NBER WORKING PAPER SERIES

FORECASTING AGGREGATE PERIOD SPECIFIC BIRTH RATES: THE TIME SERIES PROPERTIES OF A MICRODYNAMIC NEOCLASSICAL MODEL OF FERTILITY

James J. Heckman

James R. Walker

Working Paper No. 3133

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1989

Funding for this research was provided by national Institute of Child Health and Development, Grant HD-19226 and National Science Foundation Grant SES 87-10145. The authors thank Joseph Sedransk and an anonymous referee for helpful comments. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Working Paper #3133
October 1989

FORECASTING AGGREGATE PERIOD SPECIFIC BIRTH RATES: THE TIME SERIES PROPERTIES OF A MICRODYNAMIC NEOCLASSICAL MODEL OF FERTILITY

ABSTRACT

This article demonstrates the value of microdata for understanding the effect of wages on life cycle fertility dynamics. Conventional estimates of neoclassical economic fertility models obtained from linear aggregate time series regressions are widely criticized for being nonrobust when adjusted for serial correlation. Moreover, the forecasting power of these aggregative neoclassical models has been shown to be inferior when compared with conventional time series models that assign no role to wages. This article demonstrates, that when neoclassical models of fertility are estimated on microdata using methods that incorporate key demographic restrictions and when they are properly aggregated, they have considerable forecasting power.

James J. Heckman
Department of Economics
Yale University
New Haven, CT 06520

James R. Walker
Department of Economics
University of Wisconsin
Madison, WI 53706

1. INTRODUCTION

This article demonstrates the value of microdata for understanding the effect of wages on fertility dynamics. Neoclassical economic models of fertility assign a central role to male income and female wages, yet empirical evidence on the importance of these variables as predictors of fertility is weak. Much of the available evidence is based on linear time series regression analyses of aggregate data (Butz and Ward 1979, Ward and Butz 1980, Wilkinson 1973). McDonald (1981) noted that when serial correlation is properly accounted for in such models, the coefficients of the wage and income variables often become statistically insignificant. McDonald also noted that aggregative neoclassical models are inferior in their forecasting power when compared with conventional time series models that assign no role to wages and incomes. This negative assessment of the value of neoclassical fertility models is widely held in the demographic community (Little 1981).

Our article argues that this assessment may be premature. The apparent inferiority of aggregate neoclassical time series models of fertility may reflect only the informational loss from using aggregate data and may not provide evidence about the value of economic models for forecasting. Analysis of aggregate data in this context is especially dangerous, because the underlying microrelationships between the economic variables and fertility are intrinsically nonlinear. Wage and income coefficients of an aggregate linear time series regression are not simply related to the wage and income coefficients of the underlying microequations. This article demonstrates that when neoclassical models of fertility are estimated from microdata and are properly aggregated, they have considerable forecasting power.

This article draws its microestimates from a companion paper (Heckman and Walker [1989]) in which we estimate the effects of wages and income on life cycle fertility dynamics using microdata for four five-year birth cohorts of Swedish women. In our companion paper we formulate and estimate the parameters of a micro stochastic process of individual fertility dynamics. From our model we are able to determine what features of fertility are affected by economic variables. From a common set of estimated parameters that characterize the full reproductive process, we can predict any dimension of life cycle fertility: completed fertility, length of interbirth intervals, and parity-, age-, or period-specific birth rates. In this article, we simulate the best fitting microfertility model estimated in our earlier work to predict aggregate annual birth rates for each cohort. We use the predicted aggregate rates to take up the challenge implicit in McDonald's work and ask how well our aggregated nonlinear neoclassical model of fertility forecasts aggregate time series and how well it competes with linear neoclassical and autoregressive models fit on aggregate data.

In our companion paper we find strong evidence of wage and income effects consistent with neoclassical models of the timing and spacing of births. In this article we find that linear neoclassical models fit on aggregate data are sensitive to the treatment of serial correlation. This aspect of our work reproduces McDonald's findings. We find, however, that the best fitting micro nonlinear model of the timing and spacing of births when aggregated performs better than autoregressive models in predicting fertility rates for most cohorts.

This article proceeds in the following way. Section 2 summarizes the empirical results reported in our companion paper on the effects of wages and income on the timing and spacing of births. Section 3 presents linear regression estimates of the neoclassical model fit on aggregate time series data. Section 4 discusses the procedures used to aggregate microfertility equations into annual aggregate birth rates. The predictive ability of the micro model for forecasting aggregate period-specific birth rates is assessed. We compare the forecasting power of the aggregated model with the forecasting power of linear autoregressive models without economic variables. The article concludes with a summary.

