Birthspacing and early stopping behavior in Xiaoshan County, Zhejiang Province, China

Darryl J Holman[1,2,3], Han Hua[2], Stevan Harrell[1,2]

[1] Department of Anthropology
[2] Center for Studies in Demography and Ecology
[3] Center for Statistics and the Social Sciences
University of Washington
Seattle, WA 98195 USA

ABSTRACT

A long-standing controversy has recently reemerged over presence or absence of deliberate fertility control in late-traditional China. One side maintains that Chinese fertility patterns can be explained only by deliberate fertility control. The other side holds that fertility in traditional China was limited by Malthusian positive checks, and that there is neither statistical nor cultural support for fertility control. We contribute to this debate in the form of new empirical data and a new statistical method for evaluating birthspacing and stopping behavior. The data come from genealogical books compiled and revised between 1400 and 1930 from Xiaoshan County, Zhejiang Province. The genealogies were used to reconstruct the number and birthspacing between sons who survive to adulthood for women who had at least one son. We begin with the culturally-grounded contention that should couples limit their fertility, they will do so after reached a desired number of sons. We hypothesize that the distribution of these surviving-son birth intervals for a given maternal age and the fraction of couples who stop reproducing at a given age should systematically differ by the number of sons in the family. The presence or absence of fertility control is inferred using parametric survival analysis in which both a "sterile" fraction and a parametric distribution of birth intervals are modeled. The results suggest that with two sons, many couples will began to space their births, and that with three sons there is evidence that couples intentionally stopping reproducing. The findings are consistent with our recent ethnographic survey of fertility behavior in the 1940s and 1950s in Xiaoshan County. We conclude that there was stopping and spacing behavior according to the number of healthy sons in a family in late-traditional China.

The nature of the late-traditional Chinese demographic regime—the complex of fertility, nuptiality, and mortality that determined the rate of population growth on the East Asian Mainland between about 1600 and 1950—has been the subject of long lasting debate, if not continuously since Malthus then at least continuously since Ho Ping-ti's *Studies on the Population of China* in 1959 (Ho 1959). Recently, the smouldering embers of this controversy have been fanned into flames by a specific dispute over the presence or absence of deliberate fertility control. On one side of this debate stand the revisionists of what John Shepherd (pers. comm.) calls "The California School" of James Lee, Wang Feng, Cameron Campbell (Lee and Campbell 1997; Lee and Wang 1999), and others, along with the work of Zhao Zhongwei (1997); this school maintains that the agreed-upon facts of low marital fertility for a pre-modern population (Wolf 1985; Coale 1985; Liu 1995), along with early and universal marriage for women, can be explained only by deliberate fertility control by married couples, consisting of three elements: late starting, wide spacing, and early stopping. They also maintain that statistical evidence for such fertility control can be found in patterns of parity- and sex- dependent fertility behavior, and that such fertility control is consonant with what we know about the economic rationality and the psychological dynamics of Chinese families. They see this pattern of fertility as indicative of what Malthus called preventive checks, of rational behavior intended to limit births where they would be detrimental to a family's welfare.

Arrayed against these revisionists are what we might call the "traditionalists," whose most articulate current spokesman, Arthur Wolf, holds to the more established Malthusian position that fertility in traditional China was limited only by positive checks, that there is neither statistical evidence of deliberate fertility control nor any cultural support for it, that the traditional maxim of *duo zi, duo fu*, more sons, more wealth was the principle behind fertility behavior, and that low marital fertility was the result of poverty, high mortality, disease, and the consequent inability, rather than the disinclination, to bear more children. Wolf's recent writings are in a tradition that goes back all the way to Malthus himself, who wrote that the Chinese, uncapable of the "moral restraint" of delaying marriage, thus subjected themselves to population control by positive checks (1798/1992: 42-43). The pages of this journal have recently been rife with contributions on both sides of this debate (Wolf 2001; Campbell, Wang, and Lee 2002; Zhao 2002), hence we will not recount the positions of the disputants here in more detail. Nor will we attempt here to criticize any of their handling of their own data, reevaluate their data, or criticize their reevaluation of each other's data. Rather we wish to make a contribution in the form of new empirical data and new statistical method for evaluating birthspacing and stopping behavior. In particular, we bring to bear evidence from a study of lineage populations from Xiaoshan County, Zhejiang Province, using data extracted from genealogical books compiled and revised between 1600 and 1930.

To analyze these data, we have had to employ a novel method to assess fertility behavior from mostly male-based measures of fertility from incomplete data. We use parametric survival analyses to model the duration of birth intervals in genealogies that only include records of birth to sons. Previous historical demographic studies have used semi-parametric survival analysis to examine the effects of covariates on birth interval length in order to find evidence for fertility control or regulation (e.g. David and Maroz 1989a,b; Van Bavel 2001). We chose to use a parametric model so that an explicit model of a "sterile" fraction of individuals could be incorporated as well. By examining the distribution of birth intervals to sons along with a fraction of "sterile" individuals at different ages and parities, we find evidence for fertility control under some circumstances. Evidence is found for both parity-specific stopping behavior and parity-specific birth spacing.

On the basis of this new sample and these new methods, we conclude that there is evidence for conscious fertility control in this population, but that it occurred only in special circumstances. To explain how we split this difference, we present below first an account of the population and the new perspective it offers, then discuss our sources and methods, and finally present our results.

THE POPULATION AND THE DATA

Our data, we hope, will begin to fill a hole in the existing universe of approaches to the problem of fertility control in Late Traditional China. No study that has engaged this specific problem directly has dealt with what might be called a "majority population" of Ming-Qing East Asia: commoner Han Chinese in the heavily populated agricultural/commercial regions of "China Proper." Lee, Campbell, and Wang's empirical studies concern either the Qing imperial clan, both ethnically and economically very different from the majority Chinese population, or in most cases the Han banner populations of Fengtian, whose different household registration and tax status during mid- and late-Qing times may well have led to different fertility behavior or to different recording of demographic data than one would expect from the "majority" (Lee and Campbell 1997; Lee, Campbell, and Wang 1993). Wolf's studies have dealt mostly with Taiwan, a frontier province until about 1870 or 1880 and an area whose public health, and thus its demographic regime, was heavily affected by Japanese colonial measures after 1805 (Wolf and Huang 1980).[1] Both Wolf, on one side of the debate, and Zhao, on the other, have employed data from John Lossing Buck's survey of farm families in the 1930s (Buck 1937), a time when the traditional regime in many areas was probably little changed, but there are still widespread arguments about the quality of those data.

To add data on a part of the majority population, we analyzed a database containing about 22,000 men and about 25,000 of their spouses, drawn from the genealogies complied between about 1400 and 1900 by the Lin, Shi, Wu, Tian, and Han lineages of Xiaoshan County in Zhejiang Provinces. Xiaoshan is a prosperous county just across the Qiantang River from the longtime provincial capital of Hangzhou, and was part of the Hangzhou-Shaoxing-Ningpo axis that was one of the primary subcenters of the Jiangnan concentration of Ming-Qing gentry culture (Schoppa 1989). Powerful lineages were a part of the Xiaoshan landscape since the Song period, and the genealogies we have analyzed are but a fraction of the number catalogued by the Genealogical Society of Utah (Telford, Thatcher, and Yang 1983), which in turn, probably represents only a fraction of those compiled over the years in the county.

