COWLES FOUNDATION FOR RESEARCH IN ECONOMICS AT YALE UNIVERSITY

Box 2125, Yale University New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 975

NOTE: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than acknowledgment that a writer had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

EXACTLY UNBIASED ESTIMATION OF FIRST ORDER AUTOREGRESSIVE/UNIT ROOT MODELS

by

Donald W. K. Andrews

April 1991

EXACTLY UNBIASED ESTIMATION OF FIRST ORDER AUTOREGRESSIVE/UNIT ROOT MODELS

Donald W. K. Andrews¹

Cowles Foundation Yale University

December 1990

Revised: April 1991

ABSTRACT

This paper is concerned with the estimation of first-order autoregressive/unit root models with independent identically distributed normal errors. The models considered include those without an intercept, those with an intercept, and those with an intercept and time trend. The autoregressive (AR) parameter α is allowed to lie in the interval (-1,1], which includes the case of a unit root. Exactly median-unbiased estimators of the AR parameter α are proposed. Exact confidence intervals for this parameter are introduced. Corresponding exactly median-unbiased estimators and exact confidence intervals are also provided for the impulse response function and the cumulative impulse response. An unbiased model selection procedure is discussed. The procedures that are introduced are applied to several data series including real exchange rates, the velocity of money, and industrial production.

JEL Classification Number: 211.

Keywords: Autoregressive process, confidence interval, deterministic time trend, impulse response function, median-unbiased estimation, model selection procedure, stochastic time trend, unbiasedness, unit root model.

1. INTRODUCTION

This paper considers point and interval estimation in a first-order autoregressive/unit root (AR/UR) model that may contain an intercept and time trend and has independent identically distributed normal errors. Exactly median-unbiased estimators and exact confidence intervals are introduced for the AR parameter, the impulse response (IR) function, and the cumulative impulse response (CIR). An unbiased model selection procedure is also introduced.

A basic motivation for this paper is the emphasis placed in the unit root literature on testing for a unit root and the relative neglect of point and interval estimation. A problem that arises with hypothesis tests in the unit root context is that tests have low power in many scenarios of empirical interest. In such cases, the failure to reject the null hypothesis cannot be construed as providing evidence in favor of the null.² One needs to provide additional information. Point and interval estimators are statistics that can be used for doing so.

The problem with utilizing estimators in AR/UR models is that of bias. Standard estimators, such as the least squares (LS) estimator, are significantly downward biased in AR/UR models that contain an intercept or an intercept and time trend, especially when the AR parameter α is large. For example, in the model with intercept and time trend, sample size equal to 60, and normal errors, the LS estimator has downward median-biases when the AR parameter α equals .7, .85, and 1.0 of .08, .09, and .15 respectively. Even when α is as small as .3, the downward bias is .05. These downward biases are considerable, especially when α is large. They cause the LS estimator to be a misleading indicator of the true value of α . For example, the probabilities that the LS estimator underestimates the AR parameter α in the model above when α equals .3, .7, .85, and 1.0 are .66, .78, .87, and .995 respectively. See Table 1 for a list of such probabilities for a more complete grid of α values in [0, 1.0]. (Note that in models with no intercept or time trend, which are of relatively limited practical importance, the downward bias of the LS estimator is very much smaller than it is in models with an intercept or with an intercept and time trend. For models with no intercept, the LS estimator is not a misleading indicator of the true value of α , at least for samples of size sixty or more.)

To deal with the problem of downward bias, this paper introduces an exact bias correction for the LS estimator. The method is as follows: If the LS estimate equals .8, say, one does not use .8 as the estimate of α , but rather, one uses the value of α which yields the LS estimator to have a median of .8. In the model described above this is .9. With this bias correction, the estimator is exactly median-unbiased. Note that the magnitude of the correction can be quite large -- a LS estimate of .85 corresponds to a median-unbiased estimate of 1.0 in the model described above.

A similar procedure can be used to obtain exact confidence intervals for α . These confidence intervals are useful in their own right, as indicators of the variability of the median-unbiased estimator, or as the basis of exact hypothesis tests. We focus attention on *central* confidence intervals which by definition have equal probabilities of over-estimation and under-estimation.

The reason for emphasizing median-unbiased estimators and central confidence intervals in the present paper is because of their impartiality properties. In cases where the magnitude of a parameter is a contentious issue, as it is in (trend) stationary versus unit root debates, it is useful to have statistical procedures available that treat different parameter values on an even footing. Classical hypothesis tests do not do this and neither do proper Bayes estimators.

It is interesting to note that much of the work on Bayesian estimation of AR/UR models concerns estimation using some version of a flat or non-informative prior. For example, see DeJong and Whiteman (1991), Schotman and van Dijk (1991a, b), Sims (1988), and Sims and Uhlig (1991) for the use of a flat prior and Phillips (1991) for the use of a non-informative prior. The motivation for the use of such priors is the desire for a certain degree of impartiality in the estimation or model selection procedure. This is similar to the motivation for using median-unbiased estimators and central confidence intervals. The latter two, however, avoid the difficult question of choosing the most appropriate prior to use. They also avoid the problems that arise due to the question of whether the prior used is really non-informative, e.g., see the discussion of Phillips (1991). On the other hand, Bayes estimators do possess other attributes that are attractive, see the references above for a discussion of these.

Median-unbiased rather than mean-unbiased estimators are considered here for several reasons. First, median-unbiased estimators have the intuitively appealing property that the probability of over-estimation equals the probability of under-estimation. Second, mean-unbiased estimators do not generally exist when the parameter space is bounded. Third, the median-unbiasedness condition is not reliant on the tails of an estimator's distribution as is the mean-unbiasedness condition, which is desirable when estimators' distributions are skewed and kurtotic. Fourth, the median-unbiasedness property facilitates the construction of exact confidence intervals. Fifth, the median-unbiased estimators of one parameter, such as the AR parameter α , can be used to obtain median-unbiased estimators of other parameters, such as the IR function at different time horizons and the CIR, whereas mean-unbiased estimators cannot be so used.

The point and interval estimation procedures considered here have the following attributes. They are exact procedures rather than asymptotic ones. They exhibit a smooth transition between the (trend) stationary case when $|\alpha| < 1$ and the unit root (random walk) case when $\alpha = 1$. They utilize desirable assumptions regarding

initial conditions -- if $|\alpha| < 1$, a stationary initial condition is imposed and if $\alpha = 1$ an arbitrary initial condition is allowed. In addition, the median-unbiased estimator is unbiased whether the AR parameter α is fixed or is random, as in a Bayesian context. In the latter case, it is median-unbiased for all possible distributions of α .

A drawback of the exact procedures considered here is that they only apply to first-order AR processes and not to more general p-th order processes. The extension of the procedures of this paper to AR(p) models is considered in Andrews and Chen (1991), see also Rudebusch (1990) and Stock (1990). For AR(p) models, the procedures are no longer exact; they are only approximate.

A second drawback of the procedures considered here is that they require the specification of the distribution of the innovations in the AR/UR model. If the distribution is specified incorrectly, then the procedures will not be exact. On the other hand, it is shown below that the specification of normally distributed innovations is surprisingly robust against a variety of non-normal distributions. For such distributions, the procedures are approximately valid and the approximation error is small.

A third drawback of the median-unbiased estimator that is considered here is the lack of an explicit optimality property for it--we do not know whether it is a best median-unbiased estimator. It is possible that the estimator does not fully exploit all the information in the sufficient statistics for the parameters. Note that the LS estimator, which is widely used in practice, is also subject to the latter criticism.

There are a number of papers in the literature that are related to the present one. For stationary AR processes, Quenouille (1949, 1956) introduced the jackknife estimator of the AR parameter α that is mean-unbiased to order 1/T as $T \to \infty$. His results do not allow for a time trend or for an AR parameter of one. Marriott and Pope (1954) and Kendall (1954) established the mean-bias of the LS estimator in the same model as considered by Quenouille (1949). Orcutt and Winokur (1969) used the latter results to construct approximate mean-unbiased estimates of the AR parameter in stationary models. Independent of the present paper, Rudebusch (1990) has considered approximately median-unbiased estimators of the AR parameters in an AR(p) model with a time trend and Stock (1990) has introduced asymptotic confidence intervals for the largest root in an AR(p) model with a time trend (using largest root local to unity asymptotics). Rudebusch's median-unbiased estimators are not exact and are not provided with any formal justification of the approximation involved or with a measure of uncertainty to accompany the estimate. They do apply, however, in more general models than those considered here. Lastly, we note that there are numerous papers that consider hypothesis tests in the AR/UR models considered here. For brevity, we do not provide references.

The remainder of this paper is outlined as follows. Section 2 defines the models that are considered and points out a useful invariance property of the LS estimator for these models. Section 3 introduces the median-unbiased estimator of the AR parameter α . Section 4 provides exact confidence intervals for α . Section 5 discusses median-unbiased estimators and exact confidence intervals for the IR function and the CIR. Section 6 introduces an unbiased model selection procedure. Section 7 demonstrates the robustness of the procedures introduced in Sections 2-6 to some forms of non-normality of the innovations in the AR/UR models. Section 8 provides a discussion of some properties of the LS estimator and Bayes estimators. Section 9 describes the application of the results of the paper to a number of exchange rate series (analyzed previously by Schotman and van Dijk (1991a)) and the velocity and industrial production series of Nelson and Plosser (1982). Section 10 describes the numerical methods that were used to construct the tables given in Section 3 that enable one to compute the median-unbiased estimators and exact confidence intervals.

2. DEFINITION OF MODELS

Three models for the time series $\{Y_t: t=0, ..., T\}$ are considered. Each is based on a latent AR(1) time series $\{Y_t^*: t=0, ..., T\}$:

$$Y_t^* = \alpha Y_{t-1}^* + U_t \text{ for } t = 1, ..., T, \text{ where } \alpha \in (-1,1],$$

$$U_t \sim \text{iid } N(0, \sigma^2) \text{ for } \sigma^2 > 0, Y_0^* \sim N(0, \sigma^2/(1-\alpha^2)) \text{ if } \alpha \in (-1,1),$$
and Y_0 is an arbitrary constant or rv if $\alpha = 1$.

As defined, $\{Y_t^n\}$ is a strictly stationary, mean zero, normal, AR(1) process if $\alpha \in (-1, 1)$. If $\alpha = 1$, on the other hand, $\{Y_t^n\}$ is a normal random walk with arbitrary initial condition.

The three models for $\{Y_n\}$ are defined as follows:

Model 1:
$$Y_t = Y_t^*$$
 for $t = 0, ..., T$ and $\alpha \in (-1,1)$,

(2.2) Model 2:
$$Y_t = \mu + Y_t^*$$
 for $t = 0, ..., T$, for some $\mu \in R$ and $\alpha \in (-1, 1]$,

Model 3:
$$Y_t = \mu + \beta t + Y_t^*$$
 for $t = 0$, ..., T , for some μ , $\beta \in R$ and $\alpha \in (-1, 1]$.

Models 1-3 can be written equivalently as:

In (2.3), $\{U_t\}$ is as in (2.1), $Y_0 \sim N(0, \sigma^2/(1 - \alpha^2))$ in Model 1, $Y_0 \sim N(\mu, \sigma^2/(1 - \alpha^2))$ in Models 2 and 3 if $\alpha \in (-1, 1)$, and Y_0 is an arbitrary constant or rv in Models 2 and 3 if $\alpha = 1.3$

Model 1:
$$Y_t = \alpha Y_{t-1} + U_t$$
 for $t = 1, ..., T$ and $\alpha \in (-1,1),$

Model 2: $Y_t = \tilde{\mu} + \alpha Y_{t-1} + U_t$ for $t = 1, ..., T$, where $\tilde{\mu} = \mu(1-\alpha)$ and $\alpha \in (-1,1],$

(2.3)

Model 3: $Y_t = \tilde{\mu} + \tilde{\beta}t + \alpha Y_{t-1} + U_t$ for $t = 1, ..., T$, where $\tilde{\mu} = \mu(1-\alpha) + \alpha\beta,$
 $\tilde{\beta} = \beta(1-\alpha), \text{ and } \alpha \in (-1,1].$

Below we use the term sample size to denote the number of Y_t observations, viz., T+1. We refer to the number T of dependent variable values in Models 1-3 as the sample NDV.

In Model 1, $\{Y_t: t=1, ..., T\}$ is a strictly stationary, normal, AR(1) process with mean zero. In this model, the nonstationary parameter value $\alpha=1$ is excluded. The reason is that the LS estimator of α has distribution that depends on the initial condition Y_0 in this case and in consequence the exact procedures that are introduced below do not apply when $\alpha=1$. Since Model 1 is not of great practical importance -- one rarely assumes the mean is known and equals zero, this is not a serious restriction. Model 1 is included in the discussion for comparative purposes and for completeness, rather than for its applicability in practice.

In Model 2, if $\alpha \in (-1, 1)$, $\{Y_t : t = 1, ..., T\}$ is a strictly stationary, normal, AR(1) process with mean μ . In this model, if $\alpha = 1$, $\{Y_t : t = 1, ..., T\}$ is a normal random walk with arbitrary initial condition. In Model 3, if $\alpha \in (-1, 1)$, $\{Y_t : t = 1, ..., T\}$ is a strictly stationary, normal, AR(1) process around a deterministic trend line with intercept μ and slope β . In this model, if $\alpha = 1$, $\{Y_t : t = 1, ..., T\}$ is a normal random walk with drift β and arbitrary initial condition Y_0 . The results of this paper also apply to models with higher-order time trends. For brevity, we only consider Models 1-3 here. The extension of the results to models with higher-order time trends is straightforward.

Let $\hat{\alpha}_{LSj}$ denote the LS estimator of α corresponding to Model j given in (2.3). That is, $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ are the LS estimators of α from the regression of Y_t on Y_{t-1} , $(1, Y_{t-1})$, and $(1, t, Y_{t-1})$, respectively, for t = 1, ..., T. We note the following: For j = 1, 2, 3,

PROPERTY of $\hat{\alpha}_{LSj}$: The distribution of $\hat{\alpha}_{LSj}$ depends only on α when Model j is correct. In particular, it does not depend on σ^2 in Model 1, (σ^2, μ) in Model 2, or (σ^2, μ, β) in Model 3, nor on the value of Y_0 when $\alpha = 1$ in Models 2 and 3.

