageable life course.

Methods

Theoretical framework

The terminology 'kinetics' is borrowed from chemistry, where it refers to the response of the rate of a chemical reaction to the concentration/availability of substrates; here it describes how first marriage rate responds to the availability of preferred mates. If the marriage squeeze hypothesis is true, marriage kinetics can be derived by analysing the flow of singles in the context of mortality and migration.

Figure 1 shows the flow of singles (men or women) within a short time interval Δt for a region open to migration. Here, Δt is a short time interval within an intercensal period; P is the population size of a cohort in the region at the beginning of Δt; S_1 is the proportion single in P at the beginning of Δt; ΔM is the net migration (in-migrations minus out-migrations) during Δt (for simplicity, migration is assumed to occur at the beginning of Δt); $S_1 + \delta S$ is the proportion single in the net migration population ΔM; μ is the mean mortality rate during Δt (it is assumed that there is no mortality differential by marital status); η is the mean first marriage rate during Δt; and S_2 is the proportion single in the population (including migrants) at the end of Δt.

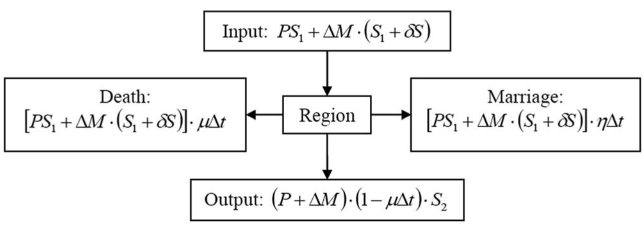

Fig. 1. The flow of singles (men or women) within a short time interval Δt for a region open to migration. See text for definition of abbreviations.

https://doi.org/10.1017/S0021932015000437 Published online by Cambridge University Press

Based on the balance of singles, it follows that:

$$\eta = \frac{PS_1 + \Delta M \cdot (S_1 + \delta S) - [PS_1 + \Delta M \cdot (S_1 + \delta S)] \cdot \mu \Delta t - (P + \Delta M) \cdot (1 - \mu \Delta t) \cdot S_2}{[PS_1 + \Delta M \cdot (S_1 + \delta S)] \cdot \Delta t}$$

The above equation reflects multiple decrements in singles: marriage, mortality and migration. If the marriage squeeze hypothesis is true, a lower availability of preferred mates brings a lower marriage rate. Furthermore, if the relationship is linear, it follows that:

$$\eta = \frac{PS_1 + \Delta M \cdot (S_1 + \delta S) - [PS_1 + \Delta M \cdot (S_1 + \delta S)] \cdot \mu \Delta t - (P + \Delta M) \cdot (1 - \mu \Delta t) \cdot S_2}{[PS_1 + \Delta M \cdot (S_1 + \delta S)] \cdot \Delta t} = k\overline{A}$$

where, $\overline{A}$ is the mean availability of preferred mates during Δt and k is a constant coefficient. Under the assumption that $\Delta M/P \to 0$, $S_2 - S_1 = \Delta S \to 0$ and $\overline{A} \to A$ when $\Delta t \to 0$, it follows that:

$$\begin{aligned}
\lim_{\Delta t \to 0} k\overline{A} = kA = \lim_{\Delta t \to 0} \eta \\
= \lim_{\Delta t \to 0} \frac{\begin{array}{c} -P \cdot (S_2 - S_1) - \Delta M \cdot (S_2 - S_1) + \Delta M \cdot \delta S \\ + (P + \Delta M) \cdot \mu \Delta t \cdot (S_2 - S_1) - \Delta M \cdot \delta S \cdot \mu \Delta t \end{array}}{[PS_1 + \Delta M \cdot (S_1 + \delta S)] \cdot \Delta t} \\
= \frac{-dS_1}{S_1 dt} - 0 + \frac{dM}{Pdt} \cdot \frac{\delta S}{S_1} + 0 - 0
\end{aligned}$$

Thus, the differential form of marriage kinetics in the context of mortality and migration is:

$$\frac{-dS}{Sdt} + \frac{dM}{Pdt} \cdot \frac{\delta S}{S} = kA \tag{1}$$

Equation (1) differentiates between nominal and actual first marriage rates. The term $-dS/Sdt$ is the nominal first marriage rate or force of nuptiality (Hajnal, 1953); over an intercensal period, the mean value of it: $\int_{t_1}^{t_2} \left(-\frac{d \ln S}{dt}\right) dt \big/ (t_2 - t_1) = -\frac{\ln S_2 - \ln S_1}{t_2 - t_1}$ depends only on the proportions single at the two ends (i.e. starting and ending censuses) of the period. The term dM/Pdt is the instantaneous net migration rate, (dM/Pdt) $(\delta S/S)$ is the migration-based adjustment to the nominal first marriage rate, and $\frac{-dS}{Sdt} + \frac{dM}{Pdt} \cdot \frac{\delta S}{S}$ is thus a representation of the adjusted or actual first marriage rate. Evidently, the nominal first marriage rate will be lower than the actual rate if the net migration population is composed of single immigrants; in extreme cases, the nominal rate could be zero or even negative, although the actual rate will always be positive.

