BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-85/04

UNIT ROOTS IN TIME SERIES MODELS: TESTS AND IMPLICATIONS

by

David A. Dickey
North Carolina State University
Dept. of Statistics
Raleigh, NC 27695-8203

William R. Bell
Statistical Research Division
Bureau of the Census
Washington, DC 20233

and

Robert B. Miller
University of Wisconsin-Madison
School of Business
1155 Observatory Drive
Madison, WI 53706

Recommended by: Kirk M. Wolter

Report completed: January, 1985

Report issued: April 30, 1985

UNIT ROOTS IN TIME SERIES MODELS: TESTS AND IMPLICATIONS

David A. Dickey*
William R. Bell**
Robert B. Miller***

January, 1985

Do not cite or quote
without permission of the
authors

*Visiting Researcher, U.S. Bureau of the Census and Associate Professor of Statistics, North Carolina State University.

**Mathematical Statistician, U.S. Bureau of the Census.

***ASA/Census Research Fellow, U.S. Bureau of the Census, and Professor of Business and Statistics, University of Wisconsin, Madison.

Address correspondence to Professor David A. Dickey, North Carolina State University, Department of Statistics, Box 8203, Raleigh, North Carolina 27695-8203.

Unit Roots in Time Series Models: Tests and Implications

ABSTRACT

The decision on whether or not to include a unit root in an AR operator has profound implications. Formal tests for the presence of unit roots give analysts objective guidance in this decision. This paper is a practical guide to the use of these tests.

1. Introduction

Let Y_t be a discrete time series following a stochastic difference equation model of the form

$$(1-\alpha_1 B- \cdots -\alpha_{p+d} B^{p+d})(Y_t-\mu) = (1-\theta_1 B- \cdots -\theta_q B^q)e_t \qquad (1.1)$$

where B is the backshift operator ($BY_t = Y_{t-1}$); e_t is a series of independent, identically distributed random shocks with mean 0 and variance σ^2; and $\alpha_1, \cdots, \alpha_{p+d}, \mu, \theta_1, \cdots, \theta_q$, and σ^2 are parameters. Now let z denote a complex variable and consider the function $\alpha(z) = (1-\alpha_1 z- \cdots -\alpha_{p+d} z^{p+d})$. The behavior of the sequence Y_t is different as the roots of the equation $\alpha(z) = 0$ (the zeroes of $\alpha(z)$) fall in different regions of the complex plane. If ξ is a root of $\alpha(z) = 0$ with $|\xi| = 1$, then ξ is called a unit root of $\alpha(z) = 0$. Unit roots are either +1, -1, or $\exp(i\lambda)$ for some λ in $[0,2\pi)$. In this paper we concentrate on unit roots of +1.

Suppose $\alpha(z)$ has d unit roots equal to 1, where $d \geq 0$. Then $\alpha(B)$ can be factored as

$$\alpha(B) = \phi(B)(1-B)^d \qquad (1.2)$$

so that writing $\nabla = (1-B)$ (1.1) becomes

$$\phi(B)\nabla^d(Y_t-\mu) = \theta(B)e_t \qquad (1.3)$$

where $\phi(B) = (1-\phi_1 B- \cdots -\phi_p B^p)$. A useful generalization of this model is

$$\phi(B)\nabla^d(Y_t - \underset{\sim}{x}_t'\underset{\sim}{\beta}) = \theta(B)e_t \tag{1.4}$$

where $\underset{\sim}{x}_t'$ is a row vector of observations on r fixed regressor variables, and $\underset{\sim}{\beta}$ is an r x 1 vector of parameters.

If all the roots of $\phi(z) = 0$ are outside the unit circle, then the model (1.3) is called ARIMA(p,d,q), a model that has become somewhat of an industry standard due to the efforts of Box and Jenkins (1970) and their followers. The quantity $\nabla^d Y_t$ is called the d-th difference of Y_t. If d = 0, (1.3) is called an ARMA(p,q) model. Often Y_t is said to be stationary when d = 0 and nonstationary when d > 0. Technically (see Appendix A) even with d = 0 additional conditions are needed for Y_t to be stationary. When d > 0, however, Y_t is nonstationary except in certain trivial cases (e.g. $\sigma^2 = 0$). For lack of any better terminology, we shall refer to d = 0 as stationarity and d > 0 as (homogeneous) nonstationarity.

In practice there can be some question about the need to include a differencing operator ∇ in (1.3), or about the need for an additional ∇ if (1.3) already has one or more. Tools often used in making the decision include examination of plots of the series and its differences for the wandering behavior characteristic of nonstationary series, inspection of the sample autocorrelation function (ACF) of the series and its differences for failure to damp out quickly, and informal inspection of a fitted model of form (1.1) for the presence of a 1-z factor in $\alpha(z)$. While these tools are useful, they all run into difficulties occasionally. In borderline cases where Y_t is stationary but $\phi(z)$ contains factor(s) close to 1-z, Y_t can wander away from its mean μ for long stretches. While the ACF of a nonstationary series should die out only slowly, the individual autocorrelations need not be large (Wichern 1973, Box and Jenkins 1970, pp. 200-201, Hasza 1980), and sampling variation in them can make nonstationary behavior difficult to detect. Finally, informal examination of a fitted model of form (1.1) is unambiguous if the fitted $\hat{\alpha}(z)$ exactly contains a 1-z factor (within computational accuracy). But if $\hat{\alpha}(z)$ has a factor $(1-\hat{\rho}z)$ close to 1-z, the question that arises is whether $1-\hat{\rho}z$ is significantly different from 1-z, i.e. whether $\hat{\rho}$ is significantly different from 1. This calls for formal statistical inference tools. Classical inference results for time series do not apply here, since they require Y_t to be stationary, but many authors have extended these results to nonstationary cases (see Fuller (1984) for a review). In particular, Fuller (1976), Dickey and Fuller (1979, 1981), Said and Dickey (1984, 1985), and Hasza and Fuller (1979) have provided formal tests for the presence of unit roots. The use of these tests and the importance of detecting unit roots are the subjects of this paper.

Before proceeding, we need to make an important qualification. The theory we shall discuss applies only to the case $d = 1$ in (1.3), that is, for deciding between one or no differences. However, Y_t can be an already differenced series, including the important case of a seasonally differenced series,

$Y_t = \nabla_s X_t = (1-B^s)X_t$ where s is the seasonal period. Thus, we assume that the other techniques mentioned (informal examination of plots, ACFs, and fitted models) or other knowledge about the series can be used to discover all differencing factors except the last ∇. Some work on dealing directly with multiple unit roots (Findley 1980, Tiao and Tsay 1983) suggests that the last ∇ factor is the most difficult to detect.

2. Implications of Unit Roots

2.1 Implications for Time Series Models

To illustrate the implications of unit roots, we begin with the model (1.1) with p+d = 1 and $\mu = 0$:

$$(1-\alpha B)Y_t = e_t. \tag{2.1}$$

Consider the three cases $0 < \alpha < 1$, $\alpha = 1$, and $\alpha > 1$. (Parallel comments would apply for $\alpha < 0$.) For these cases the root, $1/\alpha$, of $\alpha(z) = 1 - \alpha z = 0$ is greater than, equal to, and less than 1 respectively. Given any suitable starting value, Y_0 say, (2.1) can be solved to yield

$$Y_t = \alpha^t Y_0 + e_t + \alpha e_{t-1} + \cdots + \alpha^{t-1} e_1 \tag{2.2}$$

For $0 < \alpha < 1$ we see the influence of the starting value, $\alpha^t Y_0$, goes to zero as t increases, as does the influence of the shocks in the distant past. For $\alpha = 1$ the starting value and distant past shocks get the same weight, 1, as recent shocks. For $\alpha > 1$ the weights on the distant past shocks and Y_0 increase as t increases. We can summarize these results by saying that for $0 < \alpha < 1$ the present is more important than the past, for $\alpha = 1$ the past is equal in importance to the present, and for $\alpha > 1$ the past is more important than the present.

Figures 1a,b,c show 100 observations generated from the model (2.1) with $Y_0 = 0$ and the e_t's i.i.d. N(0,1) for $\alpha = .5$, 1.0, and 1.15 respectively. For $\alpha = .5$ we see the series tends to oscillate about its mean value of 0. For $\alpha = 1.0$ the series is free to wander, with no tendency for the values to remain clustered about any fixed level. For $\alpha = 1.15$ the series eventually behaves like the function $k_1 + k_2(1.15)^t$, taking off rapidly towards $-\infty$. These three types of behavior are characteristic of series following the model (1.1), according to whether $\alpha(z) = 0$ has all roots outside the unit circle (as in Fig. 1.a.), one or more roots on with the rest of the roots outside the unit circle (Fig. 1.b.), or any roots inside the unit circle (Fig. 1.c.).

If $\alpha(z) = 0$ has unit or explosive roots, Y_t following (1.1) is nonstationary; if all the roots lie outside the unit circle <u>and</u> Y_t satisfies some additional conditions, then Y_t will be stationary. (See Appendix A.) Since explosive behavior is not realistic for most series in practice, we shall not explicitly consider explosive models further. We shall concentrate on the model (1.3) with all the zeroes of $\phi(z)$ outside the unit circle.

2.2 Nonstationarity and Regression Modeling

The implications of unit root nonstationarity in series to be used in regression models have been studied by Granger and Newbold (1977, section 6.4), Plosser and Schwert (1978), and Nelson and Kang (1984). This work concentrated to an extent on three distinct possibilities in regard to dealing with nonstationarity - differencing, removal of a linear trend, and doing nothing. It was shown that doing nothing when differencing is needed can have dire consequences - frequently leading to falsely significant regressions of nonstationary series on time and on other independent nonstationary series. Linear detrending helps little in regard to the latter. The consequences of unnecessary differencing were shown to be far less serious: inefficient, though unbiased and consistent, parameter estimates. This suggested at least doing regressions with differenced as well as undifferenced data and comparing the results. One iteration of a Cochrane-Orcutt procedure (filtering the regression equation with $1-\hat{\rho}B$ where $\hat{\rho}$ is the lag 1 autocorrelation of the regression residuals) did not correct the problems noted. Iterating Cochrane-Orcutt did not correct the problem of spurious regressions on time, but it did correct the problems with relating nonstationary series. For the latter, iterating eventually pushed $\hat{\rho}$ near 1, which produced roughly the same regression results as differencing. These results were thus not overly sensitive to different values of $\hat{\rho}$ near 1.

2.3 Implications for Forecasting

Given a model (1.1), values of the parameters of the model, and data $Y_1 = y_1, \ldots, Y_n = y_n$, the minimum mean squared error forecast of a future value $Y_{n+\ell}$ is $\hat{Y}_{n+\ell} = E(Y_{n+\ell}|y_1,\ldots,y_n)$, and the forecast standard error is $[Var(Y_{n+\ell}|y_1,\ldots,y_n)]^{1/2}$. (For simplicity, we shall assume normality of Y_t.) These calculations are discussed by Box and Jenkins (1970, chapter 5).

To illustrate the impact of unit roots on forecasts we display forecasts from the two models

$$(1-1.8B + .81B^2)\dot{Y}_t = e_t \tag{2.3}$$

and

$$(1-1.8B + .80B^2)\dot{Y}_t = e_t \tag{2.4}$$

where $\dot{Y}_t = Y_t - 100$, $Var(e_t) = 2$, and the last two data values are $y_{n-1} = 128$ and $y_n = 135$. Model (2.3) can be written $(1-.9B)^2\dot{Y}_t = e_t$ and model (2.4) as $(1-.8B)\nabla\dot{Y}_t = e_t$, so the first model is stationary, the second nonstationary. The forecasts from the two models are shown in Figure 2.a and the forecast standard errors in Figure 2.b. The forecasts from the stationary model converge to the series mean of 100, and the forecast standard errors converge to the series standard deviation $\gamma_0^{1/2} = 23$. The forecasts from the nonstationary model converge to 163, and the forecast standard errors diverge to $+\infty$. This example illustrates the delicate dependence of long-term forecasts on the presence of a unit root and underlines the potential importance of correctly deciding on the presence of a unit root.

The above results generalize to higher order models of the form (1.1). When the model is stationary

$$\hat{Y}_{n+\ell} \to \mu = E(Y_t), \quad Var(Y_{n+\ell}|y_1,\ldots,y_n) \to Var(Y_t) \text{ as } \ell\to\infty,$$

For nonstationary models $\hat{Y}_{n+\ell}$ will eventually either explode or behave like a polynomial, and $Var(Y_{n+\ell}|y_1,\ldots,y_n) \to \infty$ as $\ell\to\infty$.

2.4 Differencing and Polynomial Trends

An alternative to the model (1.3) that is sometimes suggested is to regard Y_t as following a polynomial trend with stationary errors, linear trends being the most popular. However, model (1.3) is actually equivalent to

$$\phi(B)\nabla^d[Y_t-(\beta_0+\beta_1 t+\cdots+\beta_{d-1}t^{d-1})] = \theta(B)e_t$$

since $\nabla^d(\beta_0+\beta_1 t+\cdots+\beta_{d-1}t^{d-1}) = 0$. If a polynomial trend of degree $d - 1$ is present in Y_t, (1.3) automatically allows for it. For example, in forecasting Y_t, the $d - 1$ degree polynomial will be reproduced. For the particular case of $d = 1$ in (1.3) $\nabla\mu = 0$ so $\nabla(Y_t-\mu) = \nabla Y_t$ and μ, a polynomial of degree 0, is automatically included in the model. Thus, there is no need to write μ explicitly in (1.3) for any $d > 0$.

An alternative to (1.3) is

$$\phi(B)\nabla^d Y_t = c+\theta(B)e_t \tag{2.5}$$

which is equivalent to the model

$$\phi(B)\nabla^d[Y_t-(\beta_0+\beta_1 t+\cdots+\beta_{d-1}t^{d-1}+\beta_d t^d)] = \theta(B)e_t \tag{2.6}$$

where $\beta_d = c/\phi(1)d!$. For example, with $d = 1$, (2.6) becomes $\phi(B)\nabla[Y_t-\beta_0-\beta_1 t] = \theta(B)e_t$, and β_1 is the slope of a linear trend in Y_t. Polynomials of degree higher than d, d + k say, can be incorporated by replacing c in (2.5) by a polynomial of degree k.

Pierce (1975), Chan, Hayya, and Ord (1977), Abraham and Box (1978), Nelson and Kang (1981, 1984), and Nelson and Plosser (1982) discuss the issue of fitting polynomials versus differencing. From our discussion we see the critical issue in choosing between differencing and fitting polynomial trends is not whether Y_t in fact follows a polynomial trend, since both approaches allow for this, but whether the deviations of Y_t from the polynomial require differencing. For the important case of a linear trend, the question is whether we should use the model

$$\phi(B)(1-\rho B)(Y_t-\beta_0-\beta_1 t) = \theta(B)e_t \tag{2.7}$$

where $\rho < 1$, or the model

$$\phi(B)\nabla(Y_t-\beta_0-\beta_1 t) = \theta(B)e_t. \tag{2.8}$$

Model (2.8) is equivalent to $\phi(B)\nabla Y_t = c + \theta(B)e_t$ with $c = \phi(1)\beta_1$.