2. ESTIMATES FROM A NEOCLASSICAL MICRODYNAMIC MODEL OF SWEDISH FERTILITY

This section summarizes the findings reported in Heckman and Walker (1989) on the microrelationship between wages and the timing and spacing of births in Sweden. Our summary is in three parts. We first present the specification of the stochastic birth process model used to estimate life cycle fertility behavior. We then discuss the microdata analyzed and place our estimates in context with a brief overview of the relevant Swedish economic environment. We conclude this section with a brief discussion of the main empirical findings from our previous work.

2.1 A Stochastic Birth Process

A stochastic birth process model is the natural scheme for describing fertility dynamics. The model represents the reproductive process in terms of random waiting times between births. The initial conditions of the process and the probability mechanism generating the waiting times between births completely specify the model.

We use the hazard function (also called the occurrence rate) to characterize the reproductive process. We assume that the sample paths of birth durations conditional on the regressors are absolutely continuous random variables (with respect to Lebesgue measure), so they have a well-defined conditional density. To incorporate the effect of regressors, we model the conditional hazard, where the conditioning is on regressors as well as the survivor time. We parameterize the hazard function for the k^{th} birth conditional on the (possibly time varying) regressors Z in the following way:

$$h_k(t_k|Z) = \exp(\alpha_{ko} + \sum_{j=1}^{2} \alpha_{k1j}(t_k^{\eta_{kj}} - 1)/\eta_{kj} + Z'\beta_k), \quad \eta_{k1} < \eta_{k2} \qquad (1)$$

where the vector Z can include any known functions of the regressors. Yashin and Arjas (1988) demonstrated that absolute continuity of conditional duration times is sufficient to justify the widely used exponential formula connecting the integrated conditional hazard function and the survivor function. They also presented alternative criteria that justify the exponential formula. Parity dependence is incorporated by estimating a separate set of coefficients for each parity. The middle term on the right side of equation (1) captures duration dependence and encompasses common specifications frequently used in economics and demography (Heckman and Walker 1987).

In our empirical analysis we also control for potential misspecification errors caused by unobserved variables that, if measured, would be included in Z. Fecundability, which differs across women is a natural source of such variables. We estimate birth process models using an extension of the nonparametric maximum likelihood estimator of Heckman and Singer (1984) to multiple spell models. Our extension expands their estimator to include defective duration distributions (Heckman and Walker 1987). This extension allows for parity-specific sterility. Upon obtaining parity k some fraction of women (say π_k) may become or choose to become permanently sterile, whereas $1-\pi_k$ remain at risk to progress to the next parity. Although it is possible to parameterize the stopping proportions, π_k, to depend on regressors (see Heckman and Walker 1987), the models whose estimates are reported in this article do not do so.

A variety of specification tests are available to select the preferred or best fitting model. In multismultistate duration analysis, model selection is difficult because the analyst must often choose among nonnested specifications. To evaluate competing specifications we use a predictive criterion based on the classical χ^2 goodness-of-fit test. The test compares the predicted parity distribution from a fitted model with the parity distribution observed in the sample. We apply the test at four ages (20, 25, 30 and 35). Significant departures between fitted and actual parity distributions at any one of the ages tested is evidence of model misspecification. Although our test examines only one aspect of a multivariate duration distribution, we find it to be empirically powerful: only one specification with economic variables passes the full set of goodness-of-fit tests for all cohorts at all ages.

2.2 Microdata and the Swedish Context

The data analyzed in this article and our companion paper are from the Swedish Fertility Survey. It is a retrospective survey of native born Swedish women conducted by Statistics Sweden in 1981. The data consist of random samples from four five-year birth cohorts (1936-1940,...,1951-1955) with approximately 500 women in the first cohort and 1,000 women in the other three cohorts. In addition to information on the characteristics of their childhood households, respondents gave complete event histories on their births, marriages, and cohabitation (consensual unions) and on their labor force participation. Hoem and Rennermalm (1985) documented the accuracy of the data.