The differential form of the kinetics equation considers instantaneous force of nuptiality. For application to census data, it is integrated to get its integral forms. To do so, two assumptions are made. Firstly, migration rate is constant during an intercensal period. As $\overline{r} = \overline{b} - \overline{\mu} + \overline{\tau}$ and for a cohort $\overline{b} = 0$ (Preston *et al.*, 2001), it follows that $\overline{\tau} = \overline{r} + \overline{\mu}$, i.e. mean net migration rate is the sum of the mean growth rate and mortality

https://doi.org/10.1017/S0021932015000437 Published online by Cambridge University Press

rate over an intercensal period; here, $\bar{r}$ is the mean growth rate, $\bar{b}$ is the mean birth rate, $\bar{\mu}$ is the mean mortality rate, and $\bar{\tau}$ is mean net migration rate. Secondly, mate availability A is a linear function of t, i.e. $A_t = A_0 + at$.

The integration is illustrated for two typical situations of migration. If migrants are composed of singles only, $\delta S/S = (1-S)/S$. When a logistic curve is assumed for decrement of singles, $\ln[(1-S)/S]$ is a linear function of time/age. Let:

$$\ln[(1-S)/S] = u + vt,$$

i.e.

$$(1-S)/S = e^{u+vt} \Rightarrow v = \left(\ln\frac{1-S_2}{S_2} - \ln\frac{1-S_1}{S_1}\right)\Big/(t_2-t_1).$$

It follows that:

$$\int_{t_1}^{t_2}\left(-\frac{d\ln S}{dt} + \frac{dM}{Pdt}\cdot\frac{\delta S}{S}\right)dt = \int_{t_1}^{t_2} kAdt = \int_{t_1}^{t_2} k\cdot(A_0+at)dt$$

$$\Rightarrow -(\ln S_2 - \ln S_1) + (\bar{r}+\bar{\mu})\cdot\int_{t_1}^{t_2} e^{u+vt}dt = kA_0\cdot(t_2-t_1) + \frac{1}{2}ka\cdot\left(t_2^2-t_1^2\right) + C$$

$$\Rightarrow -\frac{\ln S_2 - \ln S_1}{t_2-t_1} + (\bar{r}+\bar{\mu})\cdot\left(\frac{1-S_2}{S_2} - \frac{1-S_1}{S_1}\right)\Big/\left(\ln\frac{1-S_2}{S_2} - \ln\frac{1-S_1}{S_1}\right)$$

$$= k\cdot\frac{A_{t_1}+A_{t_2}}{2} + \frac{C}{t_2-t_1} \tag{2}$$

If migrants are composed of the married only, $\delta S/S = (0-S)/S = -1$. It follows that:

$$-(\ln S_2 - S_1) + (\bar{r}+\bar{\mu})\cdot\int_{t_1}^{t_2}\frac{0-S}{S}dt = kA_0\cdot(t_2-t_1) + \frac{1}{2}ka\cdot\left(t_2^2-t_1^2\right) + C$$

$$\Rightarrow -\frac{\ln S_2 - \ln S_1}{t_2-t_1} - (\bar{r}+\bar{\mu}) = k\cdot\frac{A_{t_1}+A_{t_2}}{2} + \frac{C}{t_2-t_1} \tag{3}$$

Thus, the generalized differential form leads to various integral forms according to the marital status composition of migrants and/or empirical representation of availability. Theoretically, actual first marriage rate can also be computed by dividing total marriages by total person-years of a cohort between two censuses. However, for a cohort aged 15–19 in a census year, the count of marriages in each age from 16 to 20 years in the next year may not be available (e.g. in Japan's 1920–1940 vital statistics); in such cases, Eqns (2) and (3) have the merit of needing only statistics on marriages in each census year, but not statistics between census years. It is worth noting that if $t_2 - t_1$ is shorter, the assumptions in the integration (e.g. constant migration rate) over an intercensal period are easier to satisfy. In this sense, census data with a 5-year intercensal period (e.g. Japan and South Korea) are preferred to those with a 10-year intercensal period (e.g. Malaysia and Taiwan) in testing kinetics equations.

Through Eqns 2 and 3, the theoretical formulation can be linked to the census data via a linear regression analysis to test whether and how lower mean mate availability predicts a lower adjusted or actual mean first marriage rate in a cohort over an intercensal period.

Census data from Japan

The data used for regression analysis come from: (1) the Population Census of Japan, 1920–1940 (Statistics Bureau, 2011); and (2) Vital Statistics, 1920–1940 (Tōkei-Kyoku, 1920–1940). These data are generally in both Japanese and English, but the original Japanese version is more comprehensive and is thus used in this study. The choice of 1920–1940 census data is explained later.

The Japanese government has conducted a population census every five years since 1920, with a break in 1945 due to WWII. The census counts persons that live in local areas over 6 months, by age group (0–14, 15–19, 20–24, etc.), sex and marital status at both national and prefecture (the first administrative division level in Japan; Wikipedia Contributors, 2015) levels. Based on such data, period female SMAM can be calculated at the national level (Retherford *et al*., 2001). Until 1925, female SMAM was close to 21 years, the upper limit of SMAM typical of a non-European marriage pattern; in 1940, female SMAM was already more than 23 years, the lower limit typical of a European pattern (Hajnal, 1965). From 1920 to 1940, the proportion of single women at age 40–44 levelled off at 2%, far less than the lower limit level of 10% in a European marriage pattern (Hajnal, 1965). Thus, there was a trend towards later but not less marriage (for a discussion of the development of marriage that is not only later but also less in East Asia in recent decades, see Retherford *et al*. (2001), Chen & Chen (2014) and Raymo *et al*. (2015)); in Hajnal's terms, it might also be said that the Japanese marriage pattern shifted from a non-European one to a semi-European one from 1920 to 1940.