Genealogies of lineages wealthy enough to compile them do not provide a random

[1] Both Lee's group and Wolf have included data from various studies of our "majority population" in their comparative analyses, but their primary empirical work has not focused on these data.

sample of the whole county's population: that a lineage could compile a genealogy means that its members, rich and poor, were nevertheless probably richer on average than the population of the county as a whole, and lineage genealogy compilers, despite pious self-admonitions to do otherwise, demonstrably compiled more complete records on the rich than on the poor. Nevertheless, these were not peripheral populations such as the Han banners or the frontier Taiwanese, nor were they an ethnically different and socially overprivileged elite like the Qing imperial household.

The data contained in the genealogies are not ideal for demographic research. They contain pedigree charts that set out the father-son connections between the members recorded in the genealogies, along with an entry on each lineage member including the name of his father; his sibling order among surviving sons of that father; his date and time of birth and death; a listing for each wife or concubine who bore children, listing her surname, birth and death dates and times, and names in order of any surviving sons born to her. By means of these records, we can compile a database of every listed man and his spouses and sons, with all their respective birth and death dates.

This database would seem an ideal tool for historical demographic research, but it has its severe problems. First, we have no record, or in some genealogies only a very incomplete record, of daughters; on average only half as many daughters as sons are listed in any genealogy, and their birth and death dates are almost never given (Harrell 1985, 1987; Telford 1990, Liu 1985). This means that we cannot reconstruct total fertility, except by assuming some sex ratio, and we cannot reconstruct a sibling order that includes both brothers and sisters. Second, we have no record of infant mortality. The prefaces to most genealogies indicate that a man will get his own entry in the genealogy if he either lives to the age of 16 *sui* (which can vary from 14 to 15 years by the modern count) or marries, meaning that there are no independent entries for boys who died young and unmarried. Some of these are included as sons under their parents' entries, but even just eyeballing the data shows that their inclusion is in no way systematic. So we are doubly unable to reconstruct total fertility except by assuming a mortality schedule at early ages lifted from a model life-table or similar source. Both Liu (1985, 1995) and Telford (1995) have attempted such reconstructions, but we believe that the absolute levels of their figures are too susceptible to un-provable assumptions to be reliable. Neither can we construct total marital fertility, because we have no record of the date of marriage.[2] Nevertheless we can, as explained below, look for evidence of fertility control by comparing fertility in certain life situations with fertility in others, remembering that the only childbearing we ever get any record of is in fact son bearing.

This is not the first time genealogical materials have been used to address questions of fertility and mortality in Late Traditional China. Liu and Telford in particular have tried to derive estimates of marital fertility from such records. But neither one addressed the question of fertility limitation. Of course, if fertility were high, as Telford (1995) found, one could argue that fertility limitation was absent or minimal, but this method seems to us fraught with difficulties: since fertility is estimated based on conjectural parameters, a high estimate of fertility leads to a dismissal of the possibility of fertility control at a second degree of extrapolation. We find the more prudent course to be one of not dismissing such evidence, but at

[2] One of us (SH) has attempted to reconstruct age at marriage from genealogies, without very satisfying results. See Harrell n.d.

the same time not taking it of proof of a position it was not meant to advocate in the first place.

We are leery, in fact, of any estimates of fertility made by any participants on either side of the current debate. Working ourselves with very incomplete data, we still withhold judgment on the question of whether those who have tried to use assumptions of demographic parameters such as childhood and infant mortality, age at marriage, or sex ratio, to reconstruct such absolute rates as total marital fertility have been able to set these parameters at their proper levels. We thus do not attempt to discuss the question of fertility control by looking at reconstructed absolute fertility rates and then inferring that these rates must indicate one or the other kind of behavior. Rather we take the approach that if people are controlling fertility, they will attempt to avoid births in certain situations but not in others. In particular, we look for ways to determine whether a couple in a particular age and parity situation—say, for example, with a 35-year-old husband, a 32-year-old wife, and three living sons—will be less likely to have a son within a certain interval than another couple with one or more variables altered, say parents of the same age but only two surviving sons. This way, we do not need to worry about how many sons did not survive, how many daughters the couple has, or even their ages when they married. We thus do not, as many others have, (see, for example, Wolf 1985, Coale 1985, Telford 1995, Liu 1985, 1995, Lee 1997) look at the general outcome—the total fertility or total marital fertility—and reason backwards to the behavior in certain situations. Rather, we attempt to look at the behavior in certain situations and see if it differs from behavior in other situations of controlled comparison.

Specifically, we apply this pattern of reasoning to only two of the three elements of deliberate fertility control outlined by Lee and his colleagues—early stopping, or the voluntary termination of childbearing while a woman is still fecund because she has some self- or culturally-defined ideal number of children or of sons, and wide spacing, or slowing down childbearing after the minimum acceptable number of children is reached. It seems to us that late starting does not make much cultural sense: there was clearly a great desire for continuation of the patriline and great social pressure brought upon young couples, particularly young wives, to prove their economic and cultural worth by producing sons (See Wolf, 2001, for a particularly forceful argument on this point). There would have been no reason to wait, except in the presence of some sort of Malthusian positive check such as severe malnutrition or husband's long-term absence. Similarly, it was clear that people wanted several sons as insurance against the death or fecklessness of other sons, so there would have been very little reason to space births out until the minimum number necessary for security was reached. But there seem to us to be cultural logics for stopping, and for spacing after a certain time. We know from the work of Gates (1993) that women often desired fewer children than their mates or mothers-in-law, and this desire may have been strong enough to override men's expectations of more and more children; or, fathers with four healthy teenage sons might have been brought around to the point that they should give their wives a break since they really didn't need anymore children. Or, as Wolf states (2001), perhaps people just relaxed a little after a certain number of children, and this is reflected in wider spacing at older ages among those who already have enough children.

We are not, however, depending entirely on an observed or inferred cultural logic to explain the apparent stopping behavior in our results. One of us (HH) conducted three months of ethnographic fieldwork in late 2002 among the Lai and Han lineages of Xiaoshan. This research discovered that among the elder generation now living in these areas (people born in the 1920s and 1930s, when neither modern methods of birth control nor modern ideas about fertility limitation had reached these areas, people still sometimes attempted to limit their fertility when they had reached an optimum (or, in many cases, already too large) family size. Eight out of 66 interviewees reported attempts to stop childbearing by one of three methods: taking medicinal herbs known to induce early cessation of menses, using the heddle of a loom to pound against the lower abdomen, or stopping sexual intercourse either by the consent of the couple or by the woman's decision protested but not effectively countered by her husband and his family.

