BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-87/19

FORECASTING AGE SPECIFIC FERTILITY USING PRINCIPAL COMPONENTS

by

James E. Bozik and William R. Bell

Statistical Research Division
Bureau of the Census
Room 3000, F.O.B. #4
Washington, D.C. 20233 U.S.A.

Recommended by: Nash J. Monsour

Report completed: August 12, 1987

Report issued: August 12, 1987

FORECASTING AGE SPECIFIC FERTILITY USING PRINCIPAL COMPONENTS

by

James E. Bozik and William R. Bell

Bureau of the Census
Statistical Research Division
Washington, D.C. 20233

For Presentation at the
Annual Meetings of the American Statistical Association
San Francisco, CA
August, 1987

ABSTRACT

Recent papers by Rogers (1986) and Thompson et al. (1987) suggest fitting curves to annual age-specific fertility rates, forecasting the parameters of the curves using time series techniques, and then using the forecasted curves to generate forecasts of future age-specific fertility rates. This approach reduces the dimensionality of the forecasting problem, but the error in the fitted curves is not negligible. We present an approach based on a principal components approximation to the rates that avoids this problem to a large extent, while still reducing dimensionality. This approach is compared with direct univariate modeling of all the age-specific rates, and with the curve fitting approaches.

1. INTRODUCTION

Demographers traditionally have relied on the cohort-component method for calculating fertility projections; multiplying projected age-specific birth rates by the projected population of women in each age group to calculate a projected total number of births. Time series methods can also be used to do fertility projections. However, direct univariate or multivariate modeling of age-specific fertility rates requires working with about 36 time series, one each for women of ages between 14 and 49. (The age-specific fertility rate (denoted by f_{kt} in section 2) is defined as the number of births divided by the number of women of age k in year t.) The age-specific fertility rates for any given year nearly follow the same smooth pattern, as displayed in Figure 1 of the Appendix (the relative fertility rates g_{kt} discussed in section 2 are displayed here. They are the f_{kt}'s scaled to sum to one.) Since they follow a similar pattern each year, one might consider approximating the rates for each time point by a function involving a few parameters. "A few" means considerably less than 36, and hopefully a manageable number for multivariate time series modeling. Forecasts are obtained for these parameters, and then the forecasts of these parameters are translated into forecasts of the age-specific fertility rates.

The use of several parametric curves to approximate the age-specific rates has been discussed recently in the statistical literature. Rogers (1986) suggests using a double exponential curve, while Thompson et al. (1987) use a Gamma curve. The success of their approaches depends on how close their respective curves approximate the age-specific rates. Figure 2 in the Appendix displays the Gamma and double exponential curve approximations to the rates for 1980. Thompson et al. fit their curve to the relative fertility rates g_{kt}, while the double exponential curve Rogers suggests includes a scale factor and is fit directly to the age-specific rates. (Revised data (as of 1987) for 1980-1984 is used everywhere in this paper except in the fitting of the double exponential curve, which was done before the revised data became available. For the purposes of this comparison the revisions are slight.) It is clear from Figure 2 that there is considerable error in the approximations for some ages, especially the ages with high fertility. The pattern of the errors across ages changes slowly over the time span of the data, but significant errors are present for almost all the years. Thus, Thompson et al. (1987) use a "bias adjustment" to correct for lack of fit at some ages when the Gamma curve is used.

Our approach using principal components is similar in philosophy to fitting a parametric curve in that we reduce the dimensionality of the forecasting problem. However, we attempt to reduce the approximation error that occurs when either the double exponential or Gamma curve is used to approximate age-specific fertility rates. The principal components approach allows the data itself to select a linear function that accurately represents the curve of age-specific rates. In that respect, our approach is more flexible than one using a specific curve to approximate the rates.

In section 2 of this paper we introduce some notation and discuss the data used. In section 3 we discuss the approach using principal components. Section 4 discusses selection of the number of principal components, and section 5 presents forecasting results from a multivariate model using principal components.

2. DESCRIPTION OF DATA

Let

f_{kt} = fertility rate for women of age k in year t
$k = 14, \ldots, 49$ (see remark below)

TFR_t = total fertility rate in year t
$= \sum_k f_{kt}$

g_{kt} = relative fertility rate for women of age k in year t
$= f_{kt} / TFR_t$

Annual age-specific fertility data for white women from 1921-1984 were used in this study (Heuser (1976), Bureau of the Census (1984), and more recent unpublished data.) Following Thompson et al. (1987) we analyze and forecast the relative rates (g_{kt}). Logged relative rates ($\ln(g_{kt})$) were used to insure that forecast intervals would not drop below zero at ages with low fertility. $LTFR_t = Ln(TFR_t)$ was also used. Forecasts of f_{kt} are obtained by modeling and forecasting TFR_t, and multiplying forecasts of g_{kt} by the TFR forecasts. Because of values rounded to zero for f_{kt} in some years for some ages above 46, our analysis was cut off at age 46. Thus, 33 age groups were used.

The fertility for women over age 46 is so small, especially in recent years, that it can effectively be ignored in fertility projections.