This property is exploited in the construction of median-unbiased estimators of α in Section 3 below. The proof of this property is as follows. Consider Model 3. To obtain $\hat{\alpha}_{LS3}$, first regress Y_t on (1, t) and Y_{t-1} on (1, t) for t = 1, ..., T. Then, $\hat{\alpha}_{LS3}$ equals the LS estimate from the regression of the residuals from the former

regression on the residuals from the latter regression. Since $Y_t = \mu + \beta t + Y_t^*$ by (2.2), both sets of residuals are invariant with respect to the values of μ and β . In consequence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to (μ, β) .

Given this invariance, we can suppose $\mu = \beta = 0$ when analyzing the possible dependence of $\hat{\alpha}_{LS3}$ on σ^2 and of $\hat{\alpha}_{LS3}$ on Y_0 when $\alpha = 1$. When $\mu = \beta = 0$, $Y_t = Y_t^s$ for t = 0, ..., T. Multiplying σ^2 by a positive constant c in (2.1) causes Y_t^s and Y_t to be multiplied by the same constant c for t = 0, ..., T when $\alpha \in (-1, 1)$. In consequence, the residuals from the regression of Y_t on (1, t) and of Y_{t-1} on (1, t) are multiplied by the same constant. This constant cancels out when the former residuals are regressed on the latter to obtain $\hat{\alpha}_{LS3}$. Thus, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to σ^2 when $\alpha \in (-1, 1)$.

Next, suppose $\mu = \beta = 0$ and $\alpha = 1$. In this case, $Y_t = Y_t^* = Y_0 + \sum_{s=1}^t U_s$ for t = 0, ..., T and the residuals from the regression of Y_t on (1, t) and of Y_{t-1} on (1, t) are invariant with respect to the value of Y_0 . In consequence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to the value of Y_0 when $\alpha = 1$. Given this invariance, suppose $Y_0 = 0$. Then, the multiplication of σ^2 by a constant c causes Y_t^* and Y_t to be scaled up by the same constant. As above, this leaves $\hat{\alpha}_{LS3}$ unchanged. Hence, the distribution of $\hat{\alpha}_{LS3}$ is invariant with respect to σ^2 when $\alpha = 1$. This completes the proof for Model 3.

The proofs for Models 1 and 2 are analogous, except that for Model 1 the argument for the invariance of the LS estimator with respect to Y_0 when $\mu = \beta = 0$ and $\alpha = 1$ does not hold (because Y_t and Y_{t-1} are regressed on 0 rather than on (1, t) in this case). This does not affect the result stated in (2.4) because $\alpha = 1$ is excluded in Model 1.

3. MEDIAN-UNBIASED ESTIMATION OF THE AUTOREGRESSIVE PARAMETER

This section introduces median-unbiased estimators of α for Models 1-3.

A number m is a median of a rv X if

(3.1)
$$P(X \ge m) \ge 1/2 \text{ and } P(X \le m) \ge 1/2.$$

This definition of a median allows for non-uniqueness, but all of the medians considered here will be unique. The definition also allows for the median of X to be a probability mass point of X. This feature of the definition will be used here. If a median m of X is not a probability mass point, then P(X > m) = P(X < m) = 1/2.

Let $\hat{\alpha}$ be an estimator of the parameter α . By definition, $\hat{\alpha}$ is median-unbiased for α if the true parameter α is a median of $\hat{\alpha}$ for each α in the parameter space. The condition of median-unbiasedness has the intuitive

impartiality property that the probability of under-estimation equals the probability of over-estimation. This holds unless the true parameter value is estimated with positive probability and in this case the probabilities of under-estimation and over-estimation are each less than one half. In scenarios where the magnitude of a parameter is a contentious issue, such as in the (trend) stationary versus unit root debate, this impartiality property is quite useful. Advocates of one view are not likely to except estimates that are biased towards a different view. Median-unbiased estimators are more likely to be exceptable to a broad audience than biased estimators, because they do not favor any particular outcome.

Median-unbiasedness is a special case of the concept of *risk-unbiasedness* when loss is given by absolute error. In particular, $\hat{\alpha}$ is median-unbiased for α , if and only if it has the property that

(3.2)
$$E_{\alpha}|\hat{\alpha} - \alpha| \le E_{\alpha}|\hat{\alpha} - \alpha'| \text{ for all } \alpha, \alpha' \in A,$$

where A denotes the parameter space of α and E_{α} denotes the expectation operator when α is the true parameter value, see Lehmann (1959, p. 22). In words, $\hat{\alpha}$ is median-unbiased if and only if the distance between $\hat{\alpha}$ and the true parameter on average is less than or equal to the distance between $\hat{\alpha}$ and any other parameter value. In this sense, the value that $\hat{\alpha}$ is best at estimating is the true value α regardless of what α is.

The condition of median-unbiasedness is often more useful than that of mean-unbiasedness when the parameter space is bounded or when the distributions of estimators are skewed and/or kurtotic. When the parameter space is bounded and closed, it is impossible to have a mean-unbiased estimator because all estimators are biased at extreme boundary points. Boundary points do not present problems, however, for the condition of median-unbiasedness: If an estimator is median-unbiased for a parameter space $A \subset R$, then the estimator restricted to a closed subset A^* of A is median-unbiased for the restricted parameter space A^* . (The method of restricting the estimator, say $\hat{\alpha}$, to A^* is to set $\hat{\alpha}$ equal to the nearest element of A^* that is larger or smaller than $\hat{\alpha}$.) Next, when estimators have asymmetric distributions, there is no unambiguous measure of the centers of their distributions. In this case, the median may be a preferred measure to the mean, especially in kurtotic cases, because the median is less sensitive to the tails of the distribution.

We note that in the classical normal linear regression model with fixed regressors the LS estimator is median-unbiased. In fact, it is the best median-unbiased estimator for a wide variety of loss functions (see Andrews and Phillips (1987)). In the AR/UR model, on the other hand, the LS estimator is not median-unbiased, and hence, does not possess the same optimality properties.

We now discuss an exact method for median-bias correcting an estimator. (This method is not original to the present paper, e.g., it more or less corresponds to the method discussed by Lehmann (1959, Sec. 3.5, p. 83), but the application of the method below to the LS estimator in Models 1-3 is original.) Suppose $\hat{\alpha}$ is an esti-

mator whose median function $m(\alpha)$ (= $m_T(\alpha)$) is uniquely defined and is strictly increasing on the parameter space A which is a finite interval, say (-1, 1]. Then $\hat{\alpha}_U$ is a median-unbiased estimator of α , where $\hat{\alpha}_U$ is defined by

(3.3)
$$\hat{\alpha}_{U} = \begin{cases} 1 & \text{if } \hat{\alpha} > m(1) \\ m^{-1}(\hat{\alpha}) & \text{if } m(-1) < \hat{\alpha} \leq m(1), \\ -1 & \text{if } \hat{\alpha} \leq m(-1) \end{cases}$$

where $m(-1) = \lim_{\alpha \to -1} m(\alpha)$ and $m^{-1} : (m(-1), m(1)] \to (-1, 1]$ is the inverse function of $m(\cdot)$ that satisfies $m^{-1}(m(\alpha)) = \alpha$ for $\alpha \in (-1, 1]$. (Since the parameter space is (-1, 1) in Model 1 rather than (-1, 1], m(1) needs to be defined as $\lim_{\alpha \to 1} m(\alpha)$ in this case.)

Figure 1 illustrates the relationship between $\hat{\alpha}$ and $\hat{\alpha}_U$. The horizontal axis corresponds to different values of α . The median of $\hat{\alpha}$, $m(\alpha)$, is shown. (Its numerical computation is described in Section 10 below.) Given a value of $\hat{\alpha}$ on the vertical axis, one finds the corresponding value of $\hat{\alpha}_U$ on the horizontal axis using the function $m(\alpha)$. If $\hat{\alpha} > m(1)$ the corresponding value of $\hat{\alpha}_U$ is 1 and if $\hat{\alpha} \le m(-1)$ the corresponding value of $\hat{\alpha}_U$ is -1. In the figure, the estimator $\hat{\alpha}$ (which corresponds to $\hat{\alpha}_{LS3}$ for T+1=60) is biased downward for all values of $\alpha \in (-.7, 1]$, so if $\hat{\alpha} \in (-.7, 1]$ the bias correction increases the value of the estimate.

To show that $\hat{\alpha}_U$ is median-unbiased, we write

$$P_{\alpha}(\hat{\alpha}_{U} \geq \alpha)$$

$$= P_{\alpha}(\hat{\alpha}_{U} \geq \alpha, \, \hat{\alpha} > m(1)) + P_{\alpha}(\hat{\alpha}_{U} \geq \alpha, \, m(-1) < \hat{\alpha} \leq m(1)) + P_{\alpha}(\hat{\alpha}_{U} \geq \alpha, \, \hat{\alpha} \leq m(-1))$$

$$= P_{\alpha}(\hat{\alpha} > m(1)) + P_{\alpha}(m^{-1}(\hat{\alpha}) \geq \alpha, \, m(-1) < \hat{\alpha} \leq m(1))$$

$$= P_{\alpha}(\hat{\alpha} > m(1)) + P_{\alpha}(\hat{\alpha} \geq m(\alpha), \, m(-1) < \hat{\alpha} \leq m(1))$$

$$= P_{\alpha}(\hat{\alpha} \geq m(\alpha))$$

$$\geq 1/2,$$

where P_{α} denotes the underlying probability distribution when α is the true parameter. By an analogous argument, $P_{\alpha}(\hat{\alpha}_U \leq \alpha) \geq 1/2 \quad \forall \alpha \in (-1, 1]$. Hence, $\hat{\alpha}_U$ is median-unbiased for α .

The above method of bias correction can be applied to the estimators $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ defined in Section 2. Each of these estimators has a distribution that depends only on α (and the sample size T+1) and numerical evaluation of their median functions $m(\alpha)$ show the latter to be strictly increasing, as one would expect.⁴ In consequence, we can define three corresponding median-unbiased estimators $\hat{\alpha}_{U1}$, $\hat{\alpha}_{U2}$, and $\hat{\alpha}_{U3}$

using the formula given in (3.3) with $\hat{\alpha}$ replaced by $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ respectively. These estimators can be computed given $\hat{\alpha}_{LSi}$ if the function $m^{-1}(\cdot)$, or equivalently, $m(\cdot)$ is known.

Table 2 provides values of the median function $m(\alpha)$ of $\hat{\alpha}_{LS1}$ for a grid of α values in [0, 1) for illustrative sample sizes of 60 and 100. (Since the median function $m(\alpha)$ is odd in Model 1, i.e., $m(\alpha) = -m(-\alpha)$, only results for $\alpha \in [0, 1)$ are given in Table 2.) Tables 3 and 4 do likewise for a grid of α values in (-1, 1] for the estimators $\hat{\alpha}_{LS2}$ and $\hat{\alpha}_{LS3}$, respectively, but for a variety of different sample sizes. Table 2 is restricted to two sample sizes for brevity, since Model 1 is only of limited applicability in practice. The numerical method used to construct the tables is described in Section 10 below.

To illustrate the use of the Tables, consider the estimator $\hat{\alpha}_{LS3}$ for Model 3 and sample size T+1=60. As shown in Table 4, m(1)=.853, so any value of $\hat{\alpha}_{LS3}$ that is $\geq .853$ corresponds to $\hat{\alpha}_{U3}=1$. Similarly, m(-1)=-.997, so any value of $\hat{\alpha}_{LS3}$ that is $\leq -.997$ corresponds to $\hat{\alpha}_{U3}=-1$. For any value of $\hat{\alpha}_{LS3}$ between -.997 and .853, one finds the entry in the $m(\alpha)$ column (i.e., the .5 quantile column) that equals $\hat{\alpha}_{LS3}$ and the α value that corresponds to this entry is $\hat{\alpha}_{U3}$. That is, $\hat{\alpha}_{U3}$ is chosen such that $m(\hat{\alpha}_{U3})=\hat{\alpha}_{LS3}$. For example, if $\hat{\alpha}_{LS3}=.80$, then $\hat{\alpha}_{U3}=.90$. Since the grid of α values given in the tables is finite, interpolation between α values is often needed. Similarly, for sample sizes not given in the tables, interpolation is required. For small sample sizes this interpolation may introduce a non-negligible error. In this case, an alternative is to simulate the function $m(\alpha)$ for a particular sample size of interest. This is cheap to do in terms of computer time, but requires some programming effort, see Section 10.

4. EXACT CONFIDENCE INTERVALS FOR THE AUTOREGRESSIVE PARAMETER

This section derives exact confidence intervals for the parameter α in Models 1-3 of Section 2. These confidence intervals can be used by themselves, in conjunction with the median-unbiased estimator $\hat{\alpha}_{Uj}$ of α (to provide a measure of accuracy of $\hat{\alpha}_{Uj}$), or to construct exact one- or two-sided tests of $H_0: \alpha = \alpha_0$ for arbitrary $\alpha_0 \in (-1, 1]$.