We shall see in Sections 4 and 5 that preliminary removal of a linear trend from Y_t can make it difficult to determine whether the resulting deviations from the linear trend require differencing. For this reason, and because the decision on whether or not to difference carries profound implications for models and forecasts, we recommend against preliminary removal of linear or other polynomial trends. We shall later show how to test if a linear trend plus stationary errors model is appropriate (testing (2.7) against (2.8).) Nelson and Plosser (1982) apply this test to 14 annual U.S. macroeconomic time series and only reject (2.8) in favor of (2.7) for one series - the unemployment rate.

3. Delineation of Hypotheses

We wish to decide whether or not a process is one whose first differences follow a stationary ARMA model (null hypothesis). We shall discuss stationary alternatives to this hypothesis, but one can easily consider explosive and even two-sided alternatives in an obvious way. Different alternatives are appropriate depending on whether, under the null hypothesis, the first differences are assumed to have a zero or a non-zero mean.

The more common null case is to assume the first differences have a zero mean. Then the null and alternative hypotheses are

$$\text{(null)} \qquad H_1: \quad \phi(B)\nabla Y_t = \theta(B)e_t \tag{3.1}$$

$$\text{(alt)} \qquad H_2: \quad \phi(B)(1-\rho B)(Y_t-\mu) = \theta(B)e_t, \quad -1<\rho<1 \tag{3.2}$$

where $\mu = E(Y_t)$ when $|\rho|<1$. Both hypotheses are composite because at least one parameter, σ^2, is unspecified.

If $\phi(B) \equiv \theta(B) \equiv 1$, then the hypotheses above specialize to

$$H_1\text{: } \nabla Y_t = e_t \tag{3.3}$$

$$H_2\text{: } (1-\rho B)(Y_t-\mu) = e_t, \ -1<\rho<1 \tag{3.4}$$

The null hypothesis is a random walk. The alternative hypothesis is a stationary AR(1).

For the moment suppose we have only entertained the limited hypotheses H_1 and H_2 displayed in (3.3) and (3.4). Suppose further that we perform an appropriate test and decide in favor of H_1. At this point we would probably wish to ask if the drift is truly zero, where drift is the mean of ∇Y_t. We could do this by setting up an alternative hypothesis

$$H_3\text{: } \nabla Y_t = c + e_t \quad c \neq 0 \tag{3.5}$$

and testing it against the null hypothesis H_1.

As we have seen in Section 2.3, the model in (3.5) automatically allows for a linear trend in Y_t. A possible alternative to (3.5) is

$$H_4\text{: } (1-\rho B)(Y_t-\beta_0-\beta_1 t) = e_t, \ -1<\rho<1 \quad \beta_1 \neq 0 \tag{3.6}$$

that is, the deviations of Y_t from a linear function of time follow a stationary AR(1) model. If we set $\rho = 1$ in (3.6) we get the model in (3.5) with $c = \rho\beta_1 = \beta_1$. Given that we have accepted (3.5) instead of (3.3), we may also want to test (3.5) against (3.6). As noted in Section 2.4, we recommend against beginning a sequence of tests with hypothesis H_4, however.

Our formulation of the hypotheses in this section illustrates why the stationary, non-zero mean model is a natural alternative to the random walk with zero drift, whereas the stationary deviation from linear trend model is a natural alternative to the random walk with drift. In both cases, the alternative hypothesis reduces to the null hypothesis when we set $\rho=1$.

Our approach to deciding on the need for a differencing operator will use a framework in which models with a ∇ (such as (3.1), (3.3), and (3.5)) specify the null hypothesis, and models without a ∇ (such as (3.2), (3.4), and (3.6)), represent the alternative hypothesis. In performing a formal hypothesis test to decide on the need for differencing, we accept a ∇ in the model unless the data present (statistically significant) evidence to the contrary. We are thus expressing a preference for models with a ∇ over polynomial trend plus stationary error models, such as (3.6), and models with a $1-\rho B$ autoregressive factor with ρ near to but less than 1. We do this because of what we perceive to be the relative importance of the two possible errors in deciding on differencing. Failure to include a differencing operator when it is needed results in bounded forecast intervals that must eventually be too narrow, giving unreasonable confidence in the forecasts. This can be especially true if a polynomial trend plus stationary error model is used when differencing is needed. The polynomial trend may fit well over the span of the observed data, but extrapolating it implies a strong assumption about the future, and this may well produce highly unrealistic forecasts and forecast intervals. On the other hand, differencing when a ∇ is not needed can produce forecast results equivalent to those from a model without a ∇ (Harvey 1981). Overdifferencing can also sometimes be detected and corrected at the modeling stage (Abraham and Box 1978). At worst, use of ∇ when $1-\rho B$ with $\rho < 1$ is more appropriate will produce conservative forecast intervals that may differ greatly from those for the model with $(1-\rho B)$ only for the long term.

4. Test Statistics

4.1 A Simple Model

The simplest practical model we have introduced is (3.4), which may also be written

$$Y_t = c + \rho Y_{t-1} + e_t \qquad (4.1)$$

where $c = (1-\rho)\mu$. The model looks like a regression model. We estimate c and ρ by regression of Y_t on 1, Y_{t-1}. Denote the estimates by $\hat{c}$ and $\hat{\rho}_\mu$.

Dickey and Fuller (1979) present the distribution of $\hat{\rho}_\mu$

when $\rho = 1$ in (4.1). Percentiles of this distribution are given in Fuller (1976, p. 371). Regression estimates for coefficients in stationary time series are normalized by $\sqrt{n}$ (where n is the sample size) to obtain limit distributions. In contrast, when $\rho = 1$, $n(\hat{\rho}_\mu - 1)$ has a nondegenerate limit distribution. Thus, for large n, ρ is estimated more accurately when it is 1 than when $|\rho| < 1$. However, the fifth percentile of $n(\hat{\rho}_\mu - \rho)$ for $\rho = 1$ does not exceed the corresponding percentile of $\sqrt{n}(\hat{\rho}_\mu - \rho)$ for $\rho = 0$ until n exceeds about 70.

A statistic computed by most regression programs is the studentized coefficient, which is usually called a t test for the coefficient. We denote the statistic τ_μ, rather than t which suggests the Student t distribution (this being inappropriate under our null hypothesis). The statistic τ_μ is related to the likelihood ratio test of $(\rho, c) = (1, 0)$ in model (4.1). See Dickey and Fuller (1981).

We formally define τ_μ by letting $\underset{\sim}{Y}$ be an n dimensional vector with t^{th} entry Y_t and $\underset{\sim}{X}$ an n x 2 matrix with t^{th} row $(1, Y_{t-1})$. It follows that $(\hat{c}, \hat{\rho}_\mu)' = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}(\underset{\sim}{X}'\underset{\sim}{Y})$, $s^2 = (n-2)^{-1}(\underset{\sim}{Y}'\underset{\sim}{Y} - (\hat{c}, \hat{\rho}_\mu)(\underset{\sim}{X}'\underset{\sim}{Y}))$, and $\tau_\mu = (v_{22}s^2)^{-1/2}(\hat{\rho}_\mu - 1)$ where v_{22} is the second diagonal element of $(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$. The distribution of τ_μ is not the Student t distribution even in the limit. Fuller (1976, p. 373) published a table of selected percentiles of τ_μ, which is reproduced in Appendix C of this paper. If an experimenter had used the usual normal percentiles to test the hypotheses in (3.3) and (3.4) at a nominal significance level .01, the critical value from the normal tables would have been

-2.58, and for large n the actual significance level would have been .10! Use of the standard t tables results in spurious declaration of stationarity.

There are several ways to compute $\hat{\rho}_\mu$ and τ_μ. For example, let $\bar{Y}_{(0)} = n^{-1} \sum_{t=1}^{n} Y_t$ and $\bar{Y}_{(-1)} = n^{-1} \sum_{t=1}^{n} Y_{t-1}$. Then regressing $Y_t - \bar{Y}_{(0)}$ on $Y_{t-1} - \bar{Y}_{(-1)}$ gives $\hat{\rho}_\mu$. Under the null hypothesis, a constant added to Y_0 increases each Y_t by that same constant, and under the alternative a constant added to $E(Y_t)$ increases each Y_t by that same constant. Our current representation of $\hat{\rho}_\mu$ shows that adding a constant to all Y_t does not affect $\hat{\rho}_\mu$. An estimator with the same asymptotic distribution as $\hat{\rho}_\mu$ is obtained by regressing $(Y_t - \bar{Y})$ on $(Y_{t-1} - \bar{Y})$ where $\bar{Y} = (n+1)^{-1} \sum_{t=0}^{n} Y_t$ is the mean of the entire data set. Finally, if $\nabla Y_t = Y_t - Y_{t-1}$, then the regression of ∇Y_t on 1, Y_{t-1} gives a coefficient $\hat{\rho}_\mu - 1$ on Y_{t-1}, and the usual t statistic for the coefficient on Y_{t-1} is our τ_μ statistic. We can read the test statistic τ_μ directly off of a standard computer printout under this parameterization.

4.2 The General Model

Consider the general model (3.2). If $|\rho| < 1$ we can use any of several computer packages to estimate the parameters of $\phi(B)$ and $\theta(B)$. To perform the test of $\rho=1$ on the general model, we first subtract the series mean, $\bar{Y}$, and specify initial estimates of the parameters that are consistent under both the null <u>and</u> alternative hypotheses. We next perform a one step Gauss-Newton improvement of the initial estimates of the parameters in $\phi(B)$ and $\theta(B)$ and of $\hat{\rho} = 1$. The t-statistic associated with the one step improvement for $\hat{\rho}$ has the limit distribution of τ_μ.

There are several ways to obtain the initial estimates needed. One simple approach is to use a computer package to estimate (3.2) as a multiplicative ARMA model, with the μ term omitted since $\bar{Y}$ has already been subtracted. Or (3.2) could be rewritten as

$$\phi(B)(1-\rho B)Y_t = c + \theta(B)e_t$$

where $c = \phi(1)(1-\rho)$, and estimated directly in this form without subtracting $\bar{Y}$. Direct estimation of (3.2) with μ included is not advised since estimates of ρ and μ can be highly correlated. Estimation could be by conditional least squares or by a procedure that is exact maximum likelihood for MA models but not for AR models, such as that suggested by Hillmer and Tiao (1979). Procedures that are exact maximum likelihood for AR models should not be used since these fall apart as ρ approaches 1. Finally, for $\phi(B) \neq 1$ one must make sure that on initial estimation the possible unit root factor that is estimated is $1-\rho B$ and not part of $\phi(B)$. This could be done by starting ρ at 1 in the initial estimation as well.

Three additional comments should be made concerning the procedure. First, studies of the power function for this test emphasize the importance of obtaining good initial estimates for the parameters of $\phi(B)$ and $\theta(B)$ when $\theta(B) \neq 1$ (Said and Dickey 1985). Second, if $\theta(B) = 1$ then ordinary least squares is appropriate for estimating the parameters of $\phi(B)$ as we shall illustrate below. Third, in doing this test one step of a pure Gauss-Newton iteration is strictly required for the theory to apply. Many time series packages use Marquardt's algorithm, which is a mixture of Gauss-Newton and steepest ascent, in their iterative estimation scheme. This algorithm may give different results.

Solo (1984) suggests an alternative to the above procedure based on a Lagrange multiplier test of $\rho = 1$. His approach differs from that given above in that (i) initial parameter estimates are obtained for the "null" model (3.1) rather than (3.2), and (ii) the one step Gauss-Newton improvement is performed only on $\rho^{(0)} = 1$ (also on the constant term if $\bar{Y}$ is not removed), not on the parameters in $\phi(B)$ and $\theta(B)$. The resulting t-statistic still has the τ_μ limit distribution. The procedure amounts to regressing the residuals from the fitted null model (3.1) on their derivatives with respect to ρ. Solo's procedure does simplify the computations, but at this time we can say nothing about the power of such a test or how well it maintains its nominal significance level. Said and Dickey (1984) provide another alternative, showing that an asymptotically valid test in mixed models is obtained if the data are analyzed as though they were generated by an autoregressive model whose order is a function of n.

4.3 Autoregressive Models

We now explain how to perform the test for autoregressive models. Consider the general autoregressive process

$$\phi(B)(1-\rho B)Y_t = c + e_t \tag{4.3}$$

where $c = \phi(1)(1-\rho)\mu$ and the roots of $\phi(B) = 0$ lie outside the unit circle. If $\rho = 1$ then $c = 0$ and the regression of ∇Y_t on $\nabla Y_{t-1}, \cdots, \nabla Y_{t-p}$, yields estimates of the coefficients of the polynomial $\phi(B)$ which are $O_p(n^{-1/2})$ and whose limit distribution is multivariate normal. Thus ordinary regression testing procedures may be used for these coefficients (see Fuller 1976 chapter 8). Regressing ∇Y_t on 1, Y_{t-1}, $\nabla Y_{t-1}, \cdots \nabla Y_{t-p}$ produces consistent estimates of c and $\phi(1)(\rho-1)$ for any ρ, with the p coefficients on $\nabla Y_{t-1}, \cdots, \nabla Y_{t-p}$ providing consistent estimates of $\phi_1, \cdots, \phi_p$ when $\rho = 1$.

For example, suppose $Y_t = 50(1-\rho) + (\rho+.6)Y_{t-1} - .6\rho Y_{t-2} + e_t$. If $\rho = 1$ the intercept term is 0. For any ρ,

$$\nabla Y_t = 50(1-\rho) + .4(\rho-1)Y_{t-1} + .6\rho\nabla Y_{t-1} + e_t. \tag{4.4}$$

Here $\phi(B) = (1-.6B)$ and the regression of ∇Y_t on 1, Y_{t-1}, ∇Y_{t-1} produces a coefficient on Y_{t-1} which is a consistent estimate of $\phi(1)(\rho-1) = .4(\rho-1)$. This estimate, normalized by n, converges to the limit distribution of $.4n(\hat{\rho}_\mu-1)$. In practice the parameters in $\phi(1)$ are unknown. Under the null hypothesis $\rho = 1$ they are estimated consistently from the coefficients on $\nabla Y_{t-1}, \cdots, \nabla Y_{t-p}$ and the resulting $\hat{\phi}(1)$ could then be divided into the coefficient of Y_{t-1} to produce a test statistic approximately distributed as $n(\hat{\rho}_{\mu}-1)$.

The τ_{μ} test extends in a simpler manner. Notice that, by assumption, $\phi(1) \neq 0$ so any test of $\phi(1)(\rho-1) = 0$ is a test of $\rho = 1$. Dickey and Fuller (1979) show that the t statistic for the coefficient of Y_{t-1} in the regression of ∇Y_t on 1, Y_{t-1}, $\nabla Y_{t-1}, \cdots, \nabla Y_{t-p}$, has the τ_{μ} limit distribution previously discussed. Thus we need only run this regression and compare the t statistic for the coefficient on Y_{t-1} to the τ_{μ} tables in Appendix C.

