SOCIAL SCIENCE RESEARCH 9, 225–244 (1980)

Separating Age, Period, and Cohort Effects in White U.S. Fertility, 1920–1970

THOMAS W. PULLUM

Center for Studies in Demography and Ecology, Department of Sociology, University of Washington

Four models are developed to describe the odds transformation of period- and age-specific fertility rates as products of age, period, and cohort effects. These are applied to data for white U.S. women age 15–44 from 1920 to 1970, with equal weights given to each rate. All models which include age fit subsets of the data extremely well. Per effect, the incorporation of periods improves the fit much more than the incorporation of cohorts. It is shown that first differences are invariant in two-effect models, and second differences are invariant in the three-effect models.

The fertility of an aggregate over time is frequently described in terms of an array of age- and period-specific fertility rates. Ideally, the numerators for these rates arise in a birth registration system and the denominators come from census counts or intercensal estimates of numbers of women. When such sources are not available, the birth histories from a survey may be used, in which case the array will be missing a triangle of rates corresponding to the older ages at far-removed time periods.

The quality of estimates obtained from these two sources is a topic of considerable interest. In the present paper such questions will be put aside, and discussion will be confined to the uses of these arrays of rates, taking them at face value. It should be understood, but will not be repeated, that the accuracy of one's conclusions will be impaired to the extent that the estimates are of poor quality.

An array of fertility rates from a socially defined, relatively homogeneous population contains three structural dimensions. The first is age, which we shall vary across rows. This dimension has been focused upon in discussions of natural fertility, as by Henry (1961), and in the development of model fertility tables by Coale and Trussell (1974). Stable population theory is based on the key assumption that the age-specific schedule of fertility (and of mortality, as well) is fixed. Second, much research

Address requests for reprints to Dr. Pullum, Center for Studies in Demography and Ecology, Department of Sociology, University of Washington, Seattle, WA 98195.

0049-089X/80/030225-20$02.00/0

focuses on the time dimension, represented by the columns of the array. Through such summaries as the Total Fertility Rate (TFR), trends across time are discerned. It is probably fair to say that policy makers in nearly all countries are most interested in changes across time, and in how these may be related to family planning programs and economic indicators (see, for example, Butz and Ward, 1977).

The third dimension is the birth cohort, represented by the diagonals of the array. Some researchers, such as Lee (1974), have placed particular emphasis upon the cohort identification. The importance of the concept of the cohort as a theoretical tool has been made clear in several contexts by Ryder (1965) and has been applied to fertility by Easterlin (1973) as well.

The U.S. Department of Health, Education, and Welfare recently (Heuser, 1976) issued estimates of age- and period-specific rates for the United States for single years of age and years 1917–1973. We shall use the data for white women age 15–44 (at time of childbirth) to investigate the following questions:

1. From a statistical point of view, what is gained by employing all three dimensions simultaneously, rather than just two of them?
2. When all three dimensions are included, is it possible to disentangle their separate roles?

THE QUANTITY TO BE MODELED

The basic age- and period-specific fertility rate is defined to be the number of births in an interval of time to women in an interval of age (at date of childbirth), divided by the number of woman-years of exposure to the specified age × time interval. This rate will be termed p_{ij}, where i refers to an age interval and j to a time interval. It is approximately interpreted as the proportion of women in age interval i during time interval j who give birth. By extension, it is an approximate estimate of the probability that a woman who is at the beginning of the age interval at the beginning of the time interval will have a child in the next year. The reasons why these rates are only estimates of proportions and of probabilities are well known (see, for example, Spiegelman, 1968, pp. 254–258) and will not concern us.

In this paper, p_{ij} will be interpreted as the proportion of woman-years of exposure to age interval i and time interval j which result in a live birth. This interpretation does not involve any approximations and translates into an estimate of the probability that a woman-year of the specified type will produce a birth.

Given this interpretation, it is natural to conceptualize an $R \times C$ table of rates (with R age intervals and C time intervals) as an $R \times C \times 2$ table of frequencies, in which the two panels refer to the dichotomous response variable of whether or not a woman-year of exposure resulted in a live birth. In other words, the two-way table of rates p_{ij} may be regarded as simply a collapsed form of a three-way table of multinomially distributed frequencies n_{ijk}, such that $k = 1$ if a woman-year of exposure produced a live birth and $k = 2$ if it did not. Specifically, $p_{ij} = n_{ij1}/n_{ij\cdot}$ where $n_{ij\cdot} = n_{ij1} + n_{ij2}$ is the number of woman-years in the specified age × time category.

It is generally recognized that the proper quantity for statistical modeling in such a multinomial context will be the odds $f_{ij} = n_{ij1}/n_{ij2}$ rather than the proportion. (See, for example, Goodman, 1972a; and Bishop, Fienberg, and Holland, 1975.) Thus, we shall work with the odds, which are related to the proportion through $f_{ij} = p_{ij}/(1 - p_{ij})$, rather than with the age- and period-specific rate itself. Multiplicative models will be developed for these odds.

Because the use of the odds rather than the original rates is uncommon in fertility analysis, some justification is called for. We shall offer such a justification, which is also bound up with the use of multiplicative models on these odds. However, it is certainly not intended to be a proscription against the various alternatives.

The only three possibilities to be considered here are (a) additive models for the proportion or rate p_{ij}; (b) multiplicative models for p_{ij} (equivalently, additive models for $\log p_{ij}$); and (c) multiplicative models for the odds f_{ij} (equivalently, additive models for the logit, $\log f_{ij}$). An example of type (a) may be found in a path analysis by McKenna (1974; see also Pullum, 1976). The principal arguments against an additive model for rates are, first, that predictions outside the interval (0,1) can easily be generated and, second, that unless a complex interaction term is included, it is unrealistically assumed that the rate is linear in each predictor, with no dependence upon the levels of the other predictors.