For Model j we construct the confidence interval using $\hat{\alpha}_{LSj}$ for j=1,2,3. As shown in Section 2, the distribution of $\hat{\alpha}_{LSj}$ depends only on the parameter α for which we wish to construct the confidence interval. In addition, its distribution is absolutely continuous and has support R for all α in (-1, 1] ((-1, 1) for Model 1). Let $q_p(\alpha)$ denote the p-th quantile function of $\hat{\alpha}_{LSj}$. That is, for fixed $p \in (0, 1)$, $q_p(\alpha)$ gives the p-th quantile of $\hat{\alpha}_{LSj}$ as a function of the true parameter α . By definition,

$$(4.1) P_{\alpha}(\hat{\alpha}_{LSj} \leq q_{p}(\alpha)) = p.$$

A 100(1-p)% confidence interval (set) for α in Model j is given by the realization of the set

$$\left\{\alpha \in [-1,1]: q_{p_1}(\alpha) \leq \hat{\alpha}_{LSj} \leq q_{p_2}(\alpha)\right\},$$

where $p_1 > 0$, $p_2 > 0$, and $p_1 + p_2 = p$. This set has the correct coverage probability because for all $\alpha_0 \in (-1, 1]$

$$P_{\alpha_0} \Big| \alpha_0 \in \Big\{ \alpha \in [-1, 1] : q_{p_1}(\alpha) \le \hat{\alpha}_{LSj} \le q_{p_2}(\alpha) \Big\} \Big\}$$

$$= P_{\alpha_0} \Big| q_{p_1}(\alpha_0) \le \hat{\alpha}_{LSj} \le q_{p_2}(\alpha_0) \Big\} = 1 - p.$$

(Note that the method used here of constructing confidence intervals is time honored, only the application of it in this context is original.)

If the quantile functions $q_{p_1}(\alpha)$ and $q_{p_2}(\alpha)$ are strictly increasing in α (as they are in almost all the cases tabulated in Tables 3 and 4),⁵ then the set in (4.2) equals the interval $\{\alpha: \hat{c}_L \leq \alpha \leq \hat{c}_U\}$, where

$$\hat{c}_{L} = \begin{cases} > 1 & \text{if } \hat{\alpha}_{LSj} > q_{p_{2}}(1) \\ q_{p_{2}}^{-1}(\hat{\alpha}_{LSj}) & \text{if } q_{p_{2}}(-1) < \hat{\alpha}_{LSj} \leq q_{p_{2}}(1) \text{ and} \\ -1 & \text{if } \hat{\alpha}_{LSj} \leq q_{p_{2}}(-1) \end{cases}$$

$$\hat{c}_{U} = \begin{cases} 1 & \text{if } \hat{\alpha}_{LSj} > q_{p_{1}}(1) \\ q_{p_{1}}^{-1}(\hat{\alpha}_{LSj}) & \text{if } q_{p_{1}}(-1) < \hat{\alpha}_{LSj} \leq q_{p_{1}}(1) \end{cases}.$$

$$-1 & \text{if } \hat{\alpha}_{LS} \leq q_{p_{1}}(-1) \end{cases}$$

In (4.4), for i = 1, 2, $q_{p_i}(-1) = \lim_{\alpha \to -1} q_{p_i}(\alpha)$ and $q_{p_i}^{-1} : (q_{p_i}(-1), q_{p_i}(1)] \to (-1, 1]$ is the inverse function of $q_{p_i}(\cdot)$ that satisfies $q_{p_i}^{-1}(q_{p_i}(\alpha)) = \alpha$ for $\alpha \in (-1, 1]$. (For Model 1, $q_{p_i}(1) = \lim_{\alpha \to 1} q_{p_i}(\alpha)$ and the domain of $q_{p_i}^{-1}$ is $(q_{p_i}(-1), q_{p_i}(1))$.)

Figure 2 illustrates how the confidence intervals $[\hat{c}_L, \hat{c}_U]$ are formed. The horizontal axis corresponds to different values of α . The $p_1 = .05$ and $p_2 = .95$ quantile functions of $\hat{\alpha}_{LS3}$ are graphed for a sample size of 60. For a given value of $\hat{\alpha}_{LS3}$ on the vertical axis, the confidence interval contains all points in [-1, 1] that lie between the two quantile functions. Two examples of observed $\hat{\alpha}_{LS3}$ values and their corresponding confidence intervals are given.

Tables 2-4 provide values of the quantile functions $q_{.05}(\alpha)$ and $q_{.95}(\alpha)$ of $\hat{\alpha}_{LS1}$, $\hat{\alpha}_{LS2}$, and $\hat{\alpha}_{LS3}$ for different sample sizes. These values can be used to construct two-sided 90% confidence intervals and one-sided 95% confidence intervals for α . (For a lower bound one-sided confidence interval, one takes $p_1 = 0$ and the interval is of the form $[\hat{c}_L, 1]$. Analogously, for an upper bound one-sided confidence interval, one takes $p_2 = 1$ and the interval is of the form $(-1, \hat{c}_U)$.)

The interval endpoints \hat{c}_L and \hat{c}_U are obtained from Tables 2-4 in exactly the same way as is the median-unbiased estimator $\hat{\alpha}_{LSj}$. For example, consider a 90% two-sided confidence interval for α in Model 3 with sample size 60. As shown in Table 4, $q_{.95}(1) = .956$, so any value of $\hat{\alpha}_{LS3}$ that is $\geq .956$ corresponds to $\hat{c}_L > 1$. Similarly, $q_{.95}(-1) = -.945$, so any value of $\hat{\alpha}_{LS3}$ that is $\leq -.945$ corresponds to $\hat{c}_L = -1$. For any value of $\hat{\alpha}_{LS3}$ between -.945 and .956, one finds the entry in the $q_{.95}(\alpha)$ column (i.e., the .95 quantile column), that equals $\hat{\alpha}_{LS3}$ and the α value that corresponds to this entry is \hat{c}_L . For instance, if $\hat{\alpha}_{LS3} = .80$, then $\hat{c}_L = .74$ (utilizing an interpolation). One obtains \hat{c}_U in exactly the same manner using the column of $q_{.05}(\alpha)$ values (i.e., the .05 quantile column) rather than the column of $q_{.95}(\alpha)$ values. For example, if $\hat{\alpha}_{LS3} = .80$, then $\hat{c}_U = 1.0$, since $\hat{\alpha}_{LS3} \geq q_{.05}(1)$. Hence, if $\hat{\alpha}_{LS3} = .80$, the median-unbiased estimator $\hat{\alpha}_{U3}$ equals .90 and the 90% confidence interval $[\hat{c}_L, \hat{c}_U]$ equals [.74, 1.0] when the sample size is 60.

5. ESTIMATION OF THE IMPULSE RESPONSE FUNCTION

The impulse response function of a time series $\{Y_t: t=1, 2, ...\}$ measures the effect of a unit shock to Y_t occurring at time t (i.e., $U_t \rightarrow U_t + 1$ in Models 1-3) on the value of Y_t at the future time periods t+1, t+2, This function is of interest, because it quantifies the persistence of shocks to the time series and the latter is often of substantive importance. For example, much of the interest in the trend stationary/unit root debate for macroeconomic time series centers on the question of the degree of persistence of shocks to these series.

Given the definition of Models 1-3 in equations (2.1) and (2.2), it is clear that the effect of a unit change in U_t on future Y_t values is independent of μ and β . Hence, the impulse response functions of $\{Y_t\}$ in Models 1-3 are all equal and are equal to that of $\{Y_t^*\}$. The impulse response function of $\{Y_t^*\}$ is obtained quite simply by writing

(5.1)
$$Y_{t+h}^* = \alpha^{h+1} Y_{t-1}^* + \sum_{i=1}^h \alpha^{h-j} U_{t+j} + \alpha^h U_t \text{ for } h = 0, 1, \dots.$$

It is evident from (5.1) that the effect of a unit shock in U_t on Y_{t+h}^{\bullet} equals α^h . In consequence, the impulse response function for each of the Models 1-3 of Section 2 is given by

(5.2)
$$IR(h) = \alpha^{h} \text{ for } h = 0, 1, 2,$$

A useful scalar measure of persistence that summarizes the impulse response function is the *cumulative* impulse response (CIR). The CIR gives the total cumulative effect of a unit shock on the entire future of the time series. For Models 1-3, we have

(5.3)
$$CIR = \sum_{h=0}^{\infty} IR(h) = \frac{1}{1-\alpha} .$$

The impulse response function at horizon h and the CIR for Models 1-3 are each strictly monotone functions of α . In consequence, one can obtain median-unbiased estimators of each by plugging in the median-unbiased estimator $\hat{\alpha}_{Uj}$ in place of α in their definitions:

(5.4)
$$\hat{IR}(h) = (\hat{\alpha}_{Uj})^h \text{ for } h = 0, 1, 2, ... \text{ and } \hat{CIR} = \frac{1}{1 - \hat{\alpha}_{Uj}}.$$

One can obtain exact confidence intervals for IR(h) and CIR in a similar manner. For IR(h), the 100(1-p)% confidence interval is

(5.5)
$$[(\hat{c}_L)^h, (\hat{c}_U)^h] \text{ for } h = 1, 2, \dots.$$

where \hat{c}_L and \hat{c}_U are defined in (4.4). For CIR, the confidence interval is

$$[1/(1-\hat{c}_U), 1/(1-\hat{c}_L)].$$

In sum, it is straightforward to extend the median-unbiased estimation and exact confidence interval results for α given in Sections 3 and 4 to the impulse response function and the CIR.

6. MODEL SELECTION

The median-unbiased estimators introduced in Section 3 can be used to construct unbiased model selection procedures. By definition, a model selection procedure is unbiased if for any correct model the probability of selecting the correct model is at least as large as the probability of selecting each incorrect model. For example, in Model 2 or 3, one might want to select between the (trend) stationary model for which $\alpha \in (-1, 1)$ and the unit root (with drift) model for which $\alpha = 1$. An unbiased selection procedure in this case has the property that if $\alpha = 1$ the probability of selecting the unit root model is \geq the probability of selecting the (trend) stationary model and if $\alpha \in (-1, 1)$ the P_{α} -probability of selecting the (trend) stationary model is \geq the P_{α} -probability of selecting the unit root model for each $\alpha \in (-1, 1)$. Unbiased selection procedures exhibit an intuitive impartiality property that may be useful if the selection of one model or another is a contentious issue.

The concept of unbiased selection procedures is a special case of that of risk-unbiased decision rules, see Lehmann (1959, p. 12). For selection procedures, the space of actions is finite -- one chooses one model from a finite set of models. If the loss function equals zero when the correct model is chosen and one otherwise, then a risk-unbiased decision rule for this problem is an unbiased selection procedure.

Given Model j of Section 2 for j=1, 2, or 3, suppose the problem is to select one of two models defined by $\alpha \in I_a$ and $\alpha \in I_b$, where I_a and I_b are intervals that partition the parameter space (-1, 1] for α ((-1, 1) in Model 1). For example, one might have $I_a = (-1, 1)$ and $I_b = \{1\}$ or $I_a = (-1, .975)$ and $I_b = [.975, 1]$. (The latter are considered in DeJong and Whiteman (1991) and Phillips (1991).)

The selection procedure we consider here is

(6.1) "choose
$$I_m$$
 if $\hat{\alpha}_{Ui} \in I_m$ for $m = a, b,$ "

where $\hat{\alpha}_{Uj}$ is the median-unbiased estimator defined in Section 3.

This procedure is unbiased. To see this, suppose I_a lies to the left of I_b , then for all $\alpha \in I_a$

$$(6.2) P_{\alpha}(\hat{\alpha}_{Uj} \in I_b) \le P_{\alpha}(\hat{\alpha}_{Uj} > \alpha) \le \frac{1}{2} \le P_{\alpha}(\hat{\alpha}_{Uj} \le \alpha) \le P_{\alpha}(\hat{\alpha}_{Uj} \in I_a),$$

where the second and third inequalities use the median-unbiasedness of $\hat{\alpha}_{Uj}$. For $\alpha \in I_b$, the argument is analogous, so the selection procedure of (6.1) is unbiased. We note that the selection procedure of (6.1) is also a valid level .5 (unbiased) test of $H_0: \alpha \in I_a$ versus $H_1: \alpha \in I_b$ and of $H_0: \alpha \in I_b$ versus $H_1: \alpha \in I_a$.

7. ROBUSTNESS TO NON-NORMAL INNOVATIONS

In one respect, the methods introduced in Sections 2-6 do not rely on the assumption of normal innovations $\{U_t: t=1, ..., T\}$. Any other scale family of distributions could be used in place of the normal family (provided new tables of quantiles of $\hat{\alpha}_{LSj}$ are constructed). On the other hand, the methods of Sections 2-6 do rely on the specification of *some* scale family of distributions of the innovations and it is of interest to know whether the methods are sensitive to the specification. In consequence, we investigate here the sensitivity of the methods of Sections 2-6 to the assumption of normal innovations.

We consider the question: Suppose one assumes the innovations are normal but they actually have some other distribution, such as a t-distribution, then how close is the estimate $\hat{\alpha}_{Uj}$ to being median-unbiased and how close is the coverage probability of the confidence interval $[\hat{c}_L, \hat{c}_U]$ to the desired probability of $1-\alpha$? These questions are most easily answered by constructing analogues of Tables 3 and 4 that correspond to some non-normal distributions of interest and then assessing the difference between Tables 3 and 4 and the new tables.

This has been done, see Table 5, for the case of t-distributions with d degrees of freedom for d = 1, 2, 3, 4, and 10 and shifted chi-square distributions with 4 and 8 degrees of freedom (i.e., chi-square distributions shifted left by 4 and 8 units, respectively, so that they have mean zero).

Table 5 presents the .5, .05, and .95 quantiles of $\hat{\alpha}_{LS3}$ for the different innovation distributions considered. A corresponding table for $\hat{\alpha}_{LS2}$ was constructed, but is not reported here due to its similarity to Table 5 (in terms of the magnitudes of the effects of non-normality on the quantiles). For brevity, Table 5 only reports results for sample size sixty and positive values of α . Unlike Tables 3 and 4, Table 5 has been constructed using simulation. For each entry in the Table, ten thousand repetitions were used. For each repetition, an approximately stationary initial random variable Y_0 was obtained by setting $Y_{-201} = 0$ and generating $\{Y_t : t = -200, ..., 59\}$ according to Model 3 with $\mu = \beta = 0$ (without loss in generality) and with $\{U_t : t = -200, ..., 59\}$ being iid with the desired non-normal distribution. The final sixty observations $\{Y_t : t = 0, ..., 59\}$ were used in simulating the quantiles of $\hat{\alpha}_{LS3}$.

