New Evidence on the Timing and Spacing of Births

By JAMES J. HECKMAN, V. JOSEPH HOTZ, AND JAMES R. WALKER*

This paper is a first progress report of an ongoing empirical study of the determinants of life cycle fertility (see our 1985 paper for a more complete report). At this stage, our analysis is decidedly empirical. Unlike many other areas of knowledge in economics and social science, there are few widely accepted or carefully confirmed "stylized facts" in fertility dynamics to guide economic model builders. This vacuum has inhibited the successful development of economic models of fertility dynamics. The main objective of the early stage of our work is to codify the "facts" in a coherent statistical framework that provides the duration data analogue of the conventional simultaneous equations model.

The starting point for our analysis is the demographic literature. This literature finds that the age at marriage, the occurrence of births inside or outside of marriage, the age of the first birth and/or the durations of previous birth intervals significantly affect the timing of subsequent births over the life cycle. Studies of *marital fertility* by L. Coombs and R. Freedman (1970), L. Bumpass, R. Rindfuss, and R. Janosik (1978), J. Trussell and J. Menken (1978), F. Finnas and J. Hoem (1980), and Hoem and R. Selmer (1984) find that the younger the woman's age at marriage, the more rapid the pacing of subsequent fertility, and that those who begin childbearing early in their reproductive careers subsequently have children more rapidly.

Recent cross-country comparisons of data from the World Fertility Survey have found that the timing of marriage and the lengths of prior birth intervals directly affect the spacing of subsequent life cycle fertility. Such results have lead G. Rodriguez et al. to conclude that "birth interval lengths depend little upon birth order, but far more upon the length of the previous interval" and that the human "reproductive process can be encapsulated as an engine with its own inbuilt momentum whereby early behavior and socio-economic differences fundamentally determine (along with ageing and secular variation) the remainder of the childbearing experience" (1984, p. 5). Such a view suggests: 1) that the age of entry into marriage (or marriage-like unions, such as cohabitation) and/or the entry into parenthood are the crucial determinants of life cycle fertility and that the variation in completed fertility across the population comes primarily in these initial decisions; and 2) that subsequent childbearing is largely determined by initial events and by the lengths of preceding birth intervals. Another widely held view is that after the first birth interval, subsequent birth intervals are biologically determined until the birth process terminates.

One objective of our study is to investigate the robustness of these findings in other data sets. For a sample of Swedish women described below, we find that the conventional demographic view of the fertility process is largely confirmed on fresh data. The main objective of our study, however, is to probe the empirical findings to discriminate between "structural" and "spurious" explanations for the observed empirical regularities. As is the case in any empirical study, only a subset of plausible determinants of life cycle fertility is observed. Evidence that lagged dependent variables affect current values of those variables or evidence that initial outcomes of a life cycle process affect later outcomes may be due to serial correlation in unobservables.

In the context of fertility, one might expect a woman who desires children or who

*The University of Chicago and The Economics Research Center/NORC, Chicago, IL 60637. We thank George Yates for his help. This research was supported by Grant No. HD 19226 from the National Institute of Child Health and Human Development.

has a comparative advantage at the activity of child rearing might marry earlier, have children at an earlier age, and invest in household capital relative to a woman with a stronger motivation for market activity. Failure to control for unobserved propensities and skills may explain the regularities noted above. On the other hand, the observed empirical regularities may be the outcome of genuine behavioral influences of past events on current constraints or preferences. For example, early childbearing may cause a woman to invest less in market skills and more in nonmarket skills, leading to higher fertility rates for such women. Our study attempts to shed light on the issue of the causal significance of these demographic empirical regularities. A main conclusion is that controlling for unobservables in a robust nonparametric fashion vitally affects the sign and statistical significance of the estimated effect of early life cycle events on subsequent fertility outcomes. Some of the stylized facts of the demographic literature are not robust to controls for unobservables but others remain.

We find that for a variety of empirical specifications of the life cycle birth process and for a variety of samples that do and do not condition on marital status, in models that do not control for unobservables, the longer a preceding birth interval the longer the subsequent one. Thus, our data exhibit the "engine of fertility" phenomenon noted by demographers. Controlling for unobservables, the "well-noted empirical regularity" either vanishes or reverses in sign. For married women, it vanishes entirely. For a sample of all women in a model that controls for marital status as a covariate, controlling for unobserved heterogeneity produces a "reverse engine of fertility" phenomenon: the longer the preceding birth interval the *shorter* the subsequent one. For a sample of married women, the importance of the age of marriage on the spacing of birth intervals is considerably reduced in size and statistical significance. For a sample of all women, controlling for unobserved variables eliminates the impact of age at marriage on all but the final birth transition.