3. PRINCIPAL COMPONENTS APPROACH

Suppose we have a given set of constants λ_{kj} for each k and j = 1, ... , J, where J is the number of linear functions of the data we use to approximate $\underset{\sim}{\gamma}_t$, where $\underset{\sim}{\gamma}_t = \left[\ln(g_{14,t}), \ldots, \ln(g_{46,t})\right]$. Let $\Lambda = \{\lambda_{kj}\}$. Λ has dimension 33 x J. Our problem is to find constants $(\beta_{1t} \ldots \beta_{Jt})' = \underset{\sim}{\beta}_t$ to

$$\min_{\underset{\sim}{\beta}_t} \sum_k \left[\gamma_{kt} - (\beta_{1t}\lambda_{1k} + \cdots + \beta_{Jt}\lambda_{kJ}) \right]^2 = \min_{\underset{\sim}{\beta}_t} \left\| \underset{\sim}{\gamma}_t - \Lambda \underset{\sim}{\beta}_t \right\|^2$$

The solution is attained at $\hat{\underset{\sim}{\beta}}_t = (\Lambda'\Lambda)^{-1} \Lambda' \underset{\sim}{\gamma}_t$. There is a normalization problem here, in that for any nonsingular linear transformation of Λ one obtains the same minimum for the sum of squares by appropriately transforming $\hat{\underset{\sim}{\beta}}_t$. We can thus restrict Λ, and the most natural restriction is to require the columns of $\Lambda = \left[\underset{\sim}{\lambda}_1 \cdots \underset{\sim}{\lambda}_J\right]$ to be orthonormal. Under this restriction $\Lambda'\Lambda = I$, and the minimum of the sum of squares occurs at $\hat{\underset{\sim}{\beta}}_t = \Lambda' \underset{\sim}{\gamma}_t$, with minimum value

$$\left\| \underset{\sim}{\gamma}_t - \Lambda\Lambda' \underset{\sim}{\gamma}_t \right\|^2 = \underset{\sim}{\gamma}_t' \ (I - \Lambda\Lambda') \ (I - \Lambda\Lambda') \ \underset{\sim}{\gamma}_t = \underset{\sim}{\gamma}_t' \ (I - \Lambda\Lambda') \ \underset{\sim}{\gamma}_t .$$

The remaining problem is to choose the λ_{kj}'s. Since we have data $\underset{\sim}{\gamma}_t$ for multiple time points t, and we want our apporoximation to be good at all of these time points (and at future time points as well), this suggests using $\Lambda = (\lambda_{kj})$ that solves the following optimization problem. Note that different $\hat{\underset{\sim}{\beta}}_t$ are chosen for each time point t, but Λ remains constant over time:

$$\min_{\Lambda} \ \min_{\underset{\sim}{\beta}_t} \ \sum_t \left\| \underset{\sim}{\gamma}_t - \Lambda \underset{\sim}{\beta}_t \right\|^2 = \min_{\Lambda} \ \sum_t \underset{\sim}{\gamma}_t' \left[I - \Lambda\Lambda' \right] \underset{\sim}{\gamma}_t$$

$$= \min_{\Lambda} \ \mathrm{tr} \left[I - \Lambda\Lambda' \right] \sum_t \underset{\sim}{\gamma}_t \underset{\sim}{\gamma}_t'$$

$$= \min_{\Lambda} \ \mathrm{tr} \left[I - \Lambda\Lambda' \right] S$$

$$= \min_{\Lambda} \left[\mathrm{tr}(S) - \mathrm{tr}(\Lambda' S \Lambda) \right]$$

$$= \min_{\Lambda} \left[\sum_k \mu_k(S) - \sum_{j=1}^{J} \mu_j \ (\Lambda' S \Lambda) \right]$$

S is the sum of squares and cross products matrix of the data $(\underset{\sim}{\gamma}_t)$, tr(S) is the trace of S, and $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_J(A)$ are the ordered eigenvalues of any real, symmetric matrix A. The solution is attained by choosing $\Lambda = [\underset{\sim}{\lambda}_1 \ \ldots \ \underset{\sim}{\lambda}_J]$ such that $\underset{\sim}{\lambda}_1, \ \ldots \ , \underset{\sim}{\lambda}_J$ are the eigenvectors of S corresponding to $\mu_1(S), \ \ldots \ , \ \mu_J(S)$, in which case the minimum value obtained is $\mu_{J+1}(S) + \cdots + \mu_{33}(S)$. Note that since

$\hat{\underset{\sim}{\beta}}_t = \Lambda'\underset{\sim}{\gamma}_t$, $\hat{\beta}_{jt} = \underset{\sim}{\lambda}_j'\underset{\sim}{\gamma}_t$ is the j^{th} principal component score for the t^{th} observation.

One can also consider using weighted least squares. The problem then becomes

$$\min_{\Lambda} \min_{\underset{\sim}{\beta}_t} \sum_t (\underset{\sim}{\gamma}_t - \Lambda'\underset{\sim}{\beta}_t)' \, W_t \, (\underset{\sim}{\gamma}_t - \Lambda'\underset{\sim}{\beta}_t) =$$

$$\min_{\Lambda} \sum_t \underset{\sim}{\gamma}_t' \left[W_t - W_t\Lambda(\Lambda'W_t\Lambda)^{-1} \Lambda'W_t \right] \underset{\sim}{\gamma}_t$$

for some given symmetric positive definite weighting matrices W_t. While the above problem could be solved by numerical methods with suitable constraints imposed on Λ, it simplifies greatly if one can assume $W_t = W$ for all t. In this case we can require $\Lambda'W\Lambda = I$, and then define

$$H = W^{1/2}\Lambda \qquad \underset{\sim}{h}_t = W^{1/2} \, \underset{\sim}{\gamma}_t$$

so that H is the matrix of eigenvectors of

$$\sum_t \underset{\sim}{h}_t \, \underset{\sim}{h}_t' = \sum_t (W^{1/2}\underset{\sim}{\gamma}_t) \, (W^{1/2}\underset{\sim}{\gamma}_t)'$$

and $\hat{\underset{\sim}{\beta}}_t = H' \, \underset{\sim}{h}_t$. Our analysis was performed using weighted least squares with W a diagonal matrix with elements w_i, where $w_i = 16$ for ages 18 thru 32, and $w_i = 1$ for ages 14 thru 17 and 33 thru 46. This is the same weighting scheme used by Thompson et al. (1987), and places more importance in terms of fit on the ages with high fertility.

