Population Review

Volume 42, Numbers 1-2, 2003

Type: Article pp. 1-22

Controlling for Postpartum Amenorrhea and Heterogeneity in the Analysis of Fecundability Using Birth Interval Data: A Simulation Study with Application to Hutterite Reproductive Histories

Authors: Ulla Larsen, Ph.D., Sharon Yan, Ph.D., Anatoli Yashin, Ph.D.
Affiliations: Department of Population and International Health, Harvard School of Public Health (Dr. Larsen); Center for Drug Evaluation and Research, Food and Drug Administration, USA (Dr. Sharon Yan); Max Planck Institute for Demographic Research (Dr. Anatoli Yashin)
Corresponding author/address: Dr. Ulla Larsen, Department of Population and International Health, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115 (ularsen@hsph.harvard.edu)

Abstract

Despite numerous studies, the level and age pattern of fecundability are still under discussion in the demographic, biologic, and medical literature. Previous analyses, however, have not fully taken into account two factors affecting estimated fecundability: postpartum amenorrhea and variations in fecundability among women. This paper used simulation analysis to calibrate a new fecundability model, and it showed that fecundability by age, heterogeneity in fecundability, and postpartum amenorrhea by age could be estimated simultaneously from all birth intervals. An empirical analysis of Hutterite birth histories showed the following: that fecundability declined almost linearly from age 20 to age 40; that the fecundability of a 35 year old was half the level of fecundability of a 25 year old woman; and that fecundability varied significantly across women. For an average Hutterite woman, the waiting time to next live-birth conception was 13.1 months at age 30 and 22.2 months at age 40. These intervals included 6.5 months of postpartum amenorrhea.

Keywords

Fecundability, postpartum amenorrhea, fertility, simulation analysis, heterogeneity, Hutterite

*This work was done at the Harvard School of Public Health. Ulla Larsen was the principal investigator of the project. She directed the work and wrote the paper. Sharon Yan did the estimation of the models. Anatoli Yashin developed the theoretical models. The research was supported by the National Institute of Child Health and Human Development, Grant R01-HD34354. A preliminary version of this paper was presented at the Annual Meeting of the Population Association of America in Washington D.C., March 1997. We appreciate the useful comments and suggestions from Christopher Sinton. The views expressed in this paper are those of the authors and do not necessarily represent those of the Food and Drug Administration.

© 2003 Population Review Publications

INTRODUCTION

It is difficult to estimate the level and age pattern of fecundability, and it is no surprise that accurate information about the age schedule of fecundability is lacking (Menken 1985; Menken, Trussell and Larsen 1986; Golden and Millman 1988). Using the first birth interval in noncontracepting populations, Heckman and Walker (1990) and Wood et al. (1994) developed new models of effective fecundability. Because most couples in noncontracepting populations have their first child before age 30, however, models based on these populations cannot be used to estimate fecundability at older ages. Researchers interested in fecundability in older individuals - that is, in constructing fecundability models that include second and higher order birth intervals - face a particular problem: accounting for postpartum amenorrhea. To circumvent this problem, Larsen and Vaupel (1993) set the postpartum infecund period to a fixed constant after each birth, in spite of the fact that the duration of postpartum amenorrhea was not the same for all women and in all birth intervals. As a result, they claimed that their model matched the age pattern, but not the level of effective fecundability. D'Souza (1974) proposed a promising approach to assess the age pattern of effective fecundability by using a convolution model, in which the duration of postpartum amenorrhea was specified by a normal distribution and effective fecundability was specified by a Gompertz distribution. Two problems, however, appeared in this approach. First, D'Souza confined the analysis to couples who had at least four children and who were observed at least until their 50th birthday: hence the estimated age schedule of effective fecundability was probably overstated. Second, D'Souza assumed that there was no variation in effective fecundability across women, despite physiological evidence suggesting that fecundability varies widely (e.g., Ellison 1990; Ellison et al. 1993; Ellison, Peacock and Lager 1989). If variation in effective fecundability or unobserved heterogeneity were ignored, estimates of mean fecundability were biased downward because those women with high fecundability conceived; thus, with each additional month, the sample was increasingly selected for women with low fecundability. Yashin et al. (1998) modified and extended D'Souza's model. They proposed a new convolution model of fecundability, one based on all birth intervals. This model described the effects of postpartum infecundity on fecundability and took into account unobserved heterogeneity in fecundability; under very simplistic assumptions, the model performed well. The present study aimed at evaluating this model under more realistic, albeit more complicated, assumptions.

Fecundability is defined as the probability of conceiving during any given menstrual cycle for a sexually active woman who is not pregnant, postpartum infecund, or contracepting (Pressat and Wilson 1985). In our analysis, only live-birth conceptions have been considered and effective fecundability is estimated. To simplify presentation, the term fecundability is used throughout the paper to denote effective fecundability.

One aim of the present analysis was to evaluate the effects of unobserved or partially observed physiological processes on estimates of fecundability, and to calibrate models that took into account these unobserved processes. A prime example of such a process is the transition time from the postpartum infecund state to the fecund state. In standard reproductive histories, these transition times were usually either unavailable or incomplete. The Demographic and Health Surveys, for example, which were conducted in more than eighty developing countries, included information about the duration of postpartum amenorrhea only for children born in the last five to six years before the survey date (Institute for Resource Development, Macro International, Inc., 2001). Given this lack of data, therefore, mathematical models are needed to

evaluate the effects of postpartum amenorrhea on estimates of fecundability. It is well known that the monthly proportion of conceptions is lower in a sample of women with prolonged postpartum amenorrhea than in a sample of women with shorter durations, because during amenorrhea a woman is non-susceptible. However, it is not known how these variations in duration of amenorrhea influence estimates of fecundability. It is also well known that estimates of fecundability are downward biased in studies where heterogeneity is not taken into account (Sheps and Menken 1973). It is less well known, however, how much the baseline level of fecundability varies across women in noncontracepting populations, and how to account for unobserved variability in fecundability (Heckman and Walker 1987; Larsen and Vaupel 1993). A woman planning her childbearing needs to know not only the typical number of months it takes to conceive a child, but also how this number varies among women. Demographers who are concerned about the proximate determinants of fertility also need this information (Menken 1985).

A microsimulation model was used to evaluate the effects that different parametric specifications of unobserved reproductive characteristics had on estimates of fecundability. A wide range of reproductive characteristics, which were likely to encompass all the variation in human populations, could be postulated in this model. Therefore, the findings of this analysis should facilitate parametric specifications of fecundability models in subsequent empirical analyses. In the simulated data, we knew all the true parameters and, therefore, we could assess how changes in one reproductive characteristic affected estimated fecundability. The sensitivity analysis conducted in this paper attempted the following: 1) to find model specifications that were sufficiently flexible to capture fecundability by age, fecundability across women (i.e., unobserved heterogeneity), the average duration of postpartum amenorrhea, and variations in the duration of postpartum amenorrhea by age; and 2) to ascertain the effects of sample size and sampling variation on parameter estimates.

A second aim here was to apply the proposed model in an empirical analysis of a sample of Hutterite women's reproductive histories. The Hutterites provided an ideal basis for a baseline fecundability schedule because they appeared to be a natural-fertility population, with no evidence of deliberate fertility control (Sheps 1965). In addition, the Hutterites were a homogeneous population, and they displayed no variation in fertility according to education, occupation, income, or social status (Eaton and Mayer 1953). They were known to live a healthy lifestyle, e.g., their belief system forbade them to smoke tobacco or drink alcohol, and there were few incidences of disease that could lead to impaired fertility of pathological origin. Thus, documented variations in unobserved heterogeneity could be ascribed largely to biological factors. The Hutterites were used as a standard in numerous demographic studies, such as the Princeton European Fertility Project (Coale and Watkins 1986) and in Coale and Trussell's model of marital fertility (Coale and Trussell 1978). They have also been used to develop fecundability models (D'Souza 1974; Heckman and Walker 1987, 1990; Larsen and Vaupel 1993).