I. The Data

The data used in this analysis are from a survey of retrospective interviews conducted by the Swedish National Central Bureau of Statistics in 1981. The survey asked for complete cohabitational, marital, childbearing, and work histories as well as background information of almost 5,000 Swedish women randomly selected from five recent cohorts of women of all marital statuses. (For further description of the data set, see Hoem and B. Rennermalm, 1982.) We have restricted our analysis to a sample of women who were born between 1941–45. For this cohort, we had useable data for 990 women. In order to analyze the effects of sample conditioning in previous studies of marital fertility, we also estimated some models described below on a subset of women from this cohort who: (a) were married at least once prior to 1981, (b) experienced their first "union" (i.e., marriage or consensual union) before age 25, and (c) did not have a birth prior to their first marriage. There are 570 such women.

II. The Model

For the sake of brevity, we only sketch the statistical model used to perform the empirical work reported below. A full description of the model is available in our companion paper. The main building block is the multistate hazard rate:

$$h_{ij}(t_{ij}|\mathbf{x}'\boldsymbol{\beta}_{ij}, c_{ij}\theta), \tag{1}$$

where i denotes the origin state and j denotes the destination state, t_{ij} is the duration in state i which exits into state j, $\mathbf{x}$ is a vector of (possibly time varying) observed variables that may include lagged durations or occurrence times of previous events, $\boldsymbol{\beta}_{ij}$ is a vector of associated coefficients, θ is a scalar unobservable, and c_{ij} is a transition-specific factor loading. Specifying the functional form for the hazard and using the nonparametric maximum likelihood estimation procedure of Heckman and B. Singer (1984b), it is possible to consistently estimate the parameters of the hazard (including the

c_{ij}s) and the population distribution of the unobservables. Allowing the $h_{ij}(t_{ij})$ to be a nontrivial function of duration t_{ij} allows for spell-specific duration dependence. For a further discussion of multistate models, see Heckman and Singer (1984a).

A birth process specializes the model to transitions among successive birth states, $i = 0,\dots, I$, where I is the maximum number of births. In this specification, $j = i + 1$ for each i. In the restricted birth process model, transitions among marital states are treated as changes in exogenous variables (changes in elements of $\mathbf{x}$). Note that each duration is permitted to be governed by different parameters to explicitly allow for different birth order duration dependence and effects of the x's. Below, we test whether or not the set of parameters are significantly different for the intervals between the second and third births to examine the finding in Rodriguez et al. of no-birth-order effects on higher-order birth intervals. Finally, we note that because women are followed from the age of menarche, there is no problem of initial conditions. (See Heckman and Singer, 1984a, for a discussion of this problem.) In a more general multistate process, it is possible to introduce transitions among birth and nonbirth states. Specifically, in our companion paper, we permit transitions from the single or cohabiting state to the states of marriage and cohabitation, transitions among all three of these states and transitions to births from all possible states.

The empirical results reported below are based on a flexible model with linear and quadratic duration terms:

$$(2)\quad h_{ij}\left(t_{ij}|\mathbf{x}'\boldsymbol{\beta}_{ij}, c_{ij}\theta\right) = \exp\left(\gamma_{0ij} + \gamma_{1ij}t_{ij} + \tfrac{1}{2}\gamma_{2ij}t_{ij}^2 + \mathbf{x}'\boldsymbol{\beta}_{ij} + c_{ij}\theta\right).$$