The survey did not gather individual wage and income information. To circumvent this problem we construct for each gender a time series of income using summary measures of personal tax returns by age for selected years published by Statistics Sweden. The selected years are 1960, 1965, 1970, 1974, 1975, 1976, and 1977. We regress these income measures on calendar time toime to generate a complete age-and gender-specific time series of predicted annual income. See appendix B of Heckman and Walker (1989) for more details. The predicted male income series is used in our empirical analysis. We use the predicted female income series and annual hours of work information from the Swedish Fertility Survey to obtain an age-specific time series of female wage rates.

During the sample period (1948-1981), the Swedish labor market was characterized by centralized wage setting and explicit egalitarian wage policies. Operationally, the wage policies compressed interindividual wage variation by increasing the rate of growth of wages of low-wage workers (primarily women) while reducing the rate of growth of wages of high-wage workers and workers in profitable industries (Bjorklund 1986; Flanagan 1987). This institutional structure of wage determination has two important implications for our microanalysis. First, it implies that the observed dramatic change in relative female wages from 65% to 90% of male wages in the manufacturing sector during the period 1950-1980 is due to an exogenous force. Second, it lends credibility to our use of aggregate wages in an analysis of individual fertility histories, since aggregate wage policy uniformly applied accounts for much of the wage growth of individuals.

During the postwar period, the Swedish government enacted a variety of pronatal child care programs and other social programs designed to promote the equality of the sexes in the marketplace and in the home. Many of these programs were only available to working women (or to women who had some work experience), and the benefits of such programs became increasingly more generous over the course of our sample period. These programs tend to weaken the negative effect of female wages on fertility. Higher-wage women are more likely to work and hence qualify for child support payments.

2.3 Empirical Evidence

In Heckman and Walker (1989) we present an extensive empirical analysis of the individual fertility histories found in the 1981 Swedish Fertility Survey. The beginning of the observational period is age 13 for all women. Our baseline hazard specification includes female wage and male income variables, a set of background variables that are intended to control for the initial conditions of the fertility process and a parity-specific mover-stayer representation of unobserved heterogeneity or omitted variables. Such a model allows unobservables to assume two values. One value implies that the individual does not go on to have additional births. The other value implies that the woman may go on to have additional births. For each parity we estimate the proportion of women of both types and the value of the unobservable for the women who may go on to have births. In Heckman and Walker (1989), we determine that this specification of unobservables is the one that produces the best- fitting models. It is important to note that this specification implies that unobservables are distributed independently across spells even for the same woman.

According to neoclassical theory and common sense, a woman's access to a male partner's income is expected to have a positive effect on the probability of her giving birth. An increase in the female wage is expected to reduce the probability of a birth. In the jargon of neoclassical economics, the substitution effect dominates the income effect of a wage change on fertility.

Table 1 reports the estimated wage and income coefficients and associated asymptotic standard errors of the best-fitting version of the parameters of equation (1) for each cohort for the first three births, which effectively account for all births in Sweden and for all births in our sample. The pattern of the estimated coefficients in Table 1 is representative of the many models we fit. Notice that the estimated wage and income effects are consistent with neoclassical theory. Estimated female wage effects are statistically significant and numerically important in explaining the timing and spacing of births. Higher wages lead to fewer births and to the postponement of births. In results reported in our companion paper, we find that these estimated wage effects are robust to the inclusion or exclusion of education, marital status, time trend, age, unemployment, and policy variables. Estimated male wage effects are also consistent with neoclassical theory but are less robust to the inclusion or exclusion of marital status variables. Higher male income leads to more births and closer spacing of births. Our estimates are robust to the inclusion or exclusion of other control variables.

We also estimate models using a second time series of wages-gender-specific real average manufacturing wages. No detail by age is available for this series, but it is available on an annual basis. Wilkinson (1973) used these wages in his time series study of Swedish fertility. The manufacturing wage series is the only gender-specific wage series available from published sources for the entire period under consideration. Estimated effects of wages based on manufacturing wages are qualitatively the same as those based on the age-and period-specific tax table wages. The concordance of the estimates based on the two wage series suggests that interpolation in the first series produces little bias. Models using the second wage series, however, fail goodness-of-fit tests.

In a limited Monte Carlo study we investigate and reject the hypothesis that our estimated wage effects are the result of an errors-in-variables problem arising from using aggregated instead of individual wages and income. Our analysis also rejects the hypothesis that the estimated wage and income effects are the manifestation of a spurious regressions phenomenon due to the use of time-trended wage and income variables. Estimated wage effects on fertility are also insensitive to the specification of duration dependence and unobserved heterogeneity. See appendix C of Heckman and Walker [1989].