The present analysis extended previous work on the measurement of fecundability (see, for example, D'Souza 1974; Heckman and Walker 1987, 1990; Larsen and Vaupel 1993; Wood et al. 1994), and also provided additional substantive information about fecundability: 1) the new model made it possible to estimate the level and the age pattern of fecundability by controlling for postpartum amenorrhea and unobserved heterogeneity; 2) the model applied to first as well as to higher order birth intervals; 3) the model was calibrated in a simulation analysis prior to its application to empirical data; 4) the study evaluated the effects on estimates of fecundability by

confining analysis to closed birth intervals versus employing both closed and open birth intervals; 5) the study situated its findings about the level and age pattern of Hutterite fecundability in relation to previous findings by D'Souza (1974), Heckman and Walker (1987, 1990) and Larsen and Vaupel (1993), so assessing and illustrating the new knowledge obtained since these earlier studies. No previous study of the level and age pattern of fecundability controlled for postpartum amenorrhea and unobserved heterogeneity while at the same time providing estimates that fitted the data and encompassed all birth intervals.

FECUNDABILITY MODEL

We propose a classic heterogeneity model, based on assuming women's age trajectory of fecundability follows the pattern described by the hazard of a live-birth conception at a given age, and the level of fecundability at a given age differs across women. A woman's level of fecundability at a given age is considered an unobserved random variable (called frailty). The distribution of this random variable describes the heterogeneity in fecundability across women. The duration of postpartum amenorrhea is also considered an unobserved random variable.

Let the hazard of a woman's live birth conception at age x be denoted by $z \cdot h(x)$, where the frailty variable z measures the persistent difference, or heterogeneity, in fecundability across women and $h(x)$ describes how fecundability changes with a woman's age. For a given woman with frailty z, the probability density function of the first live-birth conception at age x can be represented as

$$f_1(x, X_0) = z \cdot h(x) \cdot exp\left(-\int_{X_0}^{x} z \cdot h(u)\, du\right), \qquad (1)$$

where X_0 is the age at marriage.

The waiting time for the i^{th} ($I > 1$) live-birth conception starts at the end of the postpartum amenorrhea period following the previous birth. Let τ be the duration of postpartum amenorrhea and $g_i(\tau)$ be the density of the duration of postpartum amenorrhea after the $i\text{-}1^{th}$ live-birth. Then, the probability density function of the i^{th} live-birth conception is

$$f_i(x, T_{i-1}) = z \cdot h(x) \cdot \int_0^{x - T_{i-1}} \exp\left(-\int_{T_{i-1}+\tau}^{x} z \cdot h(u)du\right) g_{i-1}(\tau)d\tau, \quad i = 2, \ldots, I, \qquad (2)$$

where T_i is age at live-birth i and I is the last observed live-birth.

The conditional likelihood function for one woman with reproductive history $(X_0, T_1, T_2, \ldots, T_I, x)$ given her frailty z is

$$L(Y_1, X_0; Y_2, T_1; \ldots ; Y_I, T_{I-1}; x, T_I) = f_1(Y_1, X_0) \prod_{i=2}^{I} f_i(Y_i, T_{i-1}) S(x), \qquad (3)$$

where $Y_i = T_i - G$ is the age at i^{th} live-birth conception, G is the duration of gestation, x is the censoring time, and $S(x) = \int_0^{x-T_I} \exp\left(-\int_{T_I+\tau}^{x} z \cdot h(u)du\right) g_I(\tau)d\tau$ represents that since her

I^{th} live birth the woman has not yet conceived at the time of censoring. The effect of sterility was modeled in a separate paper (Yan and Larsen 2001).

The advantage of this model is the opportunity it affords to use various parametric forms of the duration of postpartum amenorrhea, the hazard rate of a live-birth conception, and heterogeneity in fecundability to calculate the explicit form of the likelihood function.

Women's age at marriage and at all births are recorded to the day, so we treat the ages as exact ages on a continuous time scale. Ages are measured in months and fractions of months.

In the microsimulation and estimation, continuous time models of hazards are used. The fecundability, i.e., the monthly probability of live birth conception *q(x),* is given by

$$q(x)=1-exp\left(-\int_{x}^{x+1} h(u)\,du\right). \tag{4}$$

When fecundability is low, the two measures are very close.

MICROSIMULATION OF REPRODUCTIVE HISTORIES

The microsimulation modeling for the analysis of reproductive histories was first suggested by Barrett (1969, 1971), and then later revised by Trussell and Wilson (1985) and Larsen and Menken (1989). In this paper, these ideas are developed so as to model reproductive histories with unobserved heterogeneity in fecundability and hidden postpartum amenorrhea following a live birth. For generating birth histories, each woman was assigned particular reproductive characteriistics, such as fecundability, duration of postpartum amenorrhea, and age at onset of sterility. We used lunar months, i.e., 13 months per year, to approximate the menstrual cycle. Events were generated up to the date at censoring. In the first group of models, complete birth histories were generated, and all women were censored at age 50. In a second group, incomplete birth histories were generated, and a censoring date was randomly selected from the age distribution in a model life table. Table 1 lists the specific distributions used in simulating reproductive histories. Age at marriage was generated following a uniform distribution.
If censoring preceded marriage, the woman was excluded from analysis. Similarly, age at sterility was determined, and if sterility preceded marriage or censoring, the woman was excluded from analysis. Sterility was set at age 50 in one group of models. In an alternative model, sterility followed a Gompertz distribution, with a mean age of sterility of 42.3 (Pittenger 1973). Next, fecundity was chosen. Each woman had the same level of fecundability up to age 30; after age 30, fecundability declined linearly to zero at age 50. If onset of sterility preceded censoring, then fecundability of the woman dropped to zero at the month of onset of sterility, and she experienced no further reproductive events up to the month of censoring. In one group of models, fecundability was homogeneous, and in another fecundability was heterogeneous: in the latter cases the level of fecundability for each woman followed either a gamma or a two-point distribution. A random number from a uniform distribution from zero to one was generated for each month: if it did not exceed her fecundability for that month, she conceived. Otherwise, her age increased by one month and the process was repeated. Following a live birth there was a period of amenorrhea during which the woman was nonsusceptible. Thus, before a woman could conceive again, she aged the length of pregnancy plus the duration of postpartum amenorrhea. The duration of gestation was set to 10 lunar months for all pregnancies. In some models the

duration of postpartum amenorrhea was constant, in others, postpartum amenorrhea followed a gamma, normal, or negative binomial distribution, and in other models the duration of postpartum amenorrhea varied by age.

The simulation of reproductive histories continued as described above up to the date of censoring or sterility. Women who had not conceived by the age of censoring or sterility were excluded from analysis. In this model, no fetal deaths or multiple births occurred, and only reproductive characteristics of women were generated. Furthermore, all marriages were intact and all women were exposed to pregnancy from the date of marriage to the date of censoring or sterility, except during pregnancy or periods of postpartum amenorrhea.