4. HOW MANY PRINCIPAL COMPONENTS TO USE?

Using graphs of the age-specific rates by year and fits obtained from using increasing numbers of principal components, we found the fits for $J \geq 4$ or 5 difficult to distinguish from the data when plotted across age. This is illustrated for 1980 in Figure 3 of the Appendix. One can see a small improvement in the fit as the number of principal components is increased from 4 to 8. The principal components approach appears not to suffer from the more serious approximation problems that occur when the double exponential or Gamma curves are used to approximate the data.

The selection of the number of principal components is not so clear when the fits are studied for individual ages. Figures 4-6 in the Appendix illustrate this for age 25. Four principal components may not be enough, and significant improvement in the fit is seen as the number of principal components is increased to 8 or 12. A measure to judge the fit across years for individual ages was computed as:

$$s_{kJ}^2 = \frac{1}{58} \sum_t{}' \left[\text{Ln}(\gamma_{kt}) - \text{Ln}(\hat{\gamma}_{Jkt}) \right]^2 , \; J = 1, \ldots , 12$$

where k is the age group (14, ... ,46) and the $\hat{\gamma}_{Jkt}$ are the approximations to the γ_{kt} when J is the number of principal components used to obtain the fit. Σ' denotes the war years 1942-1947 were omitted from the calculation. Table 1 in the Appendix lists $\{s_{kJ}\}$ for J = 4, 8, and 12. Since the approximation is to be used in forecasting, the amount of approximation error at any age k for any given number of components J should be interpreted in relation to the magnitude of the forecast error for age k that would be likely to otherwise result. To do this, the $\{s_{kJ}\}$ were compared to the residual standard error obtained from simple univariate time series models of the γ_{kt} for each k. The model notation follows that of Box and Jenkins (1970). Substantial reduction in the $\{s_{kJ}\}$ occurs as J is increased. However, the number of principal components necessary to achieve a suitable approximation to the data remains unclear.

5. TIME SERIES MODELING OF THE PRINCIPAL COMPONENTS AND TFR

5.1 Univariate and Multivariate Models

For illustration, $\beta_{1t}, \ldots, \beta_{4t}$ and $\ln(TFR_t)$ were modeled using univariate and multivariate time series techniques (Tiao and Box, 1981). Let $LTFR_t = Ln(TFR_t)$. Inspection of the series indicated first differencing was necessary for all five series. For all ages, the data showed unusual behavior occuring between 1942 and 1947, due to the effect of World War II on fertility. The effect was removed by including indicator variables for 1942-1947 in the univariate models for $\beta_1, \ldots, \beta_4$, and LTFR. The following univariate models were identified:

SERIES	ARIMA MODEL
LTFR	(3 1 0)
β_1	(3 1 0)
β_2	(2 1 0)
β_3	(1 1 0)
β_4	(1 1 0)

Outliers were identified in the series LTFR, β_1, β_2, and β_4, but their effects were not adjusted for because of the difficulty of doing so in a multivariate context. This could be done as an enhancement to the model. The effects of the war years 1942-1947 were removed prior to the multivariate modeling. The univariate models were then used to initially identify a multivariate (3 1 0) model with strong restrictions on the second and third lag autoregressive matrices. Initially, the lag one autoregressive matrix was unrestricted, but a number of its elements corresponding to insignificant parameter estimates were set to zero.

The following multivariate (3 1 0) model was estimated:

$\hat{\Phi}_1 =$

	LTFR	β_1	β_2	β_3	β_4
LTFR	.527 (.104)	0	0	0	0
β_1	0	.320 (.111)	0	0	0
β_2	-1.195 (.254)	-1.038 (.354)	.554 (.09)	0	0
β_3	0	0	.237 (.073)	.540 (.089)	0
β_4	-1.183 (.346)	0	0	-.186 (.094)	.304 (.141)

$$\hat{\Phi}_2(3,3) = \underset{(.088)}{.244} \qquad \hat{\Phi}_3(1,1) = \underset{(.095)}{.205}$$

The standard errors for the parameters are in parentheses and italics, below the estimate. All other elements of the $\hat{\Phi}_2$ and $\hat{\Phi}_3$ matrices were constrained to zero.

5.2 Forecasts

Let $\widehat{LTFR}_{n+\ell}$ denote a forecast of LTFR for year $n+\ell$ using data through year n, and similarly for $\hat{\beta}_{1,n+\ell}, \ldots, \hat{\beta}_{4,n+\ell}$. The multivariate model can be used to obtain forecasts for $\widehat{LTFR}_{n+\ell}$, $\hat{\underset{\sim}{\beta}}_{n+\ell} = (\hat{\beta}_{1,n+\ell} \cdots \hat{\beta}_{4,n+\ell})'$ and the principal components approximation can be used to convert these into forecasts of $\underset{\sim}{\gamma}_{n+\ell}$, and then of $\ln(\underset{\sim}{f}_{n+\ell})$ as follows:

$$\begin{bmatrix} 1 & \lambda_{1,1} & \cdots & \lambda_{1,4} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & \lambda_{33,1} & \cdots & \lambda_{33,4} \end{bmatrix} \cdot \begin{bmatrix} \widehat{LTFR}_{n+\ell} \\ \hat{\beta}_{1,n+\ell} \\ \hat{\beta}_{2,n+\ell} \\ \hat{\beta}_{3,n+\ell} \\ \hat{\beta}_{4,n+\ell} \end{bmatrix} =$$

$$\begin{bmatrix} \underset{\sim}{1} & \Lambda \end{bmatrix} \cdot \begin{bmatrix} \widehat{LTFR}_{n+\ell} \\ \hat{\underset{\sim}{\beta}}_{n+\ell} \end{bmatrix} = \widehat{LTFR}_{n+\ell} \cdot \underset{\sim}{1} + \Lambda\hat{\underset{\sim}{\beta}}_{n+\ell} =$$