The intercohort pattern of the estimated wage and income effects is consistent with the introduction and enhancement of work-conditioned child benefit programs during the sample period. In Table 1 notice that the estimated female wage effect decreases (in absolute value) across cohorts for all birth orders. Wald tests using the unrestricted parameters reject the hypothesis of equality of wage coefficients across cohorts for all birth orders. Because market work entitles women in the later birth cohorts to larger child benefits, the estimated female wage coefficient reflects less of a substitution effect and more of an income effect on fertility for later cohorts of women. Moving from the earlier to the later cohorts, the estimated male income effect decreases for the first birth interval, increases for the second birth interval, and exhibits little pattern for the third birth interval. Wald tests indicate that estimated male income effects are statistically significantly different across cohorts only for the first birth interval. This evidence suggests that the growth in the availability of child care benefits permits later cohorts of women to be increasingly less dependent on their partner's income to initiate the fertility process. Because of the drift in the estimated parameters across cohorts, we are unable to pool the four cohorts into a common model of life cycle fertility. Our estimates suggest the potential danger in pooling data across cohorts of women, as is done in many aggregate fertility analyses.

These estimates and the estimates reported in Heckman and Walker (1989) provide evidence in support of neoclassical economic models of fertility. This evidence is in sharp contrast to the evidence from linear time series regression models fit on aggregated data, criticized by McDonald (1981). We compare aggregated versions of our model with aggregate linear time series models in Section 4. We first discuss estimates of linear aggregative models.

3. ESTIMATES OF LINEAR NEOCLASSICAL MODELS OF FERTILITY

In this section we report estimates of linear neoclassical models of fertility fit on aggregate data. We structure the aggregate analysis to match as closely as possible the preceding micro-analysis. Our study provides evidence on the aggregation bias that arises from using linear models to approximate an intrinsically nonlinear microdynamic birth process. To maintain comparability with the micro-analysis, we aggregate birth histories in the Swedish Fertility Survey to obtain calendar-year-specific birth rates; we use the same wage and income series as covariates and fit separate regressions for each cohort.

Table 2 reports the results of an aggregate time series analysis of the logit of year-specific birth probabilities. We follow Lee (1981) and exploit basic demographic information by estimating rates rather than levels. The number of women at risk to give birth is known and hence need not be estimated. Use of the logit transformation imposes the constraint that women cannot have more than one or fewer than zero births in any year.

The upper panel reveals that linear regression models have the expected signs and significance levels that accord with neoclassical priors. The Durbin-Watson statistics, however, indicate serial correlation problems. The lower panel shows that correcting for first-order serial correlation robs the linear neoclassical model of any statistical precision. Estimated wage coefficients are largely insignificant or "perverse." Table 3 duplicates the format of Table 2, except that for the models with estimates reported in this table male income is interacted with an aggregate marital status variable to produce a model that is similar to the birth process model fit in our microanalysis. The interaction of marital status and male income variables captures the notion that male income is relevant only when women are married. The results in Table 3 mirror those in Table 2. So do results for models fit aggregating within cohorts for each year. Our estimates of these models exhibit the same sensitivity to adjustments for serial correlation noted by McDonald (1981). These results vividly illustrate the danger of using linear models on aggregate data to test the validity of an intrinsically nonlinear microdynamic model.

4. EVALUATION OF THE TIME SERIES PROPERTIES OF THE MICROMODEL

In this section we return to the microdynamic model of fertility and ask how well it forecasts aggregate period-specific birth rates. We assess its forecasting ability both in absolute and in relative terms. In absolute terms, we compare predicted annual birth rates from the birth process model with aggregate annual birth rates observed in the sample. This comparison is an absolute measure of the performance of the birth process model, since a correctly specified micro-model of fertility should be able to predict the level and the movement over time of aggregate birth rates. To measure the forecasting ability of the micromodel in relative terms we compare the within- sample mean squared error of its predicted annual birth rates with that obtained from autoregressive models fit on aggregate data. The recent work of McDonald (1981) suggests that autoregressive moving average (ARMA) models without economic variables dominate linear neoclassical models on a mean squared error criterion. We investigate whether this dominance relationship also applies to the aggregated version of our estimated nonlinear microdynamic model of fertility.