A model of type (b) is found, for example, in Coale and Trussell's (1974) model fertility patterns. These assume that the marital fertility rate for a given age is proportional to that of a natural fertility population, and also to a damping factor which reflects the use of contraception, typically increasing with age. Osborn (1975) has applied models of type (b) to stillbirth rates which are classified according to mother's age, pregnancy order, etc. Except for a different form of dependent variable, his procedures are nearly identical to those given below for Model 1. This type of model (b) does not generate negative predictions, but it can give predictions greater than one, which will be difficult to interpret. The impact (on the rate) of a change in one predictor *does* depend upon the levels of the other predictors, but it does not depend (with type (b)) upon the nearness of the rate to the endpoints of the range (0,1).

If the rate may be clearly interpreted as an estimate of either a joint or a compound probability, and if it is decomposable as a product of independent or conditional probabilities, then a multiplicative model is clearly called for. In our particular application, that is not the case. In the Coale and Trussell (1974) application, as well, the fertility rate was not viewed as the estimate of a probability, and the product terms in the model were not viewed as probabilities. The fit which was obtained was excellent, but there is no reason to believe that a fit under type (c) would have been any less good.

Models of type (c) have been used in the context of demography by Brass *et al.* (1968) who linearly related the logit of the probability of dying in a fitted life table to the logit of the probability of dying in a standard model life table. The first advantage of a multiplicative model for $p/(1 - p)$ is that the predicted values of this ratio may be any positive number and the predicted p will still be in the range from 0 to 1. The most important advantage is a complete symmetry in the labeling of the criterion category. This advantage even applies in the context of fertility: is there any *a priori* reason why one should model the proportion of woman-years that produce a birth rather than the number that do *not* produce a birth? Models for p, whether additive or multiplicative, will not automatically bear any relationship at all to models for the complement, $1 - p$. But multiplicative models for $p/(1 - p)$ will be symmetric, so that a relabeling of the two response categories will simply result in a reciprocation of the terms in the model.

THE MODELS

We shall consider four of the seven possible multiplicative models for the odds which make use of age, period, and/or cohort effects. The three models to be omitted are those which assume only a single type of effect. Although these models will be presented in parametric form, most of the results will be independent of the selected parametrization.

The models involve constraining the frequencies in the "expected" $R \times C \times 2$ table to match those in the "observed" $R \times C \times 2$ table in certain ways. These constraints are completely analogous to the familiar procedure of matching the row totals and the column totals in an expected $R \times C$ table with those in an observed $R \times C$ table to which the hypothesis of independence is being applied. For more details, see Goodman (1972b) and Pullum (1977).

The first model, to be referred to as Model 1 (or the age × period model), is given as

$$F_{ij} = \eta\alpha_i\beta_j \tag{1}$$

for all combinations of i and j and for (unknown) parameters $\alpha_i (i = 1, \ldots, R)$ and $\beta_j (j = 1, \ldots, C)$. The α_i's and β_j's will be termed "age effects" and "period effects," respectively. F_{ij} is the expected value of the odds for combination (i,j) and η is a scaling factor of no interest. In the predicted $R \times C \times 2$ table, the following marginal frequencies will equal those in the "observed" $R \times C \times 2$ table:

(1a) the $R \times C$ case bases of women-years will be preserved;
(1b) the total number of births in each year of age (across all time periods) will be unchanged (equivalently, the overall age-specific rate for each age will be preserved); and
(1c) the total number of births in each year of time (across all ages) will be unchanged.

These conditions may be met by an iterative scaling procedure. It may be shown that when applied to an $R \times C \times 2$ table, this model will have $(R - 1)(C - 1)$ degrees of freedom. In Eq. (1), the RC odds will be expressed in terms of $RC - (R - 1)(C - 1) = R + C - 1 = 1 + (R - 1) + (C - 1)$ independent coefficients. One of these will be η. By arbitrarily assigning $\alpha_1 = 1$ and $\beta_1 = 1$, we may associate the other degrees of freedom with the estimates of the other α's and β's.

Model 2 is given as

$$F_{ij} = \eta \alpha_i \gamma_k \tag{2}$$

for all combinations of $i (i = 1, \ldots, R), j (j = 1, \ldots, C)$, and k (k = $1, \ldots, R + C - 1$), with the subscripts related by $k = j - i + R$. This will be termed the age $\times$ cohort model; the γ_k's are "cohort effects." It is fitted with these constraints:

(2a) same as (1a);
(2b) same as (1b); and
(2c) the total number of births in each cohort (across all years and ages given) will be unchanged.

This model has $(R - 1)(C - 2)$ degrees of freedom; in the parametric form, we arbitrarily assign $\alpha_1 = 1$ and $\gamma_R = 1$.

The third model (Model 3 or the period $\times$ cohort model) is included largely for the sake of completeness. It is

$$F_{ij} = \eta \beta_j \gamma_k \tag{3}$$

with i, j, k as in Model 2 and with these constraints:

(3a) same as (1a);
(3b) same as (1c); and
(3c) same as (2c).

It has $(R - 2)(C - 1)$ degrees of freedom; in the parametric form, $\beta_1 = 1$ and $\gamma_R = 1$ will permit estimation of η and the other β's and γ's.

The fourth model is the complete age $\times$ period $\times$ cohort model, given as Model 4:

$$F_{ij} = \eta \alpha_i \beta_j \gamma_k \tag{4}$$

for i, j, and k as above and these constraints:

(4a) same as (1a);
(4b) same as (1b);
(4c) same as (1c); and
(4d) same as (2c).

It has $(R - 2)(C - 2)$ degrees of freedom. A parametric form is developed by assigning four values of the parameters or functions thereof. For this purpose we assign $\alpha_1 = \beta_1 = \gamma_R = 1$ and any specific value of r in the following three equations:

$$\alpha_i(r) = \alpha_i^* r^{i-1} \text{ for all } i. \tag{5}$$

$$\beta_j(r) = \beta_j^* r^{j-1} \text{ for all } j.$$

$$\gamma_k(r) = \gamma_k^* r^{k-R} \text{ for all } k.$$

The set $\{\alpha_i^*; \beta_j^*; \gamma_k^*\}$ are obtained as the unique parameters subject to $\alpha_1 = \beta_1 = \gamma_R = \gamma_{R+1} = 1$.

The reasons for going into these details on the parametric solutions will become evident later.