III. Empirical Results

The variables utilized in our analysis are defined as follows: *DURATION* = number of months/100 spent in current spell; *DURATION*^{2} = the square of the number of months/100 spent in current spell; *AGEUNION* = the woman's age (in months/100 since her thirteenth birthday) at which she first entered either marriage or a consensual union; *AGEMAR* = the woman's age (in months/100 since her thirteenth birthday) at which she was married, if she married, and 0 otherwise; *EVERMAR* = a dummy variable = 1 at the age she gets married and = 0 prior to that age; *AGECOH* = the woman's age (in months/100 since her thirteenth birthday) at which she first cohabitated, if she cohabitated, and 0 otherwise; *EVERCOH* = a dummy variable = 1 at the age she begins cohabiting and = 0 prior to that age; *BIRTHDUR1* = the duration in months/100 of the first birth (for the ever married subsample it equals age at first birth − age at first union and for the full sample it equals number of months/100 since her thirteenth birthday); *AGE1STBIR* = age at first birth in months/100 since her thirteenth birthday: *BIRTHDUR2* = the duration in months/100 from first birth until the second birth; *LFP* = a time-varying variable = 1 if the woman is working at the current duration and 0 otherwise; *EDUC* = a time-varying = 1 if the woman is in school at the current duration and 0 otherwise; *URBAN* = a dummy variable = 1 if the woman grew up in an urban area of Sweden and 0 otherwise; *WHITECOL* = a dummy variable = 1 if the woman's father had a white-collar occupation when she was growing up and 0 otherwise; *UNIV* = a dummy variable = 1 if the woman attended a university by the time of the interview and 0 otherwise.

Table 1 reports empirical estimates of the parameters of parity specific hazard functions for *ever married* women. Panel A reports estimates for a statistical model that does not control for unobservables; panel B reports the estimates controlling for unobservables.

Panel A is largely consistent with the engine of fertility story. The age of first union (marriage or cohabitation) is positively associated with the age of first birth and negatively associated with the durations of subsequent births. The length of the previous birth interval is positively associated with the length of the subsequent interval. (This is so

TABLE 1—ESTIMATES FOR BIRTH PROCESS HAZARD RATES

Variable	*AGEUNION* to First Birth	First Birth to Second Birth	Second Birth to Third Birth
A: Not Controlling for Heterogeneity			
CONSTANT	1.77	1.05	2.01
	(0.154)	(0.245)	(0.369)
DURATION	2.15	12.00	4.84
	(0.411)	(0.650)	(0.916)
DURATION^2	−4.91	−19.8	−5.92
	(0.743)	(1.030)	(1.240)
AGEUNION	0.406	−0.349	−0.613
	(0.118)	(0.160)	(0.375)
BIRTHDUR1		−0.663	
		(0.166)	
BIRTHDUR2			−3.64
			(0.567)
LFP	−3.82	−2.62	−3.10
	(0.142)	(0.180)	(0.368)
EDUC	−3.14	−2.50	−1.94
	(0.264)	(0.574)	(0.760)
URBAN	0.476	0.109	−0.244
	(0.066)	(0.074)	(0.148)
WHITECOL	−0.055	−0.093	0.209
	(0.066)	(0.080)	(0.170)
UNIV	0.359	1.14	1.250
	(0.122)	(0.129)	(0.252)
Log Likelihood		749.8	
B: Controlling for Heterogeneity			
CONSTANT	1.04	0.350	−5.18
	(0.205)	(0.383)	(1.120)
DURATION	2.70	17.33	7.05
	(0.452)	(0.945)	(0.976)
DURATION^2	−4.98	−22.20	−6.94
	(0.793)	(1.266)	(1.282)
AGEUNION	0.697	−0.250	−1.07
	(0.160)	(0.284)	(0.455)
BIRTHDUR1		0.234	
		(0.315)	
BIRTHDUR2			1.23
			(1.003)
LFP	−4.00	−4.21	−4.44
	(0.147)	(0.205)	(0.370)
EDUC	−3.34	−3.09	−2.56
	(0.276)	(0.708)	(0.797)
URBAN	0.399	0.123	−0.296
	(0.079)	(0.131)	(0.212)
WHITECOL	−0.031	−0.150	0.125
	(0.081)	(0.141)	(0.242)
UNIV	0.217	1.37	2.15
	(0.150)	(0.230)	(0.377)
Factor Loading	1.50	4.58	8.16
	(0.145)	(0.330)	(1.042)
Log-Likelihood		942.6	

Note: Sample: Continuously married Swedish women, birth cohort 1941–45 ($N = 570$); asymptotic standard errors are shown in parentheses.

because a negative effect of preceding duration on a current hazard means the transition rate to the next child is smaller for the next duration and hence the interbirth interval tends to be larger.) Background variables such as father's occupational status play a minor role.