4.1 Aggregation of the Microdynamic Fertility Model

Before presenting our empirical results we describe the procedure used to aggregate the estimated birth process model to produce period-specific (annual) birth rates. The aggregation procedure uses the fact that our estimated birth process model provides a complete description of the stochastic process generating life cycle fertility. To obtain period-specific birth rates from our estimates we must transform the internal time of the birth process (i.e., the sum of all past and current birth intervals) into the external time of the calendar. This is accomplished once we set the calendar date of the start of the process. Recall that our data allow the observational period to begin at age 13. Our empirical results are not affected when we use later starting ages in the range 14-17 to estimate the model.

To obtain predicted annual birth rates for a cohort, we calculate for each individual in the cohort her predicted annual birth rate by integrating the estimated birth process model over the sample period using the individual's observed covariate path. We produce calendar year rates for each woman using the date of onset of the process, her previous birth intervals, and incomplete current spells to bring her to the current year. We sum these individual rates over all members of the cohort to produce the estimated aggregate rate.

More formally, let $\tau[k]$ denote the calendar date of the k^{th} birth, with $\tau[0]$ equal to the calendar date of the initiation of the process. For expositional convenience assume that the process starts at calendar time zero, $\tau[0]=0$. Consider then the probability that woman i, with observed covariate path Z, will have a birth at calendar time τ. Virtually all Swedish women have three of fewer births. By the law of total probability,

$$\hat{B}_i(\tau) = \Pr(\text{woman } i \text{ has a birth at calendar time } \tau \mid Z)$$

$$= \sum_{k=1}^{3} h_k(\tau - \tau[k-1] \mid Z) \cdot P_{k-1}(\tau \mid Z) \qquad (2)$$

where $h_k(\tau-\tau[k-1] \mid \cdot)$ is the occurrence rate of a birth of order k at duration $\tau-\tau[k-1]$ and $P_{k-1}(\tau \mid \cdot)$ is the probability that the woman is at risk to have a birth of order k at calendar time τ conditional on the observed covariate path. We use the identity that calendar time at age 13 plus the sum of previous birth intervals and the current incomplete interval give the current calendar year at which we evaluate the birth rate.

Aggregate birth rates are reported on an annual basis. Equation (1) is defined, however, for a point in time for each individual in a cohort. Therefore, for each individual we aggregate the predicted instantaneous rates into annual rates. Summation of the individual predicted birth rates yields the predicted annual birth rates for the cohort.

Evaluation at an arbitrary time τ of the probability of being at risk for the k^{th} birth is computationally demanding. For example, the probability at time τ that woman i is at risk for the third birth is

$$P_3(\tau \mid \cdot) = \int_0^{\tau} \int_0^{\tau-u_1} f_1(u_1 \mid \cdot) f_2(u_2 \mid u_1, \cdot) S_3(\tau-u_1-u_2 \mid u_1, u_2, \cdot)\, du_2\, du_1 \qquad (3)$$

where $f_k(t)$ is the conditional density function for the waiting time for the k^{th} birth and $S_k(t)$ is the conditional survivor function for parity k.

Direct numerical integration of expressions such as equation (3) is prohibitively costly. An attractive alternative to numerical integration is Monte Carlo integration. By this procedure we replicate each individual's fertility history a large number of times using the estimated birth process model and the individual's observed covariate path. Note that the best-fitting models that we simulate exclude lagged durations from the covariate set. The replications are then averaged to obtain the individual's predicted period-specific birth rates.

It is well known that Monte Carlo procedures introduce sampling error into the evaluation of the integral. By selecting the number of replications to be sufficiently large, any desired level of numerical accuracy can be obtained. We choose 100 replications per observation. This produces three decimal places of accuracy. By sequentially increasing the number of replications, we determined that after 100 replications per individual the first three decimal places do not change in estimating the cohort's parity probabilities $P_k(\tau|\cdot)$, $k = 0, 1, 2, 3$ for ages 20, 25, 30, and 35.

4.2 Tests of the Time Series Properties of the Micromodel

Tests of the time series properties of the microdynamic model of fertility use the predicted aggregate calendar year birth rates in the following way. We subtract the rates predicted from the birth process model from the sample rates for the cohort to obtain period-and cohort-specific prediction errors. We investigate autocorrelation patterns in these residuals. Our analysis does not directly respond to the recent work of McDonald (1981). His study compared the stability and quality of forecasts from linear neoclassical models of Australian first births in levels with the stability and forecast power of conventional time series models that assign no role to wages and incomes. We follow the advice of Lee (1981) and focus on forecasting birth <u>rates</u>, whereas McDonald focused on forecasting birth <u>levels</u>, ignoring the information available from the number of women alive by age. He analyzed first births over all cohorts of women, whereas we consider the first three births of four cohorts of women. Fourth births are rare in Sweden, and there are very few in the Swedish Fertility Survey. Our evidence of cohort drift in Sweden suggests that McDonald's demographic dependent variable, which is aggregated over all cohorts, is unlikely to be explained by any stable economic model based solely on wages and incomes.

