TECHNICAL WORKING PAPER SERIES

ASYMPTOTICALLY MEDIAN UNBIASED ESTIMATION OF COEFFICIENT VARIANCE IN A TIME VARYING PARAMETER MODEL

James H. Stock
Mark W. Watson

Technical Working Paper 201

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1996

This research was supported in part by National Science Foundation grant no. SBR-9409629. This paper is part of NBER's research program in Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Working Paper 201
August 1996

ASYMPTOTICALLY MEDIAN UNBIASED ESTIMATION OF COEFFICIENT VARIANCE IN A TIME VARYING PARAMETER MODEL

ABSTRACT

This paper considers the estimation of the variance of coefficients in time varying parameter models with stationary regressors. The maximum likelihood estimator has large point mass at zero. We therefore develop asymptotically median unbiased estimators and confidence intervals by inverting median functions of regression-based parameter stability test statistics, computed under the constant-parameter null. These estimators have good asymptotic relative efficiencies for small to moderate amounts of parameter variability. We apply these results to an unobserved components model of trend growth in postwar U.S. GDP: the MLE implies that there has been no change in the trend rate, while the upper range of the median-unbiased point estimates imply that the annual trend growth rate has fallen by 0.7 percentage points over the postwar period.

James H. Stock
Kennedy School of Government
Harvard University
Cambridge, MA 02138
and NBER
stock@ksg1.harvard.edu

Mark W. Watson
Woodrow Wilson School
Princeton University
Princeton, NJ 08544
and NBER
mwatson@princeton.edu

1. Introduction

Since its introduction in the early 1970's by Cooley and Prescott (1973a, 1973b, 1976), Rosenberg (1972, 1973), and Sarris (1973), the time-varying parameter (TVP), or "stochastic coefficients" regression model has been used extensively in empirical work, especially in forecasting applications; Chow (1984), Pagan (1980), Nichols and Pagan (1985), Engle and Watson (1987), Harvey (1989), and Stock and Watson (1996) provide references and discussion of the model. The appeal of the TVP model is that, by permitting the coefficients to evolve stochastically over time, it can be applied to models with parameter instability.

The TVP model considered in this paper is,

(1) $$y_t = \beta_t' X_t + u_t$$

(2) $$\beta_t = \beta_{t-1} + v_t$$

(3) $$a(L)u_t = \epsilon_t$$

(4) $$v_t = \tau \nu_t, \text{ where } \nu_t = B(L)\eta_t$$

where (y_t, X_t) are observed, X_t is an exogenous k-dimensional regressor, β_t is a $k \times 1$ vector of unobserved time-varying coefficients, τ is a scalar, a(L) is a scalar lag polynomial, B(L) is a $k \times k$ matrix lag polynomial, and ϵ_t and η_t are serially and mutually uncorrelated random disturbances. (Additional technical conditions used for the asymptotic results are given in section 2.) In empirical work it is common to use $a(L)=1$ and $B(L)=I_k$ (the $k \times k$ identity matrix). Although our results are developed for more general a(L) and B(L), this leading special case will be discussed separately below. Define Ω to be 2π times the spectral density matrix of ν_t at frequency zero, that is, $\Omega = B(1)E\eta_t\eta_t'B(1)'$. We treat Ω as fixed and chosen by the researcher, although as will be discussed the analysis carries through for cases in which Ω is consistently estimable as long as τ is identified. An important special case of this model is when $X_t=1$ and $B(L)=1$; following Harvey (1985), we refer to this case as the "local level" unobserved components model.

We consider the problem of estimation of the scale parameter τ. If (as is common) ϵ_t and η_t are assumed to be jointly normal and independent of $\{X_t\}$, then the parameters of (1)-(4) can be estimated by maximum likelihood implemented by the Kalman filter. However, the maximum likelihood estimator (MLE) has the undesirable property that if τ is small, it has point mass at zero. In the case $X_t=1$, this is related to the so-called pile-up problem that is known to occur in the first order moving average (MA(1)) model with a unit root, cf. Sargan and Bhargava (1983) and Shephard and Harvey (1990). In the general TVP model (1)-(4), the pile-up probability depends on the properties of X_t and can be large. The pile-up probability is a particular problem when τ is small and thus is readily mistaken for zero. Arguably, small values of τ are appropriate for many empirical applications; indeed, if τ is large, then the distribution of the MLE can be approximated by conventional $T^{1/2}$-asymptotic normality, but Monte Carlo evidence suggests that often this approximation is poor in cases of empirical interest (see Shephard (1993) and Davis and Dunsmuir [1996] for discussions in the case $X_t=1$.).