Of the seven cases known to have attempted to stop childbearing, we find the following ages and numbers of living children at the time of stopping or attempting to stop:

1) age 33, 1 son, 4 daughters (herbal medicine)
2) age 27, 1 son, 2 daughters (loom; unsuccessful)
3) age 39, 2 sons, 1 daughter (loom; unsuccessful)
4) age 31, 3 sons, 1 daughter (abstinence from intercourse)
5) age 33, 5 sons, no daughters (abstinence)
6) age 30, 3 sons, 2 daughters (abstinence)
7) age 33, 1 son, 5 daughters (abstinence)
8) age 37, 3 sons (heavy exercise & prolonged breastfeeding; unsuccessful)[3]

Given a cultural logic and some preliminary ethnographic evidence for early stopping among the Xiaoshan lineages, here we present our first attempt to test statistically whether such early stopping was something people in Xiaoshan in the 16th-early 20th centuries did. We perform the test by asking whether people with a larger number of sons were more likely to stop childbearing than people with a smaller number of sons, when controlling for the age of the mother or the age of the father.

METHODS

Genealogies

Genealogies for five lineages in Xiaoshan county were obtained from the Library of the Hoover Institution at Stanford University and from the Genealogical Society of Utah in Salt Lake City. Records from individual couples were transcribed from microfilm into a computer database by one of us (SH) and a variety of research assistants at different times. As mentioned above, the genealogies contain information on men who survive to adulthood, along with information on their wives and concubines. Nearly complete information is recorded for a man's date of birth, birth order, marriage (including surname of wives and concubines who had children, names of biological parents, numbers and names of his sons. The genealogies usually include the date of death for a recorded man. A wife's births and death dates are usually recorded for a woman who gives birth to at least one son. The genealogies rarely contain any systematically collected information on daughters, and we made no use of any such information that was found.

---

[3] A complete ethnographic description of this research can be found in Han n.d.

The total number of subjects and the range of dates for the five genealogies are given in Table 1. The Han genealogy is, by far, the largest.[4] It begins in 607 AD, but does not include information on wives until 1497.[5]

Statistical Methods

The information in the genealogy can be used to reconstruct the number and birthspacing of sons who survive to adulthood for women who had at least one son. Clearly, the conventional measures of fertility based on all births to women cannot be computed directly from the resulting genealogical records. Our analytic approach is to examine age and parity-specific lengths of birth interval from one son to the next, as well as termination of reproduction, for evidence of intentional fertility limitation. We begin with the culturally-grounded contention mentioned above: if couples limit their fertility through spacing and stopping, they will do so after they have reached a desired number of sons. If so, then the distribution of birth intervals for a given maternal age should systematically change by the number of sons in the family. Likewise, stopping behavior should systematically change by the number of sons. The presence or absence of fertility control is inferred using parametric survival analysis in which both a "sterile" fraction and a parametric distribution of birth intervals are modeled.

The sterile fraction: By the term *sterile fraction*, we mean the fraction of individuals who will not produce another adult son. Many of the birth intervals are right-censored because the genealogy ended, the mothers died, no more sons were born, or sons did not survive to age 16 *sui* prior to menopause. These right-censored observations provide information about both the distribution of birth intervals *and* the fraction of "sterile" individuals. We model the entire distribution of completed and right-censored birth intervals as a mixture model consisting of a distribution of birth intervals and a sterile fraction (e.g. Wood et al. 1994). We cannot assign causes of stopping for individuals (or even identify right-censored individuals who eventually would go on to reproduce). Thus, what we call "sterility" encompasses stopping for all reasons combined: natural menopause, secondary sterility, intentional stopping, accidental stopping, and, indeed, births to a series of children out of which no son survives to age 16 *sui*.

Can the fraction of sterile individuals provide any information whatsoever about stopping behavior? By examining the parity-specific sterile fraction for a group of same-aged women, the following patterns can be hypothesized:

(1) *No stopping behavior*: The sterile fraction will show no parity-specific pattern for a given age group. Across age groups, sterility will increase just by virtue of declining age-specific fecundability and eventually menopause.

(2) *Ambiguous*: The fraction of sterile women will increase by parity within an

---

[4] We are now entering data from an even larger genealogy, the Lai. Having the Lai data will nearly double the size of our database, and will enable us to refine our analysis considerably.

[5] Many genealogies begin with a long single line of descent, sometimes reaching back as far as mythical figures such as the Yellow Emperor, Yao, or Shun. Meaningful data usually begin at or after the time when the first ancestor of a lineage moved to its current place of residence.

age group, reflecting an increase in risk of secondary sterility associated with parturition. Sterility will also increase by a woman's age.

(3) *Stopping behavior*: The fraction of sterile women will increase considerably above an "ideal" number of sons. For example, if most couples have a target of three sons, then within an age group there will be little increase in the sterile fraction for women with two or fewer sons, but a large increase in the sterile fraction for women with three sons.

The distribution of birth intervals: The second component of the statistical model is a parametric distribution for surviving-son birth intervals. Simple theory is not available for constructing a proper probability distribution for birth intervals. More complex theories, treating birth intervals as convolutions of simpler etiologic processes, are found in Sheps and Menken (1973), Leridon (1973), or Wood (1994). These complex models are not helpful for analyzing the genealogy data employed here because they involve estimating many parameters, and they require information not available for this analysis, details like duration of breastfeeding. An alternative to using parametric models is to use semi-parametric methods like the Cox proportional hazards model, as has been done by a number of previous investigators (David and Maroz 1989a, b). Semi-parametric survival models, however, either place restrictions on the sterile fraction, do not provide for simultaneous estimation of covariates, or are unidentifiable (Heckman and Walker, 1984; Taylor 1995).

For this reason, we have selected a flexible parametric model that should at least approximate the distribution of birth intervals. Birth intervals are modeled as following a two-parameter Gamma distribution. A more etiologic model might be constructed from a convolution of distributions of zero or more births to boys who do not survive to 16 *sui* and births to daughters within each observed birth interval, weighted by the probability of each possible parity. The gamma distribution is somewhat compatible with this idea because convolutions of gamma distributions are themselves gamma distributions. Additionally, the Gamma distribution contains the exponential distribution as a special case, and the exponential distribution is an etiologic model of the fecund waiting time to next conception. The results are unlikely to be changed much by the particular choice of parametric distributions.

The two-parameter gamma probability density function (PDF) is given by

$$f(t) = \frac{\lambda^c t^{c-1}}{\Gamma(c) e^{\lambda t}} \tag{1}$$

Where $\Gamma()$ is the Euler's gamma function and λ (> 0) and c (> 0) are parameters of the distribution. The λ parameter can be interpreted as a hazard (sometimes the parameter $b = 1/\lambda$ is used, where b is a scale parameter), and the c is a shape parameter.

The survival function corresponding to the gamma PDF is

$$S(t) = \frac{\Gamma(c, \lambda t)}{\Gamma(c)} \tag{2}$$

and the hazard function is

$$h(t) = \frac{\lambda^c t^{c-1}}{\Gamma(c, \lambda t) e^{\lambda t}} \quad (3)$$

The mean of the distribution is c/λ and the variance is c/λ^2 (Evans et al. 2000; Johnson et al. 1994).