It makes no sense to do regression analysis at the national level, as only one value of mate availability and one first marriage rate can be arranged from each census. However, it is feasible to conduct analysis with census data from 47 prefectures. Three life-course stages (15–19 → 20–24, i.e. from age 15–19 to age 20–24; 20–24 → 25–29; 25–29 → 30–34) are analysed in the study; these covered more than 90% of marriages in single women during the period 1920–1940 (Statistics Bureau, 2011). Given that there are 47 prefectures, and for each prefecture four intercensal periods and three life-course stages are considered, there are $47 \times 4 \times 3 = 564$ data points for regression analysis; each point corresponds to one value of mean mate availability and one mean first marriage rate in a cohort over an intercensal period.

Assignment for variables in the regression analysis

Firstly, a measure of mate availability is introduced. Lampard (1993) indicated that although theoretically inferior to other more complicated measures of mate availability like iterated availability ratio (IAR), sex ratio may in practice be a better predictor of marriage rate. Thus, mate availability for single women in an age group and a given census can be represented empirically as follows (subscript e refers to *empirical*):

$$A_e = \frac{1}{2} \cdot \frac{\text{number of single men in the same age group}}{\text{number of single women in the age group}} + \frac{1}{2} \cdot \frac{\text{number of single men in the next age group}}{\text{number of single women in the age group}} \qquad (4)$$

Mate availability, as defined in Eqn (4), is a reasonable representation of the availability of mates with a preferred age difference. The survey on age preference nearest to the period 1920–1940 was the 1982 National Fertility Survey (National Institute of Population and Social Security Research, 1982), which indicated that the preferred age difference among single women roughly varied from zero to five years, with an average at about three years. The cross-country studies on mate selection preferences by Buss (1989) and Kenrick and Keefe (1992) also indicated an age difference preference by females within this range across the countries surveyed (underdeveloped, developing and developed). Thus, it is reasonable to represent single females' preferred mate set as single males in two adjacent age groups and preferred age difference as the mid-point between zero (mates in the same age group) and five years (mates in the next age group) (see also Torabi & Baschieri, 2010). Mate availability calculated in this way declined with time in the pre-WWII period but increased with time in the post-war decades after a major loss of marriageable men in the war; thus, it is better to use the pre-war data to test the marriage squeeze hypothesis, which concerns the response of marriage behaviour to a decline in mate availability, i.e. a relative, not absolute, shortage of mates, with time.

Secondly, for a cohort, the adjusted intercensal first marriage rate is obtained by adding a migration-based adjustment to the nominal marriage rate according to Eqns (2) or (3). As discussed in the theoretical formulation, the nominal mean first marriage rate over an intercensal period is $-(\ln S_2 - \ln S_1)/(t_2 - t_1)$. Here, the time interval is the intercensal period (five years) and the two proportions correspond to a cohort at the two ends of an intercensal period; for example, women aged 15–19 in 1920 were 20–24 years old in 1925 and two proportions can then be found.

In adjusting nominal rate, the intercensal growth rate is assumed to be constant and calculated as:

$$\bar{r} = (\ln P_2 - \ln P_1)/(t_2 - t_1),$$

i.e. the mean growth rate over a full intercensal period. Mortality rate is also assumed to be constant over an intercensal period and can be computed using two methods. The first method divides total deaths by total person-years during an intercensal period; this method needs annual statistics of deaths at each age (note: the females aged 15–19 in 1920 were aged 16–20 in 1921). The second method is to compute the mean value of two rates at the beginning and end of an intercensal period: for a cohort, if the mortality rate was μ_1 in 1920 (calculated by counts of deaths in each prefecture; Tōkei-Kyoku, 1920–1940) and μ_2 in 1925, the intercensal mortality rate is then calculated as $\bar{\mu} = (\mu_1 + \mu_2)/2$; generally, μ_1 was close to μ_2. The first method is the one used in the current study. However, intercensal mortality rates computed by the two methods are almost the same; for example, the coefficients of correlation between mortality rates at the life-course stage 15–19 → 20–24 calculated by the two methods are 0.943, 0.986, 0.984 and 0.982 in the four intercensal periods, respectively. When annual age-specific vital statistics are not available, the second method can be used, as it needs only age- or age-group-specific vital statistics of a cohort in each census year, but not vital statistics in the years between the two censuses.

Intercensal migration rate, i.e. mean migration rate over a full intercensal period, is computed by adding intercensal growth and mortality rates together (the results are available from the author). In general, prefecture-specific migration rates were in the interval from −40‰ to 40‰ (note: the unit of migration rate is the same as that of growth rate, year^{-1}). Most areas had a negative migration rate, i.e. out-migration; however, metropolitan areas like Tokyo, the capital of Japan, and Osaka, historically the centre of commerce in Japan, had a positive migration rate, i.e. in-migration (for a review of migration to economic centres, see Yap, 1977). This pattern arose because in that period single girls migrated from rural areas to large cities, e.g. Osaka and Tokyo, for temporal/seasonal work before marriage (note: single women after age 25 or married women were generally not favoured by employers) and married women migrated for union with their husbands (Nojiri, 1949; Taeuber, 1958). Among men, single non-heir sons migrated to large industry centres for a better subsistence chance and married men migrated for temporary or seasonal work. There was a trend towards universal emigration from 1935 to 1940, as more and more men and women migrated to other countries for military services or land opportunities during war-time (Taeuber, 1958). In the 1940 census, overseas soldiers were still included in their original hometown populations, although overseas landers were excluded (Tōkei-Kyoku, 1920–1940).