The cohort terms in Eq. (4) may be viewed as a special variety of terms representing the interaction between age and period. It is well known that an $R \times C$ table of frequencies has $(R - 1)(C - 1)$ degrees of freedom remaining after the main effect, row effects, and colunn effects have been estimated. The same is true for an $R \times C$ table of odds: a model with $(R - 1)(C - 1)$ unconstrained elements in the set $[\delta_{ij}; i = 1, \ldots, R$ and $j = 1, \ldots, C]$ can always be estimated to provide an exact fit with

$$F_{ij} = \eta\alpha_i\beta_j\delta_{ij}. \tag{6}$$

The use of cohort effects γ_k is equivalent to the assumption that the set $\{\delta_{ij}\}$ in the saturated model (6) has a simpler representation, namely, that the interaction effects are the same along each diagonal corresponding to a cohort, so that $\delta_{ij} = \gamma_k$ for all combinations of i and j and $k = j - i + R$.

APPLICATION OF THE MODELS

The present application of these models will be to population data rather than sample data. Consequently, testing will be inappropriate. χ^2 values will be reported (the maximun likelihood definition of χ^2 will be used), but only because these may be converted into certain measures of association, such as phi-square, and may describe the *relative* quality of fit of the various models. Another measure of association will be the index of dissimilarity, $\Delta = (\frac{1}{2}N)\Sigma n_{ijk} - \hat{n}_{ijk}| |$, where N is the total case base. This may be interpreted as the minimum proportion of the N cases which would have to be shifted for the observed $R \times C \times 2$ table of frequencies to perfectly match the expected table (or vice versa).

The denominators of the rates we shall use do not seem to be available in published form. Rather than forego the analysis or attempt to estimate the missing denominators, we shall follow the argument that the age- and period-specific fertility rates are entities that should be given equal weight—as indeed they are in the calculation of the Total Fertility Rate (TFR) for a specific year of time or the Cohort Total Fertility Rate (CTFR) for a specific cohort. That is, from a nonsampling point of view, the relative numbers of women for whom the rates are calculated are unimportant. We shall arbitrarily assume exactly 1000 woman-years for each cell. The procedures followed will not depend at all upon this decision, but any use whatever of the calculated χ^2 must take into account the size of the case base and the fact that it has been set arbitrarily. The choice of a uniform cell size will considerably simplify comparisons between models and between different intervals of time. For more comments on the effect of cell sizes on the estimates, see Pullum (1977).

The four models will be applied to data on white women in the United States for single years of age 15–44 (at time of childbirth) during the half-century 1920–1970 (Heuser, 1976). The data are represented graphically in Figs. 1, 2, and 3. They will first be applied to five 11-year intervals, 1920–1930, 1930–1940, 1940–1950, 1950–1960, and 1960–1970, and to four intervals of 21 years each, 1920–1940, 1930–1950, 1940–1960, and 1950–1970. Thus, for the 11-year intervals, each model is applied to a 30 × 11 (or 330-cell) table of odds, all equally weighted, that a woman-

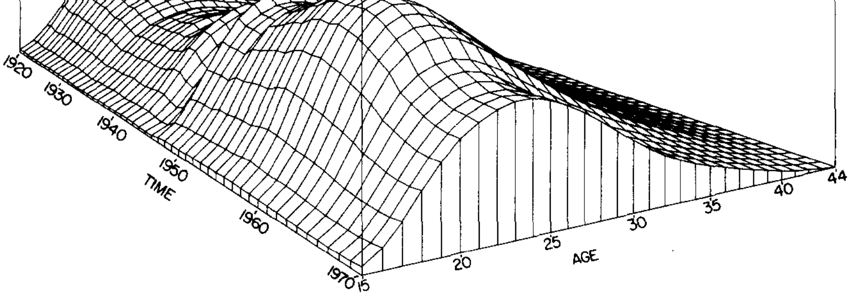

FIG. 1. First view of the plot of fertility rates of white U.S. women, age 15–44, from 1920 to 1970. The rates are proportional to the heights of the intersections above the age × time plane. (The maximum rate is .2552 for age 23 in 1957.) Source: U.S. Department of Health, Education, and Welfare, 1976.

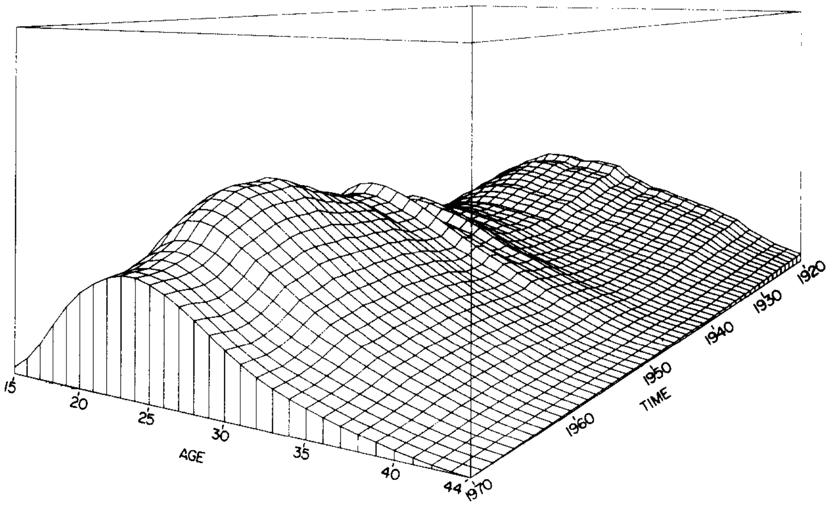

FIG. 2. Second view of the plot of fertility rates of white U.S. women, age 15–44, from 1920 to 1970. The rates are proportional to the heights of the intersections above the age × time plane. (The maximum rate is .2552 for age 23 in 1957). Source: See Fig. 1.

year of exposure will produce a live birth. The 21-year intervals yield tables of size 30 × 21, with 630 cells. The bulk of the discussion will deal with the 11-year intervals.