We therefore focus on the estimation of τ when it is small. In particular, we consider the nesting,

$$\tau = \lambda/T. \tag{5}$$

Order of magnitude calculations suggest that this might be an appropriate nesting for certain empirical problems of interest, such as estimating stochastic variation in the trend component in the logarithm of U.S. real gross domestic product (GDP), as is discussed in the context of the empirical application in section 4. This is also the nesting used to obtain local asymptotic power functions of tests of $\tau=0$; this suggests that if the researcher is in a region in which tests might yield ambiguous conclusions about the null hypothesis $\tau=0$, the nesting (5) is appropriate.

The main contribution of this paper is to develop a family of asymptotically median unbiased estimators of λ in the model (1)-(5). These estimators are obtained by inverting the asymptotic median function of statistics that test the hypothesis $\lambda=0$. The test statistics are based on generalized least squares (GLS) residuals which are readily computed under the null. These median-unbiased estimators, and the supporting asymptotic theory, are presented in section 2. As part of these calculations, we obtain asymptotic representations for a family of tests under the local alternative (5); these can be used to compute their local asymptotic power functions. Construction of asymptotically valid confidence intervals for λ is also discussed in section 2.

In section 3, numerical results are provided for the special case of the univariate local level model. Properties of the median unbiased estimators are compared to two maximum likelihood estimators, which alternatively maximize the marginal and the profile (or concentrated) likelihoods; these MLE's differ in their treatment of the initial value for β_t. Both MLE's are biased and have large pile-ups at $\hat{\lambda}=0$. When λ is small, the median unbiased estimators are more tightly concentrated around the true value of λ than either MLE.

Section 4 presents an application to the estimation of a long-run stochastic trend for U.S. postwar real per capita GDP. Point estimates from the median unbiased estimators suggest slowdown in the average trend rate of growth; the largest point estimate suggests a slowdown of .75% per annum over the postwar period. The MLE's suggest a much smaller decline, with point estimates ranging from 0 to .17%. Section 5 concludes.

2. Theoretical Results

We assume that a(L) has known finite order p and therefore consider statistics based on feasible GLS. Specifically, (i) y_t is regressed on X_t by OLS, producing residuals $\hat{u}_t$; (ii) a univariate AR(p) is estimated by OLS regression of $\hat{u}_t$ on $(1, \hat{u}_{t-1},\ldots,\hat{u}_{t-p})$, yielding $\hat{a}(L)$; and (iii) $\tilde{y}_t \equiv \hat{a}(L)y_t$ is regressed on $\tilde{X}_t \equiv \hat{a}(L)X_t$, yielding the GLS estimator $\tilde{\beta} = (T^{-1}\sum_{t=1}^{T}\tilde{X}_t\tilde{X}_t')^{-1}\, T^{-1}\sum_{t=1}^{T}\tilde{X}_t\tilde{y}_t$, residuals $\tilde{e}_t$ and covariance matrix $\tilde{V}$:

(6) $$\tilde{e}_t = \tilde{y}_t - \tilde{\beta}'\tilde{X}_t,$$

(7) $$\tilde{V} = (T^{-1}\sum_{t=1}^{T}\tilde{X}_t\tilde{X}_t')^{-1}\tilde{\sigma}_\epsilon^2$$

where $\tilde{\sigma}_\epsilon^2 = (T-k)^{-1}\sum_{t=1}^{T}\tilde{e}_t^2$. If $a(L)=1$, steps (i) and (ii) are omitted and the OLS and GLS regressions of y_t on X_t are equivalent.