Table 1. Assumptions tested in the sensitivity analysis

Variable/Assumptions	Model	
	Base	Alternative
Age at marriage		
Uniform distribution[a]	20-27	20-21, 20-23, 20-25, 20-29
Postpartum amenorrhea across women (mean in months)		
Constant	3	0 and 6
Gamma[b]		3, 6, 9, and 12
Normal[c]		3, 6, 9, and 12
Negative binomial[d]		3 and 6
Postpartum amenorrhea within a woman (mean)		
Piecewise linear[e]	3	
Fecundity within a woman (mean)		
Piecewise linear[f]	.25	
Fecundity across women		
Homogeneous (mean)	.25	
Heterogeneous (mean frailty)		
2-point[g]		1
Gamma[h]		1
Sterility (mean)		
Constant	50	
Gompertz[i]		42.3
Age at censoring		
Constant	50	
Model West[j]		Level 13

[a] The uniform distribution for the age interval 20-21 is drawn from the lunar months of (260-286).

[b] Gamma density: $g(\alpha, \beta, c) = \frac{\beta^{\alpha}}{\tau(\alpha)}(x-c)^{a-1}e^{-\beta(x-c)}$, x > c.

[c] Normal density: $\phi(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}}e^{-(x-\mu)^2/2\sigma^2}$.

[d] Negative binomial density: $h(r, p) = \binom{r+x-1}{x} p^a q^x$.

[e] Postpartum amenorrhea remains constant at 3 months for age $\leq$ 30 and constant at 5 months at age > 30.

[f] Fecundability remains constant at .25 from age at marriage to age 30, then declines linearly to zero at age 50.

[g] Probability distribution of individual frailty: $P(Z = z_1) = p_1$, $P(Z = z_2) = 1 - p_1$, and $p_1 z_1 + (1-p_1) z_2 = 1$.

[h] Individual frailty follows Gamma distribution with mean 1.

[i] Sterility density: $s(a, b, c) = ae^{b(x-c)}e^{-\frac{a}{b}(e^{b(x-c)}-1)}$ (Pittinger 1973).

[j] Coale and Demeny (1983).

To minimize random errors, large samples of 5,000 couples were generated. From the simulated reproductive histories, the sensitivity of estimated fecundability to variations in reproductive characteristics was examined. In each reproductive model, estimated fecundability was obtained from a sample of 5,000 women and 50 replications.

For each woman, the age pattern of the hazard of conception was modeled by a piecewise linear function, and support points for the estimated monthly hazard of conception were obtained at ages 20, 25, 30, 35, 40, and 45. We chose a flexible piecewise linear function because the age pattern of fecundability is not well known (Larsen and Vaupel1993).

RESULTS FROM SIMULATION STUDY

In this analysis, the effects of different assumptions on parameter estimates about a given reproductive characteristic or variable were evaluated. For example, the bias that occurred in estimates of fecundability whenever a homogeneous model was applied to a heterogeneous sample were assessed. The variables analyzed included heterogeneity in fecundability, postpartum amenorrhea, age at onset of sterility, and age of censoring; models with different assumptions about each of these variables were evaluated, and the estimated parameters produced by the models were then compared to the true parameters. Finally, the effects of sample size and of sampling variation were evaluated.

The estimation procedure used the quasi-Newton algorithm supported by the Fortran IMSL Library (Microsoft Corporation 1995).

Effect of Heterogeneity in Fecundability

In order to evaluate the effects of heterogeneity in fecundability on estimated fecundability, samples of heterogeneity were generated by multiplying the base fecundability with a frailty variable *Z*, which had a two-point distribution, and a gamma distribution with mean one. For instance, in the case of two-point frailty with $p_1 = .2$, $z_1 = .6$, 20 percent of the women in the population had fecundability of .6 × (base fecundability), and the other 80 percent of women had fecundability of z_2 × (base fecundability), such that $p_1 z_1 + (1-p_1) z_2 = 1$.

To capture heterogeneity in the sample, a two point and a gamma distribution were used in the model. A homogeneous model was also used to evaluate the bias in fecundability estimates. The analysis documented that heterogeneity generated by a two-point distribution was estimated well both by models that included a one-parameter gamma $(\alpha = \beta, c = 0)$ and by those that included a two-point distribution (Table 2 Panel a). Estimates using a two-point distribution of frailty resembled true fecundability, and this was the case not only for models with two equal sized subgroups of fecundability, but also for models with a large and a small subgroup (an 80 and a 20 percent split, for example). The proportion and the fecundability in each subgroup were also found to be close to the true values.

A one-parameter gamma model captured heterogeneity very well when heterogeneity was generated by a one-parameter gamma distribution. In contrast, estimated fecundability was significantly lower than true fecundability at all ages when it was modeled by a two-point distribution. This bias followed from the fact that the two-point distribution could not capture the range of heterogeneity in fecundability generated by a gamma distribution (Table 2 Panel b). In the case where fecundability was generated by a gamma distribution with three parameters,

Table 2. Estimated monthly hazard of live-birth conception in models with heterogeneous fecundability[1]

True frailty		Age	True	Estimated		
Two-point (p_1, z_1)				Homogeneous	Two-point	Gamma
Panel a:						
$p_1 = .2$, $z_1 = .6$		20	.250	.242	.243	.249
		25	.250	.242	.247	.249
		30	.250	.240	.248	.247
		35	.188	.180	.185	.186
		40	.125	.123	.125	.125
		45	.063	.062	.062	.063
	p_1				.187	
	z_1				.604	
	α					25.000
	Log (likelihood)			-26336.5	-26275.3*	-26274.5*
$p_1 = .5$, $z_1 = .6$		20	.250	.223	.246	.247
		25	.250	.224	.248	.250
		30	.250	.221	.249	.247
		35	.188	.168	.186	.186
		40	.125	.115	.125	.125
		45	.063	.059	.062	.062
	p_1				.499	
	z_1				.603	
	α					6.249
	Log (likelihood)			-26275.7	-25803.6*	-25878.3*
$p_1 = .8$, $z_1 = .8$		20	.250	.232	.249	.245
		25	.250	.230	.248	.248
		30	.250	.228	.250	.250
		35	.188	.173	.186	.187
		40	.125	.117	.125	.125
		45	.063	.059	.062	.063
	p_1				.798	
	z_1				.801	
	α					6.249
	Log (likelihood)			-26305.9	-26056.3*	-26165.5*
Panel b:						
$\alpha = 2$, $\beta = 2$, $c = 0$		20	.250	.186	.221	.245
		25	.250	.179	.219	.245
		30	.250	.181	.220	.247
		35	.188	.144	.168	.186
		40	.125	.105	.118	.126
		45	.063	.056	.059	.063
	p_1				.481	
	z_1				.460	
	α					2.000
	Log (likelihood)			-25747.4	-24583.9*	-24386.5*
$\alpha = 4$, $\beta = 4$, $c = 0$		20	.250	.212	.232	.247
		25	.250	.213	.237	.250
		30	.250	.210	.234	.247
		35	.188	.160	.177	.186
		40	.125	.113	.122	.125
		45	.063	.058	.061	.062
	p_1				.463	
	z_1				.576	
	α					4.004
	Log (likelihood)			-26166.9	-25672.7*	-25533.6*
$\alpha = 1$, $\beta = .1$, $c = .15$		20	.250	.228	.244	.248
		25	.250	.229	.244	.246
		30	.250	.230	.245	.246
		35	.188	.172	.184	.185
		40	.125	.118	.124	.124
		45	.063	.060	.062	.062
	p_1				.709	
	z_1				.790	
	α					6.250
	Log (likelihood)			-26311.7	-26098.2*	-26135.4*

[1] Standard errors are $\leq$ .003.
[*] .001 level of significance based on likelihood ratio test.
Two point compared to homogeneous model has 2 degrees of freedom.
Gamma compared to homogeneous model has 1 degree of freedom.

estimated fecundability from the two-point and the gamma model were almost identical, and both sets of estimated parameters were close to the true parameters (Table 2 Panel b).