Likelihoods. The observed birth intervals can be subdivided into two types for statistical analysis. The first are those for which the beginning date and the end date of the interval are observed. These are *complete* observations, because all information on the interval duration is known. The second type of birth intervals are right-censored—those for which a starting date is known but no end date is known. This type of observation arises because either the genealogy ended prior to the next birth or the individual died prior to the next birth.

Both types of observations are used to find maximum likelihood estimates of the model parameters as well as the effects of covariates. Call $f(t|\lambda,c)$ the gamma PDF and $S(t|\lambda,c)$ the corresponding survival function for which we will estimate parameters λ and c. The basic likelihood for N complete individual observations and N_r right-censored observations (right-censored after t_i days for individual i) is

$$L = \prod_{i=1}^{N} f(t_i \mid \lambda, c) \prod_{i=1}^{N_r} S(t_i \mid \lambda, c) \,. \quad (4)$$

Standard survival analysis assumes all birth intervals are ongoing and will eventually end in another birth. This assumption is violated when some fraction of couples terminate reproduction either naturally (secondary sterility or menopause) or artificially (birth control). Even in prospective studies of birth intervals on living couples, it can be difficult to decide whether open birth intervals result from chance, behavioral processes, or physiological processes. To accommodate all types of stopping, likelihood (4) is extended to provide for simultaneous estimation of p ($0 \le p \le 1$), the proportion of "sterile" couples. When sterile couples are mixed with fecund ones, the survival distribution begins at one and asymptotically approaches fraction p rather than zero. With time, the birth intervals still under observation are composed of a larger and larger fraction of the sterile couples until only sterile couples remain.

Call $S_f(t|b,c)$ and $f_f(t|b,c)$ the SDF and PDF, respectively, for the non-sterile subgroup of birth intervals. The fraction of all intervals that have not been completed by time t is equal to $S_f(t|b,c)$ weighted by the sterile fraction $1-p$, and a second fraction p of cycles that will never end. The overall survival distribution that includes both subgroups is:

$$S(t \mid \lambda, c, p) = (1-p) S_f(t \mid \lambda, c) + p \,. \quad (5)$$

Likewise, the PDF is composed of fraction $1-p$ complete intervals that end with probability $f_f(t|b, c)$ weighted by fraction p cycles that fail with probability 0, so that overall the probability density is

$$f(t \mid b,c,p) = (1-p) f_f(t \mid \lambda, c) + p \times 0 \\ = (1-p) f_f(t \mid \lambda, c) \quad (6)$$

Combining (5) and (6) with (4) gives the likelihood for all observations

$$L = \prod_{i=1}^{N} \left[(1-p) f_f(t_i \mid \lambda, c) \right] \prod_{i=1}^{N_r} \left[(1-p) S_f(t_i \mid \lambda, c) + p \right] \quad (7)$$

Covariates: Separate models are estimated for four age categories, 15 to 20, 20 to 25, 25 to 30, and 30 to 35. Each age category defines the age at which the birth interval opened. So, for example, a woman who never reproduces beyond her last confinement at age 22, will contribute an observation in the 20-25 age category, and no later age categories.

The effects of different parities are included in the model as fixed covariates. Dummy variables are created for different parities, and sometimes groups of parities. The reference category is always parity one.

The covariates are incorporated as affecting two different parts of the model. First, parity is modeled as an affect on the *hazard* of the gamma distribution. We use a loglinear specification with $\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h} = x_{1i}\beta_h 1 + x_{2i}\beta_h 2 + \ldots + x_{Mi}\beta_{hm}$ formed by M covariates $\mathbf{x}$ for the ith observation and M parameters $\boldsymbol{\beta}_\mathbf{h}$. The hazard for the i-th individual is $h_i(t_i|\lambda, c, \mathbf{x}_i\boldsymbol{\beta}_\mathbf{h}) = h(t_i|\lambda, c)\exp(\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h})$. This specification for covariates is a proportional hazards model (Kalbfleisch and Prentice 1980). Under the proportional hazards model, the effect of $\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h}$ on the PDF as $f_{f_i}(t_i \mid \lambda, c, \mathbf{x}_i\boldsymbol{\beta}_\mathbf{h}) = f_f(t_i \mid \lambda, c) S_f(t_i \mid \lambda, c)^{\exp(\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h})-1} e^{\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h}}$ and the effect of $\mathbf{x}_i\boldsymbol{\beta}_h$ on the SDF is $S_{f_i}(t_i \mid \lambda, c, \mathbf{x}_i\boldsymbol{\beta}_\mathbf{h}) = S_f(t_i \mid \lambda, c)^{\exp(\mathbf{x}_i\boldsymbol{\beta}_\mathbf{h})}$.

Covariates are also incorporated as an effect on the proportion of sterile individuals. We use a logistic specification to model effects of covariates on the overall probability of sterility. For the i-th individual, $p_i = [1 + \exp(p + \mathbf{x}_i\boldsymbol{\beta}_\mathbf{p})]$-1. Parameter p determines the baseline probability of sterility, and $\boldsymbol{\beta}_\mathbf{p}$ quantifies the effects of parity on the probability of sterility.

Maximum likelihood estimates of parameters were found by numerically evaluating likelihoods using the *mle* programming language (Holman 2003). Estimates of the standard errors for all parameters are found by the method of Nelson (1982), which involves inverting a numerical approximation of Fisher's information matrix. The most parsimonious combination of covariates was selected using minimum Akike's information criterion (AIC) (Akike, 1973, Burnham and Anderson, 1998). This criterion balances the tradeoff between goodness-of-fit and the fewest number of parameters to include in the model.

RESULTS

The numbers of birth intervals that contributed to the analysis at each surviving-son parity for each age category are given in Table 2. The number of birth intervals is smaller than the number of wives in the database. This reduction is primarily because we only included women for whom an age at death is recorded.

Parameter estimates for the most parsimonious model in each of the four age categories are given in Table 3. For each model, individual surviving-son parities were incorporated as covariates on both the hazard distribution and the sterile fraction. The highest parity category always included parities at and above that category.

Table 4 is based on the parameter estimates in Table 3, and shows the proportion of "sterile" individuals, along with the mean and standard deviation of birth interval lengths for the four age groups and by each parity category. A consistent pattern of sterility is found within all age groups. At parity 1 the estimated sterile fraction ranges from 15 to 36 percent. The sterile fraction at parity 2 drops to under 10 percent. At parity 3, the fraction sterile increases substantially and increases by parity in age groups that go beyond three.

The mean length of surviving-son birth intervals in the 15-20 and 20-25 age groups show a mean of just less than three years at parity 1. Birth intervals at parity two increase to at least 4.5 years at higher parities. The 25-30 age group, begins with a longer mean birth interval of 4.4, but increases to 5.4 at parity two. The 30-35 age group shows the opposite pattern of the previous three age groups, beginning at a mean of 7.6 years and exhibiting a shorter mean at higher parities. Birth interval survival distributions for each age group, with the estimated fraction of sterile individuals removed, are given in Figure 1 through 4.