Two test statistics are considered: Nyblom's (1989) L_T statistic (modified to use GLS residuals) and the sequential GLS Chow F-statistics, $F_T(s)$ $(0 \le s \le 1)$ which test for a break at date [Ts], where [•] denotes the greatest lesser integer. Let SSR_{t_1,t_2} denote the sum of squared residuals from the GLS regression of $\tilde{y}_t$ onto $\tilde{X}_t$ over observations $t_1 \le t \le t_2$, and let $\xi_T(s) = T^{-1/2}\sum_{t=1}^{[Ts]}\tilde{X}_t\tilde{e}_t$. The L_T and F_T statistics are,

(8) $$L_T = T^{-1}\sum_{t=1}^{T}\xi_T(t/T)'\tilde{V}^{-1}\xi_T(t/T)$$

(9) $$F_T(t/T) = (SSR_{1,T} - SSR_{1,[Ts]} - SSR_{[Ts]+1,T})/[kSSR_{1,T}/(T-k)].$$

For other tests in versions of this model, see Franzini and Harvey (1983), King and Hillier (1985), Nabeya and Tanaka (1988), Nyblom (1989), Reinsel and Tam (1996), and Shively (1988).

The F_T statistic is an empirical process and inference is performed using one-dimensional functionals of F_T. We consider three such functionals: the maximum F_T statistic (the Quandt (1960) likelihood ratio statistic), $QLR_T = \sup_{s \in (s_0, s_1)} F_T(s)$; the mean Wald statistic of Hansen (1992) and Andrews and Ploberger (1994), $MW_T = \int_{s_0}^{s_1} F_T(r)dr$; and the Andrews-Ploberger (1994) exponential Wald statistic, $EW_T = \ln\{ \int_{s_0}^{s_1} \exp(\frac{1}{2}F_T(r))dr\}$, where $0 < s_0 < s_1 < 1$.

Three assumptions are used to obtain the asymptotic results.

Assumption A. X_t is second order stationary, m-dependent, and $\sup_t EX_{it}^8 < \infty$, $i = 1, \ldots, k$.

Assumption B. $\{X_t\}$ is independent of $\{u_t, \nu_t\}$.

Assumption C. $(\epsilon_t', \eta_t')'$ is a $(k+1) \times 1$ vector of i.i.d. errors with three moments; ϵ_t and η_t are independent; a(L) has finite order p; and B(L) is one-summable with $B(1) \neq 0$.

Assumption A requires that X_t have bounded moments or, if nonstochastic, that it not exhibit a trend. The assumptions of m-dependence, second order stationarity and eight moments are made for convenience in the proofs, and presumably could be relaxed. However, the requirement that X_t not be integrated of order 1 (I(1)) or higher is essential for these results.

Assumption B requires that X_t is strictly exogenous. This assumption permits estimation of (1), under the null $\beta_t = \beta_0$, by GLS.

The assumption that a(L) has finite order p in assumption C is made to simplify estimation by feasible GLS. The assumption that ϵ_t and η_t are independent ensures that u_t and v_t have a zero cross spectral density matrix. This is a basic identifying assumption of the TVP model, cf. Harvey (1989). To construct the MLE, it is further assumed that ϵ_t and η_t are independent i.i.d. normal random variables.

The assumption that X_t is independent of the errors can be unappealing in a time series context; in particular, if the error u_t is serially uncorrelated, under constant coefficients then the regression of y_t on X_t would be valid with regressors that are predetermined but not strictly exogenous. This would allow feedback from u_t to future X_t. As an alternative to assumptions B and C, we therefore provide an assumption which permits predetermined regressors:

Assumption D. $(\epsilon_t', \eta_t')'$ is a $(k+1)\times 1$ vector of i.i.d. errors with three moments; ϵ_t and η_t are independent; $a(L)=1$; $B(L)$ is one-summable with $B(1)\neq 0$; η_t is independent of $\{X_t, X_{t\pm 1}, X_{t\pm 2},\ldots\}$; and u_t is independent of $\{X_t, X_{t-1}, X_{t-2},\ldots\}$.

This permits feedback from u_t to X_t, but not from v_t to X_t, and thus rules out X_t containing lagged y_t when $\lambda\neq 0$.

Our main theoretical results are given in the following theorem. Let "=>" denote weak convergence on D[0,1], let W_1 and W_2 be independent standard Brownian motions on $[0,1]^k$, and define $\Gamma=E\{[a(L)X_t][a(L)X_t]'\}$.