Although the null distributions of τ_{μ} and $n(\hat{\rho}_{\mu}-1)$ extend nicely to higher order autoregressive models, the distributions can be changed dramatically if the Y_t's are adjusted for effects other than an overall mean. In the first order model, for example, if we remove a linear trend from Y_t then we must use the tables of $\hat{\rho}_{\tau}$ or τ_{τ} in Fuller (1979, pp. 371, 373). The percentiles of these distributions are considerably to the left of those for $\hat{\rho}_{\mu}$ or τ_{μ}. For example, if we have 100 observations, the third table of Fuller (1976, p. 371) implies that $\hat{\rho}_{\tau}$ must be less than .73 to declare a detrended series stationary at the 1% level. Removing a time trend from a random walk makes it look stationary, so we must have a very low estimate of ρ to have statistical evidence of stationarity.

It is appparent that we cannot, a priori, use the tests with residuals from arbitrary regression adjusted series. An article dealing with cases in which regression residuals can be used is Fuller, Hasza, and Goebel (1981). We prove, in Appendix B, a special case in which a unit root time series has seasonal means removed. In this case the limit distributions of $n(\hat{\rho}_\mu-1)$ and τ_μ are not affected by the adjustment, and the tables of Dickey and Fuller provide appropriate critical values for large n.

5. Power Considerations

Dickey (1984) generates 2000 series $(Y_t)_{t=1}^n$ satisfying

$$Y_t - \beta_0 - \beta_1 t = \rho(Y_{t-1} - \beta_0 - \beta_1(t-1)) + e_t, t=1,2,..n \qquad (5.1)$$

for each combination $\beta_0 = 0,10$; $\beta_1 = 0, .1$; $n = 20, 50, 100$; and various ρ values. He computes powers of one and two sided 5% level tests for the statistics reported in this paper and some others. We summarize the one sided results $(H_1: \rho < 1)$.

The powers increase with sample size n. Let >> denote "is more powerful than". Then when $\beta_0 = \beta_1 = 0 = E(Y_0)$ in (5.1), $n(\hat{\rho}-1) >> \tau >> n(\hat{\rho}_\mu-1) >> \tau_\mu >> n(\hat{\rho}_\tau-1) >> \tau_\tau$. The powers of $n(\hat{\rho}_\mu-1)$, τ_μ, $n(\hat{\rho}_\tau-1)$ and τ_τ are the same for all β_0 in (5.1). The powers of $n(\hat{\rho}_\tau-1)$ and τ_τ are the same for all β_0 and β_1. For $\beta_0 \neq 0$ and $\beta_1 = 0$, $n(\hat{\rho}_\tau-1)$ and τ_τ have low power. For $\beta_0 \neq 0$ or $\beta_1 \neq 0$ the study shows very low (almost 0) powers for τ and $n(\hat{\rho}-1)$. For $\beta_1 \neq 0$, τ_μ and $n(\hat{\rho}_\mu-1)$ have almost 0 power as well. These are all instances of statistics being used inappropriately. Consider, however, the case where τ_μ, which adjusts only for a mean, is used when the actual process has a deterministic linear trend. Since differencing removes a linear trend, the low power of τ_μ in this case may actually be more comforting than alarming.

We advise the use of τ_μ initially. This allows us to avoid the low powers imparted to $n(\hat{\rho}_\tau-1)$, τ_τ, $n(\hat{\rho}-1)$, and τ by $\beta_0 \neq 0$. We believe alternatives with nonzero means ($\beta_0 \neq 0$) are the most common in practice, unless the series has already been differenced.

We do not advise using the $n(\hat{\rho}_\mu-1)$ test initially because in higher order and mixed models, $n(\hat{\rho}_\mu-1)$ does not remain as faithful to the nominal significance level as does τ_μ. This failure to achieve the nominal level is not too bad, for example, in the model $(1 - \phi B)\nabla Y_t = e_t$. Over the interval $-.9 < \phi < .9$ with n=50 and a nominal .05 level, the worst observed $n(\hat{\rho}_\mu-1)$ level was .085 at $\phi = .9$. The τ_μ empirical level is always closer to .05 than that of $n(\hat{\rho}_\mu-1)$. See Dickey (1984) for details. The departure from nominal level for $\nabla Y_t = (1-\theta B)e_t$ is much worse. In the range $-.8 < \theta < .8$ studied by Said and Dickey (1985), even when the true value of θ is used as an initial input to the Gauss-Newton estimation procedure ($\theta = .8$, n=50, nominal level .05), the empirical level is .18 for $n(\hat{\rho}_\mu-1)$ and .08 for τ_μ. This poor behavior is likely a result of the near cancellation of the operators $\nabla = (1-B)$ and $(1-\theta B)=(1-.8B)$. The empirical levels deteriorated to .32 and .14 respectively when a Durbin initial estimate of θ was used and further to .34 and .23 when the initial estimate was taken from the autocorrelation of the differenced data. This is the worst case reported by Dickey and Said. For example, dropping θ from .8 to .5 produced a .05 empirical level for τ_μ using the Durbin initial estimate.

Table 1 shows empirical powers for the test statistics mentioned here. The table is based on 2000 replications with $\beta_0 = \beta_1 = 0$, series length n=50,100 and using one sided 5% level tests in model (5.1). Let $\delta = \rho-1$ from the left table margin. We could compute the power of, say, the test $H_0:\rho=.6$ versus $H_1:\rho < .6$ when ρ is actually $.6+\delta$ and n=50 or 100. To do this we use the approximation $\sqrt{n}(\hat{\rho}-\rho) \overset{\bullet}{\sim} N(0,1-\rho^2)$. The power of this rather standard test is given as a reference power in the last column of the table. Our purpose is to show that the unit root tests have reasonable power when compared to a test which follows from commonly used stationary theory.

6. Examples

6.1 Iron and Steel Exports

In this example we investigate the stationarity of a series of iron and steel exports from the U.S., excluding scrap, as reported in Metal Statistics (1981), page 196. The 44 observations are yearly exports in millions of tons from 1937 through 1980. We investigate the logarithms, Y_t, of these amounts, which are plotted in Figure 3.

The first 5 autocorrelations are 1.00, .50, .09, .05, .10. The autocorrelations, partial, and inverse autocorrelations appear consistent with either an order 1 autogressive or a moving average model. We choose the autoregressive representation and regress Y_t on 1, Y_{t-1} obtaining

$$\hat{Y}_t = .695 + .510 Y_{t-1} \text{ with } \hat{\sigma}^2 = 0.12$$
$$(.196) \quad (.135)$$

Numbers in parentheses below the coefficients are standard errors and $\hat{\sigma}^2$ is the regression error mean square.

We compute $n(\hat{\rho}_\mu - 1) = 44(.5104 - 1) = -21.54$ which is less than even the 1% critical value, -18.5, interpolated from the middle display in Fuller (1976), page 371. We also compute τ_μ = (.5104-1)/.1350 = -3.63 which is compared to a 1% left tail critical value -3.62 from Appendix C. We conclude that the series is stationary.

A minor simplification would have arisen from regressing ∇Y_t on 1, Y_{t-1}. Now the estimated coefficient on Y_{t-1} is -.4896 and the computer program calculates $\hat{\tau}_\mu = -3.63$ for us. If the program reports a "P-value" (.0008 in our case), it is not correct as it has been computed from a t distribution rather than the τ_μ distribution. We have seen that the correct P-value is .01 for our data.

Finally, we could have tested for an additional lag as well as a unit root if we had regressed ∇Y_t on 1, Y_{t-1}, and ∇Y_{t-1}. For this regression one obtains

$$\nabla \hat{Y}_t = \underset{(.218)}{.867} - \underset{(.151)}{.603} Y_{t-1} + \underset{(.153)}{.224} \nabla Y_{t-1}$$

If we expect additional lags, then we are not justified in comparing $n(\hat{\rho}_\mu - 1) = 44(-.603) = -26.52$ to the table on page 371 of

Fuller (1976). We <u>are</u> justified in compairing τ_μ = -.603/.151 = -4.00 to the table in Appendix C. In the present case this makes little difference since the additional lag coefficient, .224 has a t-statistic .224/.153 = 1.47. We test this against a standard t or normal table and conclude that the additional lag is not needed.

6.2 Birth rates

Here we investigate U.S. birth rates. The 33 observations (Y_t), births per thousand women aged 20 to 24, from 1948 through 1980, are plotted in Figure 4a and their first differences are plotted in Figure 4b. The sample autocorrelation functions of Y_t and ∇Y_t and the partial autocorrelations of ∇Y_t are displayed in Table 2.

The usual identification techniques (examination of plots and autocorrelation functions) leave little doubt about the need to difference Y_t, but there is doubt about the treatment of ∇Y_t. The time series plot of ∇Y_t suggests possible nonstationarity, and the autocorrelation function of ∇Y_t dies out somewhat slowly. Moreover, the number of observations (32 for ∇Y_t) is small. The unit root test may provide useful guidance in this example.

The sample ACF and PACF of ∇Y_t suggest an AR(1) model. Fitting this yields

$$\nabla Y_t = \underset{(1.44)}{-1.079} + \underset{(.149)}{.594}\, \nabla Y_{t-1} + e_t \qquad \hat{\sigma}^2 = 60.18 \qquad (6.1)$$

To test whether we need to difference ∇Y_t we compute τ_μ = (.594-1)/.149 = -2.73, and upon entering the table in Appendix C, we find that the p-value of the test is between .10 and .05. In this small-sample situation, even $\hat{\rho}$ = .6 does not definitely rule out the possibility of ρ = 1 - the test states that differencing is plausible at the 5 percent significance level. (In the next example we conclude that for total population second differencing is necessary to achieve a stationary series. This may lend plausibility to second differencing the birth rates.)

In ambigous situations such as this, examining the implications of the models, such as by comparing forecasts, can facilitate making a choice between models. In addition to (6.1), the models

$$\nabla Y_t = \underset{(.143)}{.622}\nabla Y_{t-1} + e_t \qquad \hat{\sigma}^2 = 59.30 \tag{6.2}$$

and

$$\nabla^2 Y_t = (1-\underset{(.156)}{.459}B)e_t, \qquad \hat{\sigma}^2 = 59.30 \tag{6.3}$$

were fitted to the birth rate series. Model (6.1) is a natural alternative to a second difference model, though if a first difference model is deemed correct, then the constant term is not statistically significant and model (6.2) is acceptable. Model (6.3) comes from applying the usual identification procedures to $\nabla^2 Y_t$. Given the low value of $\hat{\rho}$ in model (6.1), it is not surprising that $\nabla^2 Y_t$ looks more like an MA(1) than a white noise series.

Table 3 displays 10 forecasts and forecast standard errors from 1980 generated from the three models. Model (6.1) forecasts a precipitous drop in the fertility rate, whereas models (6.2) and (6.3) forecast very little change. Also model (6.3) estimates much larger forecast standard errors than models (6.1) and (6.2) beyond four steps ahead. Analysts with very different opinions about the future behavior of fertility rates could pick a model corresponding to their opinion.

6.3 Population

The data, midyear resident U.S. population, 1929-1982, are plotted in Figure 5. The number of observations is n=51. Let Y_t denote the population in year t. The strong trend in the series suggests taking ∇Y_t. This is confirmed by fitting an AR(2) model to the data. The fit is

$$(1-1.8050B + .8036B^2)Y_t = e_t$$

and we see that $\hat{\phi}_1 + \hat{\phi}_2$ is very close to 1, which indicates a unit root of +1. Inspection of the sample autocorrelation function also strongly suggests the need for differencing. A plot of ∇Y_t appears in Figure 6. The sample autocorrelation functions of ∇Y_t and $\nabla^2 Y_t$ are presented in Table 4. To test whether we need to difference ∇Y_t ,we regress $\nabla^2 Y_t$ on 1, ∇Y_{t-1}, and $\nabla^2 Y_{t-1}$, obtaining

$$\nabla^2\hat{Y}_t = \underset{(119.0)}{217.8} - \underset{(.054)}{.0914}\nabla Y_{t-1} - \underset{(.138)}{.124}\nabla^2 Y_{t-1}, \quad \hat{\sigma}^2 = 73003$$

If the coefficient of ∇Y_{t-1} is not significant, then we cannot reject the hypothesis that second differencing of Y_t is appropriate. Regardless of the decision about this coefficient, the coefficient of $\nabla^2 Y_{t-1}$ is an estimate of an autoregressive parameter whose significance can be judged from standard t-tables. To test the hypothesis that a difference is needed, we use the statistic $\tau_\mu = -.0914/.054 = -1.69$. In the table in Appendix C we find the 10% critical value is approximately -2.60.

The test does not allow us to reject the hypothesis that we should difference ∇Y_t . The studentized estimate of the coefficient of $\nabla^2 Y_{t-1}$, -.124/.138=-.898, is not significant when compared to the 10% critical value of a Student-t distribution with 48 degrees of freedom. There appears to be no significant lag 1 autocorrelation in the second difference of the midyear population figures. Examination of the autocorrelation function of $\nabla^2 Y_t$ suggests that we can treat this series as if it were white noise, i.e., the model for midyear population is an ARIMA(0,2,0).

Hasza and Fuller (1979) presented a test for a double unit root (second difference). To apply the test to our population data we regress $\nabla^2 Y_t$ on ∇Y_{t-1} and Y_{t-1}. The fit is

$$\nabla^2\hat{Y}_t = \underset{(.0652)}{-.0938}\nabla Y_{t-1} + \underset{(.00083)}{.00123}Y_{t-1}$$

The F-statistic for testing for the joint significance of the coefficients of ∇Y_{t-1} and Y_{t-1} is F=83,936/75,098=1.118. Again we have a test statistic that is computed by ordinary least squares formulas, but its distribution under the null hypothesis of a double unit root is not that of Snedecor's F. The appropriate distribution is that labelled $\Phi_1(2)$ on page 1116 of Hasza and Fuller (1979). The 80th percentile is 2.03, and the median is 0.97. By interpolation, the P-value is roughly 0.46. Thus we do not reject the null hypothesis of a double unit root.

We have now presented two analyses that support the conclusion that the model for the midyear population figures is ARMA(0,2,0). The standard deviation of $\nabla^2 Y_t$ is estimated to be 0.28. The ARIMA(0,2,0) model implies that the minimum mean squared error forecast of Y_{n+1} from origin n is $\hat{Y}_{n+1} = 2Y_n - Y_{n-1}$, of Y_{n+2} from origin n is $\hat{Y}_{n+2} = 2\hat{Y}_{n+1} - Y_n = 3Y_n - 2Y_{n-1}$, and of $Y_{n+\ell}$ from origin n is $\hat{Y}_{n+\ell} = 2\hat{Y}_{n+\ell-1} - \hat{Y}_{n+\ell-2} = Y_n + (Y_n - Y_{n-1})\ell$ for $\ell > 3$. The forecast function is a straight line whose slope and intercept are determined by the values of Y_{n-1} and Y_n. The estimated forecast standard error at lead time ℓ is computed from the formula $0.28(1+2^2+3^2+\ldots+(\ell-1)^2)^{1/2}$, $\ell > 1$.