Using intercensal migration rate, nominal first marriage rate is adjusted using Eqn (2) in the case of the life-course stage 15–19 → 20–24, in that migrants at this stage were almost universally single women. As for the stage 20–24 → 25–29, the nominal rate is adjusted using Eqn (3), in that regardless of migration within Japan or migration to other countries, most (75% on average) migrants at this stage were married women (see above discussion on sex, age and spatial patterns of migration; for more details, see Taeuber, 1958). In general, the ratios of the migration adjustment-induced change in first marriage rate to nominal first marriage rate were in the interval from −10% to 10%; however, they were evidently bigger in the metropolitan areas of Tokyo and Osaka (details available from the author). Table 1 lists Pearson's coefficients of correlation between the nominal and adjusted/actual marriage rates at the stages 15–19 → 20–24 and 20–24 → 25–29 in the four intercensal periods. With migration rates in the current study, there was a high correlation between nominal and adjusted first marriage rate. A simulation with two times the migration rate in this study (i.e. a level at 80‰) indicates that, under such a condition, the correlation coefficient declines quickly from 0.954 to 0.819 at the stage 15–19 → 20–24 during the period 1920–25; a simulation with three times the migration rate (i.e. a level at 120‰) produces a lower correlation (Table 1). No adjustment is made to the nominal first marriage rate in the case of stage 25–29 → 30–34 because migration at this stage can be disregarded (Taeuber, 1958; Nakagawa, 2010).

The adjusted rates indicate that in each prefecture there was an evident secular decline in first marriage rate at the life-course stage 15–19 → 20–24 from 1920 to 1940,

Table 1. Pearson's coefficients of correlation between nominal and adjusted first marriage rates among female cohorts in four intercensal periods from 1920 to 1940 in Japan

Migration	Life-course stage	1920–1925	1925–1930	1930–1935	1935–1940
Actual[a]	15–19 → 20–24	0.954	0.971	0.970	0.980
	20–24 → 25–29	0.959	0.973	0.941	0.879
Simulation 1[b]	15–19 → 20–24	0.819	0.885	0.875	0.909
	20–24 → 25–29	0.868	0.917	0.829	0.650
Simulation 2[c]	15–19 → 20–24	0.645	0.764	0.735	0.779
	20–24 → 25–29	0.769	0.855	0.724	0.466

[a]Actual migration rate.
[b]Simulated situation at two times the actual migration rates.
[c]Simulated situation at three times the actual migration rates.
For each test, degree of freedom = 45 and $p < 0.001$.

and thus a steady marriage postponement during the period (prefecture details are available from the author; for the overall trend, see Table 2). By contrast, in almost every prefecture first marriage rate at the stages 20–24 → 25–29 and 25–29 → 30–34 only declined slightly from 1920 to 1925, but remained almost constant from 1925 to 1940 (for the overall trend, see Table 2). Thus, the transition to later marriage during the period 1920–1940 was mainly driven by a decline in first marriage rate at the stage 15–19 → 20–24.

Panel regression model

The primary analysis is a two-way fixed effects model. Prefectures are taken as panels. Given that there were only four intercensal periods from 1920 to 1940, intercensal period is included as the second fixed effect, in the form of a dummy variable (Wooldridge, 2012). Heteroskedasticity is taken into account by computing robust covariance matrices of coefficients (Zeileis & Hothorn, 2002). The modelling platform is the statistical package 'plm' (Croissant & Millo, 2008) in R 3.1.2 (R Core Team, 2014). The model structure is:

$$y_{ij} = \upsilon_i + \theta_j + \beta x_{i_j} + \varepsilon_{ij}$$

where y_{ij} is the adjusted intercensal first marriage rate in prefecture i (= 1, 2, …, 47) and intercensal period j (e.g. 1920–1925); x_{ij} is the intercensal mate availability in prefecture i and intercensal period j; β is the regression coefficient; υ_i is the fixed parameter for prefecture i; θ_j is the fixed parameter for intercensal period j; and ε_{ij} is a random variable with a distribution $\varepsilon_{ij} \sim N(0, \sigma_i^2)$.