DISCUSSION

We make the assumption that both stopping behavior and spacing behavior will be determined by the number of healthy sons a woman has. Of course, the number of healthy sons is indirectly measured through the genealogical records that only contain sons who survive to age 16 *sui*. Additionally, we are indirectly measuring both stopping behavior and spacing behavior. Stopping behavior includes all reasons that couples do not produce a son who survives to 16 *sui*, and spacing behavior includes all types of age-related and parity-related changes in spacing of births to or survival of sons. These changes undoubtedly are brought about through a mixture of physiological changes by age and parity, intentional and unintentional behaviors; they work through changes in fertility as well as changes in sub-adult mortality. In short, we are trying to uncover latent behavior using indirect and incomplete observations of the behavioral outcomes.

In the face of these daunting data deficiencies, a number of consistent and reasonably simple patterns are apparent across the four age groups. In all four age groups, baseline sterility at parity 1 is relatively high compared to sterility at parity 2 (the exception is 25-30 where sterility at parity 1 is no different for parity 2). At parity 3 and above, sterility increases to above parity 1 sterility levels. The simplest interpretation is that parity 1 and 3 are thresholds for couples who are willing to stop and capable of stopping reproduction. The subset of couples who have two sons are likely to go on to have a third son. The very large difference in sterility between parity 2 and parity 3 in all four age groups strongly suggest intentional stopping behavior at parities three and above. There is no simple physiological explanation for this pattern.[6]

---

[6] We have also considered some alternative explanations here. Perhaps a substantial portion couples who stop after one son are having a long-sought son after bearing many daughters and, relieved at the presence of an heir, stop bearing otherwise-expensive children. Additionally, a fraction of women may become physiologically sterile after one birth**.**

We use the distribution of surviving-son birth intervals as an indirect measure of spacing behavior. The results of the birth-interval analysis are a little more difficult to interpret than are the sterility results. In the first three age categories, there is a significant increase in the length of birth intervals by parity. There is a large difference between parity 1 and parity 2 and above in the two youngest age groups. In fact, the parity-specific means are very similar between the 15-20 year-old mothers and the 20-25 year-old mothers. By ages 25-30, the same parity-specific relationship holds, but the mean lengths are shifted upward and the difference in means for parity 1 and 2+ are not as large. By ages 30-35, the pattern is entirely reversed: birth interval length gets shorter at higher parities.

We suggest that this pattern reflects a selectivity bias at older ages. If some couples are able to drop out of the reproductive pool upon completing their target family composition, as appears to happen for surviving-son parities 1 and 3 at younger ages, then at later reproductive ages, parity becomes increasingly tied to fecundity and therefore birth interval length. Subfecund couples—who will typically have fewer children at later reproductive ages—that have not attained their targeted family composition, increasingly make up the pool of those trying to reproduce at later ages. In other words, low-parity, late age birth intervals become longer because of a higher fraction of less fecund couples, rather than because of intentional spacing. The remaining couples at later reproductive ages are those who cannot or chose not to limit their fertility. These couples most likely come from the entire spectrum of fecundities, so that couples of high fecundity—those with the shortest birth intervals—will achieve higher parities. From the evidence in the first three age categories, it appears that parity 2 (sons) is a threshold where many couples will begin to space births, and that parity 3 (sons) marks the beginning of intentional stopping behavior among a large portion of couples. This seems consistent with our findings from a survey of fertility behavior in the 1940s and 1950s among the descendants of the Han and Lai lineages (Han, Wang Zhou, and Harrell, n.d.). We found women born in the 1920s and 30s, whose prime childbearing years took place before the birth limitation campaigns started in 1970, parity progression ratios decline sharply once a couple has two sons, and sharply again for every number of sons thereafter. They decline much less precipitously if a family has only daughters. For example, women with three children, at least one of whom is a son, have a .65 probability of bearing another child, while those with three daughters have a probability of .875. In other words, it seems highly likely that mothers of sons can at least, in Wolf's words, "relax a little" and not continue striving to have more children. In addition, we find that all but two of the ethnographically recorded cases of intentional stopping listed above either involve couples with at least two sons, or in two cases, couples with only one son, but with at least four daughters. In other words, it appears rare for people to stop intentionally with only one son, while stopping after three sons appears quite common. This leaves unexplained the large sterile fraction found in the genealogical data at parity 1 son, described above, but is very consistent with increasing sterile fractions after 3 or more sons.

It is important to realize here, however, that from our genealogical data as well as from our survey and ethnographic observations of the present-day descendants of this population, it appears that stopping after an optimal number of sons was a culturally optimal behavior for the residents of Xiaoshan until effective birth-control campaigns came about in the 1970s. Only at very high parities did the sterile fraction become 100 percent, and these were almost certainly cases where women had reached or were at least approaching menopause. It appears from all our data that there was a cultural logic and an accepted practice of stopping after an optimal number of children, but that not all couples followed this practice. As we said at the beginning of this article, we are splitting the difference between those, such as Wolf, who maintain that couples always strove for maximum fertility and others, like Lee, who maintains that wide spacing and early stopping were pervasive in the population. In this population, it appears that stopping was one option. Ideally, we would interview these lineage members and find out who stopped and why, but they are all long dead. Failing this, we still find that the statistical patterns emerging from the long duration of genealogical records are consistent with a culturally-accepted pattern of stopping that existed in the early 20th century.

For this research we have used a new set of data, new ethnographic observations, and a new statistical approach to address long standing debate on whether the Chinese consciously regulated their fertility. The statistical evidence we present is consistent with the idea that some couples engaged in both stopping and spacing behavior according to the number of healthy sons in a family. But we are still far from knowing who stopped or spaced and why.

ACKNOWLEDGMENTS

We thank Danyu Wang for helping initiate the project and establishing contacts in Xiaoshan, as well as for invaluable advice. Bill Lavely provided valuable comments and moral support. Wang Yaosheng and Hou Huili were instrumental in the survey work mentioned as an adjunct to this genealogical project. Zhou Yingying has helped with the data analysis. Data entry was done by Chou Jen-yu, Li Meng, Cai Yong, Zhang Li, Wang Enru, Li Juan and Zhou Yingying. This work was supported by NSF (BCS-0003918)

REFERENCES

Akaike, Hirotugu. 1973. Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, ed. Boris N. Petrov and Frigyes Csaki, pp. 268-281. Budapest: Hungarian Academy of Sciences.

Akaike, Hirotugu. 1992. Information theory and an extension of the maximum likelihood principle. In *Breakthroughs in Statistics*, ed. Samuel Kotz and Norman L. Johnson, pp. 610-624. New York: Springer Verlag.

Buck, John Lossing. 1937. *Land Utilization in China*. Nanjing: University of Nanking.

Burnham, Kenneth P., and David R. Anderson. 1998. *Model Selection and Inference: A Practical Information-Theoretic Approach.* New York: Springer Verlag.