Theorem 1. Let y_t obey (1)-(5), and suppose either that assumptions A, B and C hold or that assumptions A and D hold. Then:

(a) $\tilde{V}^{-1/2}\xi_T \Rightarrow h_\lambda^0$, where $h_\lambda^0(s) = h_\lambda(s) - sh_\lambda(1)$, where $h_\lambda(s) = W_1(s)+\lambda D\int_0^s W_2(s)ds$, where $D=\Gamma^{-1/2}\Omega^{1/2}/\sigma_\epsilon$.

(b) $L_T \Rightarrow \int_0^1 h_\lambda^0(s)'h_\lambda^0(s)ds$.

(c) $F_T \Rightarrow F^*$, where $F^*(s) = h_\lambda^0(s)'h_\lambda^0(s)/(ks(1-s))$.

The proof is given in the appendix.

Limiting representations of the QLR, mean Wald, and exponential Wald statistics are obtained from part (c) of theorem 1 and the continuous mapping theorem. Specifically, $QLR_T \Rightarrow \sup_{s_0 \leq s \leq s_1} F^*(s)$, $MW_T \Rightarrow \int_{s_0}^{s_1} F^*(r)dr$, and $EW_T \Rightarrow \ln\{\int_{s_0}^{s_1} \exp(\frac{1}{2}F^*(r))dr\}$. Note also that the limiting representation for L_T can be written, $L_T \Rightarrow k \int_0^1 (r(1-r))F^*(r)dr$.

When $\lambda=0$, the process h_λ^0 is a k-dimensional Brownian bridge and the representations for the statistics L_T and F_T reduce to their well known null representations as functionals of a Brownian bridge (cf. Nabeya and Tanaka [1988], Nyblom [1989], Andrews and Ploberger [1994], Stock [1994]).

The limiting distribution of L_T and F_T evidently depend on λ and D. As mentioned in the introduction, Ω is a modeling parameter that is typically chosen by the researcher and thus is known, or is a consistently estimable function of other parameters. Although Γ and σ_ϵ^2 are unknown, they are consistently estimable. Thus D typically can be estimated consistently. Although it enters the representations in theorem 1 as a nuisance parameter, because it is consistently estimable, D can be treated as known for purposes of the asymptotic theory. In this sense, the only parameter which is not known asymptotically in these limiting representations is λ.

We will use these limiting representations for three related purposes: computation of local asymptotic power functions; construction of median unbiased estimators of λ; and construction of asymptotically valid confidence intervals for λ.

Local Asymptotic Power. The representations can be used to compute the distribution of the tests under the local alternative (5) and thus to compute the local asymptotic power of tests of the null $\tau=0$. The various test statistics have limiting representations under the local alternative that are qualitatively similar. This is interesting because the F_T-based statistics are typically motivated by considering the single break model, while the Nyblom (1989) derived the L_T statistic as the LMPI test statistic for the seemingly rather different Gaussian TVP model.

Median Unbiased Estimation of λ. Median unbiased estimators of λ can be computed from L_T or from a scalar functional of F_T. Consider for example the scalar functional $g(F_T)$, which is assumed to be continuous. By the continuous mapping theorem, $g(F_T) => g(F^*)$, the distribution of which depends on λ and D. Let $m_D(\lambda)$ denote the median of $g(F^*)$ as a function of λ for a given matrix of nuisance parameters D. Suppose that $m_D(\bullet)$ is monotone increasing and continuous in λ. Then the inverse function m_D^{-1} exists, and for D known, λ can be estimated by,

$$(9) \qquad \hat{\lambda}_g = m_D^{-1}(g(F_T)).$$

Asymptotically, $\hat{\lambda}_g => m_D^{-1}(g(F^*))$. By construction, $\Pr[\hat{\lambda}_g < \lambda] \to \Pr[m_D^{-1}(g(F^*)) < \lambda] = \Pr[g(F^*) < m_D(\lambda)] = 0.5$, so $\hat{\lambda}_g$ is asymptotically median unbiased.