Table 5 shows that the quantiles $\hat{\alpha}_{LS3}$ are not very sensitive to the underlying innovation distribution with the exception of the very heavy tailed t_1 (i.e., Cauchy) distribution. For the .5 quantile and all values of α , the maximum difference between the quantiles for the normal distribution and those for any other distribution except t_2 or t_1 is .005. For many values of α and many of the distributions the difference is less than .005. Furthermore, the standard errors on the simulated estimates of the .5 quantiles are approximately .002, so a maximum simulated difference of .005 corresponds to a noticeably smaller exact maximum difference. For the .5 quantile and the t_2 distribution, the maximum difference is .006. For the .5 quantile and the t_1 distribution, the maximum difference is .025, which is much larger than for any other distribution.

For the .05 and .95 quantiles and all values of α , the maximum difference between the quantiles of $\hat{\alpha}_{LS3}$ for the normal distribution and those for any of the other distributions except $\chi_4^2 - 4$, t_2 , and t_1 is .011, which is fairly small. For these quantiles and the $\chi_4^2 - 4$, t_2 , and t_1 distributions, the maximum differences are .015, .020, and .064 respectively. (The standard error of the simulated estimates of the .05 and .95 quantiles is approximately .001.) As above, the t_1 distribution yields by far the largest differences from the normal distribution. Also, it is clear that the effect of non-normality on the the .05 and .95 quantiles is greater than its effect on the .5 quantiles. Thus, the median-unbiased estimates of α are somewhat more robust to non-normality than are the confidence intervals.

In conclusion, the median-unbiased estimates of α (and monotone functions of α) and the corresponding exact confidence intervals for α are quite robust against substantial skewness and kurtosis in the underlying innovation distribution. Only when the innovations have very thick tails (e.g., Cauchy distributions) are the

estimates and confidence intervals in need of significant adjustment. The necessary adjustment in this case is not surprising, because the asymptotics that hold for normal innovations differ from those for Cauchy innovations. More surprising is the degree of robustness of the estimators and confidence intervals against finite variance symmetric and asymmetric innovation distributions.

Last, we note that there is an alternative available to using simulation methods to assess the effects of non-normality on the procedures suggested in this paper. Specifically, one could compute numerically the quantiles of $\hat{\alpha}_{LS3}$ for Edgeworth-type distributions using exact formulae that are available in the literature, see Subrahmaniam (1966, 1968), Davis (1976), and Knight (1985a, 1985b).

8. PROPERTIES OF THE LEAST SQUARES ESTIMATOR

In this section, we discuss various properties of the LS estimator of α . We show that some of these properties are quite sensitive to the inclusion or exclusion in the regression model of an intercept or an intercept and time trend. This sensitivity has been noted elsewhere in the literature, e.g., see Orcutt and Winokur (1969). Orcutt and Winokur (1969), Phillips (1977), and Evans and Savin (1981, 1984) provide further analysis of the finite sample properties of the LS estimator of Models 1 and 2.

We note first that the LS estimator of α is the maximum likelihood estimator if one conditions on the initial observation. It also is the Bayes estimator of α for an (improper) flat prior if one conditions on the initial observation. Use of the latter estimator (in models with or without intercepts and time trends) has been advocated by Sims (1988, p. 467) and Sims and Uhlig (1991). Some of the arguments given by Sims (1988) for use of the flat-prior Bayes estimator, however, are based on the first-order AR model with no intercept or time trend and do not carry over to the more practically relevant models that contain an intercept or an intercept and time trend.

Table 6 shows the downward bias of the LS estimator of α in Models 1-3 when the sample size T+1 equals 60. The table illustrates a striking difference between the bias for Model 1 with no intercept and the bias for Models 2 and 3. For Model 1 the bias is very small, being roughly -.01 for most values of α . Only if α is very close to one is a bias of this magnitude of import. Thus, in many cases, it does not matter appreciably if one uses the LS estimator or the median-unbiased estimator when Model 1 is employed. On the other hand, the downward bias for Models 2 and 3 is quite large, especially for α large and especially for Model 3. In these Models, there is an appreciable difference between the LS estimator and the median-unbiased estimator.

Table 7 shows the variability of the LS estimator of α as measured by its 90% range for Models 1-3 when T+1=60. (The 90% range of an estimator is the length of the interval bounded by the estimator's .05 and .95 quantiles.) For Model 1, the variability of the LS estimator is a decreasing function of α and it decreases dramatically as α approaches one. Sims (1988, p. 469) has pointed out this property of the LS estimator and argues that it counter-balances the downward bias of the LS estimator with the net result being that one is as likely to get a spuriously high LS estimate of α when α is less than 1 as one is to get a spuriously low LS estimate of α when α is close to one. Indeed, for Model 1 with T+1=60 we find that when $\alpha=.999$ the .05 quantile of the LS estimator is .94 and when $\alpha=.94$ the .95 quantile of the LS estimator is .98, so there is a near balance between spuriously high and spuriously low estimates. When $\alpha=.99$, the .05 quantile of the LS estimator is .89 and when $\alpha=.89$ the .95 quantile of the LS estimator is only .95 (rather than .99), so there is somewhat less of a balance between spuriously high and spuriously low LS estimates for these values of α .

For Models 2 and 3, the variability of the LS estimator is also a decreasing function of α . It is not, however, nearly as sensitive to the value of α as in Model 1. For example, the 90% range for $\alpha = .8$ is only 1.14 times as large as for $\alpha = .999$ in Model 3 and only 1.36 times as large in Model 2, whereas it is 3.9 times as large in Model 1. This relative insensitivity to α of the variability of the LS estimator in Models 2 and 3 coupled with its large downward biases in these models, causes the LS estimator to yield spuriously low estimates of α when α is near 1 with much greater frequency than spuriously high estimates of α when α is less than 1. For example, when $\alpha = 1$, the .05 quantile of $\hat{\alpha}_{LS3}$ is .67 for T+1=60, whereas when $\alpha = .67$, the .95 quantile of $\hat{\alpha}_{LS3}$ is only .75 (rather than 1). Similarly, when $\alpha = .9$, the .05 quantile of $\hat{\alpha}_{LS3}$ is .61, whereas when $\alpha = .61$ the .95 quantile is only .71 (rather than .9).

In fact, the probability of spuriously high LS estimates in Model 3 is very low. The .95 quantile of $\hat{\alpha}_{LS3}$ for $\alpha = .6$ is .7 when T+1=60, for $\alpha = .8$ it is only .85, for $\alpha = .9$ it is .91, for $\alpha = .95$ it is .94, and for $\alpha = 1.0$ it is .96. Thus, in this model there is a very low probability of getting a LS estimate close to or greater than 1 no matter what the value of α is within (-1, 1]. In contrast, the probability of spuriously low LS estimates in Model 3 is quite high for all values of α . For example, for $\alpha = .6$ the .05 quantile of $\hat{\alpha}_{LS3}$ is .32 when T+1=60, for $\alpha = .8$ it is .52, for $\alpha = .9$ it is .61, for $\alpha = .95$ it is .65, and for $\alpha = 1.0$ its .67.

In sum, for each value of α over a wide range of values there is a high probability of spuriously low LS estimates and a low probability of spuriously high LS estimates in Model 3. Since this is true conditional on α for each value of α , it is also true unconditionally if α is viewed as being random, as in a Bayesian context. Hence, the LS estimator of α in Model 3 is simply too small and this is true whether or not one views α as being random. Similar, but somewhat less dramatic, conclusions apply to the LS estimator in Model 2.

Next, we discuss an argument given by Sims (1988, p. 469) regarding the unbiasedness of Bayes estimators. The argument given is that the mean of a Bayes estimator equals the mean of the random parameter being estimated. This might be referred to as unconditional (on the parameter) unbiasedness of the Bayes estimator. The above argument is subject to a number of caveats, however, and at best illustrates only a very weak sort of unbiasedness that does not reflect impartiality of the estimator in our opinion.

First, the argument does not apply to Bayes estimators based on improper priors.⁸ In particular, it does not apply to the LS estimator. The LS estimator is downward mean-biased conditionally on α for each α in (0, 1] in Models 1-3, and hence, is downward biased unconditionally for any distribution of α with support in (0, 1]. Hence, the above argument has no bearing on the relative attributes of the LS estimator and the median-unbiased estimator of α .

Second, the argument relies on the prior that is used to form the Bayes estimator being equal to the marginal distribution of the parameter. If they differ, unconditional unbiasedness does not hold.

Third, the criterion of unconditional unbiasedness is very weak. To see why, suppose the parameter α has distribution $P(\alpha)$ on (-1, 1] with mean equal to .8. Then, the Bayes estimator with prior $P(\alpha)$ has unconditional mean .8 and is unconditionally unbiased. But, the constant estimator .8 also has mean .8 and is unconditionally unbiased. This constant estimator does not exhibit any sort of impartiality -- it is completely partial. Thus, the property of unconditional unbiasedness does not reflect impartiality of an estimator.

Note that, roughly speaking, the way in which a (proper) Bayes estimator achieves unconditional unbiasedness in the AR/UR model is by having a downward conditional bias for values of α greater than the mean of α and by having an upward conditional bias for values of α less than the mean of α . That is, the Bayes estimator tends to be conditionally biased towards its unconditional mean. In contrast, the median-unbiased estimator of Section 3 is median-unbiased for each value of α , and hence, is conditionally and unconditionally median-unbiased for any distribution of α on (-1, 1]. This property reflects a strong degree of impartiality of the estimator.

9. EXAMPLES

In this section, we apply the methods introduced above to a number of data series that have been analyzed in the literature. First, we consider eight real exchange rate series that have been investigated recently using Bayesian methods by Schotman and van Dijk (1991a). Model 2 is used for these series (as in Schotman and Dijk (1991a)). Second, we consider two of the series considered by Nelson and Plosser (1982), viz., the velocity of

money and industrial production in the United States. Model 3 is used for these series. (Nelson and Plosser (1982) use Model 3 for the velocity series, but allow for an AR(6) model with intercept and time trend for the industrial production series. We argue that the much more parsimonious Model 3 may provide an adequate representation for industrial production.¹⁰)

We describe the data series here only briefly, see Schotman and van Dijk (1991a) and Nelson and Plosser (1982) for more detailed information including plots of the series or summary statistics. The first six real exchange rate series are of the U.S. dollar against the currencies of France (FR), West Germany (WG), Japan (JP), Canada (CA), United Kingdom (UK), and the Netherlands (NL). The last two real exchange rate series are of the German Dmark against the currencies of France and the Netherlands, both of which are fellow members of the European Monetary System. The real exchange rate series is defined to be $Y_t = e_t - P_t + P_t^*$, where e_t is the logarithm of the nominal exchange rate expressed as the domestic price of one unit of foreign currency and P_t are the logarithms of the consumer price indices of the domestic and foreign countries respectively. The data are monthly. The nominal exchange rate and consumer price indices are from the IFS databank --series ae and 64 -- for the sample period 73:01 to 88:07, except for Canada and the United Kingdom where the terminal date is 88:06. Hence, the sample size T+1 of the series $\{Y_t: t=0, ..., T\}$ is 187 for all series except the latter two, for which it is 186.

The velocity and industrial production series are annual series covering 1869-1970 and 1860-1970 respectively. Hence, the sample sizes T+1 of these series $\{Y_t: t=0, ..., T\}$ are 102 and 111 respectively. Both series are from data files supplied by Nelson and Plosser. The velocity series is originally from Friedman and Schwartz (1963) with revisions provided by Schwartz. The industrial production series is from Long Term Economic Growth (1973).

Table 8 provides the results for the real exchange rate series. For each of the U.S. dollar exchange rates except UK/US, the median-unbiased estimate of α is 1.00. For UK/US it is .995. For the FR/WG and NL/WG exchange rates, the median-unbiased estimates of α are .968 and .965 respectively. In each case, the median-unbiased estimates of α exceed those of the LS estimates by between .017 and .022. While the magnitudes of these differences are small in absolute terms, they are large in terms of their implications for the persistence of the time series. This is illustrated by the corresponding median-unbiased and LS estimates of the impulse response function and CIR given in Table 8. Because of their impartiality properties, we view the median-unbiased estimates as providing a better "best guess" of the true model and its persistence characteristics than do the LS estimates. The 90% confidence intervals for α , IR(h), and CIR (given in square brackets below the median-unbiased estimates) indicate the degree of uncertainty of our best guesses.

Inspection of the residuals $\{Y_t - \hat{\alpha}_{U2}Y_{t-1} : t = 1, ..., T\}$ for the eight real exchange rate series reveals evidence of non-normality for some of the series. The sample coefficients of skewness and kurtosis $(\sqrt{\beta_1}]$ and β_2 for the eight series of residuals are FR/US (.09, 3.91), WG/US (-.22, 3.88), JA/US (-.27, 3.87), CA/US (.84, 5.43), UK/US (-.42, 3.51), NL/US (-.10, 3.59), FR/WG (1.16, 6.09), and NL/WG (.43, 8.36). (The population coefficients of skewness and kurtosis for normal random variables are (0, 3.0).)

When the innovations $\{U_t: t=1, ..., T\}$ are normal, the standard errors of the sample coefficients of skewness and kurtosis are approximately $\sqrt{6/(T+1)}$ and $\sqrt{24/(T+1)}$, respectively, which equal .18 and .36 in the present case. Thus, the estimated coefficients of skewness are within two standard errors of those for normal random variables for the first four data series, but not for the last four. Also, the estimated coefficients of kurtosis are within two standard deviations of those for normal random variables for the UK/US and NL/US series, but not for the remaining series. The largest deviations from normality are the high estimates of kurtosis for the CA/US, FR/WG, and NL/WG series and the high estimates of skewness for the CA/US and FR/WG series.