6.4 Solar Radiation

The data are monthly averages of daily measurements of the radiant energy per unit area received at the earth's surface (corrected for atmospheric effects) at a station in California. Data from October, 1937 through May, 1954 (200 observations) were obtained from the National Bureau of Standards in Boulder, Colorado. The data are discussed in Hoyt (1979).

The observed data are seasonal. Empirically, and from physical considerations, it makes sense to assume the seasonal effects remain fixed from year to year. We let W_t be the observed series and assume $W_t = Y_t+\mu_t$ where $\mu_{t+12} = \mu_t$ is a sequence of monthly means and Y_t is the nonseasonal series whose stationarity we would like to test. Following results in Appendix B, we subtract monthly means, $\overline{W}_{i(t)}$, (where i(t) is the month i corresponding to observation t) from W_t to get $X_t = W_t-\overline{W}_{i(t)}$. The X_t series is plotted in Figure 7. It is difficult to judge from this plot whether Y_t (estimated by X_t) is stationary.

The ACF of X_t is given in Table 5. While the autocorrelations are not large and do die out with increasing lag, their decay is not clearly exponential and they are all positive through 20 lags. The PACF of X_t (Table 5) suggests an AR(2) model. Thus the null and alternative hypotheses are

$$H_0: W_t = Y_t+\mu_t \qquad (1-\phi B)\nabla Y_t = e_t$$

$$H_1: W_t = Y_t+\mu_t \qquad (1-\phi B)(1-\rho B)Y_t = e_t$$

Notice that μ_t takes care of both seasonal effects and the overall mean level of W_t, so there is no need to include a mean in the model for Y_t in H_1.

An AR(2) model without a mean was fit to X_t giving $(1-.38B-.25B^2)X_t = (1-.72B)(1+.34B)X_t = e_t$. Testing H_0 versus H_1 will give us a formal test of whether (1-.72B) is significantly different from $\nabla = (1-B)$. We regress ∇X_t on X_{t-1}, ∇X_{t-1} (again, no constant term is needed since seasonal means have been removed) and get

$$\nabla\hat{X}_t = -.369X_{t-1} - .247\nabla X_{t-1} \qquad \hat{\sigma}^2 = .00513$$
$$\quad (.0689) \qquad (.0692)$$

The τ-statistic is $\tau_\mu = -.369/.0689 = -5.35$, which we compare to the table in Appendix C. Interpolating in the table for $n = 200$ we see our statistic is highly significant at 1%, so there is strong evidence against differencing Y_t and our model for the data would be that given under H_1.

6.5 Housing starts

The data, monthly U.S. total single family housing starts, 1/64-8/78, are plotted in Figure 8.a. Let Y_t denote the number of housing starts in month t. The seasonal behavior of the series calls for the transformation $X_t = \nabla_{12}Y_t$. X_t is plotted in Figure 8.b., and the sample ACF and PACF of X_t are given in Table 6. Initially the model

$$(1-\phi_1 B-\phi_2 B^2)X_t = c + (1-\theta B^{12})e_t \tag{6.4}$$

was fitted with $\hat{\phi}_1 = .71$, $\hat{\phi}_2 = .24$, $\hat{\theta} = .86$, and $c = .30$. The autoregressive operator on the left-hand side of (6.4) can be factored to yield approximately $(1-.96B)(1+.26B)$, so there may be a unit root. The model

$$(1-\phi B)\nabla X_t = (1-\theta B^{12})e_t \tag{6.5}$$

was fitted with $\hat{\phi} = -.26$ and $\hat{\theta} = .85$.

To test for a unit root we follow the procedure proposed by Dickey and Said (1981). Fit the model

$$(1-\rho B)(1-\phi B)X_t = c + (1-\theta B^{12})e_t \qquad (6.6)$$

but allow only <u>one</u> Gauss-Newton iteration using starting values $\rho^{(0)} = 1$, $\phi^{(0)} = -.26$, $\theta^{(0)} = .86$, and $c^{(0)} = .30$. The starting values of ϕ, θ, and c are taken from the fit of (6.4), and the starting value of ρ is the null hypothesis value 1. When the procedure is implemented, we find $\hat{\rho} = .9522(.026)$, $\hat{\phi} = -.26(.08)$, $\hat{c} = .30(.22)$, and $\hat{\theta} = .85(.06)$. The numbers in parentheses are standard errors. To test H_0: $\rho = 1$ we use the statistic $\tau_\mu = (.9522-1)/.026 = -1.84$. To determine the critical value of the test we enter the table in Appendix C to find that the 10% critical value is between -2.58 and -2.57. The test does not allow us to reject the null hypothesis of a unit root, even at the 10% significance level.

The estimates in the previous paragraph were obtained by pure Gauss-Newton iteration, as is strictly required for the theory to apply. As noted earlier, many time series estimation packages use Marquardt's algorithm, which is a mixture of Gauss-Newton and steepest ascent. For the housing starts data, one step of the Marquardt algorithm, using the starting values stated above, yields $\tau_\mu = -2.5$, which differs considerably from the value computed by the Gauss-Newton algorithm.

References

Abraham, B. and Box, G.E.P. (1978) "Deterministic and Forecast Adaptive Time-dependent Models," Applied Statistics, 27, 120-130.

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

Chan, K. H., Hayya, J.C., and Ord, J.K. (1977) "A Note on Trend Removal Methods: The Case of Polynomial Regression Versus Variate Differencing," Econometrica, 45, 737-744.

Dickey, D.A. (1984) "Powers of Unit Root Tests," Proceedings of the American Statistical Association, Business and Economic Statistics Section, to appear.

Dickey, D.A. and Fuller, W.A. (1979) "Distribution of the estimators for autoregressive time series with a unit root," Journal of the American Statistical Association, 74, 427-431.

__________ (1981) "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root," Econometrica, 49, 1057-1072.

Findley, D.F. (1980) "Large Sample Behavior of the S-array of Seasonally Nonstationary ARMA Series," in Time Series Analysis, O.D. Anderson and M.R. Perryman, eds., Amsterdam: North Holland, 163-170.

Fuller, W.A. (1976) Introduction to Statistical Time Series, New York: John Wiley and Sons.

__________ (1984). "Nonstationary Autoregressive Time Series", in Handbook of Statistics, Vol. 5: Time Series in the Time Domain, eds. P.R. Krishnaiah, M.M. Rao, and E.J. Hannan, Amsterdam: North Holland, to appear.

Fuller, W.A., Hasza, D.P., and Goebel, J.J. (1981) "Estimation of the Parameters of Stochastic Difference Equations," Annals of Statistics, 9, 531-543.

Granger, C.W.J. and Newbold, P. (1977) Forecasting Economic Time Series, New York: Academic Press.

Harvey, A.C. (1981) "Finite Sample Prediction and Overdifferencing," Journal of Time Series Analysis, 2, 221-232.

Hasza, D.P. (1980) "The Asymptotic Distribution of the Sample Autocorrelations for an Integrated ARMA Process," Journal of the American Statistical Association, 75, 349-352.

Hasza, D.P. and Fuller, W.A. (1979) "Estimation for Autoregressive Processes with Unit Roots," _Annals of Statistics,_ 7, 1106-1120.

Hillmer, S.C. and Tiao, G.C. (1979) "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," _Journal of the American Statistical Association,_ 74, 652-660.

Hoyt, D.V. (1979) "The Smithsonian Astrophysical Observatory Solar Constant Program," _Reviews of Geophysics and Space Physics,_ 17, 427-458.

Nelson, C.R. and Kang, H. (1981) "Spurious Periodicity in Innappropriately Detrended Time Series," _Econometrica,_ 49, 741-751.

__________ (1984) "Pitfalls in the Use of Time as an Explanatory Variable in Regression," _Journal of Business and Economic Statistics,_ 2, 73-82.

Nelson, C.R. and Plosser, C.I. (1982) "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," _Journal of Monetary Economics,_ 10, 139-162.

Pierce, D.A. (1975) "On Trend and Autocorrelation," _Communications in Statistics,_ 4, 163-175.

Plosser, C.I. and Schwert, G.W. (1978) "Money, Income, and Sunspots: Measuring Economic Relationships and the Effects of Differencing," _Journal of Monetary Economics,_ 4, 637-660.

Said, S.E., and Dickey, D.A. (1984) "Testing for Unit Roots in ARMA(p,q) Models with Unknown p and q" _Biometrika_, 71, 599-607.

Said, S.E., and Dickey, D.A. (1985) "Hypothesis Testing in ARIMA(p,1,q) Models," _Journal of the American Statistical Association_, 80, to appear.

Solo, V. (1984) "The Order of Differencing in ARIMA Models," _Journal of the American Statistical Association,_ 79, 916-921.

Tiao, G.C. and Tsay, R.S. (1983) "Consistency Properties of Least Squares Estimates of Autoregressive Parameters in ARMA Models," _Annals of Statistics,_ 11, 856-871.

Wichern, D.W. (1973) "The Behavior of the Sample Autocorrelation Function for an Integrated Moving Average Process," _Biometrika_, 60, 235-239.

Appendix A

Y_t is said to be strictly stationary if the joint density of any set of observations $Y_{t_1}, \ldots, Y_{t_m}$ is unaffected by a shift in time origin to $Y_{t_1+k}, \ldots, Y_{t_m+k}$. Y_t is second order stationary if its mean function is constant ($E(Y_t) = \mu$ for all t) and if its covariance function, $\mathrm{Cov}(Y_t, Y_{t+k}) = \gamma_k$, does not depend on t. The two types of stationarity are equivalent for Gaussian time series. Our remarks will refer to second-order stationarity.

We illustrate considerations for Y_t following the AR(1) model (2.1). Assuming Y_0 independent of e_t, $t \geq 1$, we have from (2.2)

$$\mathrm{Var}(Y_t) = \alpha^{2t}\mathrm{Var}(Y_0) + (1+\alpha^2+\cdots+\alpha^{2(t-1)})\sigma^2 \qquad (A.1)$$

From (A.1) we notice that Y_t cannot be stationary for $|\alpha| \geq 1$, because in this case $\mathrm{Var}(Y_t)$ increases without bound as $t\to\infty$. For Y_t to be stationary when $|\alpha| < 1$ we need $\mathrm{Var}(Y_t) = \mathrm{Var}(Y_0) = \gamma_0$. It can be shown that if $\mathrm{Var}(Y_0) = \sigma^2/(1-\alpha^2)$, then $\mathrm{Var}(Y_t) = \sigma^2/(1-\alpha^2) = \gamma_0$ and $\mathrm{Cov}(Y_t, Y_{t+k}) = \alpha^k\gamma_0$ for all t. If $\mathrm{Var}(Y_0) \neq \sigma^2/(1-\alpha^2)$, then $\mathrm{Var}(Y_t)$ will not be a constant free of t, (as long as $\sigma^2 > 0$), and Y_t will not be stationary. So for Y_t following (2.1) to be stationary, it is not enough to have $|\alpha| < 1$ and Y_0 independent of $e_1, e_2, \ldots$; we also need $\mathrm{Var}(Y_0) = \sigma^2/(1-\alpha^2)$.

Another way to look at this is the following. In obtaining (2.2) there is no reason why the starting value has to be at $t = 0$. It could equally well be at $t = -n$, say, giving

$$Y_t = \alpha^{t+n} Y_0 + e_t + \alpha e_{t-1} + \cdots + \alpha^{t+n-1} e_{1-n}$$

in place of (2.2). Letting $n \to \infty$ this becomes

$$Y_t = \sum_0^\infty \alpha^j e_{t-j} \tag{A.2}$$

and the infinite sum converges in mean square since $|\alpha| < 1$. Using (A.2) we can show that

$$\mathcal{E}(Y_t) = 0$$
$$\mathrm{Var}(Y_t) = \sigma^2/(1-\alpha^2) = \gamma_0, \quad \mathrm{Cov}(Y_t, Y_{t+k}) = \alpha^k \gamma_0$$

for all t, so that Y_t arising this way is stationary. Here we have pushed the starting value back to $t = -\infty$ where it has no effect on Y_t for $|\alpha| < 1$. For $|\alpha| \geq 1$ this cannot be done - the starting value must occur at some finite time point, and the starting value will affect Y_t for all t.

In (A.2) notice that $Y_0 = \sum_0^\infty \alpha^j e_{-j}$ has $\mathrm{Var}(Y_0) = \sigma^2/(1-\alpha^2)$, so that this scheme can be related to the previous one by thinking of (A.2) as generating Y_0 as a starting value with the correct variance.

These remarks extend to higher order models. For Y_t following a pth order model (1.1) to be stationary, we need the zeroes $\xi_1, \ldots, \xi_p$ of $\alpha(z)$ to all be such that $|\xi_i| > 1$ (analogous to $|\alpha| < 1$ in (2.1)), but this alone is not sufficient. We also need to assume that p starting values, say $Y_0, Y_{-1}, \ldots, Y_{1-p}$, are independent of shocks $e_1, e_2, \ldots$ and that $Y_0, \ldots, Y_{1-p}$ have the correct variances and covariances. Equivalently (when $|\xi_i| > 1$, $i = 1, \ldots, p$) we can assume Y_t arises from a particular infinite linear combination of the current and past shocks: $Y_t = \sum_0^{\infty} \Psi_j e_{t-j}$ (the Wold decomposition). This can be viewed as generating starting values $Y_{1-p}, \ldots, Y_0$ independent of $e_1, e_2, \ldots$ and with the correct covariance structure.

Appendix B

Here we show that removal of seasonal means from autoregressive series which are stationary in first differences has no effect on limit distributions of unit root test statistics.

Let $Y_t = Y_{t-1} + e_t$. Assume that we observe $Y_1, Y_2, \ldots, Y_n$ where n=12m and m is an integer, that is m years of monthly data. Let $\overline{Y}_i = m^{-1} \sum_{t=0}^{m-1} Y_{12t+i}$, $\overline{Y} = n^{-1} \sum_{t=1}^{n} Y_t$ and $X_t = Y_t - \overline{Y}_{i(t)}$ where i(t) is the month i in which observation t occurs. Let μ_t be a sequence of constants such that $\mu_{t+12} = \mu_t$ for all t and let $W_t = Y_t + \mu_t$. Notice that $W_t - \overline{W}_{i(t)} = Y_t - \overline{Y}_{i(t)} = X_t$. If we can show our result for Y_t it will hold for W_t.