The results of applying the four models to the successive 11-year intervals are given in Table 1. The most important observation from this

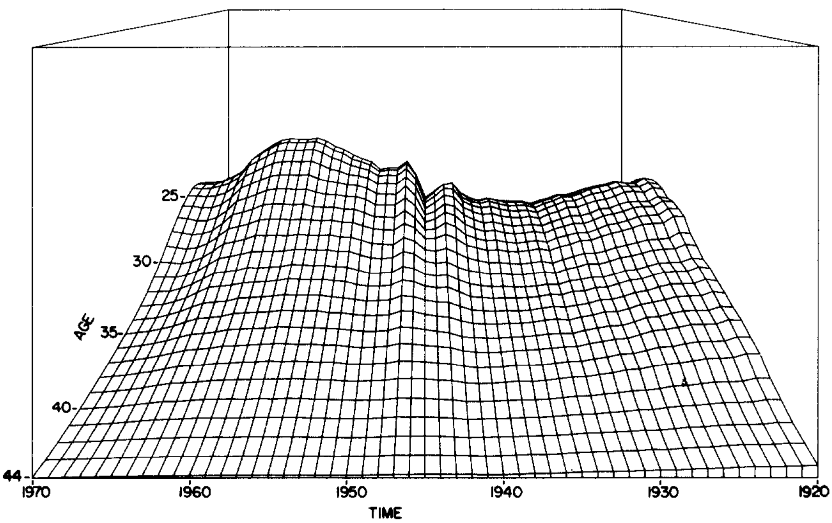

FIG. 3. Third view of the plot of fertility rates of white U.S. women, age 15–44, from 1920 to 1970. The rates are proportional to the heights of the intersections above the age × time plane. (The maximum rate is .2552 for age 23 in 1957.) Source: See Fig. 1.

TABLE 1
Measures of the Quality of Fit for Each Model and Each 11-Year Interval of Data[a]

Model	df for χ^2	Interval of time				
		1920–1930	1930–1940	1940–1950	1950–1960	1960–1970
Age × period	290	18.9	45.7	130.7	44.5	34.4
(1)		.0016	.0022	.0042	.0030	.0021
Age × cohort	261	22.8	30.8	217.7	34.5	58.0
(2)		.0020	.0019	.0058	.0026	.0030
Period × cohort	280	4775.6	3990.6	5757.4	7593.0	6581.6
(3)		.0246	.0205	.0276	.0352	.0301
Age × period	252	2.8	6.5	84.4	4.6	23.3
× cohort (4)		.0006	.0009	.0032	.0008	.0019

[a] First entry: maximum-likelihood χ^2 based on artificial case base of 330,000 woman-years for each interval. Second entry: index of dissimilarity.

table is that in all intervals Models 1, 2, and 4 furnish remarkably good fits. The values of χ^2, calculated for an artificial case base of 330,000 woman-years of exposure and 290, 280, and 252 *df*, respectively, are so low that virtually any experienced researcher would believe that a compunational error had been made. However, careful checking reveals no errors. All indicators of fit are in agreement. Consider, for example, the best fit, Model 4 for 1920–1930. When the observed odds are compared with the expected odds in each of the 330 cells, it is found that the relative deviation exceeds 2% in only 22 cells.

The Index of Dissimilarity provides a second measure of the high quality of fit by Models 1, 2, and 4. For example, Model 4 describes the data so well that no more than .23 of 1% of the outcomes in any table would have to be shifted to establish perfect agreement between the observed and expected tables.

If these were sample data based on specified sample sizes, it would be possible to use χ^2 to test for the significance of adding the third type of effect to the respective two-effect models. (See, for example, Goodman, 1972a.) A formal test is not possible here, as was mentioned earlier. However, the *relative* sizes of the calculated χ^2 values will indicate the relative importance of the effects. In this sense, the improvement in fit which results from the inclusion of each of the three types of effects will be given, per degree of freedom, by

$$M_e = \frac{\chi_m^2 - \chi_4^2}{df_m - df_4}. \tag{7}$$

Here the subscript *e* refers to the type of effect being measured (*A* for age, *P* for period, and *C* for cohort); the subscript *m* refers to the number of the model as described above (1, 2, or 3). Thus e $= A$ when $m = 3$; e $= P$

when $m = 2$; and $e = C$ when $m = 1$. M_e is the reduction in χ^2, per degree of freedom, when Model 4 is obtained by adding effect e to Model m. More precisely, it describes the importance of the explicit addition of a third set of parameters to a two-effect model.

(If the three single-effect models were calculated, then it would also be possible to evaluate the importance of period effects, for example, by comparing the age × period model, Model 1, with the age-only model, or by comparing the period × cohort model, Model 3, with the cohort-only model. The decline in χ^2 should be the same for both of these simpler comparisons, but it would differ from the comparison between the two-effect and three-effect models because the three-effect model is not directly estimable; i. e., it requires iteration. It is more appropriate for present purposes to use the comparison with the three-effect model rather than this simpler comparison.)

The calculated values of M_e are given in Table 3 for each interval of fertility data. The appropriate comparison of these values is within columns rather than within rows. It is clear that for every interval of data, the age effects are by far the most important, and the cohort effects are less important than the period effects, per degree of freedom.

To sum up our observations from Table 1:

(1) Models 1, 2, and 4 all fit the data extremely well.
(2) The period and cohort effects are not nearly as important as the age effects.
(3) For some decades (1920–1930, 1940–1950, 1960–1970) the age × period model fits better than the age × cohort model, and for alternating decades the reverse is true.
(4) As a consequence of observation (3), there is no evidence of a trend in the relative importance of period and cohort effects.
(5) *Per effect,* cohorts are statistically much less explanatory than are periods.
(6) Although all of the measures of fit for 1960–1970 are poorer than those for the first decade, 1920–1930, there is a good deal of variation from one decade to the next, suggesting that the models generally continue to be appropriate.
(7) The decade 1940–1950, when the Baby Boom began, was the only decade in which the neasures of fit for Models 1, 2, and 4 showed markedly worse values.

Overlapping 21-year intervals were also employed, with results given in Table 2 and the lower part of Table 3. Most of the observations about Table 1 continue to hold; the χ^2 values are still remarkably small for Models 1, 2, and 4.