$$\hat{\text{LTFR}}_{n+\ell} \cdot \underset{\sim}{1} + \underset{\sim}{\hat{\gamma}}_{n+\ell} = \begin{bmatrix} \ln(\hat{f}_{14,n+\ell}) \\ \cdot \\ \cdot \\ \cdot \\ \ln(\hat{f}_{46,n+\ell}) \end{bmatrix}$$

where $\underset{\sim}{1} = (1\ 1\ \ldots\ 1)'$ and $(\lambda_{1,j}\ \ldots\ \lambda_{33,j})' = \underset{\sim}{\lambda}_j$ is the j^{th} eigenvector $\Lambda = \left[\ \underset{\sim}{\lambda}_1\ \ldots\ \underset{\sim}{\lambda}_4\ \right]$. The $\ln(\hat{f}_{k,n+\ell})$ can then be exponentiated to produce forecasts of $f_{k,n+\ell}$.

Let V_ℓ denote the 5 x 5 matrix of variances and covariances of the forecast errors $\left[\ \hat{\text{LTFR}}_{n+\ell} - \text{LTFR}_{n+\ell},\ \underset{\sim}{\hat{\beta}}'_{n+\ell} - \underset{\sim}{\beta}'_{n+\ell}\ \right]'$, and let Ω_ℓ similarly denote the 33 x 33 variance covariance matrix of the forecast errors

$$\left[\ \ln(\hat{f}_{14,n+\ell}) - \ln(f_{14,n+\ell})\ \ldots\ \ln(\hat{f}_{46,n+\ell}) - \ln(f_{46,n+\ell})\ \right]'.$$

Ignoring the error in the principal components approximation of $\underset{\sim}{\gamma}_{n+\ell}$,

$$\Omega_\ell = \left[\ \underset{\sim}{1}\quad \Lambda\ \right] V_\ell \begin{bmatrix} \underset{\sim}{1}' \\ \Lambda' \end{bmatrix}.$$

We use an estimate of V_ℓ from the fitted multivariate model to calculate Ω_ℓ. Assuming approximate normality of $\ln(\underset{\sim}{f}_{n+\ell})$, we can use the standard errors from Ω_ℓ to calculate forecast intervals for $\ln(f_{k,n+\ell})$. The limits of these intervals can then be exponentiated to give forecast intervals for $f_{k,n+\ell}$.

To illustrate, the multivariate model was estimated using data through 1980, and forecasts produced for 1981-1990 for LTFR, β_1, β_2, β_3, and β_4. These forecasts were then used to obtain forecasts of the age-specific rates $f_{k,n+\ell}$ for 1981-1990. Approximate 67% forecast intervals (one standard error for each $\ln(f_{k,n+\ell})$) were also produced.

Figures 7-13 in the Appendix are graphs of the data, forecasts, and one standard error limits for the forecasts for TFR = e^{LTFR}, and the fertility rates f_k for women of age 15, 20, 25, 30, 35, and 40. The 1980 forecast origin turned out to be poor for forecasting due to shifts in fertility just after 1980 at some ages. These could not be captured by the forecasts. Thus, for ages 20 and 25, the recent decline in rates for 1981-1984 could not be predicted by our model. For the most part, the 1981-1984 data falls within the 67% intervals. The principal component approximation using only four components for age 40 was relatively poor, and this is reflected in the age 40 forecasts. Using 8 or 12 components will improve the fit at advanced ages, but if the primary objective is to forecast births, this may not be necessary since the advanced ages contribute so little to total births.

Figures 14-17 display the 1981-1984 data, forecasted fertility rates f_{kt} and 67% forecast limits across ages. The data is generally within the forecast limits. Narrow forecast limits for ages with low fertility are also apparent.

6. CONCLUSION

Forecasting age-specific fertility rates using a principal components approach appears to reduce the approximation error that occurs when the rates are approximated using either a Gamma or double exponential curve. The approach appears to have potential for producing reasonable forecasts and forecast intervals for the age-specific rates using a small number of components. We are further studying the number of components necessary for a suitable approximation to the rates, the effect of the approximation error when principal components are used, and are investigating alternative multivariate models.

7. REFERENCES

Box, G.E.P. and G.M. Jenkins (1970), Time Series Analysis: Forecasting and Control, San Francisco: Holden Day.

Bureau of the Census (1984), "Projections of the Population of the United States by Age, Sex, and Race: 1983 to 2080", Current Population Reports, Series P-25, No. 952, Washington, D.C.: U.S. Government Printing Office.

Heuser, R.L. (1976), "Fertility Tables for Birth Cohorts by Color: United States, 1917-1973", DHEW Publication No. HRA 76-1152.

Rogers, A. (1986), "Parameterized Multistate Population Dynamics and Projections", Journal of the American Statistical Association, 81, 48-61.

Thompson, P.A., W.R. Bell, J. Long, and R.B. Miller (1987), "Multivariate Time Series Projections of Parameterized Age-Specific Fertility Rates", Ohio State University Working Paper Series WPS 87-3.

Tiao, G.C. and G.E.P. Box (1981), "Modeling Multiple TIme Series with Applications", Journal of the American Statistical Association, 76, 802-816.