Controlling for unobserved heterogeneity using a nonparametric maximum-likelihood procedure causes the effect of length of the preceding birth interval on current transition rates to vanish. In fact, for the latter two transitions, the estimated effect of lagged births switches sign. The effect of age at marriage on the estimated transition rates becomes *stronger* for the transition to the first birth and for the transition from the second to the third birth. The fact that unobservables are empirically important is reflected in the great improvement in the fit of the model to the data as measured by the log likelihood.

Table 2 reports empirical estimates for a sample of all women irrespective of their marital history. The new explanatory variables reported in that table are introduced to control for marital history. Such "controls," while traditional, are *ad hoc*. In our companion paper, we estimate a multistate model that breaks out transitions from each marital state to fertility and allows for endogenous transitions among marital and fertility states.

Controlling for heterogeneity in a model estimated on this sample of women causes the effect of age at marriage to vanish from the model (except for the last transition) and causes the effect of the length of the preceding interval on the transition rate to the next birth to reverse sign. *For the full sample of women, the engine of fertility works in reverse.* Controlling for unobservables, a long first birth interval leads to a short second birth interval and a long second birth interval leads to a short third birth interval. These results are consistent with a fixed target model of fertility in which a delay in the arrival of one child is compensated for by an acceleration of the rate of arrival of the next child.

The hazard functions reported for both samples control for the woman's current labor force participation and school atten-

TABLE 2—ESTIMATES FOR BIRTH PROCESS HAZARD RATES

Variable	Age 13 to First Birth	First Birth to Second Birth	Second Birth to Third Birth
A: Not Controlling for Heterogeneity			
CONSTANT	1.43	1.49	0.062
	(0.057)	(0.226)	(0.418)
DURATION	2.67	9.34	3.73
	(0.258)	(0.470)	(0.703)
DURATION^2	−2.01	−14.22	−4.67
	(0.163)	(0.686)	(0.929)
AGEMAR	0.316	0.368	−0.814
	(0.111)	(0.156)	(0.259)
EVERMAR	1.08	0.128	1.05
	(0.136)	(0.223)	(0.472)
AGECOH	0.121	0.004	−0.232
	(0.117)	(0.131)	(0.323)
EVERCOH	0.306	−0.073	0.247
	(0.127)	(0.139)	(0.311)
AGE1STBIR		−0.646	
		(0.134)	
BIRTHDUR2			−2.06
			(0.349)
LFP	−4.14	−2.88	−2.70
	(0.106)	(0.131)	(0.234)
EDUC	−4.28	−2.64	−2.02
	(0.218)	(0.346)	(0.619)
URBAN	0.225	0.040	−0.036
	(0.047)	(0.057)	(0.118)
WHITECOL	−0.183	−0.118	−0.136
	(0.048)	(0.061)	(0.139)
UNIV	0.388	1.23	1.36
	(0.083)	(0.099)	(0.190)
Log Likelihood		748.0	
B: Controlling for Heterogeneity			
CONSTANT	2.43	3.61	1.32
	(0.096)	(0.322)	(0.522)
DURATION	3.46	14.00	6.85
	(0.328)	(0.592)	(0.783)
DURATION^2	−2.13	−16.9	−7.10
	(0.203)	(0.790)	(1.010)
AGEMAR	−0.0005	−0.087	−0.649
	(0.151)	(0.237)	(0.371)
EVERMAR	1.25	1.03	1.72
	(0.171)	(0.310)	(0.635)
AGECOH	0.232	−0.223	−0.119
	(0.149)	(0.233)	(0.437)
EVERCOH	0.272	0.429	0.426
	(0.157)	(0.243)	(0.413)
AGE1STBIR		0.360	
		(0.188)	
BIRTHDUR2			0.771
			(0.385)
LFP	−4.54	−4.44	−3.89
	(0.113)	(0.140)	(0.241)
EDUC	−4.47	−3.47	−2.47
	(0.226)	(0.391)	(0.676)
URBAN	0.182	−0.005	0.126
	(0.068)	(0.099)	(0.167)
WHITECOL	−0.144	−0.199	−0.505
	(0.074)	(0.109)	(0.201)
UNIV	0.110	1.08	1.88
	(0.127)	(0.159)	(0.265)
Factor Loading	−2.16	−4.58	−8.96
	(0.120)	(0.229)	(0.638)
Log-Likelihood		1117.9	

Note: Sample: All Swedish women, birth cohort 1941–45 (N = 990); asymptotic standard errors are shown in parentheses.

dance statuses as covariates. While some of the studies cited above include these variables, it may be argued that they are endogenous variables. In results not reported here, we reestimate the model excluding these two variables. Our conclusions are not changed when they are omitted.