The robustness of the results in Table 8 to non-normality has been checked by calculating the medianunbiased estimates of α and the corresponding exact 90% confidence intervals for α for a number of symmetric and asymmetric thick-tailed innovation distributions. The quantiles of $\hat{\alpha}_{LS2}$ that were needed for these calculations were computed using the method outlined in Section 7. The non-normal distributions used were t and shifted χ^2 distributions whose degrees of freedom yield coefficients of skewness and kurtosis that cover the range found in the residual series. The following distributions were considered (with coefficients of skewness and kurtosis given in parentheses): $(t_3(0, \infty), t_5(0, 9.0), t_{11}(0, 3.9), t_{14}(0, 3.6), \chi_4^2 - 4(1.4, 6.0), \text{ and } \chi_8^2 - 8(1.0, 4.5).$ (The t_3 distribution was considered as an extreme case even though its degrees of freedom lie outside the range found in the residual series.)

In short, the results reported in Table 8 were hardly changed at all when any of the above non-normal distributions was used to compute the bias-corrected estimates and exact confidence intervals. The median-unbiased estimates of α were unchanged except for the FR/WG and NL/WG series for which the estimates differed by at most .001 and .002, respectively, when any of the alternative distributions was used. The confidence intervals were unchanged except for the FR/WG series for which the confidence interval changed from [.92, 1.00] for normal innovations to [.93, 1.00] for all other innovations. (The latter difference is due to border-line downward rounding for the normal case and borderline upward rounding for the other distributions.) In conclusion, non-normality of the innovations does not seem to be a problem in the present case.

Schotman and van Dijk's (1991a) Bayesian estimates, based on a data-dependent uniform prior for α on $[a^*, 1)$ and no prior mass on $\alpha = 1$, lie between the LS and median-unbiased estimates. For the series ordered

as above and as in Table 8, their estimates are .982, .982, .990, .987, .977, .983, .954, and .965. If prior probability mass is placed on $\alpha = 1$, then their estimates are a corresponding weighted average of the estimates listed above and one. Depending on how much prior mass is placed on $\alpha = 1$, Schotman and van Dijk's estimates may be close to or somewhat distant from the median-unbiased estimates. An advantage of the median-unbiased estimates is that they do not rely on the resolution of the contentious question of how much prior mass should be placed on $\alpha = 1$.

Table 9 provides the results for the velocity and industrial production series. For velocity, the median-unbiased estimate of α is 1.0 compared to the LS estimate of .94. In fact, the LS estimate is barely above the bottom edge of the 90% central confidence interval of [.93, 1.0]. In consequence, this series exhibits the most dramatic differences of all the series between the LS and median-unbiased estimated persistence of shocks to the series, as measured by IR(h) and CIR. The differences are quite substantial. For velocity, the unbiased model selection procedure of Section 6 chooses the random walk with drift model over the trend stationary model.

For industrial production, the median-unbiased estimate of α is .89 as compared to the LS estimate of .84. This yields a 50% larger estimate of CIR using the unbiased estimate rather than the LS estimate. This is a fairly large difference, but it is not as dramatic as for the velocity series. The length of the 90% central confidence interval [.79, 1.0] for α indicates that α cannot be estimated very precisely in this example. Although the random walk with drift model is not rejected at level .05 (since the upper bound one-sided confidence interval for α contains 1), the unbiased model selection procedure of Section 6 chooses the trend stationary model.

The bottom half of Table 9 provides some information re the adequacy of Model 3 for the velocity and industrial production series. For industrial production, there is some evidence of skewness and kurtosis in the residuals. For velocity, the magnitude of the estimated coefficient on the time trend is small, but appears large relative to its standard error given that Model 3 requires $\tilde{\beta} = 0$ when $\alpha = 1$. The standard error estimate is spuriously precise, however, since it does not take into account the variability of $\hat{\alpha}_{U3}$.

The difference between the median-unbiased estimates, confidence intervals, and unbiased model selection results for the velocity series and those for the industrial production series indicates the value of having such statistics available. The outcome of a level .05 test of a unit root is the same for both series -- it does not reject. But the evidence for a unit root is much stronger for the velocity series than for the industrial production series, as indicated by the median-unbiased estimates, 90% confidence intervals, and unbiased model selection results.

10. COMPUTATION OF QUANTILES

In this section, we describe the numerical procedure that was used to compute the .05, .5, and .95 quantiles of the LS estimators $\hat{\alpha}_{LSj}$, j=1, 2, 3 which are reported in Tables 2-4. The basic idea is to write $P_{\alpha}(\hat{\alpha}_{LSj} \leq c)$ as the probability that a quadratic form in standard normal variates is less than or equal to zero, to apply the Imhof (1961) algorithm to compute such probabilities, and to compute the desired quantiles using an iterative procedure that involves computing such probabilities for different values of c until the value of c is found that yields the desired probability .05, .5, or .95.

First, we define the LS estimator $\hat{\alpha}_{LSj}$ as

(10.1)
$$\hat{\alpha}_{LSj} = Y'D_0'(I - P_j)D_TY/Y'D_T'(I - P_j)D_TY \text{ for } j = 1, 2, 3,$$

where $Y = (Y_0, ..., Y_T)'$, $D_0 = [0 : T] \in R^{T \times (T+1)}$, $D_T = [I : 0] \in R^{T \times (T+1)}$, I equals the T dimensional identity matrix, $P_1 = 0 \in R^{T \times T}$, $P_j = X_j (X_j X_j)^{-1} X_j'$ for $j = 2, 3, X_1 = (1, ..., 1)' \in R^T$, $X_2 = (X_1 : x_2) \in R^{T \times 2}$, and $X_2 = (1, 2, ..., T)'$.

Next, we write Y in terms of the underlying errors $\{U_t\}$. By the invariance properties of $\hat{\alpha}_{LSj}$ described in Section 2, it suffices to consider the case where $\mu = \beta = 0$ and $\sigma^2 = 1$ in Models 1-3 and $Y_0 = 0$ when $\alpha = 1$ in Models 2 and 3. In this case, we have

$$Y = R_{\alpha}U, \text{ where } U = (U_0, ..., U_T)', U \sim N(0, I_{T+1}),$$

$$R_{\alpha} = \begin{bmatrix} b & 0 & 0 & \cdots & 0 & 0 \\ b\alpha & 1 & 0 & \cdots & 0 & 0 \\ b\alpha^2 & \alpha & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b\alpha^T & \alpha^{T-1} & \alpha^{T-2} & \cdots & \alpha & 1 \end{bmatrix}, b = 1/(1 - \alpha^2)^{1/2} \text{ when } \alpha \in (-1, 1), \text{ and } b = 0 \text{ otherwise.}$$

We now can write

(10.3)
$$P_{\alpha}(\hat{\alpha}_{LSj} \leq c) = P(\underline{U'W_{\alpha jc}U} \leq 0), \text{ where}$$

$$W_{\alpha jc} = R'_{\alpha}[D'_{0}(I-P_{j})D_{T}/2 + D'_{T}(I-P_{j})D_{0}/2 - cD'_{T}(I-P_{j})D_{T}]R_{\alpha}.$$

(Note that, as defined, the weight matrix $W_{\alpha jc}$ is symmetric.) Since $U'W_{\alpha jc}U$ is a quadratic form in standard normal variates, $P_{\alpha}(\hat{\alpha}_{LSj} \leq c)$ can be computed using Imhof's (1961) algorithm. This was done by employing the FORTRAN subroutines given by Koerts and Abrahamse (1971, Ch. 9).¹¹

To obtain the p-th quantile of $\hat{\alpha}_{LSj}$, we need to find the value c_p such that $P_{\alpha}(\hat{\alpha}_{LSj} \leq c_p) = p$. Since we can compute $P_{\alpha}(\hat{\alpha}_{LSj} \leq c)$ for arbitrary c and $P_{\alpha}(\hat{\alpha}_{LSj} \leq c)$ is increasing in c, a simple iterative procedure was used to determine c_p . The results given in Tables 2-4 were obtained in this manner.

We note that two alternative methods of calculating the quantiles of $\hat{\alpha}_{LSj}$ are numerical complex integration and simulation. These methods are particularly useful when the sample size is large -- say, greater than one hundred, and especially when it is greater than two hundred. Simulation avoids the use of FORTRAN subroutines, which may be an advantage in some cases. Of course, many repetitions and a good random number generator need to be used when the simulation method is employed.

FOOTNOTES

¹I would like to thank Hong-Yuan Chen, Eric Zivot, and especially Zhaoliang Zhu for their excellent research assistance, Peter Phillips, Glenn Rudebusch, Chris Sims, Bent Sorensen, and Jim Stock for helpful comments, and Willem Buiter for supplying the exchange rate data analyzed in Section 9. I gratefully acknowledge the research support of the National Science Foundation via grant no. SES-8821021. I dedicate this paper to my wife Anne Kao Andrews whose continual support is responsible for any and all research success I have had to date (and who has promised unspeakable repercussions if I do not dedicate the paper thus).

²See Andrews (1989) for further discussion of this problem in the context of nonlinear models without deterministic or stochastic trends.

³Model 3 is quite similar to that considered by Bhargava (1986) who is concerned with tests of a unit root. Model 3 is less restrictive, however, in that it allows the initial observation Y_0 to be arbitrary when $\alpha = 1$, whereas Bhargava requires it to have a particular normal distribution. On the other hand, Bhargava considers cases where $\alpha > 1$. The approach of the present paper could also be extended to cover cases where $\alpha > 1$, but only by imposing a very specific assumption on the initial observation Y_0 (as in Bhargava (1986)) and this assumption seems to be somewhat arbitrary.

It would be useful to have an analytic proof that the median of $\hat{\alpha}_{LSj}$ is strictly increasing for $\alpha \in (-1, 1]$ for all $T \ge 4$ and j = 1, 2, 3, but we do not have such a proof at present. We do know, however, that the more general proposition that all of the quantiles of $\hat{\alpha}_{LSj}$ are strictly increasing for $\alpha \in (-1, 1]$ for all $T \ge 4$ and j = 1, 2, 3 is false. Numerical calculations of the .95 quantile function of $\hat{\alpha}_{LS1}$ for sample sizes T+1=10, ..., 60 reveal that the function dips very slightly for α very close to one, see Table 2. For T+1=10, where the dip is the largest, the numerical calculations have been substantiated by simulation results. For sample sizes of 70 or more the dip disappears. The fact that monotonicity is not a universal feature of the quantiles suggests that its proof for those estimators, quantiles, and sample sizes for which it holds may be more difficult than it appears it ought to be.

⁵As mentioned in footnote 4, the .95 quantiles of $\hat{\alpha}_{LS1}$ are not increasing over a tiny range of α values near one for some sample sizes, e.g., see Table 2 with T+1=60, $\alpha=.995$, and $\alpha=.999$. For the estimators $\hat{\alpha}_{LS2}$ and $\hat{\alpha}_{LS3}$, the .5 and .95 quantiles of $\hat{\alpha}_{LSj}$ are monotone for all sample sizes in Tables 3 and 4. For these estimators, however, a dip occurs in the .05 quantile function of $\hat{\alpha}_{LSj}$ for values of α near -1 for sample sizes of 60 or less for $\hat{\alpha}_{LS2}$ and for sample sizes of 70 or less for $\hat{\alpha}_{LS3}$. Except for sample size 10, the dip always occurs in the interval (-.999, -.99]. For practical purposes, this region of α values is very rarely of interest. In consequence, non-monotonicity of the quantiles of $\hat{\alpha}_{LSj}$ is of little or no practical import for the estimators, quantiles, sample sizes, and α values of interest.

⁶As defined, if $\hat{\alpha}_{LSj} > q_{p_2}(1)$, then $\hat{c}_L > 1$, $\hat{c}_U = 1$, and the confidence interval $[\hat{c}_L, \hat{c}_U]$ equals the null set. This definition guarantees that the coverage probability of the random interval $[\hat{c}_L, \hat{c}_U]$ is exactly p when $\alpha = 1$ just as it is for all other values of α . An alternative, and perhaps more natural, definition of \hat{c}_L is to set it equal to 1 when $\hat{\alpha}_{LSj} > q_{p_2}(1)$ and then the confidence interval is the single point $\{1\}$ in this case. If this is done, the coverage probability of the random interval $[\hat{c}_L, \hat{c}_U]$ is exactly 1-p for all $\alpha \in (-1, 1)$, but is $1-p_1$ (> 1-p) for $\alpha = 1$ (since there is no way for the confidence interval to miss the point $\alpha = 1$ by being to the right of 1). This asymmetry between the coverage probabilities of the confidence interval for the (trend) stationary and non-stationary cases may be deemed undesirable if attention is focussed on which of the two models is correct. If it is deemed to be undesirable, then one can define the confidence interval as in (4.4) and view an occurrence of a null confidence set as evidence against the model restriction that $\alpha \in (-1, 1]$, since the test "reject H_0 : $\alpha \in (-1, 1]$ in favor of H_1 : $\alpha > 1$ if $\hat{\alpha}_{LSj} > q_{p_2}(1)$ " is a valid level p_2 significance test.

⁷As pointed out by Peter Phillips (personal communication), the results of Knight (1985b) suggest that the estimators $\hat{\alpha}_{Uj}$ for j=1, 2, 3 and the confidence intervals $[\hat{c}_L, \hat{c}_U]$ may be robust against some non-normal innovation distributions. Knight found a lack of sensitivity of the moments of the LS estimator and the two stage LS estimator of linear simultaneous equations models to the introduction of some skewness and kurtosis in the equation errors. His results are relevant, because the estimators that he considers, like $\hat{\alpha}_{LSj}$, are ratios of quadratic forms in the errors.

⁸The argument that establishes the unconditional unbiasedness of a Bayes estimator is as follows: Let $\hat{\alpha}_B(Y)$ denote the Bayes estimator of α as a function of the data Y. Then,

$$E_{Y}\hat{\alpha}_{B}(Y) = E_{Y}E(\alpha|Y) = E_{Y,\alpha}\alpha = E_{\alpha}\alpha,$$

where $E_{Y,\alpha}$, E_Y , and E_{α} denote expectations with respect to (Y, α) , Y, and α respectively. This argument relies on $E_{\alpha}|\alpha|$ being finite, which does not hold if the prior on α is improper. In fact, the posterior mean of α does not necessarily exist if the prior on α is improper, e.g., it does not exist for a Jeffrey's prior, see Phillips (1991).