Now

$$X_t = Y_t - \overline{Y} - \overline{Y}_{i(t)} + \overline{Y},$$

and $Y_t = e_t + e_{t-1} + \ldots + e_1$ so $\overline{Y} = n^{-1}(ne_1 + (n-1)e_2 + \ldots + e_n)$. Furthermore,

$$\overline{Y}_i = m^{-1}(Y_i + Y_{i+12} + \ldots + Y_{n-12+i}) =$$

$$n^{-1}(n(e_1 + e_2 + \ldots + e_i) + (n-12)(e_{i+1} + e_{i+2} + \ldots + e_{i+12})$$

$$+ \ldots + (e_{n-23+i} + e_{n-22+i} + \ldots + e_{n-12+i})).$$

Finally, we see that $n(\overline{Y}_i - \overline{Y}) = \sum_{t=1}^{n} C_{it} e_t$ where $-11 < C_{it} < 11$ for all t,i. It follows that $(\overline{Y}_i - \overline{Y}) = 0_p(n^{-1/2})$ so that

$$n^{-2}\Sigma X_t^2 = n^{-2}\Sigma(Y_t-\overline{Y})^2 + O_p(n^{-2})$$

and

$$n^{-1}\Sigma X_{t-1}e_t = n^{-1}\Sigma(Y_{t-1}-\overline{Y})e_t + O_p(n^{-1}).$$

Let $\hat{\rho}_s$ denote the regression estimate of ρ obtained by regressing X_t on X_{t-1}. We have shown that $n(\hat{\rho}_s-1) = n(\hat{\rho}_\mu-1) + O_p(n^{-1})$ where $n(\hat{\rho}_\mu-1) = [n^{-2}\Sigma(Y_{t-1}-\overline{Y})^2]^{-1}[n^{-1}\Sigma(Y_t-\overline{Y})(e_t-\overline{e})]$ so the limit distribution is not affected by removal of seasonal means. Using the results of Dickey and Fuller (1979) the same holds true for τ_μ and for higher order autoregeressions providing they have only 1 unit root.

Appendix C

Table C1. Empirical Cumulative Distribution of $\hat{\tau}_\mu$ for $\rho = 1$.

Sample Size n	Probability of a Smaller Value 0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
				$\hat{\tau}_\mu$				
25	-3.75	-3.33	-3.00	-2.63	-0.37	0.00	0.34	0.72
50	-3.58	-3.22	-2.93	-2.60	-0.40	-0.03	0.29	0.66
100	-3.51	-3.17	-2.89	-2.58	-0.42	-0.05	0.26	0.63
250	-3.46	-3.14	-2.88	-2.57	-0.42	-0.06	0.24	0.62
500	-3.44	-3.13	-2.87	-2.57	-0.43	-0.07	0.24	0.61
∞	-3.43	-3.12	-2.86	-2.57	-0.44	-0.07	0.23	0.60

Source: Fuller (1976), used with permission.

APPENDIX D

Data and sources are listed here for the five examples.

Example 1. -- U.S. Iron and Steel Exports, Excluding Scrap, 1937-80
(Weight in millions of tons)

Source: Metal Statistics, New York: American Metal Market, 1982.

3.89	2.41	2.80	8.72	7.12	7.24	7.15	6.05	5.21	5.03
6.88	4.70	5.06	3.16	3.62	4.55	2.43	3.16	4.55	5.17
6.95	3.46	2.13	3.47	2.79	2.52	2.80	4.04	3.08	2.28
2.17	2.78	5.94	8.14	3.55	3.61	5.06	7.13	4.15	3.86
3.22	3.50	3.76	5.11						

Example 2. -- Births per Thousand U.S. Women Aged 20-24, 1948-80

Source: Bureau of the Census

192.4	194.1	192.8	207.1	213.5	220.5	231.7	236.3	248.5	257.0
252.0	252.7	246.1	253.0	243.2	227.6	215.4	190.0	178.9	170.2
163.6	162.8	163.1	149.1	128.8	119.4	117.7	113.6	112.1	115.2
112.3	115.7	115.1							

Example 3. -- Midyear Resident Population of the U.S., 1929-82
(in thousands of persons)

Source: Bureau of the Census

121767	123077	124040	124840	125579	126374	127250	128053
128825	129825	130880	132594	133894	135361	137250	138916
140468	141936	144698	147208	149767	152271	154878	157553
160184	163026	165931	168903	171984	174882	177830	180671
183691	186538	189242	191889	194303	196560	198712	200706
202677	205052	207661	209896	211909	213854	215973	218035
220239	222585	225055	227704	229849	232057		

Example 4. -- Solar Radiation Corrected for Atmospheric Effects, October 1937 through May 1954 (measured at Table Mountain, CA)

Source: National Bureau of Standards, Boulder, CO

1.9454	1.9475	1.9484	1.9405	1.9475	1.9489
1.9383	1.9466	1.9425	1.9444	1.9393	1.9459
1.9465	1.9502	1.9523	1.9487	1.9392	1.9413
1.9287	1.9495	1.9480	1.9473	1.9454	1.9514
1.9512	1.9527	1.9554	1.9474	1.9454	1.9440
1.9419	1.9417	1.9421	1.9441	1.9397	1.9454
1.9419	1.9450	1.9450	1.9369	1.9386	1.9423
1.9358	1.9428	1.9486	1.9463	1.9457	1.9483
1.9494	1.9532	1.9525	1.9472	1.9501	1.9448
1.9570	1.9501	1.9530	1.9482	1.9433	1.9416
1.9437	1.9487	1.9446	1.9313	1.9457	1.9423
1.9451	1.9381	1.9496	1.9464	1.9487	1.9460
1.9486	1.9403	1.9430	1.9506	1.9545	1.9534
1.9487	1.9476	1.9424	1.9478	1.9503	1.9414
1.9391	1.9412	1.9429	1.9310	1.9396	1.9143
1.9307	1.9470	1.9354	1.9338	1.9402	1.9439
1.9482	1.9531	1.9521	1.9516	1.9408	1.9380
1.9466	1.9499	1.9514	1.9470	1.9463	1.9492
1.9490	1.9552	1.9571	1.9539	1.9521	1.9412
1.9425	1.9527	1.9552	1.9463	1.9500	1.9439
1.9537	1.9530	1.9597	1.9505	1.9546	1.9502
1.9510	1.9543	1.9608	1.9560	1.9554	1.9548
1.9506	1.9560	1.9533	1.9573	1.9576	1.9487
1.9499	1.9557	1.9452	1.9498	1.9500	1.9509
1.9484	1.9516	1.9559	1.9572	1.9520	1.9460
1.9447	1.9457	1.9455	1.9485	1.9468	1.9463
1.9523	1.9505	1.9505	1.9427	1.9487	1.9439
1.9326	1.9530	1.9490	1.9504	1.9462	1.9448
1.9477	1.9480	1.9450	1.9467	1.9411	1.9421
1.9486	1.9508	1.9474	1.9381	1.9381	1.9437
1.9399	1.9410	1.9492	1.9566	1.9482	1.9524
1.9439	1.9428	1.9556	1.9467	1.9524	1.9407
1.9474	1.9507	1.9386	1.9410	1.9430	1.9329
1.9480	1.9514				

Example 5. -- U.S. Monthly Single Family Housing Starts, January 1964 through August 1978 (in thousands)

Source: Bureau of the Census

58.008	62.448	82.180	94.927	98.230	100.875
89.885	91.988	79.757	89.435	67.514	55.227
52.149	47.205	82.150	100.931	98.408	97.351
96.489	88.830	80.876	85.750	72.351	61.198
46.561	50.361	83.236	94.343	84.748	79.828
69.068	69.362	59.404	53.530	50.212	37.972
40.157	40.274	66.592	79.839	87.341	87.594
82.344	83.712	78.194	81.704	69.088	47.026
45.234	55.431	79.325	97.983	86.806	81.424
86.398	82.522	80.078	85.560	64.819	53.847
51.300	47.909	71.941	84.982	91.301	82.741
73.523	69.465	71.504	68.039	55.069	42.827
33.363	41.367	61.879	73.835	74.848	83.007
75.461	77.291	75.961	79.393	67.443	69.041
54.856	58.287	91.584	116.013	115.627	116.946
107.747	111.663	102.149	102.882	92.904	80.362
76.185	76.306	111.358	119.840	135.167	131.870
119.078	131.324	120.491	116.990	97.428	73.195
77.105	73.560	105.136	120.453	131.643	114.822
114.746	106.806	84.504	86.004	70.488	46.767
43.292	57.593	76.946	102.237	96.340	99.318
90.715	79.782	73.443	69.460	57.898	41.041
39.791	39.959	62.498	77.777	92.782	90.284
92.782	90.655	84.517	93.826	71.646	55.650
53.997	72.585	92.443	107.804	112.242	119.627
112.807	112.798	108.038	109.114	89.368	71.584
55.746	87.172	125.802	138.772	152.198	149.061
138.181	140.527	131.644	135.398	109.310	87.123
63.349	72.800	121.391	139.857	154.928	154.278
139.219	140.106				

Table 1 Empirical Power

ρ	$\hat{\rho}$	τ	$\hat{\rho}_\mu$	τ_μ	$\hat{\rho}_\tau$	τ_τ	Reference
(n = 50)							
1.000	.06	.05	.05	.05	.04	.05	.05
.999	.05	.05	.03	.04	.03	.04	.05
.990	.06	.07	.06	.06	.05	.06	.06
.980	.07	.07	.07	.06	.05	.05	.07
.950	.15	.16	.10	.07	.05	.06	.13
.900	.31	.31	.19	.11	.10	.08	.24
.850	.55	.57	.29	.20	.15	.12	.38
.800	.77	.77	.46	.30	.24	.20	.54
.700	.97	.97	.77	.61	.48	.39	.81
.500	1.00	1.00	1.00	.98	.91	.87	.99
(n = 100)							
1.000	.04	.04	.05	.05	.05	.06	.05
.999	.05	.05	.05	.04	.04	.05	.05
.990	.08	.08	.07	.06	.05	.05	.07
.980	.13	.12	.09	.06	.06	.05	.09
.950	.31	.31	.19	.12	.09	.08	.16
.900	.78	.78	.47	.31	.23	.19	.36
.850	.96	.96	.77	.60	.46	.37	.58
.800	1.00	1.00	.94	.86	.75	.65	.77
.700	1.00	1.00	1.00	1.00	.98	.95	.96
.500	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 2. Sample Autocorrelations for Y_t and ∇Y_t, and Partial Autocorrelations for ∇Y_t; Birth Rates for U.S. Women, Ages 20-24, 1948-80

	Lag									
	1	2	3	4	5	6	7	8	9	10
ACF of Y_t	.95	.89	.80	.70	.59	.47	.34	.21	.08	-.03
ACF of ∇Y_t	.59	.48	.36	.17	.23	.10	.10	-.01	-.14	-.14
PACF of ∇Y_t	.59	.19	.02	-.16	.20	-.11	.04	-.18	-.10	-.05

NOTE: Standard errors for the above are all .18 or larger.

Table 3. Forecasts (and Standard Errors) of Birth Rates by Three Models from Forecast Origin 1980

Using Model

	(6.1)		(6.2)		(6.3)	
Lead Time	Forecast	Std. Error	Forecast	Std. Error	Forecast	Std. Error
1	113.7	7.8	114.7	7.7	115.3	7.7
2	111.7	14.6	114.5	14.7	115.5	14.1
3	109.5	21.0	114.4	21.3	115.7	21.4
4	107.1	26.9	114.3	27.5	115.9	29.4
5	104.6	32.2	114.2	33.1	116.2	38.2
6	102.0	37.0	114.2	38.3	116.4	47.7
7	99.4	41.4	114.1	43.0	116.6	57.8
8	96.8	45.5	114.1	47.4	116.8	68.6
9	94.2	49.3	114.1	51.5	117.0	79.9
10	91.5	52.8	114.1	55.3	117.2	91.8

Table 4. Sample Autocorrelation Functions of First and Second Differences of U.S. Midyear Resident Population, 1929-1982.

	Lag											
Variable	1	2	3	4	5	6	7	8	9	10	11	12
∇Y_t	.91	.84	.77	.71	.62	.53	.47	.38	.29	.18	.09	.01
$\nabla^2 Y_t$	-.16	-.06	-.09	.14	.10	-.26	.27	-.01	.13	-.01	-.20	.00

Table 5

Sample autocorrelations and partial autocorrelations for the solar radiation series with seasonal means removed

Lag	1	2	3	4	5	6	7	8	9	10
ACF	.49	.42	.35	.28	.24	.22	.12	.14	.07	.11
PACF	.49	.24	.10	.03	.03	.04	-.07	.05	-.05	.06

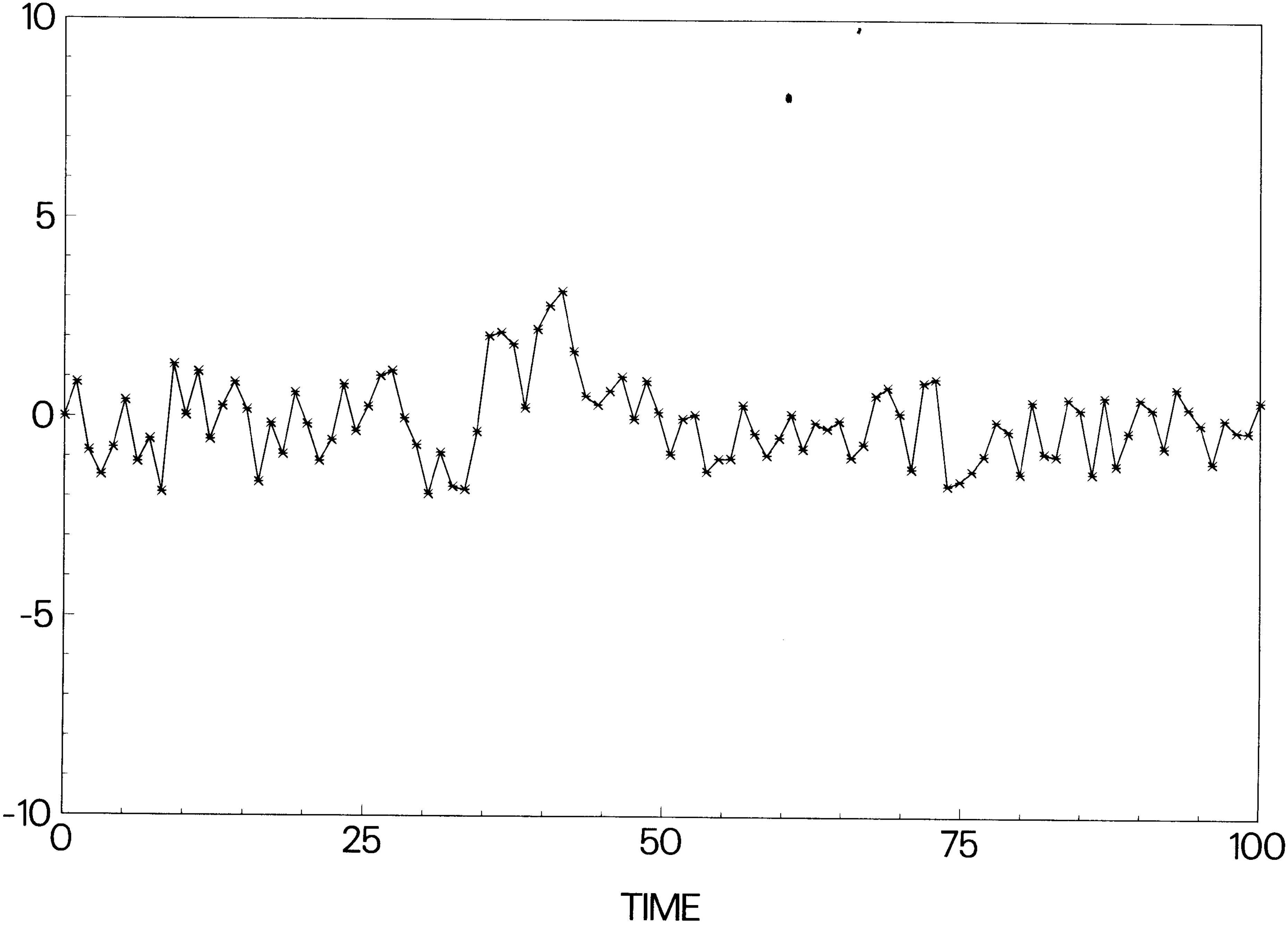
Lag	11	12	13	14	15	16	17	18	19	20
ACF	.08	.03	.01	.04	.01	.10	.04	.07	.08	.08
PACF	.00	-.05	-.03	.04	.00	.10	-.03	.03	.03	.02