Campbell, Cameron, Wang Feng, and James Z. Lee. 2002. Pretransitional Fertility in China. *Population and Development Review* 28, 4: 735-47.

Coale, Ansley J. 1985. Fertility in rural China: a confirmation of the Barclay reassessment. In *Family and Population in East Asian History,* ed. Susan B. Hanley and Arthur P. Wolf, pp. 186-95. Stanford: Stanford University Press.

Coale, Ansley J., and James T. Trussell. 1974. Model fertility schedules: Variations in the age structure of childbearing in human populations. *Population Index* 40(2):185-258.

David, Paul A., Thomas A Mroz. 1989a. Evidence of fertility regulation among rural French villagers, 1749-1789. A sequential econometric model of birth-spacing behavior (part 1). *European Journal of Population* 5(1):1-26.

David, Paul A., Thomas A. Mroz. (1989b) Evidence of fertility regulation among rural French villagers, 1749-1789. A sequential econometric model of birth-spacing behavior (part 2). *European Journal of Population* 5(2):173-206.

Gates, Hill. 1993. Cultural support for birth limitation among urban capital-owning women. In Deborah Davis and Stevan Harrell, eds., *Chinese Families in the Post-Mao Era*, pp. 251-74. Berkeley and Los Angeles: University of California Press.

Guinnane, Timothy W., Barbara S. Okun, James Trussell. 1994. What do we know about the timing of fertility transitions in Europe. *Demography* 31(1):1-20.

Han, Hua. n.d. Under the shadow of the collective good: An ethnographic analysis of deliberate fertility control prior to mandatory birth-planing programs in Xiaoxhan, Zhejiang. Submitted to *Modern China*.

Han, Hua, Wang Yaosheng, Zhou Yingying, and Stevan Harrell. n.d. The three-step fertility decline in Xiaoshan, Zhejiang. Paper in preparation.

Harrell, Stevan. 1995. Introduction: Microdemography and the modeling of population processes in late imperial China. In S Harrell (ed.) *Chinese Historical Microdemography*, Berkeley: University of California Press. pp. 1-20.

_____. n.d. Figuring age at marriage from Chinese genealogies. Unpublished paper: http://faculty.washington.edu/stevehar/ageatmarriage.html

Ho, Ping-ti. 1959. *Studies on the Population of China, 1368-1953.* Cambridge, MA: Harvard University Press.

Holman, Darryl J. 2003. *mle*: *A Programming Language for Building Likelihood Models.* Version 2.1 (Software and manual). http://faculty.washington.edu/~djholman/mle/

Lee, James, and Cameron Campbell. 1997. *Fate and Fortune in Rural China.* Cambridge: Cambridge University Press.

Lee, James Z, Cameron Campbell, and Wang Feng. 1993. The last emperors: An introduction to the demography of the Qing (1644-1911) imperial lineage. In Divid S. Richer and Roger Schofield, eds., *Old and New Methods in Historical Demography*. Oxford: Oxford University Press, pp. 361-82.

Lee, James, and Wang Feng. 1999. *One Quarter of Humanity: Malthusian Mythology and Chinese Realities*. Cambridge, MA: Harvard University Press.

Leridon, Henri. 1977. *Human Fertility: The Basic Components*. Chicago: University of Chicago Press.

Liu Tsui-jung. 1985. The demography of two Chinese clans in Hsiao-shan, Chekiang, 1650-1850. In Susan B. Hanley and Arthur P. Wolf , eds., *Family and Population in East Asian History*, pp. 13-61. Stanford: Stanford University Press.

______. 1995. A comparison of lineage populations in South China, ca. 1300-1900. In S Harrell (ed.) *Chinese Historical Microdemography*, Berkeley: University of California Press. pp. 94-121.

Malthus, Thomas Robert. 1798/1992. *An Essay on the Principle of Population, Second Edition* ed. Donald Winch. Cambridge: Cambridge University Press.

Nelson, Wayne. 1982. *Applied Life Data Analysis.* New York: John Wiley and Sons.

Schoppa, R. Keith. 1989. *Xiang Lake—Nine Centuries of Chinese Life*. New Haven: Yale University Press.

Sheps, Mindel C., Jane A. Menken. 1973. *Mathematical Models of Conception and Birth*. Chicago: University of Chicago Press.

Jeremy M. G. Taylor. 1995. Semi-parametric estimation in failure time mixture models. *Biometrics* **51**:899-907.

Telford, Ted A. 1990. Patching the holes in Chinese Genealogies. *Late Imperial China* 11, 2: 116-35.

______1995. Fertility and population growth in the lineages of Tongcheng County, 1520-1661. In S Harrell (ed.) *Chinese Historical Microdemography*, Berkeley: University of California Press. pp. 48-93.

Telford, Ted A., Melvin Thatcher, and Basil P.N. Yang. 1983. *Chinese Genealogies at the Geneaological Society of Utah: An Annotated Bibliography xi.* Taipei: Ch'eng-wen Publishing Company.

Van Bavel, Jan. 2001. Family control, bridal pregnancy, and illegitimacy. An event history analysis in Leuven Belgium, 1846-1856. *Social Science History* 25(3):451-482.

Wolf , Arthur P. 1985. Fertility in prerevolutionary rural China. In *Family and Population in East Asian History,* ed. Susan B. Hanley and Arthur P. Wolf, pp. 154-85. Stanford: Stanford University Press.

______. 2001. Is there evidence of birth control in late imperial China? *Population and*

Development Review **27**(1):133-54.

Wolf, Arthur P., and Chieh-shan Huang. 1980. *Marriage and Adoption in China, 1845-1945.* Stanford: Stanford University Press.

Wood, James W. 1994. *Dynamics of Human Reproduction: Biology, Biometry, Demography*. Hawthorne, NY: Aldine de Gruyter.

Wood, James W., Darryl J. Holman, Anatoli Yashin, Raymond J. Peterson, Maxine Weinstein, and Ming-Cheng Chang. 1994. A multistate model of fecundability and sterility. *Demography* **31**(3):403-426.

Zhao Zhongwei. 1997. Deliberate birth control under a high-fertility regime: Reproductive behavior in China before 1970. *Population and Development Review* 23: 729-768.

______. 2002. Fertility control in China's past. *Population and Development Review* 28, 4: 751-57.

Table 1. Number of records and dates in the five genealogies.

Family name	Earliest date	Latest date	Men	Women	Total
Shi	1256	1888	1,995	2,230	4,225
Lin	1425	1897	457	560	1,017
Tian	1420	1848	538	675	1,213
Wu	1529	1908	1,097	1,374	2,471
Han	607 (1497)[a]	1892	17,477	19,785	37,262
Total			21,564	24,624	46,188

[a] First year a wife's age was recorded

Table 2. Number of birth intervals by parity and age (years) for sons that survived to age 16 sui.

Parity	15 to 20	20 to 25	25 to 30	30 to 35	Total
1	460	1581	1294	835	4170
2	29	390	849	834	2102
3	12	140	406	569	1127
4	1	18	118	265	402
5	0	1	18	83	102
6	0	0	1	19	20
Total	502	2130	2686	2605	7923

Table 3. Parameter estimates for four age groups.