The preceding comments refer to arrays which may be described as rectangles, drawn from the full rectangular data set with 30 rows representing ages 15–44, 51 columns representing years 1920–1970, and 80

TABLE 2
Measures of the Quality of Fit for Each Model and Each 21-year Interval of Data[a]

Model	*df* for χ^2	Interval of time: 1920–1940	1930–1950	1940–1960	1950–1970
Age × period	580	148.1	330.2	411.3	93.3
(1)		.0028	.0048	.0054	.0027
Age × cohort	551	90.2	397.8	411.6	800.2
(2)		.0026	.0053	.0057	.0088
Period × cohort	560	13,306.8	14,952.4	20,603.5	21,130.1
(3)		.0301	.0323	.0419	.0427
Age × period ×	532	21.6	133.1	98.0	40.8
cohort (4)		.0011	.0028	.0023	.0017

[a] First entry: maximum likelihood of χ^2 based on an artificial case base of 630,000 cases for each interval. Second entry: index of dissimilarity.

diagonals (aligned from upper left to lower right) representing the birth cohorts 1886–1955. We shall not consider any rectangles with a width greater than two decades because of the evidence of increasing interactions between age and period which cannot be described as cohort effects.

The final application will be to the largest parallelogram which can be drawn from the full 30 × 51 array, comprising the 22 birth cohorts (1905–1926) on which we have complete data. This parallelogram is obtained by removing the lower-left and upper-right triangles from the full array, and consists of 30 × 22 = 660 fertility rates or odds, with an artificial case base of 660 × 1,000 = 660,000 woman-years of exposure. A parallelogram of this sort is fully oriented toward the cohort as a dimension for structuring the collection, representation, and analysis of time-series data.

TABLE 3
The Reduction in χ^2, *per Degree of Freedom*, When the Age × Period × Cohort Model is Obtained by the Addition of Age, Period, or Cohort to Models 3, 2, or 1, Respectively

Effect	Interval of time: 1920–1930	1930–1940	1940–1950	1950–1960	1960–1970
Age	170.5	142.3	202.6	271.0	234.2
Period	2.2	2.7	14.8	3.3	3.9
Cohort	0.6	1.4	1.7	1.4	0.4

Effect	1920–1940	1930–1950	1940–1960	1950–1970
Age	474.5	529.3	732.3	753.2
Period	3.6	13.9	16.5	40.0
Cohort	2.6	4.1	6.5	1.1

Note. Case bases are artificially fixed.

TABLE 4
Measures of the Quality of Fit When the Models Are Applied to the Parallelogram of Complete Data on the 1905–1926 Birth Cohorts[a]

Model	df for χ^2	Measure of fit	
		χ^2	Δ
Age × period (1)	580	96.1	.0023
Age × cohort (2)	609	730.4	.0075
Period × cohort (3)	588	15,846.8	.0325
Age × period × cohort (4)	560	87.6	.0021

[a] Based on an artificial case base of 660,000 woman-years.

The measures of fit for the four models when applied to this array, are given in Table 4. Again, Models 1, 2, and 4 fit the data well, considering the large size of the artificial case base. However, in contrast to the previous applications to rectangular data, the inclusion of period effects appears *much* more useful than the inclusion of cohort effects. This is ascertained by comparing the three-factor model with the first two two-factor models. Thus, when Model 4 is obtained by adding cohort effects to Model 2, χ^2 is reduced from 96.1 to 87.6, a relative decline (which is independent of the artificial case base) of only 8.8%, alongside a cost of 580–560 = 20 df. However, when Model 4 is obtained by adding period effects to Model 2, χ^2 declines from 730.4 to 87.6, a relative decline of 88.0%, at a cost of 609–560 = 49 df. Although we cannot attempt to assess levels of statistical significance, it is at least clear that the fertility of these 22 cohorts varied much more along a period dimension than along a cohort dimension. There are far fewer cohorts than periods represented in the parallelogram, to be sure. But the measure M_C, given by Eq. (7), which takes the degrees of freedom into account, is only 0.4 for Table 4. That is, by any measure, the cohort identification adds virtually nothing to the age × period model for the 1905–1926 cohorts. The age × cohort model gives a much worse fit than the age × period model, primarily for the later births to the later cohorts, but also in all four corners of the parallelogram.

INTERPRETING THE PARAMETER ESTIMATES

Earlier researchers (see, for example, Glenn, 1976, 1977; and Palmore, 1978) have emphasized the identification problem associated with estimating the parameters in the age × period × cohort model. We intend to show here that the seriousness of this problem is not so great as to preclude all uses of the estimates. The invariance properties to be described in this section were alluded to in Pullum (1977) and were developed in detail (independently) by Fienberg and Mason (1978). Our purpose here is to restate these properties for the present context.

In the three-effect model, unique estimates may be obtained using iterative scaling if arbitrary values are assigned to four coefficients: one age effect, one period effect, one cohort effect, and one other effect from any of the three sets. In computational work, we have arbitrarily assigned $\alpha_1 = \beta_1 = \gamma_R = \gamma_{R+1} = 1$ for this purpose.

The estimates of all other parameters in all of the models will depend upon the choice of the constraints. However, in the two-effect models (1, 2, and 3), which only require two constraints, it is easily shown that ratios of the form $\alpha_{i'}/\alpha_i$, $\beta_{j'}/\beta_j$ and $\gamma_{k'}/\gamma_k$ do not depend at all upon the form of the constraints. Thus, in these models, the *change* from year to year, age to age, and cohort to cohort is *invariant* with respect to the form of the constraints. For example, working with the logarithms of the effects and adjacent values, in the two-way models *the first differences are invariant:*

$$\begin{aligned} \Delta\alpha_i &\equiv \log \alpha_{i+1} - \log \alpha_i, \quad \text{for } i=1,\ldots, R-1, \\ \Delta\beta_j &\equiv \log \beta_{j+1} - \log \beta_j, \quad \text{for } j=1,\ldots, C-1, \\ \text{and } \Delta\gamma_k &\equiv \log \gamma_{k+1} - \log \gamma_k, \quad \text{for } k=1,\ldots R+C-2, \end{aligned}$$

do not depend at all upon how the constraints are made.