In summary, a wide range of heterogeneous fecundability was accurately estimated by both a one-parameter gamma and by a two-point distribution. Heterogeneous fecundability was consistently approximated more closely by a heterogeneous model than by a homogeneous model, even when the wrong model of heterogeneity was used. Each of the heterogeneous models fitted the data significantly better than did the homogeneous models based on the likelihood ratio test. Furthermore, we found that a one-parameter gamma performed better than a two-point distribution because: 1) estimated fecundability from a one-parameter gamma was closer to true fecundability, and 2) in a one-parameter gamma, only one parameter needed to be estimated, whereas a two-point distribution required two estimated parameters.

Effect of Postpartum Amenorrhea

Subsequent analysis examined the sensitivity of various parametric models to different distributions and different mean durations of postpartum amenorrhea. First, the duration of postpartum amenorrhea was modeled by a single parameter distribution. We used a normal distribution following D'Souza's (1974) seminal work. Because of its flexibility, a gamma distribution was also used. Samples of postpartum amenorrhea were generated by a constant and from a normal, a gamma, and a negative binomial distribution. Table 3 presents the estimated parameters from a normal and from a gamma model of postpartum amenorrhea, and the sample mean duration of postpartum amenorrhea was 3 months.

Table 3. Estimated monthly hazard of live-birth conception and postpartum amenorrhea by different distributions of postpartum amenorrhea[1]

Distributions of P.A.[2]		P.A.	Age					
Sample	Model	(months)	20	25	30	35	40	45
True parameters		3.0	.250	.250	.250	.188	.125	.062
Constant	Normal	3.3	.247	.266	.272	.199	.132	.062
	Gamma	3.7	.246	.278	.292	.212	.138	.064
Normal	Normal	3.0	.253	.248	.249	.186	.125	.061
	Gamma	3.4	.255	.263	.271	.199	.132	.062
Gamma	Normal	2.6	.239	.230	.229	.174	.120	.060
	Gamma	3.1	.243	.246	.249	.187	.126	.061
Negative binomial	Normal	2.3	.237	.219	.213	.167	.116	.059
	Gamma	2.7	.241	.232	.229	.177	.120	.060

[1] Standard errors are $\leq$ .003.
[2] P.A. is postpartum amenorrhea.

As expected, the estimated parameters were very close to the true parameters in the models where postpartum amenorrhea was geenerated and estimated by the same distribution, either a

normal or a gamma. However, neither the normal nor the gamma model could fully describe the wide range of postpartum amenorrhea generated by different distributions. In the one-parameter normal model, the standard deviation was set to be one. The normal model could not capture durations of postpartum amenorrhea beyond three standard deviations of the mean. The gamma and negative binomial distributions had variances of 3 and 6, respectively. Thus, the long durations of postpartum amenorrhea in the samples generated by the gamma and negative binomial distributions could not be taken into account in the normal model. Consequently, the duration of postpartum amenorrhea was underestimated. In contrast, when the true standard deviation was less than one, as in the case of the constant model, the duration of postpartum amenorrhea was overestimated.

Similar results were obtained when postpartum amenorrhea was modeled by a gamma distribution. In the gamma model, the scale parameter was set at one, so that the variance equaled the mean, which was the parameter to be estimated. We found that the duration of postpartum amenorrhea was underestimated when the true variance was larger than the mean, as in the sample generated by a negative binomial distribution; conversely, it was overestimated when the true variance was less than the mean, as in the samples generated by a constant or a normal distribution. Furthermore, bias in the estimate of postpartum amenorrhea would in turn bias the estimate of fecundability: fecundability would be underestimated if postpartum amenorrhea was underestimated, and overestimated if postpartum amenorrhea was overestimated. This was because, for a given number of children a woman had at a given age, the shorter the estimated duration of postpartum amenorrhea, the longer the estimated waiting time to next conception, i.e., the lower the estimated fecundability. The difference of the estimated postpartum amenorrhea from the normal model and the gamma model was about 0.35 month across the range of different samples analyzed. This discrepancy was because the variance of postpartum amenorrhea in the normal model was less than the variance in the gamma model. If the difference in standard deviations was taken into account, the normal model and the gamma model performed equally well. In conclusion, a misspecification of the variance of postpartum amenorrhea in the one-parameter model would cause bias in the estimated parameters of postpartum amenorrhea and fecundability by age.

Because of the finding suggesting the importance of the variance of postpartum amenorrhea, two-parameter models of postpartum amenorrhea were examined. However, neither a two-parameter gamma nor a two-parameter normal model could pick up the mean and the variance of postpartum amenorrhea, even in samples where postpartum amenorrhea was generated and estimated by the same distribution (results not shown). We conjectured that the two-parameter models could not distinguish the variance of postpartum amenorrhea from the variance in waiting time. As a consequence, when the model failed to identify the variance, it also failed to identify the mean duration of postpartum amenorrhea.

The subsequent analysis addressed whether one-parameter models of postpartum amenorrhea (e.g., a normal or a gamma model) could capture postpartum amenorrhea generated by a wide range of different distributions when the true variance of postpartum amenorrhea was specified in the model. A duration of 6 months of postpartum amenorrhea was used to allow for a wider variance. The normal model and gamma model performed equally well, and the normal model was chosen because of its computational convenience. The results are presented in Table 4, and they confirm the findings in the previous analysis of postpartum amenorrhea (Table 3). That is, estimated postpartum amenorrhea was not sensitive to different distributions of postpartum amenorrhea, but it was sensitive to the variance. Irrespective of the distribution used

to generate postpartum amenorrhea, the estimated parameters were very close to the true parameters when the true variance was specified in the model. In contrast, the estimated parameters deviated from the true parameters when the variance was misspecified. The parameters were overestimated if the variance was overstated, and underestimated if the variance was understated.

Table 4. Estimated monthly hazard of live-birth conception and postpartum amenorrhea by different mean durations and variances of postpartum amenorrhea[1]

Distributions of P.A.		P.A.	Age					
Sample	Model	(months)	20	25	30	35	40	45
True parameters			.250	.250	.250	.188	.125	.062
Postpartum amenorrhea 6months								
Constant $\sigma^2 = 0.0$	$\sigma^2 = 0.1$	6.2	.247	.259	.262	.192	.128	.061
	$\sigma^2 = 4.0$	6.5	.244	.271	.284	.205	.134	.062
Normal $\sigma^2 = 1.0$	$\sigma^2 = 1.0$	5.8	.248	.240	.240	.180	.123	.060
	$\sigma^2 = 4.0$	6.3	.250	.258	.267	.196	.131	.062
Gamma $\sigma^2 = 6.0$	$\sigma^2 = 6.0$	5.9	.243	.240	.238	.182	.124	.060
	$\sigma^2 = 4.0$	5.5	.242	.227	.220	.171	.118	.059
Negative binomial	$\sigma^2 = 12.0$	6.2	.256	.253	.256	.197	.132	.062
$\sigma^2 = 12.0$	$\sigma^2 = 4.0$	4.8	.245	.203	.192	.154	.110	.058
Postpartum amenorrhea 9months								
Normal $\sigma^2 = 9.0$	$\sigma^2 = 4.0$	7.6	.240	.200	.185	.149	.107	.057
	$\sigma^2 = 9.0$	8.6	.246	.231	.225	.175	.121	.060
	$\sigma^2 = 16.0$	9.3	.243	.251	.257	.197	.132	.063
Postpartum amenorrhea 12months								
Normal $\sigma^2 = 9.0$	$\sigma^2 = 4.0$	10.4	.246	.196	.180	.144	.106	.056
	$\sigma^2 = 9.0$	11.6	.250	.230	.224	.172	.121	.060
	$\sigma^2 = 16.0$	12.1	.246	.244	.248	.189	.129	.061

[1] Standard errors are $\leq .01$.