	Age 15 to 20		Age 20 to 25		Age 25 to 30		Age 30 to 35	
Parameters[a]	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
p	0.566	0.121	1.021	0.066	1.726	0.072	1.184	0.097
β_{p2}	1.845	0.800	2.750	0.416	—	—	4.048	0.671
β_{p3}	-1.759[c]	0.675[c]	-2.176	0.215	-1.691	0.129	-0.776	0.133
β_{p4}	—	—	-3.148[d]	0.783[d]	-2.705	0.225	-1.593	0.164
β_{p5+}	—	—	—	—	-4.571	1.032	-2.124	0.245
a	6.710	0.245	4.487	0.128	1.050	0.041	0.481	0.024
b	18.148	1.423	12.99	0.557	4.641	0.212	3.758	0.165
β_{h2}	-2.023[b]	0.187[b]	-1.462	0.050	-0.426[b]	0.047[b]	—	—
β_{h3}	—	—	-1.374[c]	0.199[c]	—	—	0.550	0.073
β_{h4}	—	—	—	—	—	—	1.039	0.129
β_{h5+}	—	—	—	—	—	—	1.466	0.222

[a] Parameters: p is a logistic parameter that determines the baseline fraction of "sterile" individuals; $\beta_{p2}, \ldots, \beta_{p5+}$ are the parameter estimates of parity 2, …, 5+ on the fraction of "sterile" individuals; a is the location parameter of a gamma distribution; b is the shape parameter for the gamma distribution; $\beta_{h2}, \ldots, \beta_{h5+}$ are the parameter estimates for the effect of being parity 2..5+ on the hazard distribution until birth of the next surviving son. Parity denotes the number of previous sons surviving to age 16 *sui*.

[b] Includes parities 2 and higher

[c] Includes parities 3 and higher

[d] Includes parities 3 and higher

Table 4 Fraction "sterile", mean, and standard deviation of surviving-son birth intervals by age and parity at the start of the birth interval. Parity denotes the number of previous sons surviving to age 16 sui.

	15-20			20-25			25-30			30-35		
Parity	% sterile	Mean	SD	% sterile	Mean	SD	% sterile	Mean	SD	% sterile	Mean	SD
1	36.2	2.71	1.65	26.5	2.90	0.95	15.11	4.42	2.05	23.4	7.59	3.73
2	8.24	4.97	2.23[a]	2.25	4.64	1.86	—	5.40[a]	2.33[a]	0.53	—	—
3	76.3[b]	—	—	76.0	4.49[b]	1.75[b]	49.1	—	—	39.9	6.00	2.83
4	—	—	—	89.4[c]	—	—	72.7[c]	—	—	60.1	4.83	2.15
5+	—	—	—	—	—	—	—	—	—	71.9	4.05	1.71

[a]Includes parities 2 and higher.
[b]Includes parities 3 and higher.
[c]Includes parities 4 and higher.

FFigure 1. Survival distribution for intervals to birth of the next son that survives to age 16 *sui* for mother's ages 15 to 20. Parity denotes the number of previous sons surviving to age 16 *sui*. The curves are gamma distributions are based on parameter estimates in Table 3.

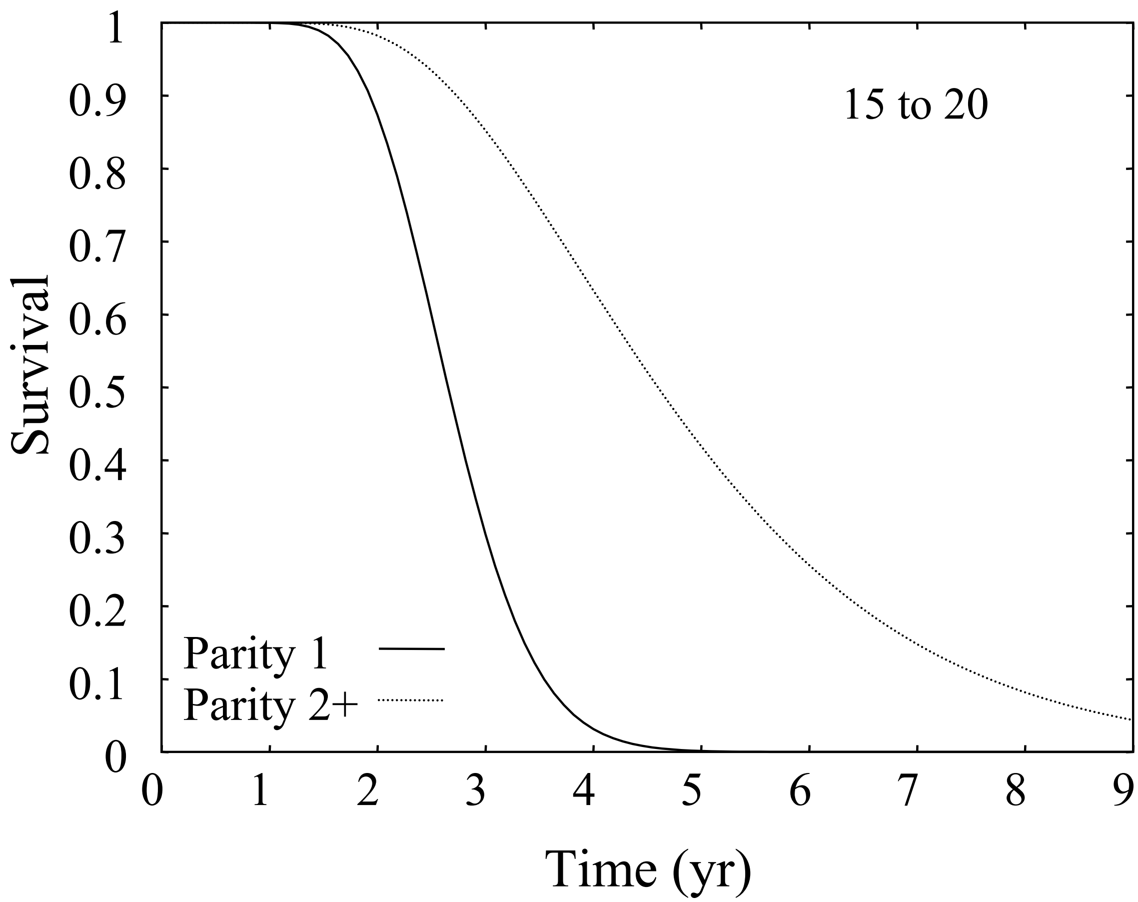

Figure 2. Survival distribution for interval to the birth of the next son that survives to age 16 *sui* for mother's ages 20 to 25. Parity denotes the number of previous sons surviving to age 16 *sui*. The curves are gamma distributions are based on parameter estimates in Table 3.