In other contexts, analysts are content to work with first differences or equivalent sorts of deviations. For example, in many-way analysis of variance or multiple classification analysis, the estimated main effects are expressed as deviations from an overall mean. In multiple regressions in which the C categories of a classificatory variable are represented by a set of $C-1$ binary or dummy variables, the estimated regression coefficients are well known to be simply deviations between pairs of effects which cannot themselves be estimated. In short, a limitation to differences and deviations is common with other procedures and has not prevented useful interpretations.

With the three-effect model, the first differences no longer satisfy the condition above, but Theorem 1 in the Appendix shows that *the second differences are invariant:*

$$\begin{aligned} \Delta^2\alpha_i &\equiv \Delta\alpha_{i+1} - \Delta\alpha_i = \log \alpha_{i+2} - 2\log \alpha_{i+1} + \log \alpha_i, \\ \Delta^2\beta_j &\equiv \Delta\beta_{j+1} - \Delta\beta_j = \log \beta_{j+2} - 2\log \beta_{j+1} + \log \beta_j, \\ \text{and } \Delta^2\gamma_k &\equiv \Delta\gamma_{k+1} - \Delta\gamma_k = \log \gamma_{k+2} - 2\log \gamma_{k+1} + \log \gamma_k, \end{aligned}$$

$$\begin{aligned} &\text{for } i=1,\ldots, R-2, \\ &\text{for } j=1,\ldots, C-2, \\ &\text{for } k=1,\ldots, R+C-3, \end{aligned}$$

do not depend on the form of the constraints.

Analysts are only rarely limited to second differences or their equivalent, and may well feel uncomfortable with them. Nevertheless, second differences are interpretable. Goldberg (1958), for example, provides examples and discussion of differences of various orders, and relates them to corresponding derivatives of continuous functions.

Our own discussion will employ a variant of the usual form (given above) for second differences. Above, $\Delta^2\beta_j$ was given as the change in period effects from year $j+1$ to year $j+2$ *less* the change from year j to year $j+1$. Thus, $\Delta^2\beta_j > 0$ implies acceleration in the period function (concave upward) and $\Delta^2\beta_j < 0$ implies deceleration (concave downward). The (first) difference between two adjacent period effects is compared with the immediately preceding difference.

We believe that it is easier to interpret the following alternative: the differences between two adjacent period (or other) effects as compared with the *average* annual change across a long interval such as a decade. This is also a second difference, and by Theorems 1 and 2 in the Appendix may be shown to be invariant across alternative parametrizations. We shall call such a quantity a *relative difference,* and define

$$D(\alpha_i) = \Delta\alpha_i - (1/t)\sum_{m=i_0}^{i_0+t-1} \Delta\alpha_m$$

$$= \Delta\alpha_i - [\log(\alpha_{i_0+t}) - \log(\alpha_{i_0})]/t \text{ for } i_0 \le i \le i_0 + t - 1$$

and similarly

$$D(\beta_j) = \Delta\beta_j - [\log(\beta_{j_0+t}) - \log(\beta_{j_0})]/t \text{ for } j_0 \le j \le j_0 + t - 1$$
$$D(\gamma_k) = \Delta\gamma_k - [\log(\gamma_{k_0+t}) = \log(\gamma_{k_0})]/t \text{ for } k_0 \le k \le k_0 + t - 1.$$

Here t refers to a unit of age or time, as appropriate. Note that

$$\sum_{m=i_0}^{i_0+t-1} D(\alpha_m) = \sum_{m=j_0}^{j_0+t-1} D(\beta_m) = \sum_{m=k_0}^{k_0+t-1} D(\gamma_m) = 0.$$

Thus, for example, $D(\beta_j) > 0$ implies that the change from year j to the year $j + 1$ was more positive than the average annual change across the interval j_0 to $j_0 + t$. The largest value of $D(\beta_j)$ in the t-year range will indicate that the most positive change occurred from year j to year $j + 1$. The smallest value of $D(\beta_j)$ in the t-year range will indicate that the smallest positive change (or most negative change) occurred from year j to year $j + 1$, and so on. Analogous statements apply to the values of the $D(\gamma_k)$ differences within a t-year range of cohorts.

Despite this invariance property of the second differences in the age x period x cohort model, there need not be a direct correspondence across different data sets. For example, if Model 4 is applied to the arrays of rates for 1920–1940, 1930–1940, and 1930–1950, and the second differences of period effects are calculated, then one has three different sets of estimates for the years 1930–1940. These three sets will agree if and only if Model 4 gives a perfect fit for the entire combined array of data for

1920–1950. As we have seen, the age x period x cohort model does provide an excellent fit across this range, but it is not perfect; there is some interaction between ages and periods which cannot be expressed in terms of cohorts.

Similarly, for example, the 1900–1910 birth cohorts are represented in the arrays for 1920–1930 (when they were included in ages 15–30), for 1930–1940 (in ages 20–40), and for 1940–1950 (in ages 30–44). The three sets of estimates of second differences in effects for the 1900–1910 cohorts will not agree because, again, Model 4 does not give a perfect fit for the entire 1920–1950 array.

Because the model does fit quite well for the U.S. data, different estimates of the relative differences for specific periods, $D(\beta_j)$, are generally strongly positively correlated, regardless of the arrays from which they were estimated. They are also strongly correlated with fluctuations in the TFR for the specific periods (or in TFR/(30-TFR), the marginal odds to which the fitting was done). These fluctuations in the TFR give an approximation of any estimates of fluctuations in the period effects, at least in terms of direction and relative magnitude. Similarly for the adjusted second differences across specific cohorts, $D(\gamma_k)$; different estimates are correlated with one another and with the diagonal sums of the arrays (partial cohort fertility totals) and the corresponding marginal odds.

Because of this variability, and because the selection of arrays to begin and end on multiples of ten years is essentially arbitrary, we shall not present any estimates here. The purpose of this section has been simply to show that any parametrization of Model 4 for a given array of data will lead to a unique set of second differences in the age, period, and cohort effects, and to describe their interpretation.