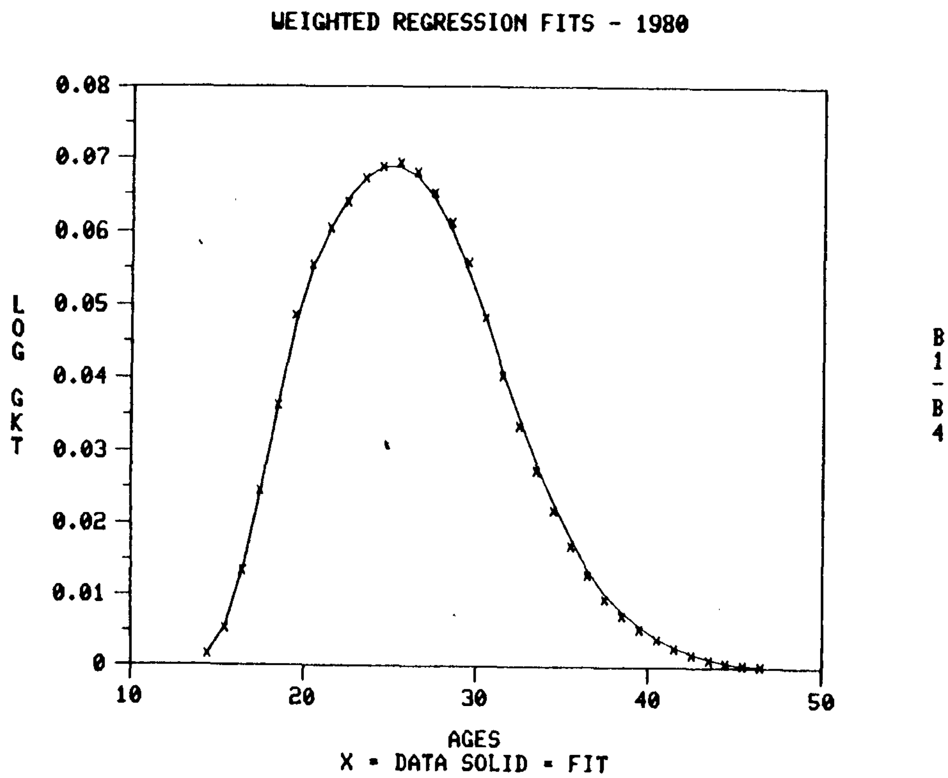
TABLE 1

LOGGED RELATIVE RATES

PRINCIPAL COMPONENTS APPROXIMATION ERRORS
(AS A PERCENTAGE OF UNIVARIATE ARIMA RESIDUAL STD. ERRORS)

AGE	s_{kj} / RSE k=4	s_{kj} / RSE k=8	s_{kj} / RSE k=12	UNIVARIATE MODEL	RESIDUAL STD. ERROR
14	.98	.32	.15	(2 1 0)	0.0544
15	1.28	.33	.30	(1 1 0)	0.0366
16	1.49	.36	.29	(1 1 0)	0.0298
17	1.37	.32	.19	(1 1 0)	0.0254
18	1.17	.45	.09	(1 1 0)	0.0184
19	.90	.23	.19	(2 1 0)	0.0138
20	.64	.35	.14	(2 1 0)	0.0132
21	.68	.27	.16	(1 1 0)	0.0124
22	.72	.27	.21	(1 1 0)	0.0121
23	1.07	.56	.20	(1 1 0)	0.0097
24	1.52	.57	.35	(1 1 0)	0.0075
25	1.74	.53	.34	(2 1 0)	0.0077
26	1.45	.53	.27	(1 1 0)	0.0088
27	1.30	.39	.35	(1 1 0)	0.0090
28	.99	.40	.26	(1 1 0)	0.0098
29	.71	.48	.18	(1 1 0)	0.0110
30	.44	.31	.24	(1 1 0)	0.0126
31	.55	.20	.17	(2 1 0)	0.0140
32	.78	.36	.15	(1 1 0)	0.0141
33	.91	.44	.34	(2 1 0)	0.0174
34	1.03	.58	.37	(1 1 0)	0.0176
35	1.07	.65	.39	(1 1 0)	0.0179
36	1.09	.60	.35	(1 1 0)	0.0197
37	1.28	.67	.44	(1 1 0)	0.0195
38	1.04	.47	.32	(2 1 0)	0.0236
39	1.05	.50	.34	(1 1 0)	0.0246
40	.90	.51	.30	(2 1 0)	0.0261
41	.92	.51	.28	(1 1 0)	0.0322
42	.87	.46	.31	(2 1 0)	0.0297
43	.56	.36	.20	(4 1 0)	0.0633
44	.64	.41	.14	(3 1 0)	0.0666
45	.67	.23	.02	(1 1 0)	0.0993

FIGURE 3


WEIGHTED REGRESSION FITS - 1980
LOG GKT
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
10
20
30
40
50
AGES
X = DATA SOLID = FIT
B1-B4