In practice D is not known, so the estimator (9) is infeasible. However, as discussed above, in general D can be consistently estimated for a given choice of Ω. If in addition $m_D^{-1}(\bullet)$ is continuous in D (which it is for the functionals discussed in this paper), then (9) can be computed with D replaced by a consistent estimator $\hat{D}$ and the same asymptotic distribution obtains. Note however that this is computationally cumbersome for it requires computing the inverse median function $m_{\hat{D}}^{-1}(\bullet)$ for every estimate $\hat{D}$ under consideration.

The conventional approach to this nuisance parameter problem is simpler, and entails choosing Ω so that $D=I_k$, that is, choosing $\Omega=\sigma_\epsilon^2\Gamma^{-1}$ (cf. Nyblom [1989]). In this case, the limiting distributions of L_T and F_T depend only on λ and k under the local alternative. When $X_t=1$, setting $\Omega=\sigma_\epsilon^2\Gamma^{-1}$ amounts to a normalizing assumption that is made without loss of generality. For general X_t, this assumption has a certain appeal, for it insures that the scale of the innovations to β matches the scale of the GLS transformed regressors X_t. Whether this assumption is desirable for general X_t is a matter of modeling strategy in a particular empirical application.

It would be of interest to obtain theoretical results comparing the efficiency of median-unbiased estimators based on the various functionals of F_T. However, the limiting distributions are nonstandard and do not appear to have any simple relation to each other. The efficiency comparisons are therefore undertaken numerically and are reported in the next section.

Confidence Intervals for λ. Suppose that $D = I_k$, in which case the local asymptotic representations in theorem 1 depend only on λ and k. For a given scalar test statistic, its representation can then be used to compute a family of asymptotic 5% critical values of tests of $\lambda = \lambda_0$ against a two-sided alternative, and these critical values can in turn be used to construct the set of λ_0 which are not rejected. This set constitutes a 95% confidence set for λ_0. This process of inverting the test statistic can be done graphically by the method of confidence belts or by interpolation of a lookup table. The details parallel those for construction of confidence intervals for autoregressive roots local to unity, cf. Stock (1991), and are omitted.

3. Numerical Results for the Univariate Local Level Model

In the univariate local level model, $X_t = 1$ and $B(L) = 1$, so that y_t is the sum of an I(0) component and an independent random walk, which under the parameterization (5) has a small disturbance variance. In this model, Δy_t follows a MA process, with largest moving average root $-(1-\lambda/T)^{-1} + o(T^{-1})$. In this section, we first compare numerically the power of the tests of section 2, plus some other tests previously proposed in the literature, against the local alternative. We then turn to an analysis of the properties of median unbiased estimators. All computations of asymptotic distributions are based on simulation of the limiting representations, with $T=500$ and 5,000 Monte Carlo replications.

3.1. Asymptotic Power of Tests

There has been a great deal of work on tests for $\lambda=0$ in the local levels model and of a unit moving average root in the related MA(1) model; references include Nyblom and Mäkeläinen (1983), Shively (1988), Tanaka (1990), Saikkonen and Luukonen (1993) (also see the review article by Stock (1994)). In addition to asymptotic powers based on theorem 1, as a basis for comparison we report the asymptotic power envelope and the asymptotic power of two point optimal invariant (POI) tests, the test which is POI against $\lambda=7$ (denoted POI(7)) and the POI(17) test (Shively (1988), Saikkonen and Luukkonen (1993)).

Asymptotic powers of various 5% tests are summarized in figure 1. Evidently, for small values of λ all tests effectively lie on the asymptotic power envelope. For more distant alternatives, the MW and L tests lose power and, to a lesser degree, so do the EW and QLR. The asymptotic power functions of the EW and QLR tests are essentially indistinguishable, consistent with findings elsewhere (Andrews, Lee and Ploberger [1996], Stock and Watson [1996]) that these tests perform similarly.