We then addressed the degree of bias in the parameter estimates when a misspecified variance of postpartum amenorrhea was used in the model. The bias was trivial in models with mean durations of postpartum amenorrhea of 3 and 6 months (Table 3 and Table 4). Subsequently, we estimated parameters from samples with mean durations of postpartum amenorrhea of 9 and 12 months. In these models, only a normal distribution was used to generate postpartum amenorrhea, because, as noted above, the estimated parameters were not sensitive to the underlying distribution. To evaluate the bias of the estimated parameters from a misspecified variance of postpartum amenorrhea, three different values of the variance were specified: one true variance, one understated, and one overstated.

Table 4 shows that models with postpartum amenorrhea of 9 or 12 months provided estimates that were very close to the true parameters when the true variance was specified. As expected, the parameters were slightly underestimated when the variance was understated, and conversely, when the variance was overstated, the parameters were slightly overestimated. However, the

variances specified cover a wide range of postpartum amenorrhea, and the bias of the resulting estimates was relatively small.

To summarize, one-parameter postpartum amenorrhea models performed better than two-parameter models. Both the one-parameter normal and the gamma models of postpartum amenorrhea provided good estimates of the parameters when the true variance was specified, regardless of the underlying distribution of postpartum amenorrhea. When the mean postpartum amenorrhea was 6 months or less, the bias of the estimates was minor, even when the variance was misspecified, due to the relatively small range within which the variance could vary. For longer durations of postpartum amenorrhea (9 months or longer), the estimates provided an accurate approximation of the true parameters when the true variances were specified, and a relatively small error arose for a wide range of variances.

Estimated Fecundability with Postpartum Amenorrhea, Heterogeneity and Sterility

The subsequent model was calibrated by estimating fecundability, postpartum amenorrhea, and heterogeneity simultaneously. The model used a one-parameter gamma distribution for heterogeneity and a one-parameter normal distribution for postpartum amenorrhea, as discussed above. Sample frailty was generated by a gamma distribution with various parameters to address different degrees of heterogeneity. Duration of postpartum amenorrhea was generated by a normal distribution with a mean of 3 months and a variance of 1 month.

Table 5 presents the estimates of frailty, postpartum amenorrhea and fecundability by age. This analysis demonstrated that postpartum amenorrhea, fecundability by age, and heterogeneity in fecundability could be estimated simultaneously in a convolution model. All the reproductive characteristics were captured well by the model.

Table 5. Estimated monthly hazard of live-birth conception and postpartum amenorrhea in models with heterogeneous fecundability[1]

Sample frailty	Estimated frailty Gamma (α, α) $\sigma^2 = 1/\hat{\alpha}$	P.A. (months)	Age					
			20	25	30	35	40	45
True parameters		3.0	.250	.250	.250	.188	.125	.062
Gamma (2, 2, 0) $\sigma^2 = .50$	.424	2.9	.236	.234	.234	.180	.125	.060
Gamma (4, 4, 0) $\sigma^2 = .25$	.235	3.0	.238	.241	.241	.183	.126	.061
Gamma (1, 2.5, .6) $\sigma^2 = .16$	.129	3.0	.245	.243	.245	.184	.125	.060

[1] Standard errors are $\leq .006$.

Next, we examined whether the model could also capture variations in the duration of postpartum amenorrhea by age. In this analysis the true duration of postpartum amenorrhea was set to 3 months before age 30 and to 5 months after age 30 (Table 6). In Model 1, one parameter was used to estimate the duration of postpartum amenorrhea, i.e., we assumed that the duration of postpartum amenorrhea was constant for all ages. This model misspecification resulted in an estimate of postpartum amenorrhea of 3.3 months, which is higher than the true 3 months before age 30, and lower than the true 5 months after age 30. Furthermore, this error in the estimate of postpartum amenorrhea resulted in an expected error in the estimates of fecundability by age. More specifically, fecundability was overestimated at age 25 (because estimated postpartum

amenorrhea was overstated), and underestimated above age 30 (because estimated postpartum amenorrhea was understated). In Model 2, two parameters were used to estimate the average duration of postpartum amenorrhea. In this model the estimated duration of postpartum amenorrhea was 3.0 months before age 30 and 4.3 months after age 30 suggesting that the model captured variations in the duration of postpartum amenorrhea. The duration of postpartum amenorrhea was slightly underestimated above age 30 resulting in slightly reduced estimates of fecundability above age 35. In Model 3, the true duration of postpartum amenorrhea was set at 3 months for all ages, but two parameters were used to estimate the duration of postpartum amenorrhea by age. In this case, estimates of postpartum amenorrhea and fecundability by age were very close to the true parameters of amenorrhea and fecundability. Thus, modeling variations in the duration of postpartum amenorrhea by age, when postpartum amenorrhea was truly constant, resulted in unbiased estimates of postpartum amenorrhea and fecundability by age. In summary, the model was able to measure variability in fecundability by age and across women, as well as variability in duration of amenorrhea by age and across women.

Table 6. Estimated monthly hazard of live-birth conception and postpartum amenorrhea in models with heterogeneous fecundability[1]

Model	Estimated frailty Gamma (4, 4) $\sigma^2 = 1/\hat{\alpha}$	P.A. (months) Age ≤ 30	P.A. (months) Age > 30	Age 20	Age 25	Age 30	Age 35	Age 40	Age 45
True parameters	.250	3.0	2.0	.250	.250	.250	.188	.125	.062
Model 1	.212	3.3		.249	.261	.228	.117	.080	.000
Model 2	.220	3.0	1.3	.250	.237	.253	.138	.084	.000
True parameters	.250	3.0	0.0	.250	.250	.250	.188	.125	.062
Model 3	.236	3.0	0.0	.255	.244	.235	.167	.088	.000

[1] Standard errors are ≤ 0.007.

In order to asses the effect of sterility on parameter estimates, samples were generated with sterility obtained from a Gompertz distribution with a mean age of sterility of 42.3 years (Table 1). When the true age at sterility was known, we set *x* to age at sterility in the likelihood function described by equation (3). However, age at sterility was usually an unobserved event. Two ways were considered to circumvent this limitation. First, all women were analyzed up to the age of their last birth, i.e., the analysis was limited to closed birth intervals. Second, all women were included up to a given age *C*, set to the oldest age of last birth in the sample analyzed. In the simulated populations, the oldest age at last birth changed in each replication, and we set *C* to 46.

The simulation analysis suggested that, in a population with little sterility before the late 30s, the model estimated fecundability accurately up to age 40, both from all birth intervals and from closed birth intervals (results available from author). However, it was difficult to separate the effects of selection on estimates of fecundability from closed birth intervals. In contrast, the downward error in estimates of fecundability from all birth intervals was obvious.