Results

The fixed effects model (Model 1) is run, using mate availability as defined in Eqn (4), separately for each life-course stage, i.e. three model runs in total (Table 2). At the stage

https://doi.org/10.1017/S0021932015000437 Published online by Cambridge University Press

Table 2. Regression coefficients from models with intercensal first marriage rate as the response variable

Life-course stage	Predictor	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
15–19 → 20–24	1920–1925[e]	0.0993	0.0963	0.0823***[h]	0.136
	1925–1930[f]	−0.0264***	−0.0261***	−0.0248***	−0.0332***
	1930–1935	−0.0470***	−0.0465***	−0.0442***	−0.0550***
	1935–1940	−0.0786***		−0.0754***	−0.0865***
	Mate availability[g]	0.0850***	0.0872***	0.0976***	0.0657***
20–24 → 25–29	1920–1925	0.277	0.240	0.302***	0.270
	1925–1930	−0.0284***	−0.0316***	−0.0262***	−0.0265***
	1930–1935	−0.0261***	−0.0254***	−0.0265***	−0.0254***
	1935–1940	−0.0249***		−0.0273***	−0.0318***
	Mate availability	0.0025	0.0236	−0.0116	0.0036
25–29 → 30–34	1920–1925	0.182	0.182	0.183***	0.182
	1925–1930	−0.0389***	−0.0389***	−0.0388***	−0.0389***
	1930–1935	−0.0354***	−0.0354***	−0.0352***	−0.0354***
	1935–1940	−0.0376***		−0.0376***	−0.0376***
	Mate availability	0.0022	0.0019	0.0013	0.0022

[a]Fixed effects model: intercensal first marriage rate is corrected using Eqn (4); both prefecture and intercensal period fixed effects are included.
[b]Fixed effects model: intercensal first marriage rate is corrected using Eqn (4); both prefecture and intercensal period fixed effects are included; 1935–1940 intercensal period is not considered.
[c]Random effects model: first marriage rate is corrected using Eqn (4); intercensal period dummy variable is included.
[d]Fixed effects model: intercensal first marriage rate is not corrected; both prefecture and intercensal period fixed effects are included.
[e]Regression coefficient of first marriage rate on the intercensal period 1920–1925 is treated as 'intercept' (not given by model but computed manually), which is the mean of individual intercepts for prefectures.
[f]Regression coefficients of first marriage rate on the intercensal periods 1925–1930, 1930–1935 and 1935–1940 are given by the model as the difference from that for the period 1920–1925.
[g]Regression coefficient of first marriage rate on mate availability.
[h]The intercept is given by model.
R^2 values for the three life-course stages are: Model 1, 0.957, 0.469 and 0.761; Model 2, 0.947, 0.693 and 0.794; Model 3, 0.943, 0.401 and 0.706; and Model 4, 0.960, 0.539 and 0.761.
***$p < 0.001$.

15–19 → 20–24, mate availability significantly affected first marriage rate; the regression coefficient is 0.0850, i.e. every 0.1 unit decrease in mate availability caused a 0.0085 decrease in first marriage rate. Mate availability did not significantly affect the first marriage rate at the other two life-course stages; in other words, change of mate availability was not responsible for the dynamics of first marriage rate at these two stages from 1920 to 1940.

The change in first marriage rate at the life-course stage 15–19 → 20–24 with time could be via two paths: first, mate availability changed with time and this change affected first marriage rate; second, some other unspecified factors changing with time affected first marriage rate. The contribution of mate availability to change in first marriage rate can then be calculated as the ratio of change in first marriage rate that was due to change in mate availability to total change in first marriage rate. At the stage 15–19 → 20–24, with mate availability controlled for, intercensal first marriage rate declined by 0.0786 from 1920 to 1940 (Table 2); the model with intercensal availability as the response variable and intercensal period as the predictor shows that mate availability declined by 0.251 from 1920 to 1940. Thus, the contribution to the decline in first marriage rate from the decline in mate availability was $0.251 \times 0.0850/(0.251 \times 0.0850 + 0.0786) = 21.3\%$.

Some parallel models are conducted as a contrast. Firstly, as mentioned above, the 1940 census included overseas soldiers so that the 1935–1940 intercensal mate availability could be overestimated and, consequently, the effect of mate shortage or marriage squeeze could be underestimated. Model 2 shows that if the 1940 census is excluded, the estimated effect of mate availability on the force of nuptiality at the life-course stage 15–19 → 20–24 would be slightly higher (Table 2); in such a case, the contribution of a decline in mate availability to the decline in the force of nuptiality from 1920 to 1935 is estimated at $0.225 \times 0.0872/(0.225 \times 0.0872 + 0.0465) = 29.7\%$. Secondly, Model 3 is a random effects model and gives a result qualitatively consistent with that of fixed effects Model 1. The estimated effect of mate availability at the stage 15–19 → 20–24 is higher than that from Model 1; presumably, some unobserved prefecture effect (e.g. local cultural norm encouraging early marriage among single women) contributed to both higher mate availability and marriage rate. As a result of the higher estimated effect of mate availability, the contribution would be estimated at 24.3% (Table 2). In both Model 2 and Model 3, the estimated effect of mate availability is not significant at the stages 20–24 → 25–29 and 25–29 → 30–34. Finally, to make a contrast with the above models, where force of nuptiality is adjusted by migration rate, Model 4 is run using nominal force of nuptiality. Still, the above conclusion does not change. However, the estimated effect of mate availability regarding the stage 15–19 → 20–24 is only about 0.0657, which means a contribution of 16.0%. In other words, the contribution of decline in mate availability to the decline in first marriage rate would be underestimated if intercensal marriage rate is not corrected for migration.