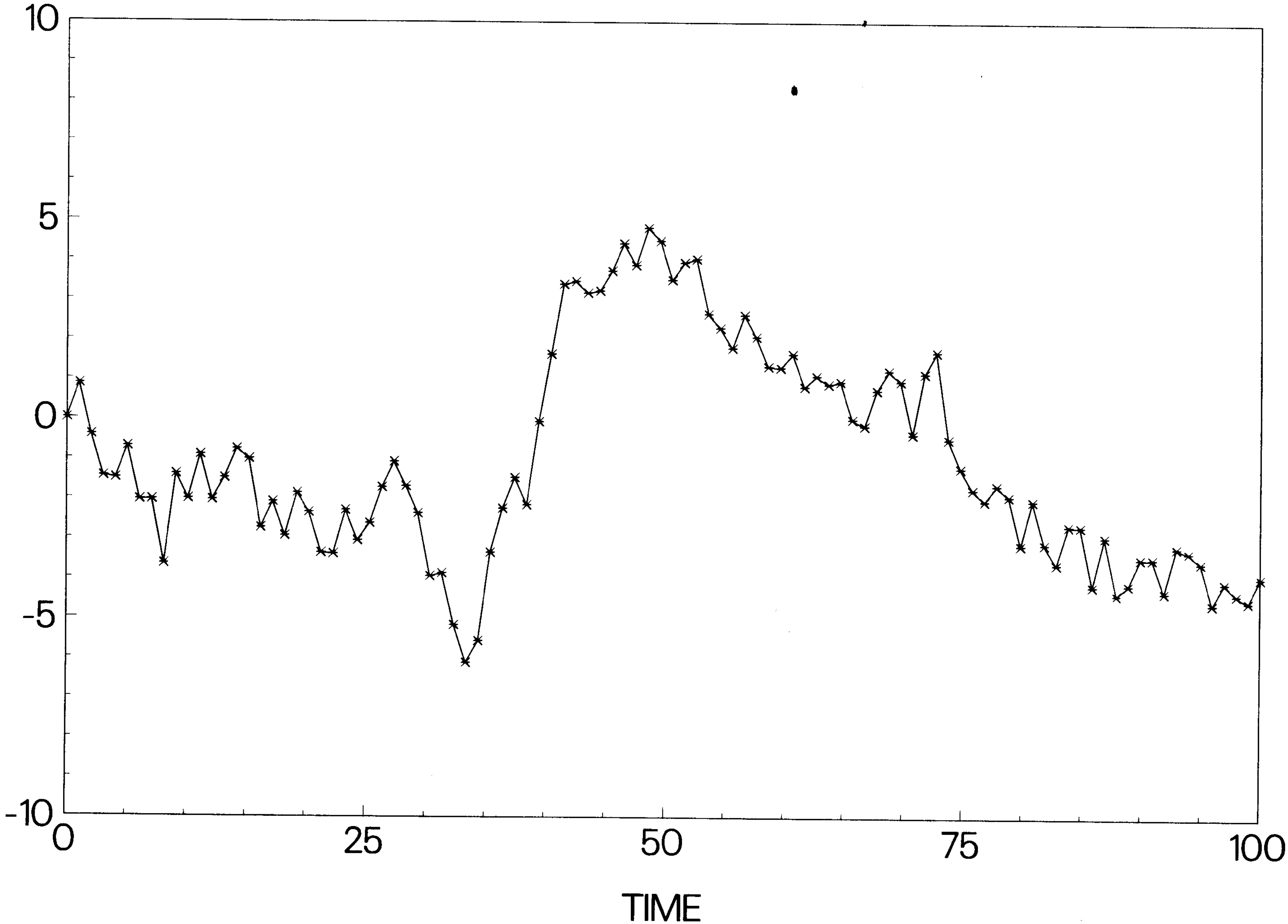
Table 6. Sample ACF and PACF of $X_t = \nabla_{12} Y_t$ and ACF of ∇X_t for Y_t = Monthly U.S. Single Family Housing Starts, January 1964 through August 1978

	Lag 1	2	3	4	5	6	7	8	9	10
ACF of X_t	.87	.81	.76	.70	.64	.56	.48	.42	.35	.26
PACF of X_t	.87	.23	.04	.00	-.05	-.10	-.12	-.01	-.01	-.17
ACF of ∇X_t	-.29	.02	-.01	.02	.05	.04	-.07	.00	.12	-.12

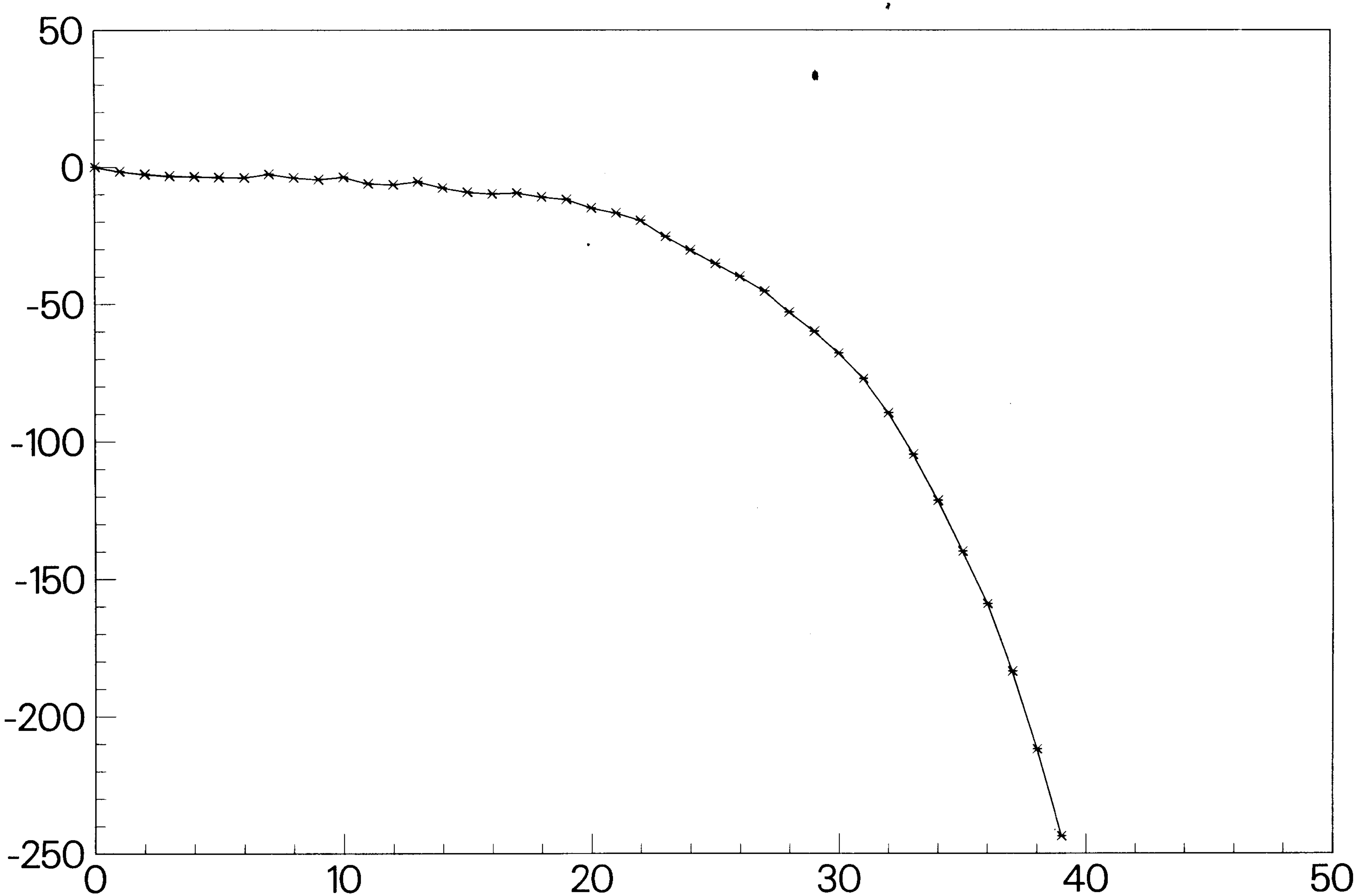
	Lag 11	12	13	14	15
ACF of X_t	.20	.10	.10	.06	-.01
PACF of X_t	.01	-.15	.32	-.04	-.17
ACF of ∇X_t	.13	-.40	.20	.09	-.04

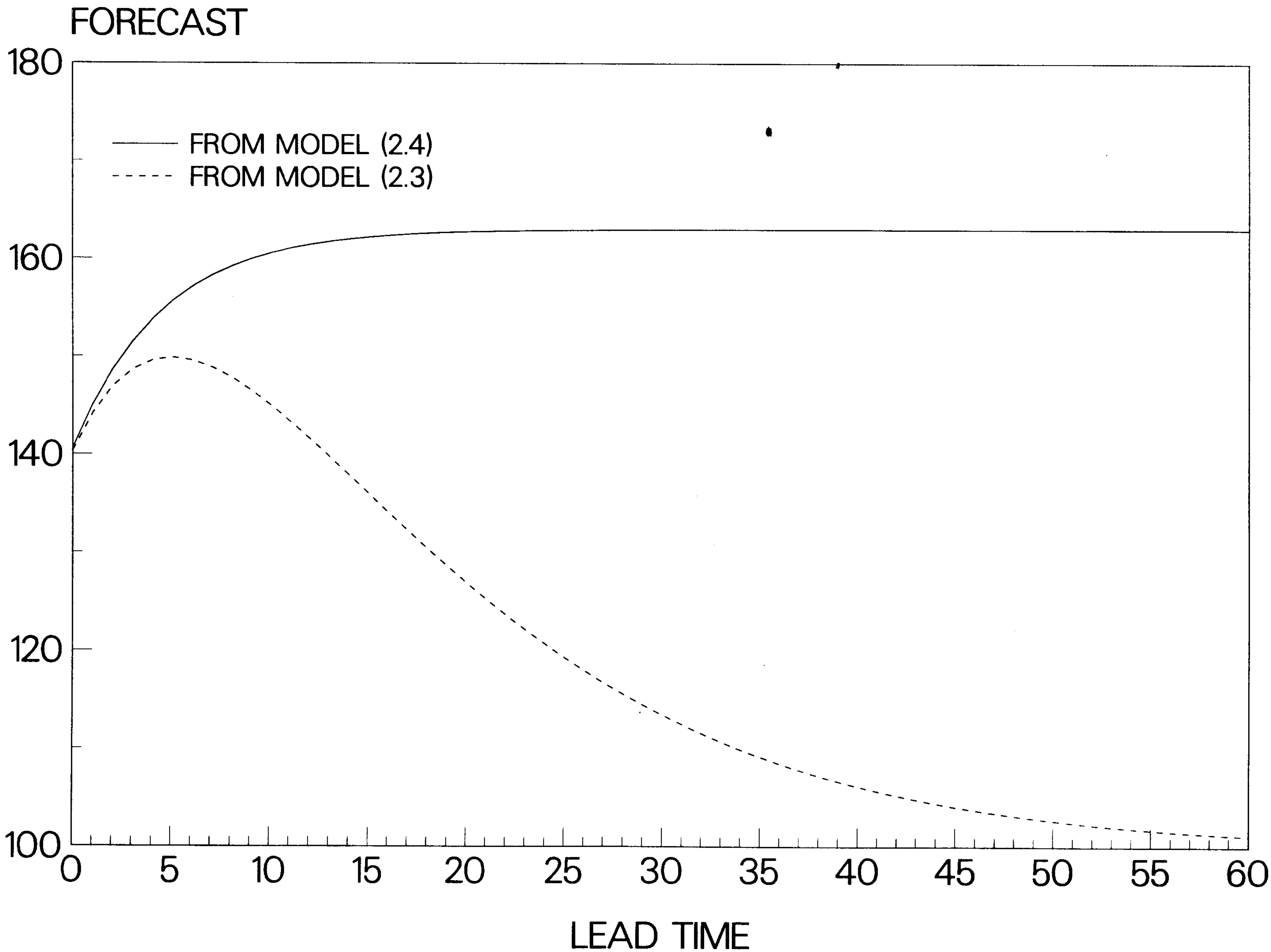
NOTE: Standard errors for the above are all .08 or larger.