Effect of Censoring and Sample Size Evaluation

The parameter estimates from complete and incomplete birth histories were compared in order to evaluate the effects of different censoring schemes (Table 1). In samples with incomplete birth histories, the number of women in the sample at older ages was reduced, resulting in underestimates of heterogeneity. The error in heterogeneity was relatively small from all birth intervals, though it was greater when only closed birth intervals were analyzed. Furthermore, estimates of fecundability from closed birth intervals were biased by selection. Therefore, we concluded that analyses should be based on all birth intervals both when birth histories were censored and when they were observed up to age 50 (results available from author). The effects of sample size and sampling variation on parameter estimates were evaluated in samples of 500, 250, and 100 women with complete and incomplete birth histories. The model was not sensitive to sample size, and there was no significant difference in the estimated parameters from samples with at least 100 women (results available from author).

A CASE STUDY OF THE HUTTERITES

In the empirical analysis, the model was applied to Hutterite reproductive histories. (For more details about the Hutterites and the data set, see Sheps (1965).) The sample included 7224 unions. In this sample, 10 women had been married twice. Data related to second unions, as well as those 194 unions with missing birth, marriage, or death dates were not included in the analysis. The illustrative example was also confined to fecundability above age 20 to eliminate the effect of adolescent subfecundity, and 137 women who married before age 20 were not analyzed. The sample analyzed contained 370 women who had 2,049 live births.

The Hutterite data did not include any information about postpartum amenorrhea or about age at onset of sterility. Estimated parameters were obtained from models with different variances of postpartum amenorrhea to determine the value of the variance that best fitted the data. Two censoring schemes were used to circumvent the fact that information about age at onset of sterility was not available from the Hutterite data, and that sterility was not modeled. In one group of models all intervals were used, and in a second group only closed intervals were used. Finally, the age schedule of Hutterite fecundability was estimated from models where heterogeneity in fecundability was specified either by a gamma or by a homogeneous distribution. Parameter estimates are presented in Table 7.

Table7. Parameter estimates for Hutterite women

Model	Variance of frailty	P.A. (months)	Monthly hazard by age					
			20	25	30	35	40	45
All intervals								
$\sigma^2 = 6.0$	.255	6.2	.391	.245	.152	.117	.064	.004
$\sigma^2 = 7.0$	.275	6.4	.399	.260	.159	.122	.065	.004
$\sigma^2 = 8.0$	.286	6.6	.400	.273	.164	.123	.066	.004
All intervals								
$\sigma^2 = 6.0$		6.3	.316	.199	.123	.101	.065	.003
$\sigma^2 = 7.0$		6.3	.344	.201	.122	.098	.068	.004
$\sigma^2 = 8.0$		6.3	.308	.203	.126	.099	.062	.004
Closed intervals								
$\sigma^2 = 6.0$	.173	6.2	.360	.250	.177	.148	.091	.111
$\sigma^2 = 7.0$	.183	6.4	.364	.264	.185	.154	.093	.112
$\sigma^2 = 8.0$	.193	6.6	.366	.278	.193	.160	.095	.113

The results of this illustrative analysis demonstrated that the model proposed could be applied to empirical analyses of fecundability. Fecundability, heterogeneity in fecundability, and postpartum amenorrhea could be estimated simultaneously, and estimates of fecundability reflected the underlying level and age pattern of fecundability, accounting for the effects of heterogeneity and postpartum amenorrhea. The mean duration of postpartum amenorrhea was estimated to be about 6.5 months, the duration of amenorrhea did not vary significantly by age of the women, and information about postpartum amenorrhea could be obtained only by using a modeling approach.

The model with heterogeneous fecundability fitted the data significantly better than did the model with homogeneous fecundability based on the likelihood ratio test (Table 7). As expected, fecundability was underestimated in the model with homogeneous fecundability. For instance, at age 25 the monthly hazard of a live birth was .273 in the heterogeneous model compared to .203 in the homogeneous model, when the variance of postpartum amenorrhea was set at 8, (or fecundability was .238 versus .184).

To assess the value that best approximated the true variance of postpartum amenorrhea, and to determine the censoring scheme that resulted in the best parameter estimates, goodness of fit tests were performed. For each set of parameters listed in Table 7, births were simulated for 370 women (the same number of observations as in the Hutterite sample used). The empirical distributions of waiting times from one birth to next live-birth conception after age 20, 25, 30, 35, and 40 were calculated from the Hutterite data and from the simulated data. Very few such intervals were longer than 26 lunar months at age 20, so the intervals of 19.5-25 and 26+ were combined at age 20. At age 25, 30, 35, and 40 the intervals of < 3 and 3-5 were combined because most such intervals were of second or higher order, and they included a period of postpartum amenorrhea. The distribution at age 45 was not analyzed because the Hutterite data contained only five births after age 45. The simulation was conducted 1,000 times, and χ^2 values from the difference of empirical distributions of waiting times from one birth to next live-birth conception between actual and simulated data were obtained.

The goodness of fit analysis underlined the usefulness of the model and the importance of using all birth intervals. Models using all birth intervals and heterogeneous fecundability fitted the data well. The simulated waiting times to live-birth conception were in good agreement with the actual waiting times when the variance of postpartum amenorrhea was specified as 6, 7 or 8, respectively. Models with varying duration of postpartum amenorrhea by parity did not fit as well as models with constant duration of postpartum amenorrhea (Larsen and Yan 2001). Table 8 presents the distributions of actual and simulated waiting times by age from the best fitting model with the variance of postpartum amenorrhea specified as 8 and fecundability as heterogeneous. The goodness of fit analysis suggested that, regardless of the specified value of the variance of postpartum amenorrhea, the model from closed birth intervals did not fit the Hutterite data (results not shown).

The findings from the analysis of Hutterite birth histories were consistent with the results from the simulation analysis. In analyses restricted to closed birth intervals, estimated heterogeneity was substantially lower, and estimated fecundability increased above age 40, reflecting selection bias. Figure 1 illustrates the level and age pattern of estimated fecundability from closed and from all birth intervals of Hutterite women, when the variance of postpartum amenorrhea was specified to be 8 and heterogeneous fecundability. Thus, when analysis was restricted to closed birth intervals, the estimated age pattern of fecundability became distorted because of selection bias.

Table 8. Distribution of waiting times from birth to next live-birth conception for Hutterite women[1]

	Waiting times in lunar months								
Mean	< 3	3 - 5	6 -8	9 -12	13 - 19.4	19.5 - 25	26+	χ^2	P-value
Age 20 (316 observations)[2]									
Actual									
3.66	.611	.174	.085	.079	.035	.016			
Simulated									
3.62	.616	.208	.085	.047	.028	.016		9.21	.101
(.008)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)			
Age 25 (284 observations)[3]									
Actual									
10.70		.232	.218	.243	.204	.081	.021		
Simulated									
11.02		.204	.232	.281	.187	.058	.039	8.22	.145
(.012)		(.001)	(.001)	(.001)	(.000)	(.000)	(.000)		
Age 30 (195 observations)									
Actual									
13.63		.108	.190	.251	.272	.133	.046		
Simulated									
13.65		.113	.198	.275	.235	.097	.083	7.40	.193
(.018)		(.001)	(.001)	(.001)	(.001)	(.001)	(.001)		
Age 35 (127 observations)									
Actual									
16.02		.063	.110	.283	.299	.118	.126		
Simulated									
15.50		.090	.165	.246	.246	.122	.131	5.53	.355
(.026)		(.001)	(.001)	(.001)	(.001)	(.001)	(.001)		
Age 40 (50 observations)									
Actual									
17.45		.060	.160	.140	.220	.220	.200		
Simulated									
17.79		.065	.130	.210	.248	.151	.196	3.29	.655
(.047)		(.001)	(.002)	(.002)	(.002)	(.002)	(.002)		
Total χ^2 value from combined test								33.65	.116
5% critical value of χ^2 test by age (df = 5)								11.07	
5% critical value of combined χ^2 test (df = 25)								37.65	

[1] Standard errors are presented in parentheses.
[2] At age 20 waiting times of 19.5 – 25 months and 26+ lunar months are combined.
[3] At age 25 and older waiting times of < 3 lunar months and 3 – 5 lunar months are combined.