Discussion

This study provides a framework to estimate the effect of the availability of preferred mates on marriage formation in terms of marriage rate/incidence rather than prevalence in a cohort in the context of mortality and migration. The marriage kinetics Eqns (1)–(3) indicate that it makes more sense to check the association between mate availability and migration-adjusted rather than nominal first marriage rate. If no adjustment is made for migration, the mate availability effect could be underestimated, although the conclusion in significance terms from analysing nominal first marriage rate will be similar to that from analysing the adjusted one, provided migration rate is low, e.g. within ±40‰, as analysed in this study (cf. Table 2). When migration rate is high, nominal first marriage rate could be a biased measure of marriage formation (cf. Table 1) and the conclusion from analysing it might be misleading.

https://doi.org/10.1017/S0021932015000437 Published online by Cambridge University Press

The fixed effects models that are used to give an empirical test of kinetics equations with the Japanese demographic data indicate that mate availability was positively associated with first marriage rate and a decline in the former was partly responsible for the decline in the latter at the early stage of the marriageable life course, i.e. 15–19 → 20–24 years, in 1920–1940 Japan. Similar results were found in the 1946–1948 cohort of women in England & Wales (Ní Bhrolcháin, 2001) and the 1891–1895 cohort of women in France (Henry, 1966); both cohorts suffered severe marriage squeeze due to an excess of male deaths in the two World War periods, and as a result their first marriages were delayed or temporarily forsaken. Thus, at young ages, female mate choice was not flexible, the marriage squeeze hypothesis held, and the dynamics of mate availability played an important role in Japan's transition to later marriage by affecting first marriage rate. On the other hand, consistent with Bergstrom and Lam (1991), Ní Bhrolcháin (2001) and Choo and Siow (2006), the study also indicates that there was some flexibility in female choice at later stages of the marriageable life course, i.e. 20–24 → 25–29 and 25–29 → 30–34: in the period 1920–1940, there was no significant association between mate availability and first marriage rate at these two stages. Presumably, if mates with preferred age difference were in short supply, single women at these ages chose to marry less-preferred men. The Japanese vital statistics for 1920–1940 supported this inference: for females in age groups 25–29 and 30–34, 20% to 30% of all marriages were with men older by more than 10 years; in contrast, less than 15% of females in the age group 20–24 married men older by more than 10 years. Henry (1966) also found that at older ages, French women of the 1891–1895 cohort were more likely to marry too old or too young men rather than men with a 'normal' age difference, divorced men or widowed men. A similar phenomenon was observed in Romania's 1967 female cohort, which was much larger than cohorts both before and after it (Bradatan, 2009). South (1991) indicated that divorced or widowed men were more acceptable for older women, although these men were not highly valued, i.e. preferred.

It is worth discussing some points that appear to contradict the present finding. Firstly, the arranged marriage in Japan (so-called *miai*) does not affect the above interpretation of the results, in that *miai* was essentially just a manner of meeting, i.e. arranged meeting in contrast to *ren'ai* or love meeting; the final decision of marriage lay largely with the single females themselves (Hendry, 1981). Secondly, the contrasting pattern of female choosiness at different ages may not be observed, if the effect of mate availability at different life-course stages is not analysed separately, i.e. the interaction between mate availability and age is not analysed. Without considering the interaction effect, Torabi and Baschieri (2010) found that from the mid-1970s to 2000, the sex ratio calculated in the same way as in the current study was not significant in predicting female first marriage hazard, a measure closely related to first marriage rate, in Iran. Brandt *et al.* (2008) found that small size – i.e. an advantageous marriage market – of the female cohort born in China's 1958–1962 famine, did not bring a higher marriage rate to cohort members aged 27 or more in 1990; however, the study did not tell the case before age 27. Thirdly, as non-marital consensual union or pregnancy are different from marriage, which is generally accompanied by long-term commitment, the association between low sex ratio or mate availability and increased non-marital pregnancy rate in teenage girls found in Guttentag and Secord (1983) and Barber (2001) is not in conflict with the finding here, i.e. low mate availability brought about a decreased first marriage rate in young females.

Relaxation of female choosiness in selecting a marriage partner along the marriageable life course has also been noticed in other human societies. Analysis of Lonely Hearts advertisements in the US and UK found that, compared with young women, older women were less demanding in terms of a number of traits – including age difference – demanded of prospective long-term partners (Waynforth & Dunbar, 1995; Pawlowski & Dunbar, 1999). A similar result was found in a study of marriage or stable partner selection based on 7415 advertisements published in Spanish newspapers (Gil-Burmann *et al*., 2002). Thus, the change of female choosiness in long-term mate choice with age could be genuine and some evolutionary mechanism seems to underlie the cross-cultural phenomenon. Presumably, in selecting a marriage partner, females face a trade-off between the benefit of waiting for an ideal partner – e.g. a male who displays a preferred age difference, a trait signalling both sound investment ability and survivorship in the future – and the cost of rapid decline in reproductive opportunity with age, i.e. reproductive senescence (Pawlowski & Dunbar, 1999; Moore & Moore, 2001). Natural selection thus favours moderate choosiness regarding age difference at young ages but reduced choosiness at older ages in females.

To conclude, this test of the marriage kinetics equations indicates that a shortage of preferred mates partly accounts for why Japanese single women chose to postpone their marriage at young ages during the period 1920–1940, reflected in a consistently declining first marriage rate at the life-course stage 15–19 →20–24 (see Methods and Table 2), which left naturally a higher proportion of single women in age group 20–24. From 1920 to 1940, this proportion increased from 31% to 54% (Retherford *et al*., 2001). In this period, single women presumably did not want to forsake their marriage, as mate availability did not affect their marriage rate at later ages (Table 2), and the proportion of single women in the age group 30–34 just increased from 4% to 5%. Such findings on micro-level marriage behaviours shed light on the trend towards late but not less marriage in Japan from 1920 to 1940.