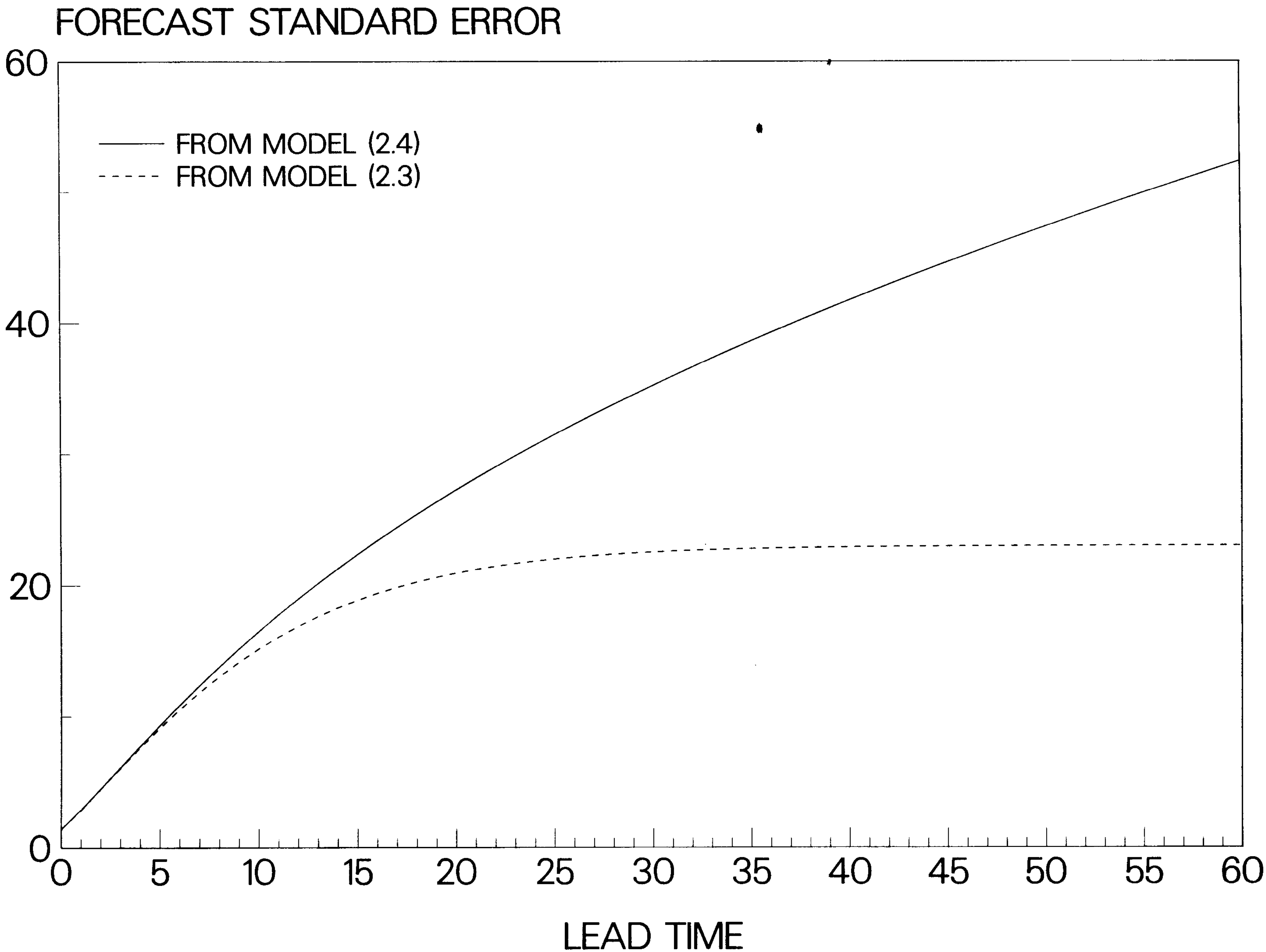

10
5
0
-5
-10
0
25
50
75
100
TIME

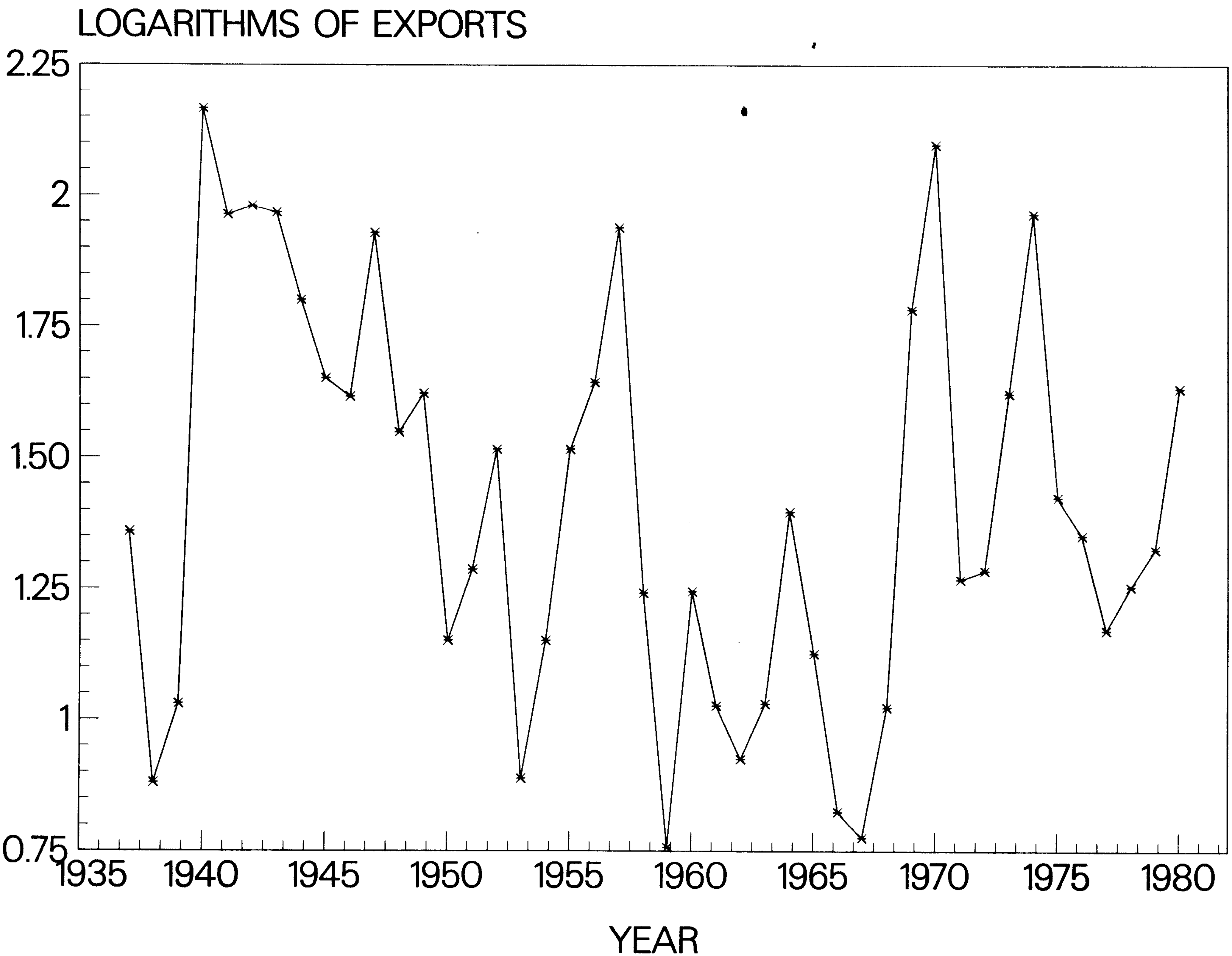


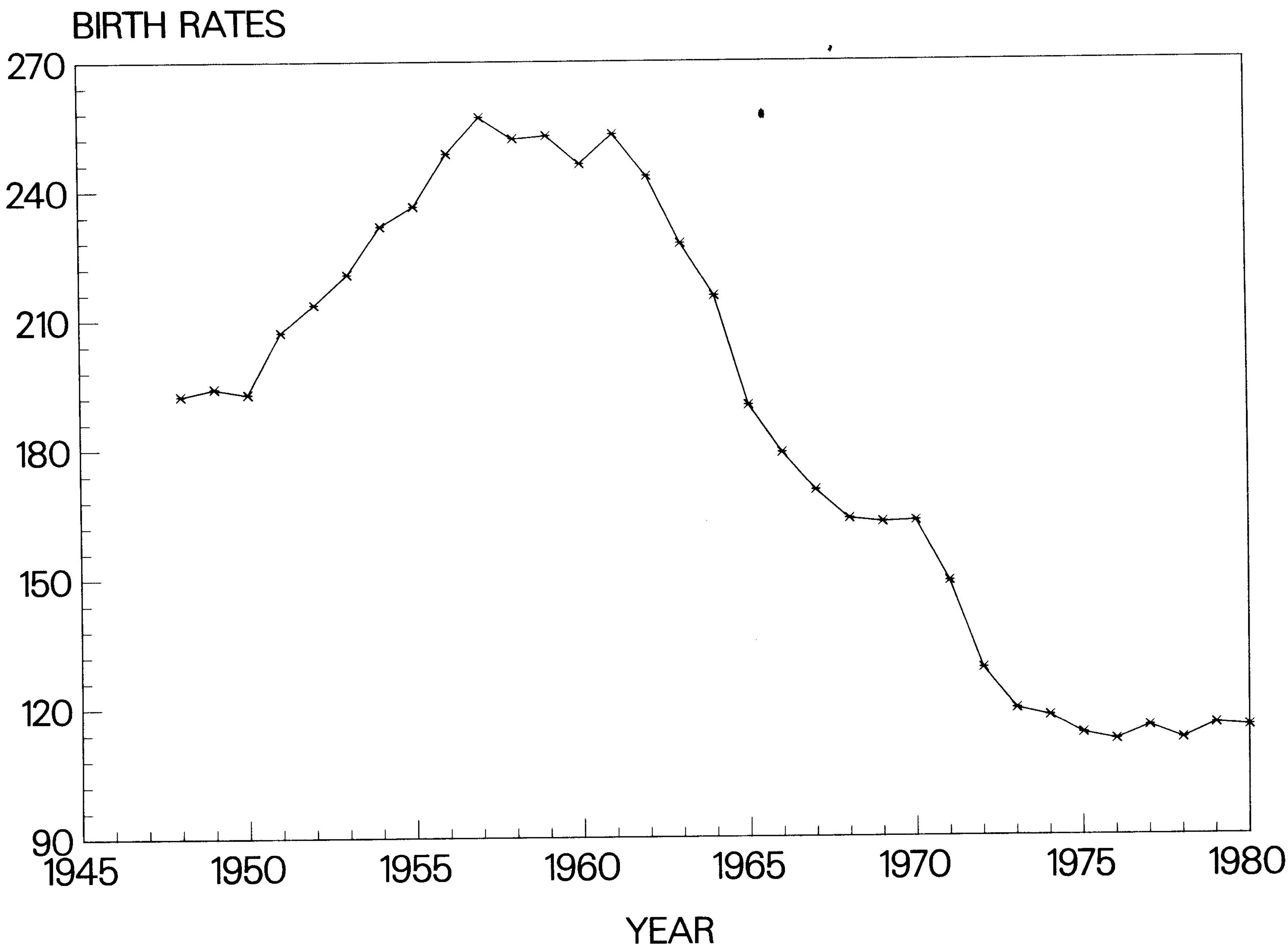
10
5
0
-5
-10
0
25
50
75
100
TIME


50
0
-50
-100
-150
-200
-250
0
10
20
30
40
50


FORECAST
180
160
140
120
100
FROM MODEL (2.4)
FROM MODEL (2.3)
0
5
10
15
20
25
30
35
40
45
50
55
60
LEAD TIME

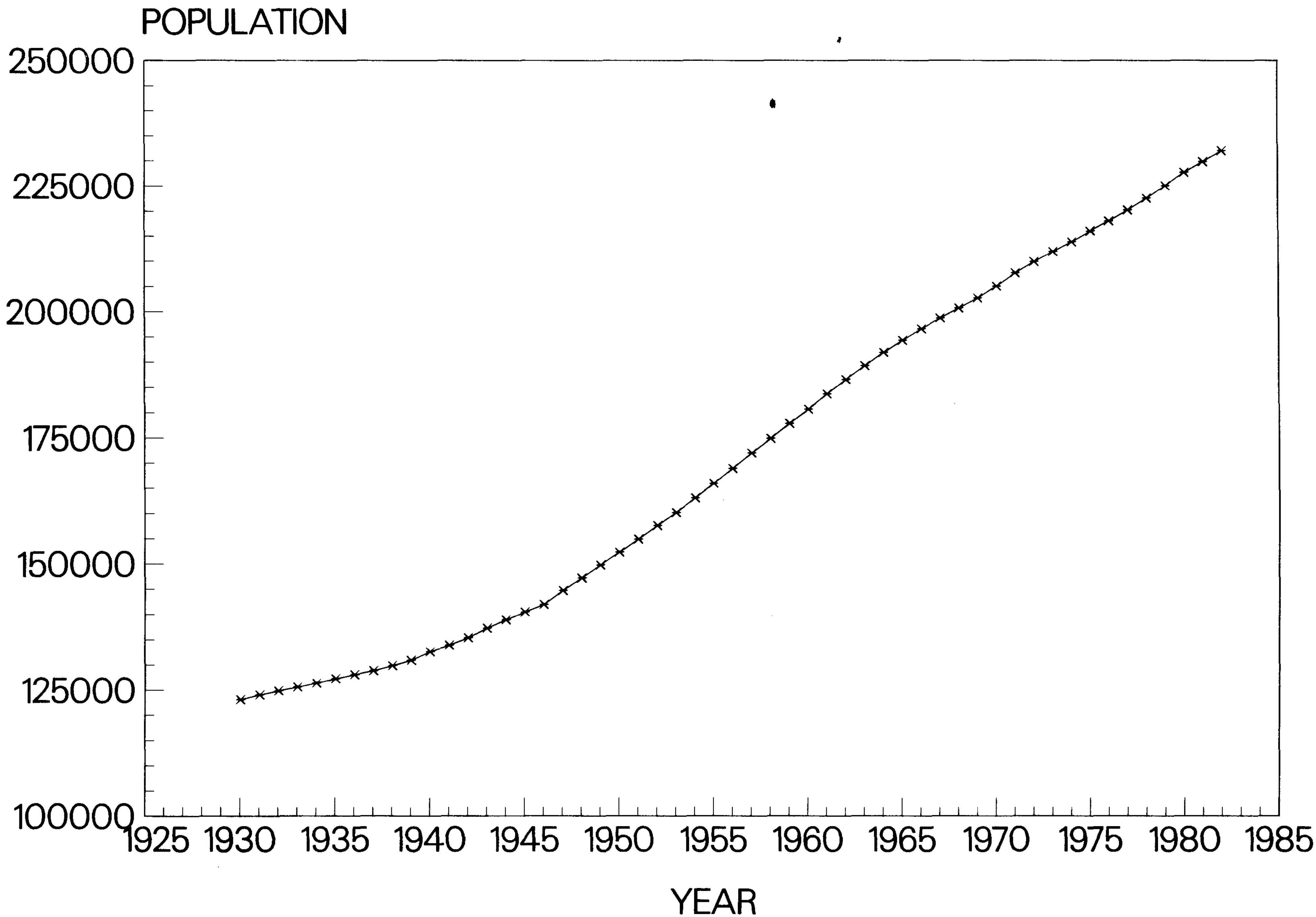

FORECAST STANDARD ERROR
FROM MODEL (2.4)
FROM MODEL (2.3)
0
20
40
60
0
5
10
15
20
25
30
35
40
45
50
55
60
LEAD TIME


LOGARITHMS OF EXPORTS
2.25
2
1.75
1.50
1.25
1
0.75
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
YEAR