3.2. Estimators of λ

Each of the tests examined in figure 1 has a power function that depends only on λ and has a median function which is monotone and continuous in λ. Asymptotically median unbiased estimators of λ based on each of these tests can therefore be constructed as described in section 2. In addition, results are reported for two versions of the Gaussian maximum likelihood estimator that differ in their assumptions concerning the initial value of β_0. The first estimator, the maximum profile (or concentrated) likelihood estimator (MPLE), treats β_0 as an unknown nuisance parameter which is concentrated out of the likelihood. The second estimator, the maximum marginal likelihood estimator (MMLE), treats β_0 as a $N(\bar{\beta},\kappa)$ random variable that is independent of $\{u_t,v_t\}_{t=1}^{T}$, so that β_0 is integrated out of the likelihood. When $\kappa \to \infty$, this produces the "diffuse prior" likelihood function (see Shephard and Harvey (1990) and Shephard (1993)). The MMLE is equivalent, after reparameterization on a restricted parameter space, to the MA(1) MLE analyzed by Davis and Dunsmuir (1996), and their local-to-unity asymptotic results apply here.

Pile-up probabilities that λ is estimated to be exactly zero are reported in table 1. The mass of the median unbiased estimators at zero is similar for all estimators. The pile-up probability for the MPLE remains large as λ increases, both in absolute terms (it is above 50% for $\lambda \leq 6$) and relative to the median unbiased estimators. As pointed out in Shephard and Harvey (1990), the pile-up probability for the MMLE is smaller than for MPLE.

Cumulative distribution functions of the various estimators are plotted in figure 2 for $\lambda=5$. As expected, both MLE's are biased and median biased. 77% of the mass of the distribution of the MPLE is below the true value $\lambda=5$, and the median of MPLE is 0. MMLE performs better, with 64% of its mass below the $\lambda=5$ and a median bias of approximately -1. The cdf's of the median unbiased estimators are fairly similar to each other, but markedly different than the MLE. One evident cost of unbiasness is their longer right tail relative to the MLEs.

We compare the estimators by computing their asymptotic relative efficiencies (ARE's). Because the distributions are nonnormal and not proportional, conventional methods of computing ARE's do not apply. Instead, we measure the ARE ($ARE_{i,MMLE}$) of the ith estimator $\hat{\tau}_i$ relative to the MMLE, $\hat{\tau}_{MMLE}$, as the limit of the ratio of observations T_{MMLE}/T_i needed for $\Pr[\hat{\tau}_i \in \mathcal{T}(\tau);T_i] = \Pr[\hat{\tau}_{MMLE} \in \mathcal{T}(\tau);T_{MMLE}]$, where T_i and T_{MMLE} denote the number of observations used to compute $\hat{\lambda}_i$ and $\hat{\tau}_{MMLE}$. The ARE's reported here were for the sets $\mathcal{T}(\tau)=\{x: 0.5\tau \leq x \leq 1.5\tau\}$, so $\Pr[\hat{\tau}_i \in \mathcal{T}(\tau);T_i]=\Pr[|T_i\hat{\tau}_i-T_i\tau| \leq 0.5T_i\tau] \to p_i(T_i\tau)$, say, and similarly for $\hat{\tau}_{MMLE}$. Using (5), set $\lambda = T_{MMLE}\tau$; then $ARE_{i,MMLE} = \lim T_{MMLE}/T_i$ can be computed by solving, $p_i(\lambda/ARE_{i,MMLE}) = p_{MMLE}(\lambda)$. In general the ARE depends on λ.

Table 2 reports these ARE's for the MPLE and six median unbiased estimators for various values of λ; all ARE's are relative to the MMLE. For example, when $\lambda=4$, the ARE of the QLR-based median unbiased estimator, relative to the MMLE, is 1.02, which indicates that, in large samples, the MMLE requires 1.02 times as many observations as the QLR-based estimator to achieve the same probability of falling in the set $T(\tau)$. Evidently, MMLE dominates MPLE for all values of λ shown, and is considerably more efficient for small to moderate values of λ. In contrast, the median unbiased estimators perform slightly better than MMLE for small values of λ ($\lambda \leq 4$) and comparably for moderate values of λ ($5 \leq \lambda \leq 8$). Their performance deteriorates however for large values of λ ($\lambda > 10$).