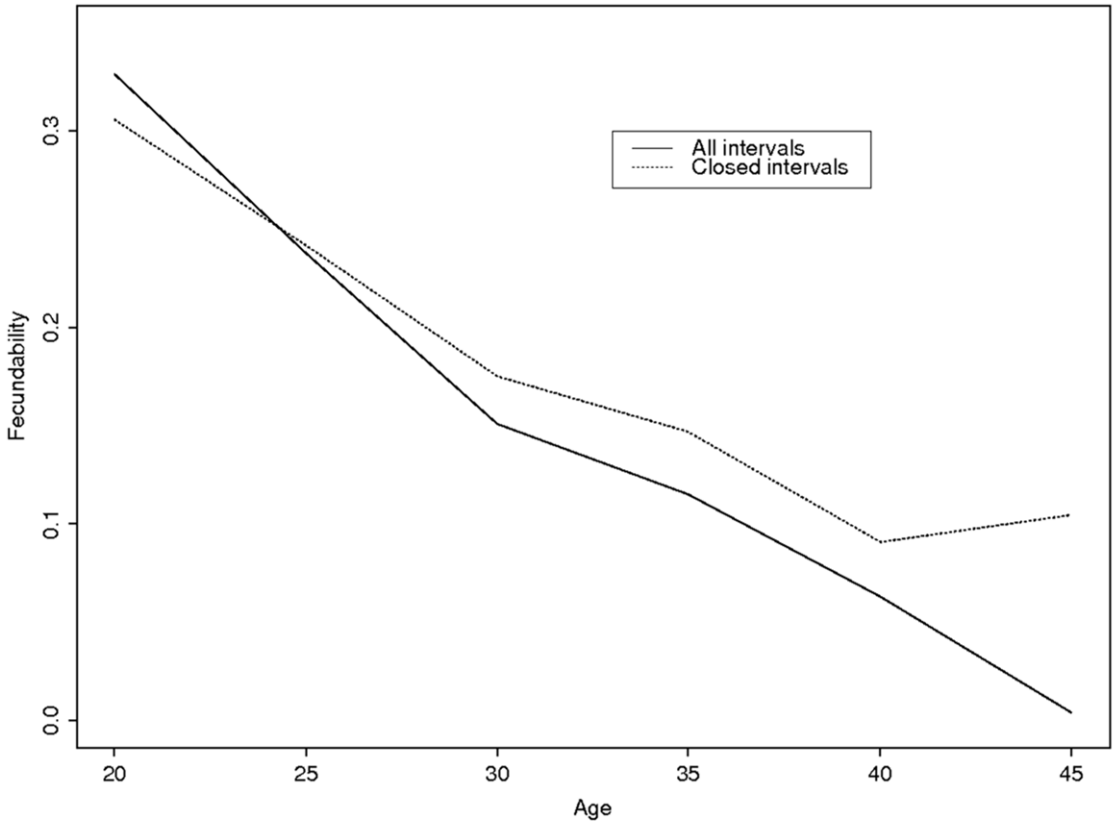

Figure 1. The age pattern of fecundability for Hutterite women from closed birth intervals and all intervals
Note: Fecundability *=1-exp(-h(x+.5))*, where *h(x)* is the monthly hazard of conception at age *x*.

DISCUSSION AND CONCLUSION

The present analysis extended previous work on the measurement of fecundability (see, for example, D'Souza 1974; Heckman and Walker 1987, 1990; Larsen and Vaupel 1993; Wood et al. 1994), but also provided additional substantive information about fecundability.

A convolution model of postpartum amenorrhea and fecundability was calibrated in a simulation analysis prior to its application to empirical data. This simulation showed that the model with a piecewise linear function of fecundability by age, a normal distribution for postpartum amenorrhea, and a gamma distribution for heterogeneity, accurately captured a wide range of variations of fecundability within a woman and across women. For this model, eight parameters needed to be estimated: six of them characterized the age pattern and level of fecundability, one represented the mean duration of postpartum amenorrhea, and one measured the degree of heterogeneity.

Simulation also showed that variations in fecundability across women could be summarized and quantified by a single value: the variance of frailty. Different models were applied to different samples of frailty, and the one-parameter gamma model of frailty was chosen as the final model, because it was simple, flexible, and able to capture a wide range of heterogeneity.

This analysis also established the accuracy of a simple model of postpartum amenorrhea, and it showed that a wide range of distributions of postpartum amenorrhea could be captured by a normal model. Variation in duration of postpartum amenorrhea by age could be captured by a piecewise linear function.

Empirical analysis of Hutterite birth histories documented that the duration of postpartum amenorrhea was about 6.5 months for women of all ages; that fecundability declined almost linearly from age 20 to 40; and that the fecundability of a 35-year-old was one half that of a 25-year-old woman. Fecundability could not be estimated beyond age 40 because the model did not take sterility into account. Variations in fecundability across Hutterite women were considerable. For example, women one standard deviation above the mean had a 2.9 time higher hazard of a live-birth conception than did women one standard deviation below the mean. Finally, the empirical analysis confirmed the results of the simulation analysis, in particular the suggestion that fecundability was downward biased in homogeneous models, and that the age pattern of fecundability was distorted when only closed birth intervals were analyzed.
Some of this decline in fecundability was related to changes associated with age and/or duration of marriage. These changes include a decline in the quality of ova and thus in the chance of conception; an increased risk of spontaneous abortion; the possibility of outright sterility; and reduced coital frequency (Wood, Holman and O'Connor 1997). The documented variability in fecundability across women might be due primarily to biological factors, such as the risk of conception and fetal loss, because there was virtually no observed variation in fertility of Hutterite women (Eaton and Mayer 1953; Sheps 1965).

The proposed model builds on previous work by simultaneously modeling postpartum amenorrhea by age, fecundability by age, and heterogeneity in fecundability. To illustrate the advances made by this model, substantive findings from Hutterite birth histories about the duration of postpartum amenorrhea and fecundability by age were compared with results from previous research (D'Souza 1974; Heckman and Walker 1987, 1990; Larsen and Vaupel 1993). Heckman and Walker (1987) focused on developing approaches to choose among alternative models, and they did not provide any direct information about the age schedule of fecundability. Thus, in their analysis of Hutterite birth history data, Heckman and Walker (1987) ignored postpartum amenorrhea, and they did not present parameter estimates in a form that made them interpretable or informative about the age pattern of fecundability. Heckman and Walker (1990) confined their analysis to the first birth interval and thereby circumvented, but did not solve, the problem of controlling for postpartum amenorrhea. Heckman and Walker's (1990) work focused on measurement issues and presented no substantive findings about the age pattern of fecundability. In fact, the age pattern of fecundability and sterility could not be obtained from first birth intervals alone for two reasons. First, because many women became sterile after they had a child, and second, because of selection bias.

According to D'Souza (1974, p.121), the duration of postpartum amenorrhea was 7 to 8 months. This author was not able to capture the age pattern of fecundability because he restricted the analysis to closed birth intervals, resulting in upward biased estimates of fecundability with age. D'Souza's model was also unable to capture the level of fecundability because it overestimated the duration of postpartum amennorea, it assumed homogeneous fecundability, and only couples with at least four children were analyzed. In fact, the caveats in D'Souza's model are reflected in the finding that the goodness of fit tests suggested that the model did not fit the Hutterite data.