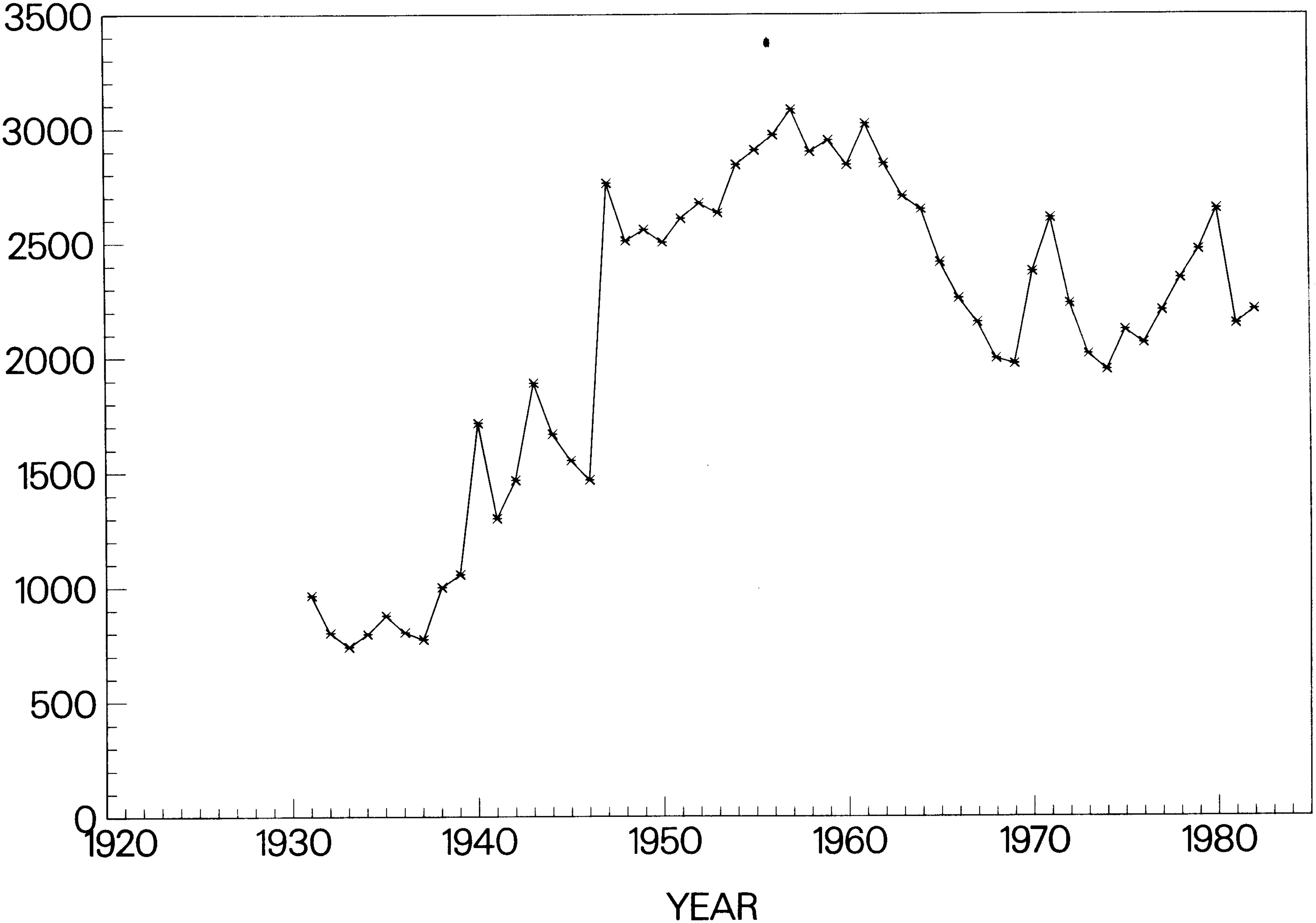


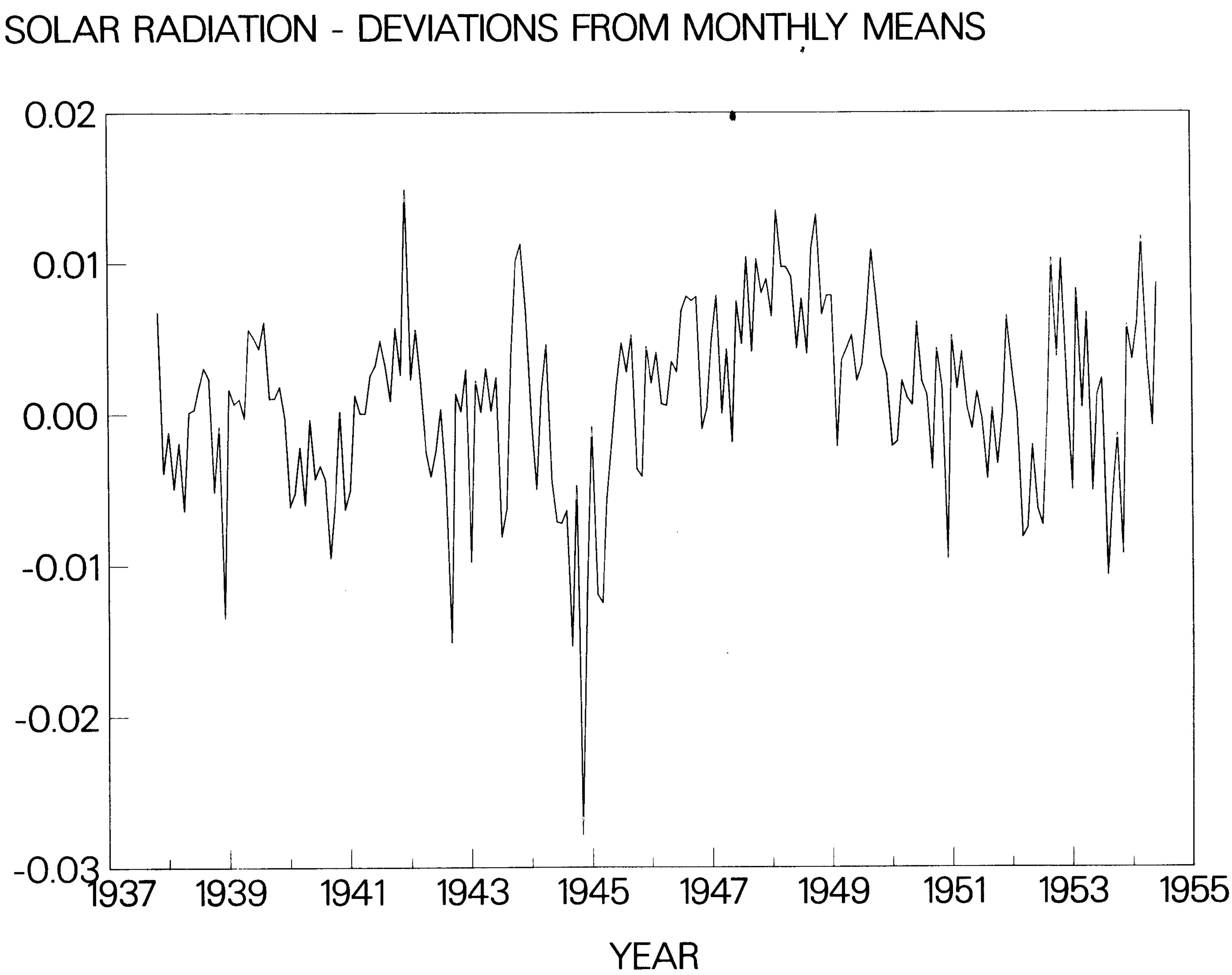
BIRTH RATES
270
240
210
180
150
120
90
1945
1950
1955
1960
1965
1970
1975
1980
YEAR

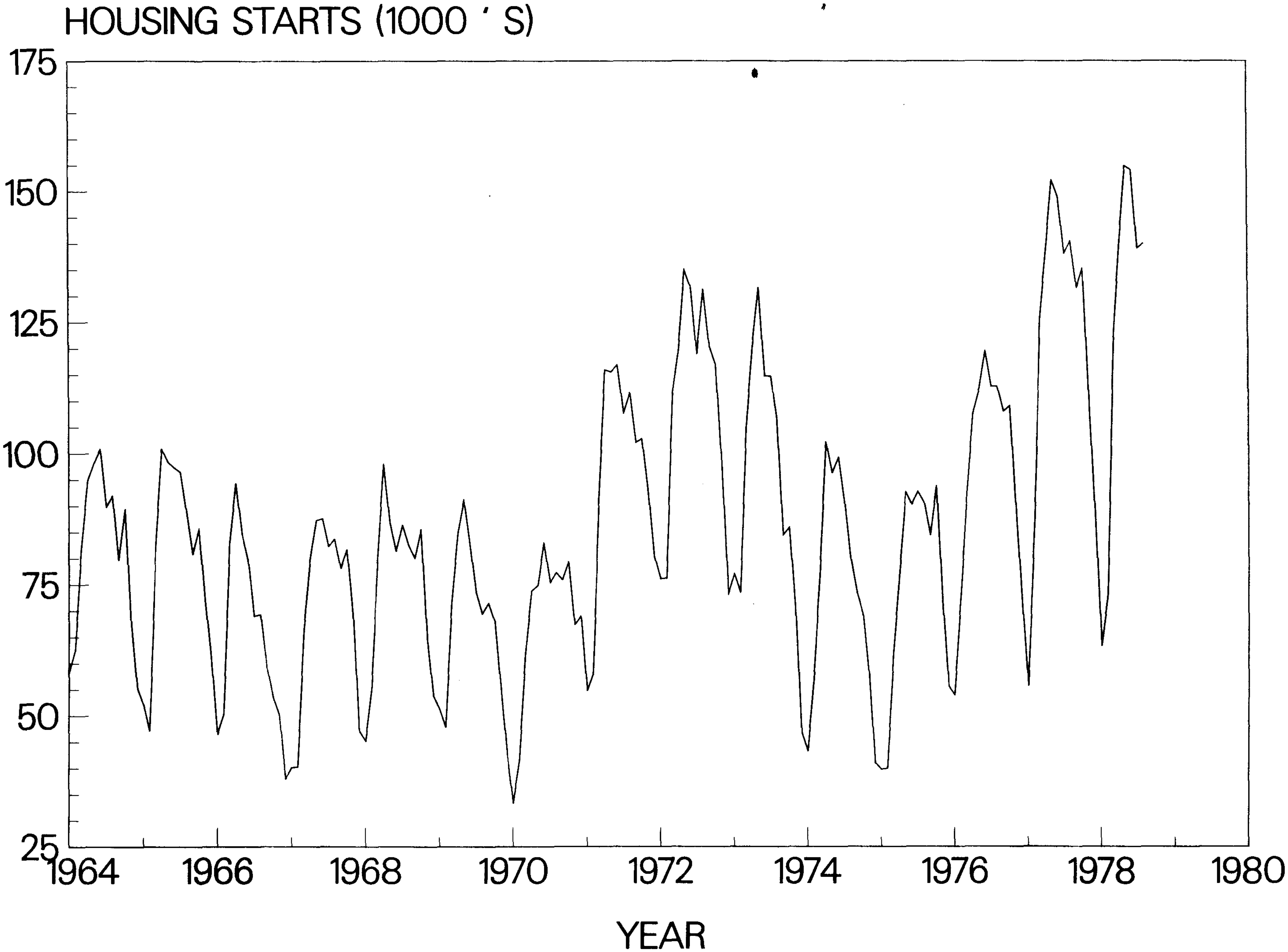
FIRST DIFFERENCES OF BIRTH RATES
30
20
10
0
-10
-20
-30
1945
1950
1955
1960
1965
1970
1975
1980
YEAR

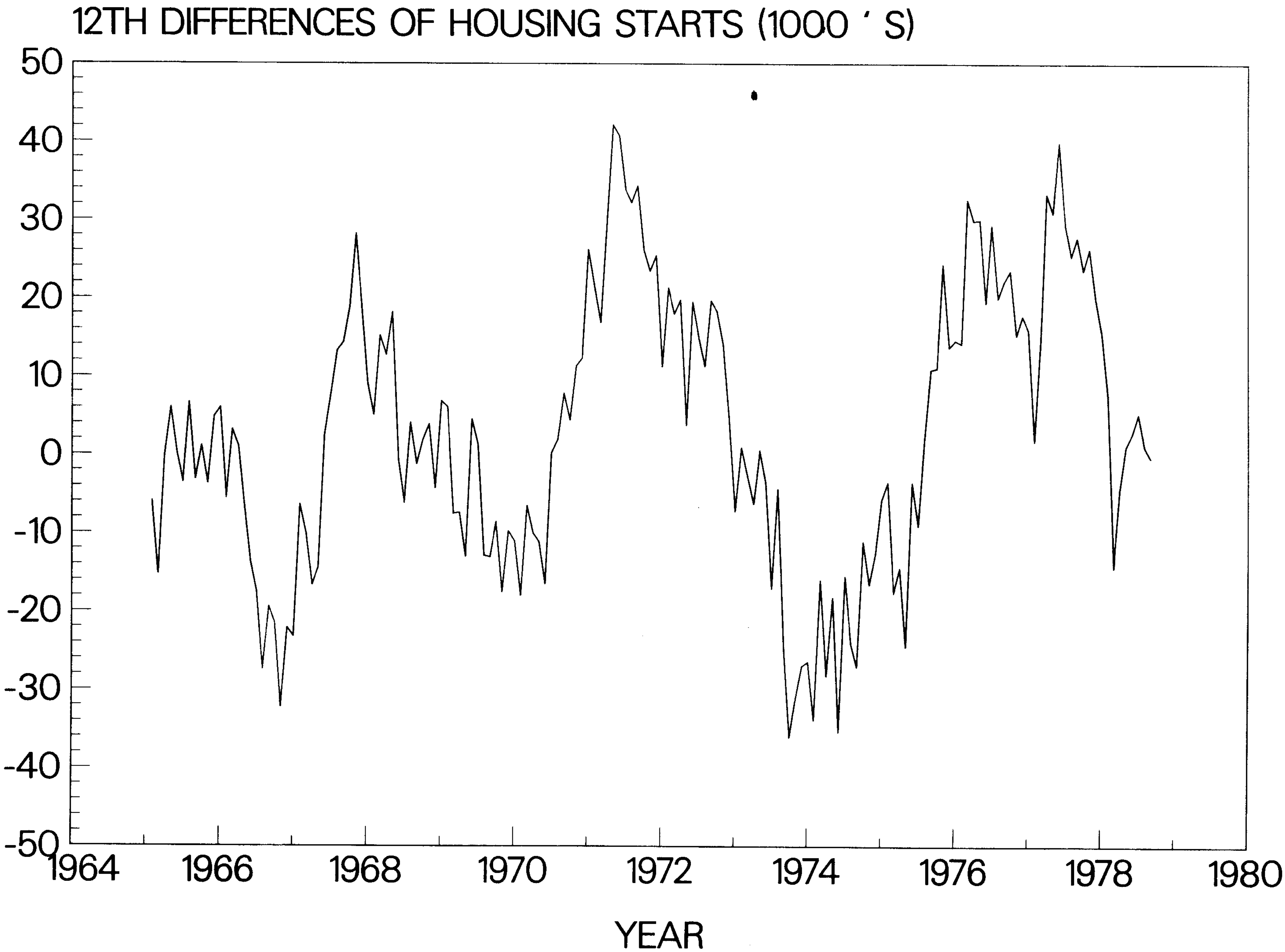

POPULATION
250000
225000
200000
175000
150000
125000
100000
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
YEAR

FIRST DIFFERENCES OF POPULATION




SOLAR RADIATION - DEVIATIONS FROM MONTHLY MEANS
0.02
0.01
0.00
-0.01
-0.02
-0.03
1937
1939
1941
1943
1945
1947
1949
1951
1953
1955
YEAR


HOUSING STARTS (1000 ' S)
175
150
125
100
75
50
25
1964
1966
1968
1970
1972
1974
1976
1978
1980
YEAR


12TH DIFFERENCES OF HOUSING STARTS (1000 ' S)
50
40
30
20
10
0
-10
-20
-30
-40
-50
1964
1966
1968
1970
1972
1974
1976
1978
1980
YEAR