MaCurdy (1986) established that under the null hypothesis of no serial correlation, valid large-sample tests of serial dependence in the residuals can be conducted ignoring the contribution of estimation error to the fitted residuals. His results apply without modification to our model, because our estimated parameters are $\sqrt{N}$ consistent where N is cross section sample size.

Table 4 presents an analysis of the residuals defined as the difference between predicted and actual birth rates by cohort. For each cohort first-and second-order autoregressive models are fit. Except for Cohort 2, there is strong evidence that the best-fitting model predicts the level of fertility since the estimated intercept terms are not statistically significantly different from 0. Only the model for Cohort 2 exhibits first-order serial correlation in the residuals. There is no evidence of serial correlation in the residuals from the models fit for the other cohorts.

Table 5 reports mean squared error calculations for the birth process model. Table 6 reports mean squared error calculations for autoregressive models, with one and two lags estimated in the logits of the birth rates. The mean squared error in both tables is defined for the <u>logits</u> of the birth rates. For our data, N is large relative to T (the number of time series observations) and assuming both N and T become large, under standard conditions estimation error can be ignored in comparing aggregate time series mean squared errors across cohorts.

Compare the mean squared error from the birth process model presented in Table 5 with the mean squared error for the time series model in Table 6. Except for the Cohort 2 model, the aggregated neoclassical nonlinear models have <u>lower</u> mean squared error than their autoregressive counterparts.

These results indicate that the aggregated micro birth process model survives conventional time series specification tests except for the model fit for Cohort 2. That model is noted in our companion paper to <u>fail</u> our goodness-of-fit test. The aggregated nonlinear neoclassical model generally outperforms autoregressive models in making forecasts. Our results illustrate the value of having access to disaggregated data in estimating and testing nonlinear microdynamic models. They demonstrate the danger in ignoring fundamental nonlinearities intrinsic to the theory in assessing the value of the neoclassical theory of fertility. Micromodels that pass goodness-of-fit tests, when aggregated, produce reliable neoclassical forecasting equations.

5. SUMMARY AND CONCLUSIONS

We examine the ability of microdynamic models to account for time series variation in cohort fertility. This exercise is especially interesting in light of recent negative assessments of the value of neoclassical theory for predicting births (McDonald 1981). We find that, for most cohorts of Swedish women, our estimated models pass the time series specification tests used by McDonald (1981) to discredit linearized versions of neoclassical models applied to Australian data. The micromodel that fails to predict the aggregate time series also fails to pass micro goodness-of-fit tests. Aggregated neoclassical microdynamic models explain the time series better, in the sense of mean squared error of forecast, than do time series autoregressions. The poor previous performance of neoclassical linear time series models may well be a consequence of the poor quality of linearized aggregate models as approximations to an intrinsically nonlinear microdynamic model.

Table 1
Estimated Wage and Income Coefficients
for a Stochastic Birth Process Model of Individual Fertility

Birth Interval	Cohort 1 1936-40 Est.	S.E.	Cohort 2 1941-45 Est.	S.E.	Cohort 3 1946-50 Est.	S.E.	Cohort 4 1951-55 Est.	S.E.
<u>First</u>:								
male income	1.09	.072	.842	.047	.993	.060	.881	.059
female wage	-4.74	.574	-3.96	.307	-3.38	.273	-2.69	.258
<u>Second</u>:								
male income	.577	.096	.656	.078	.771	.093	.848	.123
female wage	-3.03	.443	-2.60	.265	-1.95	.247	-1.82	.298
<u>Third</u>:								
male income	.506	.271	.236	.111	.062	.119	.812	.312
female wage	-4.87	.995	-3.07	.375	-1.99	.370	-2.90	.866

Notes:

Source: Table 3 of Heckman and Walker (1989).

Weibull Birth Process Model, ($\eta_{k11} = 0$, $\alpha_{k12} = 0$, k=1,2,3 in equation 1) with mover-stayer unobserved heterogeneity controls.

Wage and income variables are age and period specific (described in text).