Given that other East Asian and South-East Asian countries also had a proportion of single women far less than 5% at age 35–39 when SMAM increased to 23 years (Jones & Gubhaju, 2009), it is clear that late marriage developed earlier than less or non-universal marriage, and population growth or wars inducing a shortage of preferred mates among females may partly explain the trend towards later marriage in these countries. For example, a study on historical Indian population supported this conclusion (Mari Bhat & Halli, 1999). It would be interesting to test the conclusion using data from Asian countries other than Japan and India. Additionally, investigation of the contributions of different factors – such as mate availability, urbanization, cultural change, female employment – to the nuptiality transition in Asia is warranted; some studies have produced illuminating qualitative results on the issue (e.g. Smith, 1980; Retherford *et al*., 2001; Jones & Gubhaju, 2009).

Acknowledgments

The author received financial support from a postdoctoral fellowship provided by the Max Planck Society. The author sincerely thanks Professor Mikko Myrskylä for suggestions on the research and comments on an early version of the work. The author would also like to thank Dr Akihiko Kato for help in reading the Japanese data and references and comments on an early version of the work, Dr Yu-Hua Chen for suggestions on data selection and revising the manuscript for publication, Tianyu Cui for collecting some of the data and the anonymous referees for their constructive comments.

References

Akers, D. S. (1967) On measuring the marriage squeeze. *Demography* **4**, 907–924.

Barber, N. (2001) On the relationship between marital opportunity and teen pregnancy – the sex ratio question. *Journal of Cross-Cultural Psychology* **32**, 259–267.

Bergstrom, T. & Lam, D. (1991) The two-sex problem and the marriage squeeze in an equilibrium model of marriage markets. *CREST Working Paper* 91–7, Center for Research on Economic and Social Theory, Department of Economics, University of Michigan, Ann Arbor, MI.

Bradatan, C. (2009) Large, but adaptable? A successful population policy and its long term effects. *Population Research and Policy Review* **28**, 389–404.

Brandt, L., Siow, A. & Vogel, C. (2008) Large shocks and small changes in the marriage market for famine born cohorts in China. *Working Paper* 334, Department of Economics, University of Toronto, Toronto, Canada.

Buss, D. M. (1989) Sex-differences in human mate preferences – evolutionary hypothesis tested in 37 cultures. *Behavioral and Brain Sciences* **12**, 1–14.

Chen, Y.-H. & Chen, H. (2014) Continuity and changes in the timing and formation of first marriage among postwar birth cohorts in Taiwan. *Journal of Family Issues* **35**, 1584–1604.

Choo, E. & Siow, A. (2006) Estimating a marriage matching model with spillover effects. *Demography* **43**, 463–490.

Croissant, Y. & Millo, G. (2008) Panel data econometrics in R: the plm package. *Journal of Statistical Software* **27**, 1–43.

Dixon, R. B. (1971) Explaining cross-cultural variations in age at marriage and proportions never marrying. *Population Studies* **25**, 215–233.

Fossett, M. A. & Kiecolt, K. J. (1991) A methodological review of the sex-ratio – alternatives for comparative research. *Journal of Marriage and Family* **53**, 941–957.

Gil-Burmann, C., Peláez, F. & Sánchez, S. (2002) Mate choice differences according to sex and age – an analysis of personal advertisements in Spanish newspapers. *Human Nature* **13**, 493–508.

Glick, P. C., Heer, D. M. & Beresford, J. C. (1963) Family formation and family composition: trends and prospects. In Sussman, M. B. (ed.) *Sourcebook of Marriage and the Family*. Houghton Mifflin Co., New York, pp. 30–40.

Goodkind, D. (1997) The Vietnamese double marriage squeeze. *International Migration Review* **31**, 108–127.

Guilmoto, C. Z. (2012) Skewed sex ratios at birth and future marriage squeeze in China and India, 2005–2100. *Demography* **49**, 77–100.

Guttentag, M. & Secord, P. F. (1983) *Too Many Women? The Sex Ratio Question*. Sage Publications, Beverly Hills, CA.

Hajnal, J. (1953) Age at marriage and proportions marrying. *Population Studies* **7**, 111–136.

Hajnal, J. (1965) European marriage patterns in perspective. In Glass, D. V. & Everskey, D. E. C. (eds) *Population in History: Essays in Historical Demgraphy*. Edward Arnold, London, pp. 101–143.

Hendry, J. (1981) *Marriage in Changing Japan: Community and Society*. Croom Helm Ltd, London.

Henry, L. (1966) Perturbations de la nuptialité résultant de la guerre 1914–1918. *Population* **21**, 273–332.

Jones, G. W. & Gubhaju, B. (2009) Factors influencing changes in mean age at first marriage and proportions never marrying in the low-fertility countries of East and Southeast Asia. *Asian Population Studies* **5**, 237–265.

Kenrick, D. T. & Keefe, R. C. (1992) Age preferences in mates reflect sex-differences in reproductive strategies. *Behavioral and Brain Sciences* **15**, 75–91.