THE USE OF COHORT EFFECTS TO IMPROVE PROJECTIONS

The cohort effects provide a continuity across time which may be useful in improving the quality of short-term projections of fertility rates for those cohorts represented in the earlier observations. We shall briefly indicate the nature of such a refinement in the following context: 11 years of observed rates (10 year-to-year changes) are to be used to project 5 years ahead, extrapolating across the last 6 years (5 year-to-year changes) of observations. It would be possible, of course, to use a longer or shorter period of observation, to have a different base for the extrapolation, to assume nonlinear trends, etc. We simply wish to indicate how Model 4 may be employed as a refinement of Model 1.

Note that the actual projection, in all cases, will be of the odds rather than the rates.

From Eq. (1) for Model 1, we have

$$F_{i,j+1} = F_{ij}(\beta_{j+1} / \beta_j).$$

Let r_j be defined to be the ratio β_{j+1}/β_j. Proceeding as in Theorem 1 of the Appendix, it may be shown that the ratio r_j is invariant across all parametrizations of Model 1. However, it cannot be estimated for values of j beyond the last year of observation, except by some form of extrapolation. For example, one could calculate the geometric mean of the last five values of r_j during the period of observation, defined to be $\bar{r}$, and let $\hat{r}_j = \bar{r}$ for the following 5 years. That is,

$$F^*_{i,C+t} = f_{iC}(\bar{r})^t, \; i = 1,\ldots,R; \; t = 1,\ldots,5$$

would be the projected odds for age i and year $C + t$ extrapolating from the final year (C) of observed rates, f_{iC}. The age-specific rate for year j would be $F_{ij}/(1 + F_{ij})$.

How might the improved fit of Model 4 over Model 1 be incorporated into such a projection? From Eq. (4) for Model 4, we have

$$F_{i,j+1} = F_{ij}(\beta_{j+1}/\beta_j)(\gamma_{k+1}/\gamma_k).$$

The product of the two ratios on the right hand side is invariant across parametrizations, as shown in the Appendix. This product may be extrapolated by separate extrapolations of the β's and γ's under any parametrization. However, consider only the continuation of cohorts for which we have been able to estimate the effects during a period of observation; then, as above, only the β's require extrapolation. As before, let $\bar{r}$ be the geometric mean of the last five values of $r_j = \beta_{j+1}/\beta_j$ (estimated with Model 4) and let $\hat{r}_j = F$ for years j beyond the period of observation. Then

$$F^{**}_{i,C+t} = f_{iC}(\bar{r})^t(\gamma_{C+t-i+R})/(\gamma_{C-i+R}),$$
$$i = t + 1,\ldots,R; \; t = 1,\ldots,5$$

would be the projection based on Model 4 which corresponded to $F^*_{i,C+t}$ based on Model 1.

Projections employing cohort effects will be superior to those not employing them, in the same sense that calculations with a calculator are superior to those with a slide rule: they make more thorough use of the data. However, there is little reason to expect that the refined projections will be any closer to actual subsequent fertility.

As an exercise to evaluate the refinement, each of the 11-year applications—1920–1930, 1930–1940, 1940–1950, 1950–1960, and 1960–1970—was projected to the following 5 years (excluding 1974 and 1975), for a total of 23 pairs of projections. Comparisons between the projections and with observations were limited to cohorts which were included during the years of observations. Comparisons were based on partial total fertility rates for the projected years. Generally, the two projections were within 1% of each other. For 11 out of the 23 pairs the refined projection was closer. If the comparison was further limited to cohorts which appeared in the whole decade of observation, the refined projection was closer in 18 out of the 23 pairs. The simpler projection was an average of

8.8% away from the observed partial total fertility rate (ignoring signs) and the refinement was off by an average of 8.0%.

Although certainly not conclusive, this exercise suggests that more elaborate projection procedures may also be improved marginally (and at trivial cost) by incorporating effects for those cohorts which have been observed for several years.

CONCLUSION

We have offered two types of answers to the underlying, motivating question, "Did cohort effects apply to white U.S. fertility during the half-century 1920–1970?" By comparing age × period, age × cohort, and age × period × cohort multiplicative models on the odds, it was inferred that the cohort identification was less important than the period identification. Considering rectangles of data corresponding to decades of time, there is no evidence of a long-term trend, such as any increase or decrease in the relative importance of the cohort identification. Within a parallelogram corresponding to the complete histories of the 1905–1926 cohorts, the identification of the cohorts added virtually nothing to our ability to replicate the specific rates. The probable explanation of the nore extreme conclusions from the parallelogram than from the rectangles is the disruption in cohort continuities stemming from the postwar baby boom.

We also referred to specific parametrizations of these models, and described interpretations of the parameters. In particular, each two-effect model has invariant first differences in log effects, and the three-effect model has invariant second differences in log effects. Projection models may be improved slightly by the use of cohort parameters.

Despite the great theoretical appeal of the notion that continuities exist in the behavior of cohorts, we have found that the explanatory gain per cohort parameter is far less than the gain per period parameter. This statement applies approximately equally well to both the 1920s and the 1960s. The implications for our understanding of U.S. fertility are that, as a set, changes in those temporal variables which cut across cohorts, such as economic cycles, appear to be more important than changes in those variables which distinguish cohorts, such as shared socializing experiences. In extending these conclusions, it would be useful to apply these models to order-specific fertility rates.

COMMENT ON DATA PROCESSING

All calculations were performed by a Fortran program written by the author, but may be replicated with more general programs for log-linear models. ECTA was used to verify the low values of χ^2.

The so-called parallelogram of data on the 1905–1926 cohorts was processed by conforming it to a rectangular shape, in which the rows represent age, the columns represent cohorts, and the diagonals aligned from lower left to upper right represent periods.

APPENDIX

THEOREM 1. *Let the expectations* F_{ij},

$$F_{ij} = \eta\alpha_i\beta_j\gamma_k, \qquad \text{for all } i \text{ and } j,\ k = j - i + R,$$

be any parametric fit to the data. Then

$$\text{(a)}\ \frac{\alpha_{i+1}\alpha_{i-1}}{\alpha_i^2}, \qquad \text{for all } i,$$

$$\text{(b)}\ \frac{\beta_{j+1}\beta_{j-1}}{\beta_j^2}, \qquad \text{for all } j,$$

and

$$\text{(c)}\ \frac{\gamma_{k+1}\gamma_{k-1}}{\gamma_k^2}, \qquad \text{for all } k,$$

will be invariant with respect to any alternative parametrizations of the model.