Other covariates included in the specification: urban (a dummy variable which equals one if the respondent was raised in an urban area) and white collar (a dummy variable which equals one if the respondent's father had a white collar job).

Table 2
Aggregate Time Series Regression
Dependent Variable = logit (probability of a birth)[a]
Ordinary Least Squares

	Cohort							
	1936-40		1941-45		1946-50		1951-55	
Variable	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
const	-2.31	.131	-2.50	.131	-2.79	.097	-3.53	.078
female wage	-.280	.024	-.127	.030	-.009	.030	-.011	.300
male income	.069	.009	.044	.013	-.017	.013	.043	.016
Mean Square Error	.507		.448		.196		.082	
Durbin-Watson[d]	.606		.437		.488		.735	
# of obs[c]	140		115		90		65	

First Order Autocorrelation Correction

	Cohort							
	1936-40		1941-45		1946-50		1951-55	
Variable	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
const	-2.43	.308	-3.27	.361	-3.20	.209	-3.70	.128
female wage	-.127	.032	.048	.030	.048	.024	-.008	.028
male income	.019	.012	-.010	.011	-.001	.010	.038	.014
ρ[b]	.809	.048	.894	.039	.821	.059	.658	.096
Mean Square Error	.232		.149		.082		.049	
Durbin-Watson[d]	1.88		1.62		1.63		1.56	
# of obs[c]	140		115		90		65	

Notes:
[a]dependent variable = $\log(\pi[a,t]) - \log(1-\pi[a,t])$
$\pi(a,t)$ = probability of a birth by a woman age a in year t
[b] ρ = first order autocorrelation coefficient
[c]# of obs = number of age and period specific observations
The age and period specific birth probabilities are derived from the 1981 Swedish Fertility Survey. The age-specific wage and income series are described in the text.
[d] = the same inference is obtained from the theoretically appropriate Durbin h statistic for each case it is defined.

Table 3
Aggregate Time Series Regression
Dependent Variable = logit (probability of a birth)[a]
Male Income Interacted with Proportion Married or Cohabiting
Ordinary Least Squares

Variable	1936-40 Est.	1936-40 S.E.	1941-45 Est.	1941-45 S.E.	1946-50 Est.	1946-50 S.E.	1951-55 Est.	1951-55 S.E.
const	-1.74	.118	-2.28	.142	-2.87	.133	-3.60	.155
female wage	-.225	.026	-.045	.031	.005	.025	.077	.023
male income *prop[c]	.044	.009	.009	.011	-.011	.010	-.002	.011
Mean Square Error	.609		.491		.197		.091	
Durbin-Watson[e]	.497		.364		.448		.734	
# of obs[d]	140		115		90		65	

First Order Autocorrelation Correction

Variable	1936-40 Est.	1936-40 S.E.	1941-45 Est.	1941-45 S.E.	1946-50 Est.	1946-50 S.E.	1951-55 Est.	1951-55 S.E.
const	-2.35	.332	-3.34	.351	-3.30	.219	-3.84	.201
female wage	-.075	.030	.072	.026	.048	.069	.008	.023
male income *prop[c]	-.003	.011	-.020	.009	-.010	.007	.002	.010
ρ[b]	.85	.04	.90	.04	.83.	.06	.74	.09
Mean Square Error	.233		.142		.080		.041	
Durbin-Watson[e]	1.86		1.60		1.61		1.54	
# of obs[d]	140		115		90		65	

Notes:
[a]dependent variable = $\log(\pi[a,t]) - \log(1-\pi[a,t])$
$\pi[a,t]$ = probability of birth to a woman age a in year t.
[b] ρ = first order autocorrelation coefficient
[c]Male income * prop = estimated male income interacted with proportion women married or cohabiting.
[d]# of obs = number of age and period specific observations
The age and period specific birth probabilities are derived from the 1981 Swedish Fertility Survey. The age-specific wage and income series are described in the text.
[e] = the same inference is obtained from the theoretically appropriate Durbin h statistic for each case it is defined.