The main point to extract from our analysis is the fragility of the "empirical regularities" to the introduction of unobservables. The "stylized facts" currently offered by demographers do not recognize the importance of unobservables. It would be unfortunate if economists were to take as their mission the formulation of economic models to explain phenomena generated by the simple fact that people are different.

An encouraging feature of these empirical results is that, for the Swedish data, we decisively reject the "biological determination" view that higher parity hazards are identical. Controlling for heterogeneity that could lead to spurious differences in parity specific hazards even if the conditional hazards (given the unobservable) were equal across parities, we reject the hypothesis that the transition rate from the first to the second birth is identical to the transition rate from the second birth to the third birth. The χ^2 statistics for the hypothesis that the corresponding coefficients of these two transitions are the same are: 581.9 with 11 degrees of freedom for the ever married sample of women and 351.0 with 14 degrees of freedom for the full sample of women.

IV. Conclusions

Our results demonstrate the importance of controlling for unobservables in the analysis of life cycle birth processes. We conclude this paper by noting an additional issue that is also of potential importance in the analysis of fertility data and which we investigate in our companion paper.

Variations in the sample selection rules used can affect the inference from estimates, such as those presented above, in a nontrivial way. All of the studies cited above are for marital fertility. None correct for the sample selection bias that arises from using such behaviorally conditioned samples. Corrections for selection have been shown to be empirically important in many studies (see Mark Killingsworth, 1983; James Smith, 1980). The appropriate generalization of methods for dealing with selection bias in a dynamic setting is a multistate duration model. In our companion paper, we introduce transitions among birth and nonbirth states using a general multistate process. Specifically, we permit transitions from the single or cohabiting state to the states of marriage and cohabitation, transitions among all three of these states and transitions to births from all possible marital states. We examine the impact of correct conditioning on estimated marital fertility hazard rates, utilizing a multistate specification. Explicitly accounting for the transitions among these life cycle states in estimation of the parameters of the fertility processes corrects for the selection on marital status.

REFERENCES

Bumpass, L., Rindfuss, R. and Janosik, R., "Age and Marital Status at First Birth and the Pace of Subsequent Fertility," *Demography*, February 1978, *12*, 75–86.

Coombs, L. and Freedman, R., "Premarital Pregnancy, Childspacing and Later Economic Achievement," *Population Studies*, November 1970, *24*, 389–412.

Finnas, F. and Hoem, J., "Starting Age and Subsequent Birth Intervals in Cohabitational Unions in Current Danish Cohorts, 1975," *Demography*, August 1980, *17*, 275–95.

Heckman, J. and Singer, B., (1984a) "Econometric Duration Analysis," *Journal of Econometrics*, January 1984, *24*, 63–132.

______ **and** ______, (1984b) "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, March 1984, *52*, 271–320.

______, **Hotz, V. J. and Walker, J.**, "The Determinants of Birth Intervals: A Multistate Approach," unpublished manuscript, University of Chicago, January 1985.

Hoem, J. and Rennermalm, B., "Cohabitation, Marriage, and First Birth among Never-Married Swedish Women in Cohorts Born 1936–1690," unpublished manuscript, University of Stockholm, December 1982.

______ **and Selmer, R.**, "The Negligible Influence of Premarital Cohabitation on Marital Fertility in Current Danish Cohorts, 1975," *Demography*, May 1984, *21*, 193–206.

Killingsworth, M., *Labor Supply*, Cambridge, England: Cambridge University Press, 1983.

Marini, M. and Hodsdon, P., "Effects of the Timing of Marriage and First Births on the Spacing of Subsequent Births," *Demography*, November 1981, *18*, 529–48.

Rodriguez et al., G., "A Comparative Analysis of the Determinants of Birth Intervals," *Comparative Studies*, No. 30, London: World Fertility Survey, April 1984.

Smith, J., "Introduction," in his *Female Labor Supply: Theory and Evidence*, Princeton: Princeton University Press, 1980.

Trussell, J. and Menken, J., "Early Childbearing and Subsequent Fertility," *Family Planning Perspectives*, July/August 1978, *10*, 209–18.