Kim, D. S. (2004) The deficit of women in South Korea: evolution, levels and regional variations. *Population* **59**, 983–997.

Lampard, R. (1993) Availability of marriage partners in England and Wales: a comparison of three measures. *Journal of Biosocial Science* **25**, 333–350.

Lichter, D. T., LeClere, F. B. & McLaughlin, D. K. (1991) Local marriage markets and the marital behavior of black and white women. *American Journal of Sociology* **96**, 843–867.

Mari Bhat, P. N. & Halli, S. S. (1999) Demography of brideprice and dowry: causes and consequences of the Indian marriage squeeze. *Population Studies* **53**, 129–148.

Matthews, A. P. & Garenne, M. L. (2013) A dynamic model of the marriage market – Part 1: matching algorithm based on age preference and availability. *Theoretical Population Biology* **88**, 78–85.

Moore, P. J. & Moore, A. J. (2001) Reproductive aging and mating: the ticking of the biological clock in female cockroaches. *Proceedings of the National Academy of Sciences of the USA* **98**, 9171–9176.

Nakagawa, S. (2010) Internal migration and potential out-migrants in the 1920s and 1930s. In Takahashi, S. & Nakagawa, S. (eds) *Regional Population and the Demographic Transition in Japan. Kokon, Tokyo*, pp. 193–210.

National Institute of Population and Social Security Research (1982) *Attitudes Toward Marriage and Family Among the Unmarried Japanese Youth*. Ministry of Health, Labour and Welfare of Japan, Tokyo.

Ní Bhrolcháin, M. (2001) Flexibility in the marriage market. *Population-E* **13**, 9–47.

Ní Bhrolcháin, M. & Sigle-Rushton, W. (2005) Partner supply in Britain and the US: estimates and gender contrasts. *Population* **60**, 37–64.

Nojiri, S. (1949) *Nomin rison no jisshoteki kenkyu*, 3rd edn. Iwanami Shoten, Tokyo.

Pawlowski, B. & Dunbar, R. I. M. (1999) Impact of market value on human mate choice decisions. *Proceedings of the Royal Society Series B: Biological Sciences* **266**, 281–285.

Preston, S. H., Heuveline, P. & Guillot, M. (2001) *Demography: Measuring and Modeling Population Processes*. Blackwell Publishers Inc., Malden, MA.

R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. URL: http://www.R-project.org/

Raymo, J. M. (1998) Later marriages or fewer? Changes in the marital behavior of Japanese women. *Journal of Marriage and Family* **60**, 1023–1034.

Raymo, J. M., Park, H., Xie, Y. & Yeung, W. J. (2015) Marriage and family in East Asia: continuity and change. *Annual Review of Sociology* **41**, 471–492.

Retherford, R. D., Ogawa, N. & Matsukura, R. (2001) Late marriage and less marriage in Japan. *Population and Development Review* **27**, 65–102.

Saxena, P. C., Kulczycki, A. & Jurdi, R. (2004) Nuptiality transition and marriage squeeze in Lebanon: consequences of sixteen years of Civil War. *Journal of Comparative Family Studies* **35**, 241–258.

Schoen, R. & Baj, J. (1985) The impact of the marriage squeeze in five western countries. *Sociology and Social Research* **70**, 8–19.

Smith, P. C. (1980) Asian marriage patterns in transition. *Journal of Family History* **5**, 58–96.

South, S. J. (1991) Sociodemographic differentials in mate selection preferences. *Journal of Marriage and Family* **53**, 928–940.

South, S. J. & Lloyd, K. M. (1992) Marriage opportunities and family formation: further implications of imbalanced sex-ratios. *Journal of Marriage and Family* **54**, 440–451.

Statistics Bureau (2011) *Marital Status of the Population by Age in 5-Year Groups and Sex*. Ministry of Internal Affairs and Communications of Japan, Tokyo.

Taeuber, I. B. (1958) *The Population of Japan*. Princeton University Press, Princeton, NJ.

Tōkei-Kyoku (1920–1940) *Nihon teikoku jinkō dōtai tōkei*. Nihon Naikaku, Tokyo.

Torabi, F. & Baschieri, A. (2010) Ethnic differences in transition to first marriage in Iran: the role of marriage market, women's socio-economic status, and process of development. *Demographic Research* **22**, 29–62.

Waynforth, D. & Dunbar, R. I. M. (1995) Conditional mate choice strategies in humans: evidence from 'Lonely Hearts' advertisements. *Behaviour* **132**, 755–779.

Wikipedia Contributors (2015) *Prefectures of Japan*. Wikipedia, The Free Encyclopedia. URL: https://en.wikipedia.org/wiki/Prefectures_of_Japan (accessed 3rd August 2015).

Wooldridge, J. M. (2012) *Introductory Econometrics: A Modern Approach*, 5th edn. South-Western, Mason, OH.

Yap, L. Y. L. (1977) The attraction of cities: a review of the migration literature. *Journal of Development Economics* **4**, 239–264.

Zeileis, A. & Hothorn, T. (2002) Diagnostic checking in regression relationships. *R News* **2**, 7–10.

Zhang, K. H. & Song, S. (2003) Rural–urban migration and urbanization in China: evidence from time-series and cross-section analyses. *China Economic Review* **14**, 386–400.