Larsen and Vaupel (1993), like D'Souza (1974), could not capture the age pattern of fecundability because they restricted their analyses to closed birth intervals to avoid error introduced by not modeling sterility. The level of fecundability was also inaccurate for these authors because they set duration of postpartum amenorrhea to a constant of 5 months for all women, and Hutterite postpartum amenorrhea has a distribution with a mean of about 6.5 months, as noted here. Thus, Larsen and Vaupel (1993) underestimated fecundability up to the early 30s, and at older ages the bias from analyzing only closed birth intervals resulted in further distortions of the fecundability estimates. Thus, goodness of fit tests suggested that the models by Larsen and Vaupel (1993), like the model by D'Souza (1974), did not fit the Hutterite data. We believe that the fecundability estimates from the model proposed are better, not only because the goodness of fit tests suggest that these estimates fit the Hutterite data; but also because prior simulation analysis documented that the model accurately captured a wide range of variations in fecundability within a woman and across women. None of the previous models accurately fitted the Hutterite data. To enhance the understanding of the baseline level and age pattern of fecundability further studies of other non-contracepting populations are in order. We are currently applying the model proposed here to a number of different historical non-contracepting populations. Preliminary findings suggest that there may not be one baseline level and age pattern of fecundability.

The model proposed here answers a need in the methods currently available to estimate fecundability from birth history data. It includes estimation of postpartum amenorrhea by age, fecundability by age and heterogeneity in fecundability from all birth intervals. These modifications facilitate estimation of the level and age pattern of fecundability and estimation of the variability of fecundability across women. The model does not, however, include sterility, and therefore fecundability is underestimated at older ages of the reproductive age span. The magnitude of this error depends on the age pattern of sterility. In populations such as the Hutterites, with little disease-induced sterility, the model accurately captures the age pattern of fecundability up to age 40, as well as unobserved heterogeneity in fecundability and the duration of postpartum amenorrhea. For populations with elevated levels of sterility at young ages, analyses of the age pattern of fecundability could be extended to include sterility. The model proposed here can be extended to include sterility.

In conclusion, the model proposed made it possible to estimate the level and the age pattern of fecundability by controlling for postpartum amenorrhea and unobserved heterogeneity; the model applied to first as well as to higher order birth intervals. Finally, the study evaluated the effects on estimates of fecundability by confining analysis to closed birth intervals versus employing both closed and open birth intervals.

REFERENCES

Barrett, J. 1969. “A Monte Carlo Simulation of Human Reproduction." *Genus* 25:1-22.

Barrett, J. C. 1971. ”Use of a Fertility Model to Refine Measurement Techniques." *Demography* 8: 481-490.

Coale, A.J. and P. Demeny. 1983. *Regional Model Life Tables and Stable Populations.* 2nd ed. New York: Academic Press.

Coale, A. and J.Trussell. 1978. "Technical Note: Finding the Two Parameters that Specify a Model Schedule for Marital Fertility." *Population Index* 44:203-213.

Coale, J. and S.Watkins. 1986. *The Decline of Fertility in Europe: The Revised Proceedings of a Conference on the Princeton European Fertility Project.* Princeton: Princeton University Press.

Dunson, D.B. and H. Zhou. 2000. "A Bayesian Model for Fecundability and Sterility." *Journal of the American Statistical Association* 95:1054-1062.

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

Eaton, J.W. and A.J. Mayer, 1953. “The Social Biology of Very High Fertility among the Hutterites: The Demography of a Unique Population." *HumanBiology* 25:206-264.

Ellison, P.T. 1990. “Human Ovarian Function and Reproductive Ecology: New Hypotheses." *American Anthropologist* 92:933-952.

Ellison, P.T., C. Panter-Brick, S.F. Lipson, and M.T. O'Rouke. 1993. “The Ecological Context of Human Ovarian Function." *Human Reproduction* 8:2248-2258.

Ellison, P.T., N.R. Peacock, and C. Lager. 1989. “Ecology and Ovarian Function among Lese Women of the Ituri Forest, Zaire." *American Journal of Physical Anthropology* 78:519-526.

Golden, M. L. and S. R. Millman. 1988. Comparison of Fecundability Models and Their Estimates. Working Paper, Population Studies and Training Center, Brown University.

Heckman, J. and J. Walker. 1987. “Using Goodness of Fit and Other Criteria to Choose among Competing Duration Models: A Case Study of Hutterite Data." Pp. 247-309 in *Sociological Methodology*, edited by C. C. Clogg. San Francisco: Jossey-Bass.

Heckman, J. and J. Walker. 1990. “Estimating Fecundability from Data on Waiting Times to First Conception." *Journal of the American Statistical Association* 85:283-294.

Institute for Resource Development, Macro International, Inc. 2001. "Demographic and Health Surveys". MEASURE Measure DHS+.

Larsen, U. and B. Desjardins. 1997. "Marital Fertility in a Natural Fertility Population: a Case Study of the French-Canadians, 1660-1719." *Proceedings of the IUSSP General Population Conference Beijin, China, October 1997*.

Larsen, U. and J. Menken. 1989. “Measuring Sterility from Incomplete Birth Histories." *Demography* 26:185-201.

Larsen, U. and J. Vaupel. 1993. “Hutterite Fecundability by Age and Parity: Strategies for Frailty Modeling of Event Histories." *Demography* 30:81-102.

Larsen, U. and S. -. 2001. "The Age Pattern of Fecundability: An Analysis of French Canadian and Hutterite Birth Histories." Social Biology 47:34-50.

Menken, J.A. 1985. "Age and Fertility: How Late Can You Wait?" Demography 22:469-483.

Menken, J., J. Trussell, and U. Larsen. 1986. “Age and Infertility." *Science* 233:1389-1394.

Microsoft Corporation. 1995. *Fortran PowerStation Programmer's Guide*. Microsoft Corporation. United States of America.

Pittenger, D. B. 1973. "An Exponential Model of Female Sterility." *Demography* 10:113-121.

Pressat, R. and C. Wilson. 1985. *The Dictionary of Demography*. New York: Blackwell.

Sheps, M. 1965. "An Analysis of Reproductive Patterns in an American Isolate." *Population Studies* 19:65-80.

Sheps, M.C. and J.A. Menken. 1973. *Mathematical Models of Conception and birth*. Chicago: The University of Chicago Press.

Trussell, J. and C. Wilson. 1985. “Sterility in a Population with Natural Fertility." *Population Studies* 39:269-286.

Wood, J.W., D.J. Holman, A.I. Yashin, R.J. Peterson, M. Weinstein, and M.C. Chang. 1994. “A Multistate Model of Fecundability and Sterility." *Demography* 31:403-426.

Wood, J.W., D.J. Holman, and K.A. O'Connor. 1997. Age and Female Reproductive Function: Identifying the Most Important Biological Determinants. Paper presented at the IUSSP meeting on ”Age: Between Nature and Culture" in Rostock, May, 1997.

Yan, S. and U. Larsen. 2001. “A Non-linear Model of Fecundability, Postpartum Amenorrhea and Sterility.” *Mathematical Population Studies* 9:143-164.

Yashin, A.I., I.A. Iachine, K.F. Andreev, and U. Larsen. 1998. “Multistate Models of Postpartum Infecundity, Fecundability and Sterility by Age and Parity: Methodological Issues." *Mathematical Population Studies* 7:1-28.