Table 4
Autoregressions of the Difference Between Calendar Year Sample and Predicted Birth Probabilities by Cohort

	First Order[a]		Second Order[b]	
Cohort 1 (Born 1936-40)	Est.	S.E.	Est.	S.E.
α	-.001	.002	-.001	.002
$\rho(1)$	.049	.189	.045	.196
$\rho(2)$	———		.077	.196
DW[c]	2.00		1.99	
σ^2[d]	.0091		.0094	
N[e]	30		29	
Cohort 2 (Born 1941-45)				
α	-.118	.002	-.000	.002
$\rho(1)$	.464	.185	.497	.220
$\rho(2)$	———		-.069	.223
DW[c]	1.92		2.01	
σ^2[d]	.0102		.0107	
N[e]	25		24	
Cohort 3 (Born 1946-50)				
α	-.001	.002	-.001	.002
$\rho(1)$	-.385	.218	-.363	.250
$\rho(2)$	———		.056	.250
DW[c]	1.95		2.01	
σ^2[d]	.0088		.0094	
N[e]	20		19	
Cohort 4 (Born 1951-55)				
α	-.001	.002	-.001	.002
$\rho(1)$	-.151	.277	-.182	.294
$\rho(2)$	———		-.236	.307
DW[c]	2.05		1.89	
σ^2[d]	.0067		.0071	
N[e]	15		14	

Notes
[a]First Order Model: $e_t = \alpha + \rho(1)e_{t-1} + u_t$
[b]Second Order Model: $e_t = \alpha + \rho(1)e_{t-1} + \rho(2)e_{t-2} + u_t$
[c]DW = Durbin-Watson Statistic. The same inference is obtained from the theoretically appropriate Durbin h statistic for each case it is defined.
[d]σ^2 = Estimated Var(u_t)
[e]N = Number of years (observations)

Table 5
Mean Square Error of Sample Versus Predicted Calendar Birth Probabilites

Panel A

	Cohort Birth Years	Number of Periods	MSE[a]
Cohort 1	1936-40	29	0.105
Cohort 2	1941-45	23	0.060
Cohort 3	1946-50	19	0.023
Cohort 4	1950-51	14	0.012

Panel B

ARMA Models

	Number of Periods	MSE One Lag	MSE Two Lags
Cohort 1	29	0.216	0.233
Cohort 2	23	0.047	0.042
Cohort 3	19	0.048	0.035
Cohort 4	14	0.035	0.031

Note: [a]This MSE is computed by taking the sum of the squared difference between the logit of sample birth rates and the logit of predicted birth rates and dividing by the number of annual observations. Periods with no sample births are deleted to make the logit transform well defined. They are always initial periods for each cohort.

References

Bjorklund, A. (1986), "Assessing the Decline of Wage Dispersion in Sweden," unpublished manuscript, Industrial Institute For Economic and Social Research, Stockholm.

Butz, W., and Ward, M. (1979), "The Emergence of Countercyclical U.S. Fertility," American Economic Review, 6, 318-328.

Flanagan, R. (1987), "Efficiency and Equality in Swedish Labor Markets," in The Swedish Economy, eds. B. Bosworth and A. Rivlin, Washington, DC.: Brookings Institute, pp. 125-184.

Heckman, J. and Singer, B. (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica, 52, 63-132.

Heckman, J., and Walker, J. (1987), "Using Goodness of Fit and Other Criteria to Choose Among Competing Duration Models: A Case Of Hutterite Data," in Sociological Methodology 1987, ed. C. Clogg, Washington, DC.: American Sociological Association, pp. 247-307.

———, (1989), "Life Cycle Fertility Dynamics: Evidence From Swedish Data," unpublished manuscript, Department of Economics, Yale University

Hoem, J. and Rennermalm, B. (1985), "Modern Family Initiation In Sweden: Experience of Women Born Between 1936 and 1960," European Journal of Population, 1, 81-112.

Lee, R. (1981), Comment on "Modeling Demographic Relationships," by J. McDonald, Journal of the American Statistical Association, 76, 793-795.

Little, R. (1981), Comment on "Modeling Demographic Relationships," by J. McDonald, Journal of the American Statistical Association, 76, 795-796.

MaCurdy, T. (1986), "A Guide To Applying Time Series Models To Panel Data," unpublished manuscript, Stanford University, Department of Economics, Palo Alto, California.

McDonald, J. (1981), "Modeling Demographic Relationships: An Analysis of Forecast Functions for Australian Births," Journal of the American Statistical Association, 76, 782-793.

Ward, M., and Butz, W. (1980), "Completed Fertility and Its Timing," Journal of Political Economy, 88, 917-940.

Wilkinson, M. (1973), "An Econometric Analysis of Fertility in Sweden, 1870-1965," Econometrica, 41, 633-42.

Yashin, A., and Arjas, E. (1988), "A Note on Random Intensities and Conditional Survivor Functions," Journal of Applied Probability, 25, 630-635.