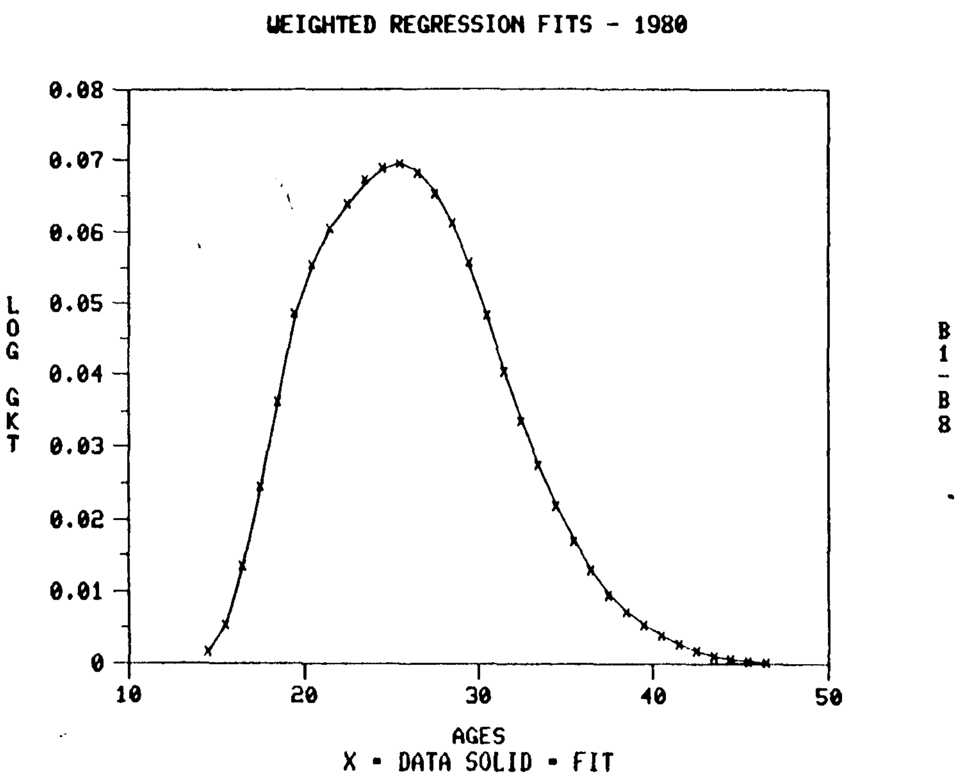
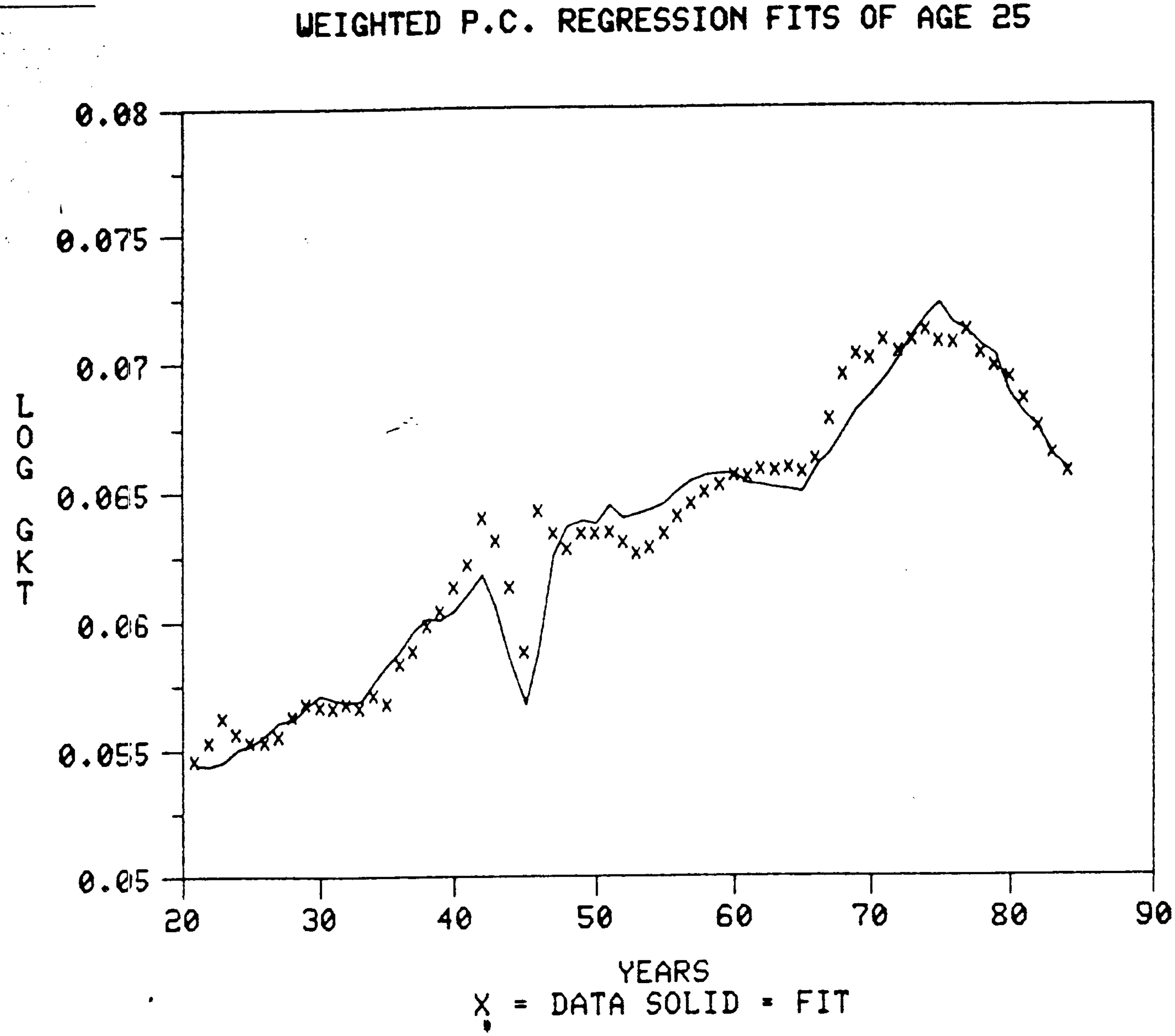