Schotman and van Dijk (1991a) use Model 2 but condition on Y_0 . We view it to be an attribute of our procedure that we do not need to condition on Y_0 , since Y_0 contains information about α when $\alpha \in (-1, 1)$.

¹⁰The LS estimates of the five regressors ΔY_{t-1} , ..., ΔY_{t-5} that appear in Nelson and Plosser's (1982) model for industrial production but do not appear in Model 3 equation (2.3) are .10, -.04, .04, -.05, -.22, respectively, each with a standard error estimate of .10. This pattern of coefficients seems somewhat implausible, since there is no good reason to expect a significant coefficient on ΔY_{t-5} , but zero coefficients on the other lags. In fact, the *F*-statistic for testing jointly whether the coefficients on all five ΔY_{t-j} regressors are zero is .2, which is very small (standard asymptotics are applicable for this test statistic whether or not $\alpha = 1$). This and other evidence suggest that Model 3 with no ΔY_{t-j} regressors may be appropriate for modelling the industrial production series, although one cannot be certain of this.

¹¹In Koerts and Abrahamse's (1971) subroutine FQUAD, the tolerance parameters EPS1 and EPS2 were set equal to 10⁻⁶ and 10⁻³⁶ respectively. This condition was found to provide a good tradeoff between speed and accuracy.

REFERENCES

- Andrews, D. W. K. (1989): "Power in Econometric Applications," Econometrica, 57, 1059-1090.
- Andrews, D. W. K. and P. C. B. Phillips (1987): "Best Median-Unbiased Estimation in Linear Regression with Bounded Asymmetric Loss Functions," *Journal of the American Statistical Association*, 82, 886-893.
- Andrews, D. W. K. and H. Y. Chen (1991): "Unbiased Estimation of and Confidence Intervals for the Cumulative Impulse Response of U.S. Macroeconomic Time Series," manuscript in preparation, Cowles Foundation, Yale University.
- Bhargava, A. (1986): "On the Theory of Testing for Unit Roots in Observed Time Series," Review of Economic Studies, 53, 369-384.
- Davis, A. W. (1976): "Statistical Distributions in Univariate and Multivariate Edgeworth Populations," Biometrika, 3, 661-670.
- DeJong, D. and C. H. Whiteman (1991): "Trends and Random Walks in Macroeconomic Time Series: A Reconsideration Based on the Likelihood Principle," Journal of Monetary Economics, 19, forthcoming.
- Evans, G. B. A. and N. E. Savin (1981): "Testing for Unit Roots: I," Econometrica, 49, 753-779.
- (1984): "Testing for Unit Roots: 2," Econometrica, 52, 1241-1269.
- Friedman, M. and A. J. Schwartz (1963): A Monetary History of the United States, 1867-1960. Princeton, NJ: Princeton University Press.
- Imhof, J. P. (1961): "Computing the Distribution of Quadratic Forms in Normal Variates," Biometrika, 48, 419-426.
- Kendall, M. G. (1954): "Note on Bias in the Estimation of Autocorrelation," Biometrika, 41, 403-404.
- Knight, J. L. (1985a): "The Joint Characteristic Function of Linear and Quadratic Forms of Non-normal Variables," Sankhya, Series A, 47, 231-238.
- (1985b): "The Moments of OLS and 2SLS When the Disturbances Are Non-normal," Journal of Econometrics, 27, 39-60.
- Koerts, J. and A. P. J. Abrahamse (1971): On the Theory and Application of the General Linear Model. Rotter-dam: Rotterdam University Press.
- Lehmann, E. L. (1959): Testing Statistical Hypotheses. New York: John Wiley and Sons.
- Long Term Economic Growth (1973): Washington, D.C.: U.S. Department of Commerce.
- Marriott, F. H. C. and J. A. Pope (1954): "Bias in the Estimation of Autocorrelations," Biometrika, 41, 390-402.
- Nelson, C. R. and C. I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series," Journal of Monetary Economics, 10, 139-162.
- Orcutt, G. H. and H. S. Winokur, Jr. (1969): "First Order Autoregression: Inference, Estimation, and Prediction," Econometrica, 37, 1-14.

Phillips, P. C. B. (1977): "Approximations to Some Finite Sample Distributions Associated with a First-Order Stochastic Difference Equation," Econometrica, 45, 463-485. (1991): "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," Journal of Applied Econometrics, 6, forthcoming. Quenouille, M. H. (1949): "Approximate Tests of Correlation in Time-Series," Journal of the Royal Statistical Society Series B, 11, 68-84. (1956): "Notes on Bias in Estimation," Biometrika, 43, 353-360. Rudebusch, G. D. (1990): "Trends and Random Walks in Macroeconomic Time Series: A Re-examination," unpublished manuscript, Division of Research and Statistics, Board of Governors of the Federal Reserve System, Washington, D.C. Schotman, P. and H. K. van Dijk (1991a): "A Bayesian Analysis of the Unit Root in Real Exchange Rates," Journal of Econometrics, forthcoming. (1991b): "Posterior Analysis of Possibly Integrated Time Series with an Application to Real GNP," in New Directions in Time Series Analysis, IMA Volume in Mathematics and its Applications, ed. by E. Parzen et al. New York: Springer Verlag, forthcoming. Sims, C. A. (1988): "Bayesian Skepticism on Unit Root Econometrics," Journal of Economic Dynamics and Control, 12, 463-474. Sims, C. A. and H. Uhlig (1991): "Understanding Unit Rooters: A Helicopter Tour," Econometrica, 59, forthcoming. Stock, J. H. (1990): "Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series," unpublished manuscript, Department of Economics, University of California, Berkeley.

Subrahmaniam, K. (1966): "Some Contributions to the Theory of Non-normality - I (Univariate Case)," Sankhya,

(1968): "Some Contributions to the Theory of Non-normality - II," Sankhya, Series A, 30, 411-432.

Series A, 28, 389-406.

TABLE 1 PROBABILITY OF UNDERESTIMATION OF α BY THE LEAST SQUARES ESTIMATOR IN MODEL 3 FOR SAMPLE SIZES 60, 100, AND 150 a

	Probabi	ility of Underestin	mation
α	T+1=60	T+1 = 100	T+1=150
0.0	.61	<i>.</i> 58	.57
.05	.61	.5 9	.58
.10	.62	.60	.5 8
.15	.63	.60	. 5 8
.20	.63	.61	.58
.25	.65	.61	.59
.30	.6 6	.61	.5 9
.35	.68	.63	.60
.40	.69	.63	.61
.45	.69	.65	.63
.50	.70	.65	.63
.55	.72	.67	.65
.60	.74	.68	.65
.65	.76	.71	.67
.70	.78	.72	.67
.75	.81	.75	.69
.80	.84	.78	.73
.85	.87	.82	.77
.90	.92	.87	.81
.95	.97	.94	.90
.99	.99	.99	.99
1.00	.995	.995	.996

^aThe table gives $P_{\alpha}(\hat{a}_{LS3} < \alpha)$ for Model 3 (which contains an intercept and time trend) of Section 2. This table was constructed via simulation with 10,000 repetitions for each entry.

TABLE 2 $\label{eq:QUANTILES OF THE LEAST SQUARES ESTIMATOR OF α FOR MODEL 1^{α}}$

		T+1 = 60			T+1 = 100)
$\alpha \backslash Quantile$.05	.5	.95	.05	.5	.95
.00	211	.000	.211	164	.000	.164
.05	164	.049	.258	115	.050	.212
.10	116	.098	.304	067	.099	.259
.15	068	.148	.350	017	.149	.307
.20	019	.197	.395	.032	.198	.353
.25	.030	.246	.440	.082	.248	.400
.30	.080	.295	.484	.132	.297	.446
.35	.130	.344	.528	.183	.347	.491
.40	.181	.393	.571	.235	.39 6	.536
.45	.233	.443	.614	.286	.446	.581
.50	.285	.492	.656	.339	.4 95	.625
.55	.338	.541	.697	.392	.54 5	.669
.60	.392	.590	.738	.445	.594	.711
.65	.446	.640	.777	.500	,644	.754
.70	.502	.689	.816	.555	.693	.795
.75	.559	.738	.854	.611	.743	.835
.80	.618	.787	.890	.669	.792	.874
.85	.678	.837	.924	.728	.842	.912
.90	.719	.886	.957	.790	.891	.948
.93	.784	.917	.975	.829	.921	.968
.95	.813	.937	.987	.844	.941	.980
.97	.846	.958	.99 8	.886	.962	.992
.99	.891	.981	1.009	.920	.983	1.003
.995	.910	.9 89	1.011	.935	.990	1.006
.999	.944	.997	1.010	.955	.997	1.007

^aNumerical calculations show the distribution of $\hat{\alpha}_{LS1}$ to be an odd function of α . Thus, the .05, .5, and .95 quantiles of $\hat{\alpha}_{LS1}$ for $\alpha \in (-.999, 0)$ equal the .95, .5, and .05 quantiles, respectively, of $\hat{\alpha}_{LS1}$ for $-\alpha \in (0, .999)$. In consequence, Table 1 provides the quantiles of $\hat{\alpha}_{LS1}$ for all values of $\alpha \in (-.999, .999)$.

QUANTILES OF THE LEAST SQUARES ESTIMATOR OF α FOR MODEL 2 FOR SAMPLE SIZES (T+1) 10-200

TABLE 3

α/Quantile 999 995 996 950 950 650 650 7	
	T+1=10
888374784266771694678668484868746483787767377689	T+1=20
	T+1 = 30
-1.014 - 997 - 928 -1.017 - 988 - 889 -1.017 - 986 - 858 -1.017 - 986 - 858 -1.017 - 986 - 858 -1.017 - 986 - 933 - 770 -1.984 - 913 - 739 -1.967 - 884 - 697 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 572 -1.987 - 787 - 787 -1.987 - 787 - 787 -1.987 - 787 - 787 -1.987 - 787 - 787 -1.987 - 787 - 787 -1.987 - 787 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.987 - 787 -1.988 - 789 -1.988 - 789 -1.988 - 789 -1.988 - 998 -1.998 - 998 -1.998 - 998 -1.999	T+1 = 40
-1. 012 - 997 - 938 -1. 013 - 981 - 899 -1. 013 - 981 - 899 -1. 013 - 981 - 891 -1. 090 - 936 - 797 - 990 - 936 - 726 - 931 - 838 - 662 - 884 - 741 - 544 - 828 - 789 - 662 - 898 - 789 - 662 - 898 - 789 - 662 - 898 - 789 - 662 - 787 - 787 - 377 - 787 -	T+1 = 50

- 995 - 995 - 996 - 996 - 996 - 996 - 156 - 156	α/Quantile	
-1.010 -1	.05	7
	, _{(,1}	T+1 = (
		8
-1.000 -1	. 05	
		T+1:
8 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	ٔ ن	= 70
роди и продержите в под в в в в в в в в в в в в в в в в в в в	Š	
-1.008 -1	.05	T
	'n	$T+1=\{$
	. 95	80
-1.007 -1	.0	
94111111111111111111111111111111111111		T+1 =
		8
999986556637442566633666336663744956663366633666336663366636663746666366636	95	
-1.006 -1	.05	7
	·.	T+1 =
	. 95	100
- 999 - 995 - 996 - 996 - 996 - 996 - 100 - 100	ð	

1. 999 1. 990 1. 990 1. 990 1. 990 1. 750 1. 750	α/Quantile	
-11.005 -11.00	.05	7
		T+1=1
	. 95	110
88881388138131111111111111111111111111	.05	T+1
9997 9997 9985 9985 9987 9989 9924	.5 .95	= 125
-1.004991973 -1.003991993 -1.9899859919769859919769859819639258879647967068247476447476965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965994276965963752692721119352220121056030123105030135260333206141271092092294006141271007127008919319596685137634721686596137634721687988968997908997908971999	.05 .5 .95	T+1=150
-1.003 - 998 - 976 -1.002 - 992 - 961 -1.987 - 966 - 951 -1.987 - 966 - 952 -1.974 - 946 - 893 -1.988 - 197 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.858 - 177 - 175 -1.121 - 105 -1.121 - 105 -1.121 - 105 -1.121 - 105 -1.121 - 105 -1.121 - 105 -1.121 - 105 -1.121 - 106 -1.126 - 1.137 -1.121 - 105 -1.121 - 105 -1.121 - 106 -1.126 - 1.137 -1.121 - 106 -1.126 - 1.137 -1.121 - 106 -1.126 - 1.137 -1.121 - 106 -1.127 - 107 -1.121 - 108 -1.121	.05 .5 .95	T+1=200
	Ω	

QUANTILES OF THE LEAST SQUARES ESTIMATOR OF α FOR MODEL 3 FOR SAMPLE SIZES (T+1) 10-200

TABLE 4

1. 995 1. 995 1. 995 1. 995 1. 995 1. 995 1. 1995 1. 1995	α/Quantile
-1.050 -1.1051	.05
	$T+1 = \frac{1}{2}$
	10
-1.037 -1.037 -1.037 -1.028 -1	. 05
	T+1 =
	.95
-1.019 -1.024 -1.024 -1.024 -1.025 -1.025 -1.025 -1.027 -1.027 -1.027 -1.027 -1.027 -1.027 -1.027 -1.027 -1.028 -1	.05
1133	T+1=3
	30
-1.018 -1.018 -1.019 -1	.05
	T+1=40
	.95
-1.012 -1.013 -1.013 -1.003 -1.003 -1.003 -1.003 -1.003 -1.003 -1.103 -1.003 -1.103 -1	T.
	T+1=50
	.95
- 999 - 990 - 990 - 990 - 990 - 990 - 990 - 990 - 150	Ω

1. 9990 1. 9990 1. 9990 1. 1000 1.	α/Quantile
-1.010 -997 -1.010 -998 -1.999 -1.991 -1.999	T+1 = 60 .05 .5
-1.000 -1.0000 -1.000 -1.000 -1.000 -1.000 -1.000 -1.000 -1.000 -1.00	.95 .05
991 982 983 984 987 988 991 988 991 991 991 991 991	T+1 = 70 .5 .95
-1.008997955 -1.008997955 -1.00699399199399198399192281699399192399399192399399192399399192399399479699399665999479665999555538560850736218152136218152136218152136218152911220631150741312591122063162026159162026159162026159162026159172359172359173391520211406563321550606321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608321550608331550 .	T+1 = 80
-1.007 -993 -1997 -1.005 -1993 -1993 -1992 -1992 -1992 -1992 -1992 -1992 -1993 -1992 -1993	T+1 = 90 .05 .5
958 -1.006995 9134 -1.006997 9139 -1.006997 9139 -1.006997 1852961 1852961 1852961 1852961 1963963 1853963 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1786973 1787 1788973 1788973 1789973 1	T+1 = 100
97 1.938995 84924995 85924995 86924995 87924995 88924995 96924995 97924995 98924995 98924995 98924995 98924925 98924925 98925925 98925925 98925925 99925925 99925925 99925925 99925925 99925925 99925925 99925925 99925925 99925925 99925925 99925925	= 100 5 .95 α

TABLE 4 (cont.)