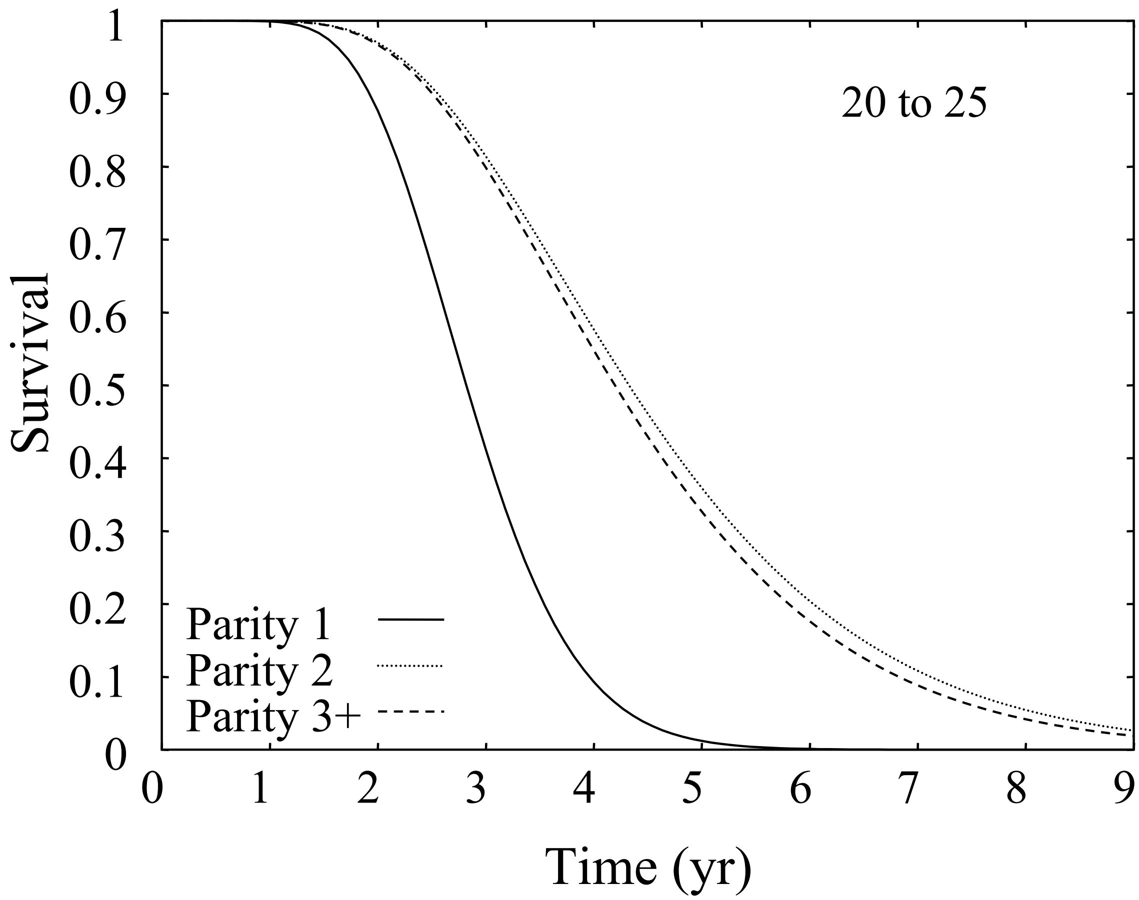

Figure 3. Survival distribution for interval to the birth of the next son that survives to age 16 *sui* for mother's ages 25 to 30. Parity denotes the number of previous sons surviving to age 16 *sui*. The curves are gamma distributions are based on parameter estimates in Table 3.

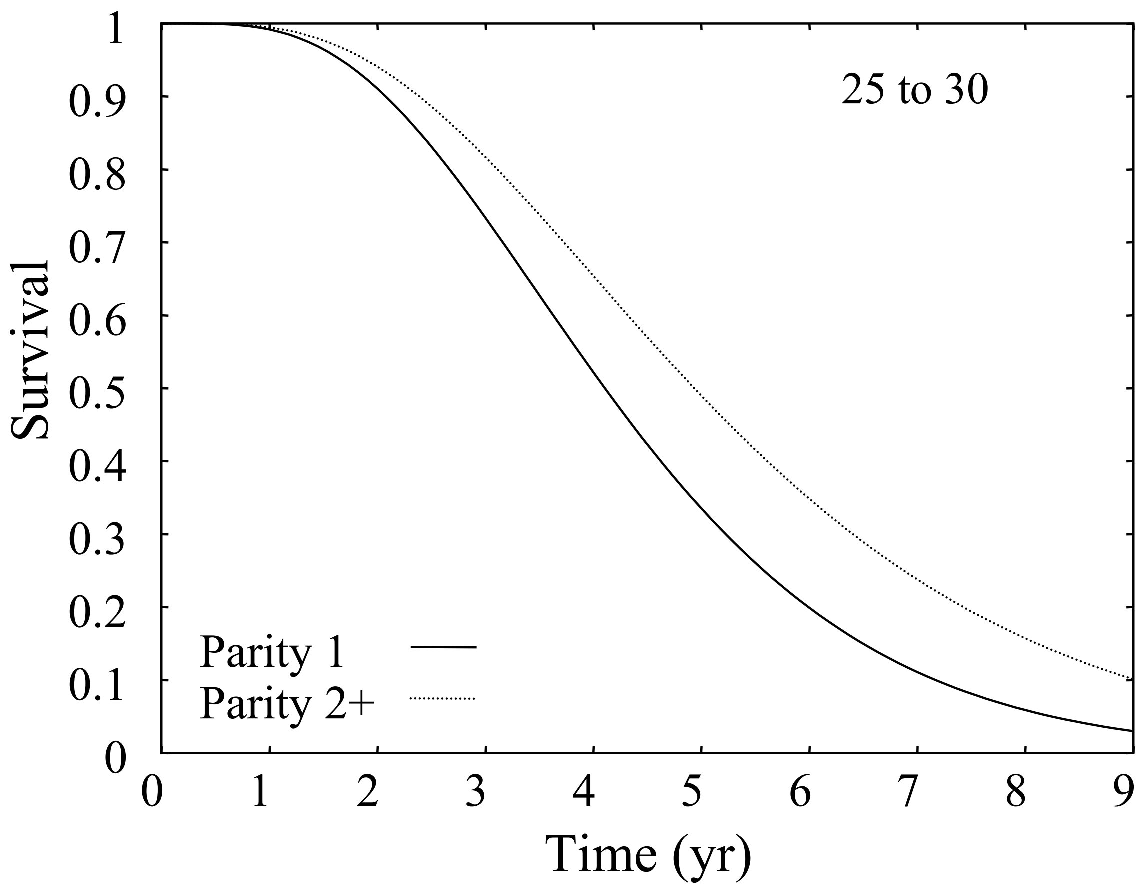

Figure 4. Survival distribution for interval to the birth of the next son that survives to age 16 *sui* for mother's ages 30 to 35. Parity denotes the number of previous sons surviving to age 16 *sui*. The curves are gamma distributions are based on parameter estimates in Table 3.