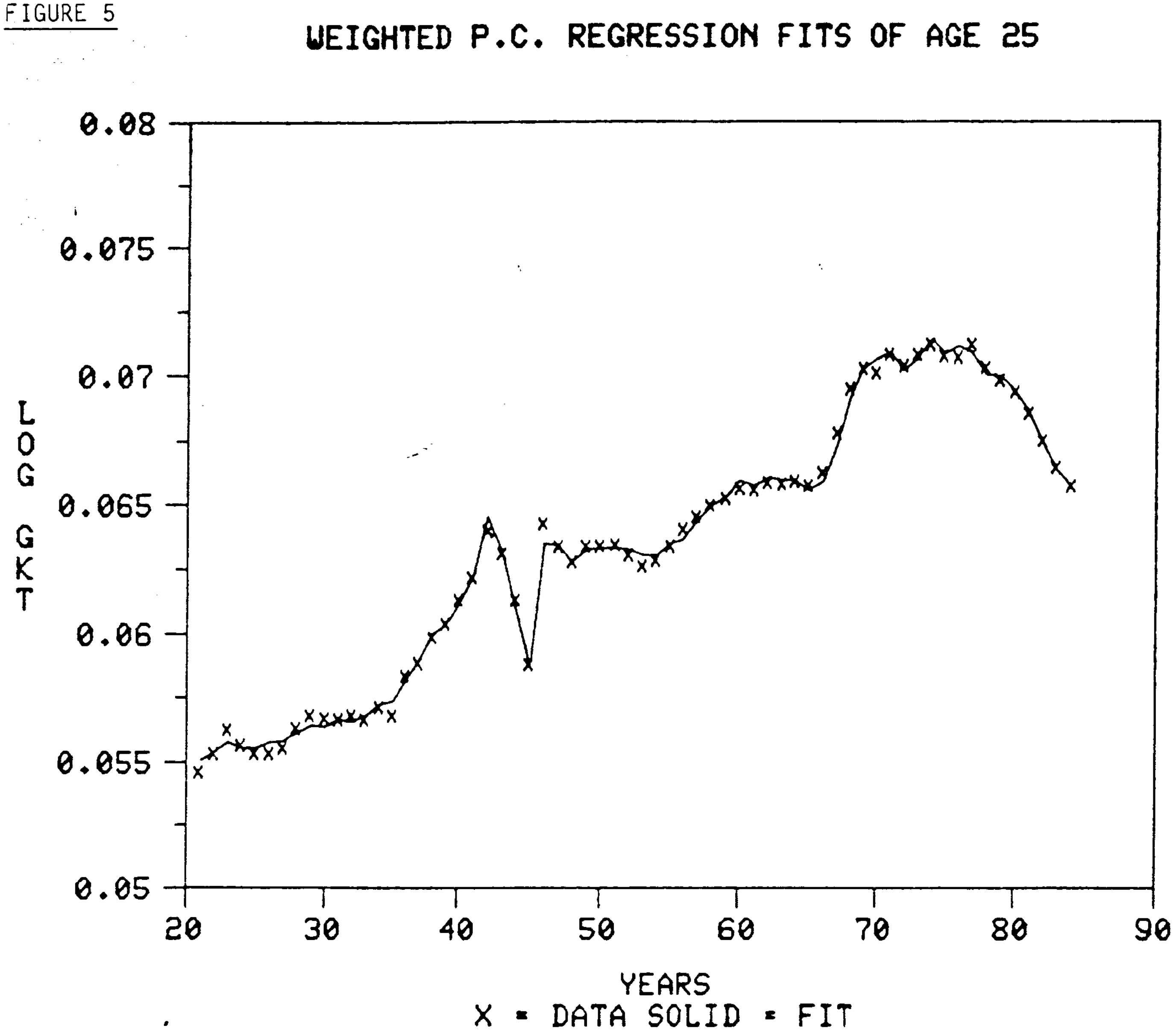