	7	+1 = 1	110	7	T+1 =	125	7	r +1 = 1	150		T+1=2	200	
α /Quantile	.05	· . 5	. 95	.05	.5	.95	.05	. 5	.95	.05	. 5	.95	α
999	-1.006	997	963	-1.005		966	-1.004	997	- . 970	-1.003	-,998	- . 976	999
995	-1.005	990	−. 942	-1.004	−. 991	 947	-1.003	- . 991	 953	-1.002		- 962	995
990	-1.003	984	- 929	-1.002	 985	-, 935	-1.000	985	942	- . 999	986	 951	990
970 950	991 979	963	894	- .990	964	- ,901	 . 989	965	- . 909	987	966	921	970
930 930	967	943 924	865 838	978 965	944 924	一. 872 一. 846	- .976 - .963	—. 945 —. 925	8B2	, 974	946	 894	950
900	- 947	894	800	 945	895	808	- . 942	- .896	856 818	960		869	930
850	- . 912	845	740	909	846	748	905	847	- 759	938 938	897	- 832	900
800	- 874	797	- 682	871	797	690	866	 798	- 702	899 858	一. 847 一. 798	774 717	850 800
750	 835	 748	625	 831	 748	- . 634	825	748	- 646	 816	749	- 661	750
700	 . 795	699	570	- .790	- . 699	 579	- . 784	699	- 591	-: 774	- 700	607	- 700
650	- . 755	- .651	- .516	 . 749	- .651	 525	- . 741	−. 650	 537	 730	650	553	650
600	713	₹. 602	- 463	- . 707	602	- .472	698	601	 484	686	 601	500	600
550	 671	553	410	- . 664	553	419	654	552	431	 641	- . 552	448	550
500 450	628 584	505 456	358 306	-: 620 -: 576	一. 504 一. 455	367 315	—. 610 —. 566	T. 503	- 379 - 322	596		 396	500
400	T. 540	407	-, 306 -, 255	532	407	315 264	- . 521	一. 454 一. 405	327 376	- .551	 453	- 344	450
350	496	359	-, 205	- .487	358	214	475	356	276 225	 505		- 293	~.400
300	- . 451	310	−. 155	442	309	- 163	129	307	175	T. 458	355 355	- 242	350
	406	262	105	396	260	- .113	383	- 258	125	 412 - . 365	T. 306	- 191	300 250
250 200	360	213	056	- .350	211	- 064	337	209	075	- .318	─.256 ─.207	- 141 - 091	200
150	- . 314	- 164	007	 304	163	- .015	 290	160	026	270	158	041	150
100	268	 . 116	.042	− .257	- .114	.034	- . 243	- .111	.023	222	- .109	.008	100
050	221	- 067	.090	210	065	.082	- 195	~ : 062	.072	174	059	.057	050
- 000	174	018	.138	- .163	016	.131	- .147	- .013	. 121	− .126	010	.106	.000
.050 .100	一. 127 一. 079	.030	.186	115 - 067	.033	.178	 099	.036	.169	~. 077	.039	.155	.050
.150	031	.079 .127	.233 .280	一.067 一.019	.081	.226 .273	051 002	.084	.217	029	.088	.203	.100
.200	.017	.176	.326	.030	.179	.320	.047	.133 .182	.264 .311	.021	.138	. 251	- 150
.250	.066	.224	.372	.079	.228	.366	.096	.231	.358	.070	-187	.299	.200
.300	.115	.273	418	.128	. 276	.412	. 146	.280	.405	.120	.236 .285	.346 .393	.250 .300
.350	. 164	.322	.463	.178	.325	. 458	.196	.329	.451	.220	.335	. 440	.350
- 400	.214	.370	.508	.228	. 374	. 503	. 246	.378	. 497	.271	.384	.487	.400
.450	. 264	-419	.553	- 278	.422	. 548	. 297	.427	. 542	. 321	.433	. 533	.450
.500 .550	. 314	. 467	. 597	- 329	471	. 593	- 348	476	. 587	. 373	.482	.579	.500
.600	.365 .417	.516 .564	.640	.380 .432	.520 .568	. 637	. 399	.525	. 632	- 424	. 531	. 624	.550
.650	.468	-612	.683 .725	.484	.617	.680 .723	.451 .503	.574 .623	.676 .719	.476	. 581	.669	.600
.700	. 521	.660	.767	.536	.665	.765	.556	.671	.719	. 529	.630	.713	.650 .700
.750	. 573	.708	.808	. 589	.714	.806	. 609	.720	.702 .804	. 582 . 635	. 679	.757	.750
.800	. 626	.756	. 847	. 643	762	. 847	. 663	769	.845	.689	.728 .777	.800 .843	.800
-850	. 679	.803	. 885	.696	.809	.886	. 717	817	.885	.744	.826	.884	.850
- 900	. 731	.849	. 922	.749	.856	. 923	. 771	.864	. 923	.799	.874	. 923	.900
.930	. 762	.876	. 942	.780	.883	. 943	. 803	.B92	. 945	.032	902	946	.930
.950 .970	. 781	.892	. 954	.800	.900	- 956	.821	909	. 958	.854	. 921	. 960	. 950
.990	.797	.906	. 966	.818	.915	. 968	. 843	. 925	.970	.874	. 938	. 973	.970
.995	.809 .811	.917 .918	.974	. 831 . 832	.926 .928	.977 .978	. 857	938	- 980	. 891	952	. 984	.990
1.000	.811	.919	.975 .976	.833	.928	.978	.859 .860	.939 .910	. 982	. 893	. 954	. 986	.995
	4		. ,						. 982	. 894	. 955	. 997	1.000

TABLE 5 COMPARISON OF THE QUANTILES OF THE LEAST SQUARES ESTIMATOR OF α FOR DIFFERENT INNOVATION DISTRIBUTIONS FOR MODEL 3 AND SAMPLE SIZE $60\,$

			(a) 5 QU	ANTILE				
α /Distribution	N(0,1)	<i>t</i> ₁₀	<i>t</i> ₄	<i>t</i> ₃	t_2	t_1	χ_8^2 - 8	$\chi_4^2 - 4$
.00	034	031	039	033	035	032	037	038
.05	.013	.017	.008	.015	.013	.017	.011	.009
.10	.061	.064	.056	.062	.061	.064	.058	.009 .057
.15	.108	.112	.104	.111	.108	.113	.106	.104
.20	.155	.159	.151	.158	.156	.161	.153	.152
.25	.202	.207	.199	.158 .206 .253	.156 .203 .251	.209	.200	.198
.30	.250	.253	.247	.253	.251	.257	.247	.246
.35	.297	.300	.294	.301 .348	.298 .345	.305	.295	.294 .340
.40	.344	.348	.342	.348 205	.343 303	.353	.343	.34U 200
.45	.391	.395 .442	.390	.395 .442	.393 .441	.401 .449	.390	.388 .435
.50 .55	.438	. 44 2 .489	.436 .483	. 44 2 .489	.487	. 41 9 .497	.437 .483	.433 192
.60	.485 .532	.535	.530	.535	.534	. 49 7 .544	.530	530
.65	.578	.581	.577	.583	.581	.592	.530 .577	.550 576
.70	.624	.627	.623	.628	.627	.639	.623	.622
.75 .75	.670	.672	.669	.675	.673	.685	.669	.482 .530 .576 .622 .667 .712
.80	.715	.717	.715	.719	.718	.730	.714	.712
.85	.758	.760	.759	.761	.762	.775	.758	.757 .797
.90	.79 9	.802	.801	.803	.803	.819	. 7 99	<i>.7</i> 97
.93	.821	.824	.824	.825	.826	.843	.820	.821
.95	.834	.837	.838	.838	.839	.858	.833	.821 .834
.97	.845 .851	.848	.84 8	.848 .855	.851 .857	.87 0	.843	.845 .851
.99	.851	.854	.856	.855	.857	.876	.850	.851
.995	.852	.855	.856	.855	.858	.875	.851	.852
1.00	.853	.855	.857	.856	.859	.873	.851	.852
			(b) .05 QU	JANTILE				
α /Distribution	N(0,1)	t ₁₀	<i>t</i> ₄	<i>t</i> ₃	t_2	<i>t</i> ₁	χ_8^2 - 8	$\chi_4^2 - 4$
.00	244	244	238	234	226	189	242	232
.05	199	199	193	190	181	141	196	186
.10	154	153	147	145	134	~.094	152	140
.15	108	107	101	099	088	048	107	094
.20	062	060	~.055	054	043	001	061	048
.25	016	014	009	007	.003	.046	014	001
.30	.031	.035	.038	.039	.050	.093	.032	.046
.35	.078	.081	.084	.088	.097	.141	.081	.092
.40	126	177	171	136		100	.128	.140
.45	.126	.127	.131	.136	.145	.190	177	10/
	.173	.178	.180	.184	.190	.237	.177	.186
.50	.173 .222	.178 .228	.180 .228	.184 .230	.190 .238	.237 .284	.177 .224	.232
.55	.173 .222 .270	.178 .228 .278	.180 .228 .275	.184 .230 .278	.190 .238 .288	.237 .284 .331	.177 .224 .273	.232 .280
.55 .60	.173 .222 .270 .319	.178 .228 .278 .327	.180 .228 .275 .324	.184 .230 .278 .328	.190 .238 .288 .337	.237 .284 .331 .379	.177 .224 .273 .322	.232 .280 .329
.55 .60 .65	.173 .222 .270 .319 .368	.178 .228 .278 .327 .375	.180 .228 .275 .324 .371	.184 .230 .278 .328 .376	.190 .238 .288 .337 .386	.237 .284 .331 .379 .428	.177 .224 .273 .322 .370	.232 .280 .329 .377
.55 .60 .65 .70	.173 .222 .270 .319 .368 .417	.178 .228 .278 .327 .375 .423	.180 .228 .275 .324 .371 .422	.184 .230 .278 .328 .376 .425	.190 .238 .288 .337 .386 .434	.237 .284 .331 .379 .428 .476	.177 .224 .273 .322 .370 .419	.232 .280 .329 .377 .425
.55 .60 .65 .70 .75	.173 .222 .270 .319 .368 .417	.178 .228 .278 .327 .375 .423 .473	.180 .228 .275 .324 .371 .422 .472	.184 230 .278 .328 .376 .425 .473	.190 .238 .288 .337 .386 .434	.237 .284 .331 .379 .428 .476	.177 .224 .273 .322 .370 .419	.232 .280 .329 .377 .425
.55 .60 .65 .70 .75	.173 .222 .270 .319 .368 .417 .466	.178 .228 .278 .327 .375 .423 .473 .521	.180 .228 .275 .324 .371 .422 .472 .521	.184 .230 .278 .328 .376 .425 .473	.190 .238 .288 .337 .386 .434 .482 .531	.237 .284 .331 .379 .428 .476 .524 .570	.177 .224 .273 .322 .370 .419	.232 .280 .329 .377 .425 .474 .522 .569
.55 .60 .65 .70 .75 .80	.173 .222 .270 .319 .368 .417 .466 .515	.178 .228 .278 .327 .375 .423 .473 .521	.180 .228 .275 .324 .371 .422 .472 .521 .569	.184 .230 .278 .328 .376 .425 .473 .523 .571	.190 .238 .288 .337 .386 .434 .482 .531 .578	.237 .284 .331 .379 .428 .476 .524 .570 .613	.177 .224 .273 .322 .370 .419 .467 .516 .564	.232 .280 .329 .377 .425 .474 .522 .569
.55 .60 .65 .70 .75 .80 .85 .90	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607	.232 .280 .329 .377 .425 .474 .522 .569 .613
.55 .60 .65 .70 .75 .80 .85 .90 .93	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607 .632	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615 .638	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615 .640	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615 .638	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620 .641	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646 .665	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607 .631	.232 .280 .329 .377 .425 .474 .522 .569 .613 .634
.55 .60 .65 .70 .75 .80 .85 .90 .93 .95	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607 .632 .646	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615 .638 .654	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615 .640 .655	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615 .638 .651	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620 .641 .653	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646 .665 .675	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607 .631 .644	.232 .280 .329 .377 .425 .474 .522 .569 .613 .634 .646
.55 .60 .65 .70 .75 .80 .85 .90 .93 .95	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607 .632 .646 .658	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615 .638 .654	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615 .640 .655	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615 .638 .651 .663	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620 .641 .653 .664	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646 .665 .675 .686	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607 .631 .644 .654	.232 .280 .329 .377 .425 .474 .522 .569 .613 .634 .646 .656
.55 .60 .65 .70 .75 .80 .85 .90 .93 .95 .97	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607 .632 .646 .658	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615 .638 .654 .666	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615 .640 .655 .666	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615 .638 .651 .663	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620 .641 .653 .664	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646 .665 .675 .686	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607 .631 .644 .654	.232 .280 .329 .377 .425 .474 .522 .569 .613 .634 .646 .656
.55 .60 .65 .70 .75 .80 .85 .90 .93 .95	.173 .222 .270 .319 .368 .417 .466 .515 .562 .607 .632 .646 .658	.178 .228 .278 .327 .375 .423 .473 .521 .568 .615 .638 .654	.180 .228 .275 .324 .371 .422 .472 .521 .569 .615 .640 .655	.184 .230 .278 .328 .376 .425 .473 .523 .571 .615 .638 .651 .663	.190 .238 .288 .337 .386 .434 .482 .531 .578 .620 .641 .653 .664	.237 .284 .331 .379 .428 .476 .524 .570 .613 .646 .665 .675 .686	.177 .224 .273 .322 .370 .419 .467 .516 .564 .607 .631 .644 .654	.232 .280 .329 .377 .425 .474 .522 .569 .613 .634 .646 .656