One way to calibrate the magnitude of λ is to compare it to the asymptotic powers given in figure 1. When $\lambda=4$, the tests have rejection probabilities of approximately 25%; when $\lambda=7$, the rejection probabilities are approximately 50%. For $\lambda \geq 14$, the power exceeds 80%. As an empirical guideline, this suggests that the median unbiased estimators will be roughly as efficient as the MMLE when the results of stability tests are ambiguous; when there is substantial instability, the MMLE will be more efficient than the median unbiased estimators.

Table 3 is a lookup table that permits computing median unbiased estimates, given a value of the test statistic. The normalization used in table 3 is that $D=1$, and users of this lookup table must be sure to impose this normalization when using the resulting estimator of λ.

4. Application to Trend Growth of U.S. GDP

The question of whether there has been a decline in the long run U.S. GDP growth rate and whether there has been a recent increase in this trend growth rate is of considerable practical and policy interest. In this section we follow Harvey (1985) and use the local levels model to estimate trend GDP growth in the postwar United States.

The data used are real quarterly values of per capita GDP from 1947:II-1995:IV. The data from 1959:I-1995:IV are the GDP chain-weighted quantity index, quarterly, seasonally adjusted (Citibase Series GDPFC). The data from 1947:I-1958:IV are real GDP in 1987 dollars, seasonally adjusted (Citibase Series GDPQ, in releases prior to 1996) proportionally splicea to the GDP chain-weighted quantity index in 1959:I. These series were deflated by the civilian population (Citibase series P16). This GDP series was transformed to (approximate) percentage growth at an annual rate, GY_t, by setting $GY_t=400\Delta\ln$(real per capita GDP).

Following Harvey (1985), we consider a model in which the growth rate of real GDP is allowed to have a small random walk component; this admits the possibility of a persistent decline in mean GDP growth, consistent with the productivity slowdown. The specific model considered is,

(10) $$GY_t = \beta_t + u_t$$

(11) $$\Delta\beta_t = (\lambda/T)\eta_t$$

(12) $$a(L)u_t=\epsilon_t$$

where the order p=4 is used for a(L). (The results are insensitive to choice of the AR order or to substituting an ARMA(2,3) parameterization for a(L), the latter being consistent with Harvey's (1985) original unobserved components formulation.)

It is worth digressing to discuss the implications of this model for orders of integration and unit roots. If there is a random walk component in GY_t, (10)-(12) imply that the logarithm of real per capita GDP is I(2). This hypothesis is rejected by unit root tests in our data and in the literature more generally. However, for τ small, the model implies that ΔGY_t has a nearly unit MA root. Because it is well known that unit root tests have high rejection rates in the presence of large MA roots (Schwert [1989], Pantula [1991]), these rejections are consistent with the postulated model.

Test statistics, median unbiased estimates, and equal-tailed confidence intervals for λ and the standard deviation of $\Delta\beta$ are presented in table 4. None of the tests reject at the 10% level. Of course this could mean that the tests have insufficient power to detect a small but nonzero value of λ, and indeed the median-unbiased estimates are, with only one exception, nonzero. The median unbiased estimates of λ are all small, ranging from 0 (POI(17)) to 4.1 (L). These correspond to point estimates of τ, the standard deviation of $\Delta\beta_t$, ranging from 0.0 and 0.13%. This range of estimates is consistent with intuition. For example, a value of $\sigma_{\Delta\beta}=.1$ corresponds to a standard deviation of $\beta_{1995:IV}-\beta_{1947:II}$ of 1.4 percentage points.

Estimates of the model parameters are presented in table 5, panel a, for various values of λ: (i) MPLE and MMLE; (ii) the median-unbiased estimate based on the L (which is the upper range of the point estimates); (iv) the upper end of the 90% confidence interval for λ based on L (the upper range of 90% confidence intervals). $\hat{\lambda}_{MPLE}=0$, consistent with the large pile-up probability discussed in section 3. MMLE produces a small but non-zero estimate of τ equal to .04%, which corresponds to a point estimate of λ of 1.4. Estimates of parameters of the u_t process change little for this range of value of τ, however estimates of the initial value of the trend growth rate increase (as do their standard errors) as τ increases.