WEIGHTED REGRESSION FITS - 1980
LOG GKT
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
10
20
30
40
50
AGES
X = DATA SOLID = FIT
B1-B8

FIGURE 4



B1-B4

FIGURE 5



B 1 - B 8

FIGURE 6

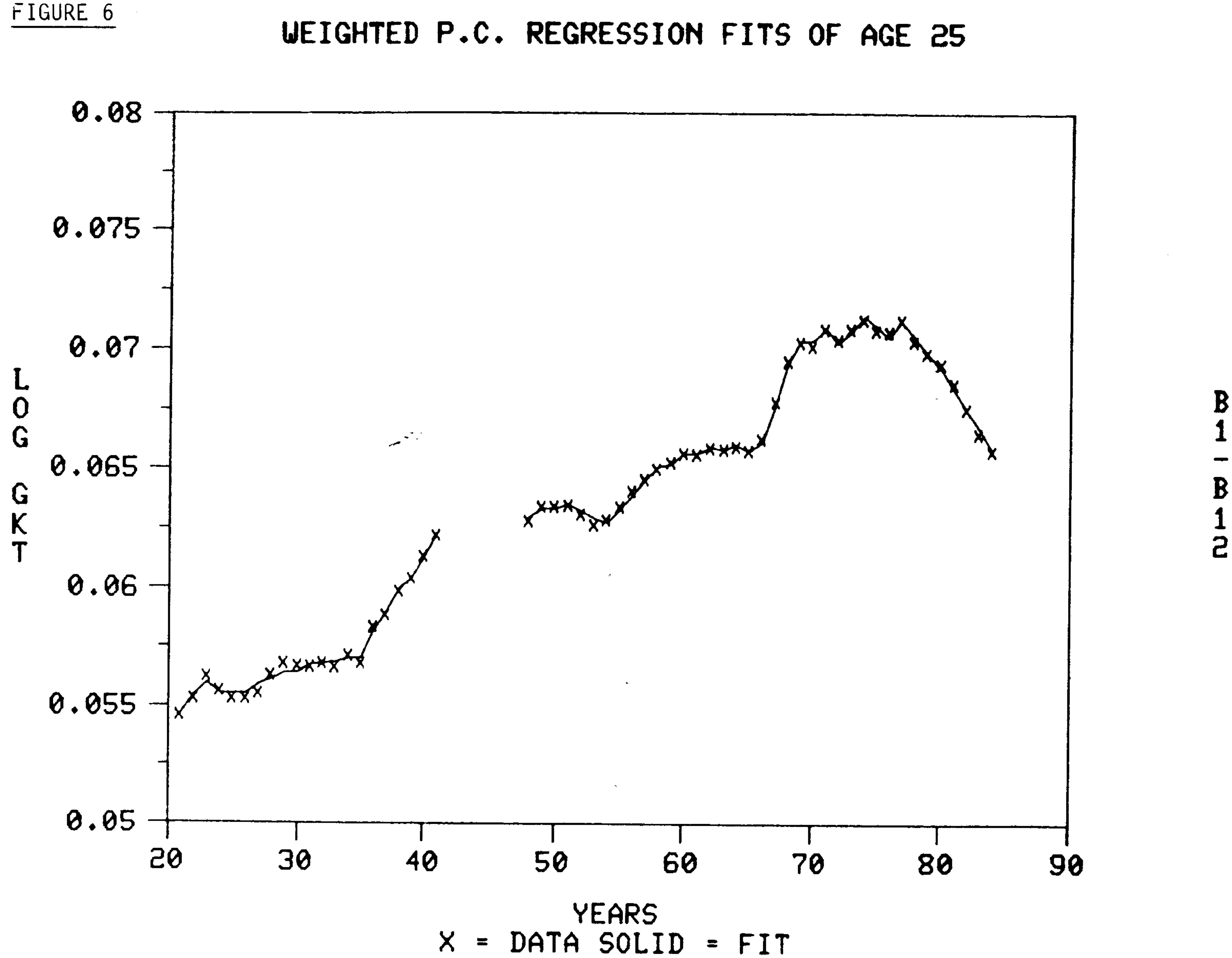
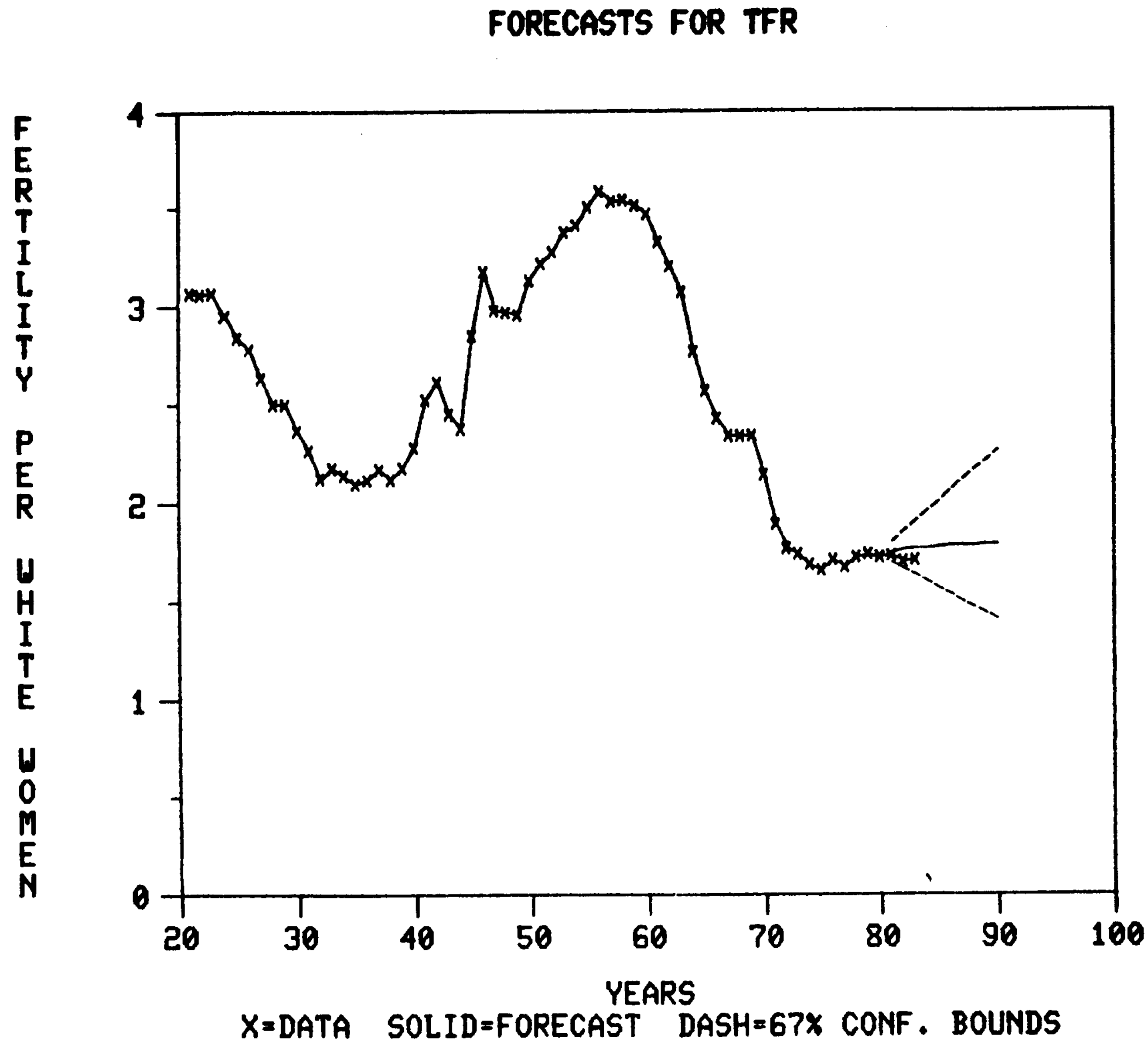

WEIGHTED P.C. REGRESSION FITS OF AGE 25
0.08
0.075
0.07
0.065
0.06
0.055
0.05
LOG GKT
20
30
40
50
60
70
80
90
YEARS
X = DATA SOLID = FIT
B1-B12

FIGURE 7


FORECASTS FOR TFR
FERTILITY PER WHITE WOMEN
4
3
2
1
0
20
30
40
50
60
70
80
90
100
YEARS
X=DATA SOLID=FORECAST DASH=67% CONF. BOUNDS