TABLE 5, continued

(c) .95 QUANTILE

α /Distribution	N(0,1)	t ₁₀	<i>t</i> ₄	<i>t</i> ₃	t_2	t_1	$\chi_8^2 - 8$	$\chi_4^2 - 4$
.00	.177	.175	.166	.169	.158	.117	.180	.186
.05	.223	.220	.214	.214	.205	.165	.226	.232
.10	.269	.266	.2 61	.260	.249	.210	.269	.278
.15	.313	.311	.306	.306	.294	.257	.314	.322
.20	.358	.356	.351	.350	.339	.303	.359	.366
.25	.402	.400	.396	.393	.384	.348	.404	.409
.3 0	.445	.445	.439	.436	.428	.394	.447	.452
.3 5	.488	.487	.483	.480	.472	.439	.489	.495
.40	.531	.531	<i>-</i> 527	<i>5</i> 23	.514	.484	.531	.535
.45	.573	<i>.57</i> 3	<i>.</i> 570	.5 65	.556	.529	.574	.57 7
.5 0	.614	.615	.612	.607	.599	.575	.615	.618
.5 5	.655	.655	.653	.649	.642	.619	.655	.658
.60	.695	.696	.693	.68 9	.683	.661	.696	.697
.65	.734	.735	.733	.728	.723	.705	.734	.737
.70	.773	.774	.770	.766	.763	.747	.772	.774
.75	.810	.811	.807	.805	.802	.789	.810	.812
.80	.846	.847	.844	.842	.840	.828	.846	.84 8
.85	.880	.880	.880	.877	.875	.868	.880	.881
.90	.912	.912	.912	.910	.909	.907	.913	.912
.93	.929	.929	.929	.928	.928	.932	.930	.929
.95	.940	.941	.939	.939	.93 9	.950	.940	.939
.97	.94 9	.949	.948	.949	.949	.966	.949	.94 8
.9 9	.955	.955	.956	.955	.955	.964	.954	.953
.9 95	.956	.95 6	.957	.956	.955	.957	.955	.954
1.00	.956	.956	.957	.956	.955	.950	.955	.955

TABLE 6 DOWNWARD MEDIAN-BIAS OF THE LEAST SQUARES ESTIMATOR OF α FOR SAMPLE SIZE 60

α	0	.3	.5	.7	.8	.9	.95	.99	.999	1.0
Model 1	0	.005	.008	.011	.013	.014	.013	.009	.002	_
Model 2	.02	.03	.03	.04	.05	.05	.06	.07	.07	.07
Model 3	.03	.05	.06	.08	.08	.10	.12	.14	.15	.15

TABLE 7 $90\% \ RANGE \ OF \ THE \ LEAST \ SQUARES \ ESTIMATOR \ OF \ \alpha \ FOR \ SAMPLE \ SIZE \ 60^{2}$

α	0	.3	.5	.7	.8	.9	.95	.99	.999	1.0
Model 1	.42	.40	.37	.32	.27	.24	.17	.12	.07	_
Model 2	.42	.41	.39	.34	.30	.26	.24	.23	.22	.22
Model 3	.42	.41	.39	.35	.33	.30	.29	.29	.29	.29

^aThe 90% range of an estimator is the length of the interval bounded by the estimator's .05 and .95 quantiles.

TABLE 8

REAL EXCHANGE RATE SERIES^a

Data Series	Estimator	α	IR(3)	IR(6)	IR(12)	IR(24)	IR(36)	CIR
FR/US	LS	.982	.947	.897	.804	.647	.520	55.6
	Median-	1.00	1.00	1.00	1.00	1.00	1.00	∞
	Unbiased	[.97,1.0]	[.91,1.0]	[.83,1.0]	[.69,1.0]	[.48,1.0]	[.33,1.0]	[33.3,∞]
WG/US	LS	.982	.947	.897	.804	.647	.520	55.6
	Median-	1.00	1.00	1.00	1.00	1.00	1.00	∞
	Unbiased	[.97,1.0]	[.91,1.0]	[.83,1.0]	[.69,1.0]	[.48,1.0]	[.33,1.0]	[33.3,∞]
JP/US	LS	.983	.950	.902	.814	.663	.539	58.8
	Median-	1.00	1.00	1.00	1.00	1.00	1.00	∞
	Unbiased	[.97,1.0]	[.91,1.0]	[.83,1.0]	[.69,1.0]	[.48,1.0]	[.33,1.0]	[33.3,∞]
CA/US	LS	.983	.950	.902	.814	.663	.539	58.8
	Median-	1.00	1.00	1.00	1.00	1.00	1.00	∞
	Unbiased	[.97,1.0]	[.91,1.0]	[.83,1.0]	[.69,1.0]	[.48,1.0]	[.33,1.0]	[33.3,∞]
UK/US	LS	.973	.921	.849	.720	.518	.373	37.0
	Median-	.995	.985	.970	.942	.887	.835	200
	Unbiased	[.96,1.0]	[.88,1.0]	[.78,1.0]	[.61,1.0]	[.38,1.0]	[.23,1.0]	[25,∞]
NL/US	LS	.981	.944	.891	.794	.631	.501	52.6
	Median-	1.00	1.00	1.00	1.00	1.00	1.00	∞
	Unbiased	[.97,1.0]	[.91,1.0]	[.83,1.0]	[.69,1.0]	[.48,1.0]	[.33,1.0]	[33.3,∞]
FR/WG	LS	.950	.857	.735	.540	.292	.158	20.0
	Median-	.968	.907	.823	.677	.458	.310	31.3
	Unbiased	[.93,1.0]	[.80,1.0]	[.65,1.0]	[.42,1.0]	[.18,1.0]	[.07,1.0]	[14.3,∞]
NL/WG	LS	.947	.849	.721	.520	.271	.141	18.9
	Median-	.966	.901	.813	.660	.436	.288	29.4
	Unbiased	[.92,1.0]	[.78,1.0]	[.61,1.0]	[.37,1.0]	[.14,1.0]	[.05,1.0]	[12.5,∞]

^aThe results of this table are based on Model 2 of Section 2. IR(h) denotes the impulse response function at time horizon h. CIR denotes the cumulative impulse response. The entries in the rows labeled LS are the estimates of α , IR(h), and CIR obtained using the least squares estimates of α from the regression equation (2.3) Model 2. The entries in the rows labeled median-unbiased are the median-unbiased estimates $\hat{\alpha}_{Uj}$, IR(h), and CÎR of α , IR(h), and CIR, respectively, defined in Sections 3 and 5. The intervals in square brackets below the median-unbiased estimates are the 90% central (i.e., $p_1 = p_2 = .05$) confidence intervals introduced in Sections 4 and 5.

The median-unbiased estimates and exact confidence intervals given in this table were determined using quantiles generated by simulation (with 10,000 repetitions) for the exact sample size of each series. Interpolation of the quantiles given in Table 3 for sample sizes 150 and 200 produces nearly the same results.

TABLE 9
VELOCITY AND INDUSTRIAL PRODUCTION SERIES^a

Data Series	Estimator	α	IR(2)	IR(4)	IR(8)	IR(16)	IR(32)	CIR
Velocity	LS Median- Unbiased	.94 1.0 [.93,1.0]	.88 1.0 [.86,1.0]	.78 1.0 [.75,1.0]	.61 1.0 [.56,1.0]	.37 1.0 [.31,1.0]	.14 1.0 [.10,1.0]	16.7 ∞ [14.3,∞]
Industrial Produc- tion	LS Median- Unbiased	.84 .89 [.79,1.0]	.71 .79 [.62,1.0]	.50 .63 [.39,1.0]	.25 .39 [.15,1.0]	.06 .15 [.02,1.0]	.004 .024 [.00,1.0]	6.3 9.1 [4.8,∞]
		μ	β	σ^2	Durbin- Watson	Skewness	Excess Kurtosis	
Velocity	LS Median- Unbiased	.052 (.035) 033 (.014)	00032 (.00049) .00040 (.00023)	.0045	1.78 1.82	-,49 -35	.35	
Industrial Produc- tion	LS Median - Unbiased	.043 (.019) .044 (.019)	.0066 (.0022) .0045 (.00029)	.0096	1.79 1.86	-1.01 96	1.62 1.33	

^aThe top half of Table 9 is analogous to Table 8 except that Model 3 equation (2.3) is estimated rather than Model 2 equation (2.3). The bottom half of Table 9 provides additional information. The entries in the rows labeled LS give the LS estimates of the parameters $\tilde{\mu}$, $\tilde{\beta}$, and σ^2 (with standard error estimates in parentheses) as well as the Durbin-Watson statistic and the sample coefficients of skewness and excess kurtosis from the regression of Y_t on $(1, t, Y_{t-1})$. The rows labeled median-unbiased report the corresponding statistics from the regression with α restricted to equal its median-unbiased estimate, i.e., the regression of $Y_t - \hat{\alpha}_{U3}Y_{t-1}$ on (1, t).

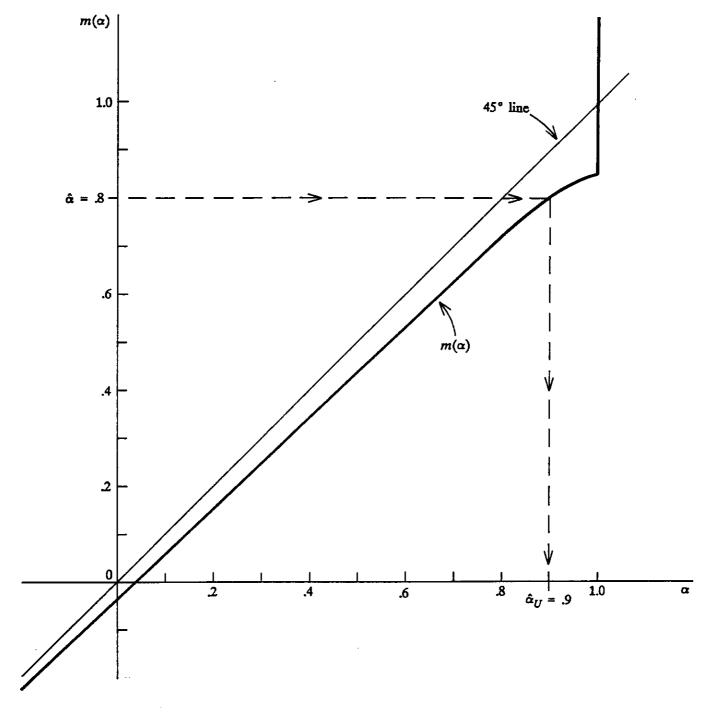


FIGURE 1 — Calculation of the Median-Unbiased Estimate $\hat{\alpha}_U$ Given the Estimate $\hat{\alpha}$ for Model 3 and Sample Size 60

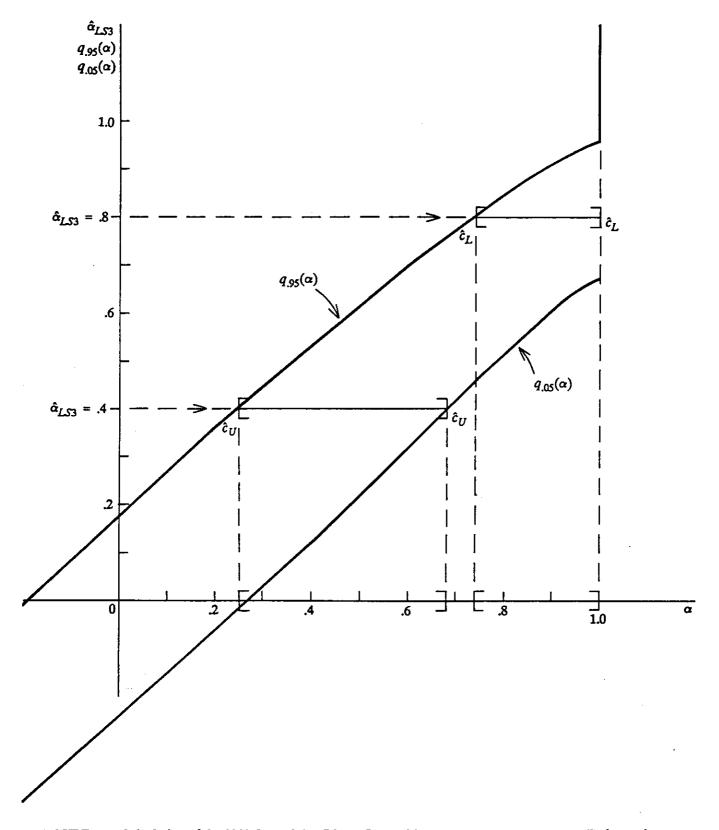


FIGURE 2 — Calculation of the 90% Central Confidence Interval $[\hat{c}_L, \hat{c}_U]$ Given a Least Squares Estimate \hat{a}_{LS3} for Model 3 and Sample Size 60