Proof. The model implies the following, for all i, j, and k:

$$a = F_{i-1,j-1} = \eta\alpha_{i-1}\beta_{j-1}\gamma_k,$$
$$b = F_{i-1,j} = \eta\alpha_{i-1}\beta_j\gamma_{k+1},$$
$$c = F_{i,j-1} = \eta\alpha_i\beta_{j-1}\gamma_{k-1},$$
$$d = F_{ij} = \eta\alpha_i\beta_j\gamma_k,$$
$$e = F_{i,j+1} = \eta\alpha_i\beta_{j+1}\gamma_{k+1},$$
$$f = F_{i+1,j} = \eta\alpha_{i+1}\beta_j\gamma_{k-1}.$$

The letters a, b, c, d, e, and f are used to simplify the notation. These odds correspond to six cells in a 3×3 subtable of the table of expected odds, as follows:

	$j-1$	j	$j+1$
$i-1$	a	b	
i	c	d	e
$i+1$		f	

Now, the ratio

$$\frac{af}{cd} = \frac{\alpha_{i+1}\alpha_{i-1}}{\alpha_i^2} \qquad \text{for all } i.$$

Since the left hand side will be the same for all parametrizations, the right hand side is invariant. Similarly,

$$\frac{ae}{bd} = \frac{\beta_{j+1}\beta_{j-1}}{\beta_j^2} \qquad \text{for all } j$$

and

$$\frac{ad}{bc} = \frac{\gamma_{k+1}\gamma_{k-1}}{\gamma_k^2} \qquad \text{for all } k.$$

THEOREM 2. *Functions of the parameters of the form* $(\alpha_{i_1+h})(\alpha_{i_2})/(\alpha_{i_1})(\alpha_{i_2+h})$ *will be invariant across parametrizations for all choices of subscripts* i_1, i_2, *and* h *such that* i_1, $i_1 + h$, i_2, *and* $i_2 + h$ *are in the range* $1,\ldots,R$. *Similarly, for*

$$\frac{(\beta_{j_1+h})(\beta_{j_2})}{(\beta_{j_2})(\beta_{j_2+h})} \qquad \text{and} \qquad \frac{(\gamma_{k_1+h})(\gamma_{k_2})}{(\gamma_{k_1})(\gamma_{k_2+h})}.$$

Proof. By applications of Theorem 1; products of invariant functions of the parameters will also be invariant.

REFERENCES

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975), Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge, Mass.

Brass, William, Coale, A. J., Heisel, D. F., Lorimer, F., Romaniuk, A., and van de Walle, E. (1968), The Demography of Tropical Africa, Princeton Univ. Press, Princeton, N. J.

Butz, W. P., and Ward, M. P. (1977), The emergence of Countercyclical U.S. Fertility, Rand Corporation Report No. R-1605-NIH.

Coale, A., and Trussell, J. T. (1974), "Model fertility schedules: variations in the age structure of childbearing in human populations," Population Index **40**(2), 185–261.

Easterlin, R. A. (1973), "Relative economic status and the American fertility swing," *in* Family Economic Behavior (E. B. Sheldon, Ed.), pp. 170–223, Lippincott, Philadelphia.

Fienberg, S. E., and Mason, W. M. (1978), "Identification and estimation of age-period-cohort models in the analysis of discrete archival data," *in* Sociological Methodology 1979 (K. F. Schuessler, Ed.), pp. 1–67, Jossey-Bass, San Francisco.

Glenn, N. D. (1976), "Cohort analysts' futile quest: statistical attempts to separate age, period and cohort effects," American Sociological Review **41**, 900–904.

Glenn, N. D. (1977), Cohort Analysis, Sage Univ. Paper No. 5.

Goldberg, S. (1958), Introduction to Difference Equations, John Wiley and Sons, Inc., New York.

Goodman, L. A. (1972), "A general model for the analysis of surveys," American Journal of Sociology **77**, 1035–1086. (a)

Goodman, L. A. (1972), "Some multiplicative models for the analysis of cross-classified data," *in* Proceedings, Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 649–696, Univ. of California Press, Berkeley. (b)

Goodman, L. A. (1975), A Note on Cohort Analysis Using Multiplicative Models and the Modified Multiple Regression Approach, Unpublished paper, University of Chicago, Department of Sociology.

Henry, L. (1961), "Some data on natural fertility," Eugenics Quarterly **8**, 2.

Heuser, R. L. (1976), Fertility Tables for Birth Cohorts by Color: United States, 1917–73. DHEW Publication No. (HRA) 76-1152, U.S. Govt. Printing Office, Washington, D.C.

Lee, R. (1974), "The formal dynamics of controlled populations and the echo, the boom and the bust," Demography **11,** 563–585.

Mason, K. O., Mason, W. M., Winsborough, H. H., and Poole, W. K. (1973), "Some methodological issues in cohort analysis of archival data," American Sociological Review **38,** 242–258.

McKenna, E. E. (1974), "Marriage and fertility in postfamine Ireland: a multivariate analysis," American Journal of Sociology **80,** 688–705.

Osborn, J. (1975), "A multiplicative model for the analysis of vital statistics rates," Journal of Applied Statistics **24,** 75–84.

Palmore, E. (1978), "When can age, period, and cohort be separated?" Social Forces **57,** 282–295.

Pullum, T. W. (1976), "Comment on McKenna," American Journal of Sociology **81,** 1472–1476.

Pullum, T. W. (1977), "Parametrizing age, period, and cohort effects: an application to U.S. delinquency rates, 1964–1973," *in* Sociological Methodology1978 (K. F. Schuessler, Ed.), Jossey-Bass, San Francisco.

Ryder, N. B. (1965), "The cohort as a concept in the study of social change," American Sociological Review **30,** 843–861.

Spiegelman, M. (1968), Introduction to Demography, Harvard Univ. Press, Cambridge, Mass.