Estimates of the trend growth rates $\beta_{t|T}$ based on these models over various time spans (computed using the Kalman smoother) are given in the second panel of table 5, and these cumulated trend growth rates, which are estimates of the stochastic trend component of the logarithm of real per capita GDP, are plotted in figure 3. The estimate of trend GDP based on MPLE is, of course, a straight line that essentially connects the endpoints of the sample values of the data. In contrast, the other estimates reflect, to a varying degree, a slowdown in mean GDP growth over this period. The point estimate based on L implies a slowdown in the annual trend growth rate of approximately .7% from the first half to the second half of the postwar period. Finally, none of the methods detect any substantial increase in trend GDP growth over the 1990's so far, relative to the 1980's; indeed all of the point estimates suggest a modest decrease.

5. Discussion and Conclusions

The median unbiased estimators developed here provide empirical researchers with a device to circumvent the undesirable pile-up problem and bias of the MLE when coefficient variation is small. The L_T and F_T-based test statistics are easily computed using statistics from the GLS regression of y_t on X_t. Given these statistics, the median unbiased estimates can be obtained by interpolating the entries in a lookup table. Such a lookup table is provided here for the univariate local levels model (table 3), and lookup tables in higher dimensions for the normalization $D=I_k$ are available from the authors upon request.

In the special case of the univariate local levels model, we examined six asymptotically median unbiased estimators and two MLEs, and found considerable differences among them. The MLEs were badly biased, particularly MPLE. When the variance of the coefficients is small, the median unbiased estimators based on the QLR and POI(17) test statistics had good ARE's. Because no asymptotic theory for the POI tests in the TVP model appears to be available outside the case $X_t = 1$, and because even in the local level model the POI tests are somewhat cumbersome to compute, these results provide support for using the QLR-based median unbiased estimators in the general TVP model, at least when the coefficient variance is small.

Appendix

Before proving theorem 1, we state and prove two preliminary lemmas. Let $\hat{U}_{t-1} = (\hat{u}_{t-1},\ldots,\hat{u}_{t-p})'$, $A=(-a_1, -a_2,\ldots, -a_p)'$, and $\hat{A}=(\hat{U}_{-1}'\hat{U}_{-1})^{-1}(\hat{U}_{-1}'\hat{u})$ using the usual matrix notation.

Lemma A1. Under assumptions A-C, $T^{1/2}(\hat{A}-A) = O_p(1)$.

Proof

The result follows by showing $T^{1/2}(\hat{A}-\bar{A})\xrightarrow{p}0$, where $\bar{A}=(U_{-1}'U_{-1})^{-1}(U_{-1}'u)$, where $U_{t-1}=(u_{t-1},\ldots,u_{t-p})'$. After straightforward algebra, it is seen that this follows if $\mu_{1T}\xrightarrow{p}0$ and $\mu_{2T}\xrightarrow{p}0$, where μ_{1T} and μ_{2T} are matrices with (i,j) elements, $\mu_{1t,ij}=T^{-1/2}\sum_{t=1}^{T}u_{t-j}X_{t-i}'(\beta_{t-i}-\hat{\beta})$ and $\mu_{2t,ij}= T^{-1/2}\sum_{t=1}^{T}(\beta_{t-i}-\hat{\beta})'X_{t-i}X_{t-j}'(\beta_{t-j}-\hat{\beta})$. These limits follow using the Markov and Chebyschev inequalities and applying assumptions A-C, assuming $T^{1/2}(\hat{\beta}-\beta_0)=>O_p(1)$. An $O_p(1)$ limiting representation for $T^{1/2}(\hat{\beta}-\beta_0)$ can be obtained using the methods in the proof of theorem 1, but showing the $T^{1/2}$ rate (which is all that is required here) can be verified directly using Chebyschevs' inequality.

Lemma A2. Let z_t be a scalar random variable with $E(z_t|z_{t-m-1},z_{t-m-2},\ldots)=0$ for some $m<\infty$, and let w_t be either a scalar nonrandom sequence or a random variable which is independent of z_t. Further suppose that $\lim_{T\to\infty}\max_{1\leq t\leq T}E|z_t|^4\leq c$ and $\lim_{T\to\infty}\max_{1\leq t\leq T}E|w_t|^4\leq c$ for some $c<\infty$. Then $T^{-1}\sum_{t=1}^{[Ts]}z_tw_t \xrightarrow{p} 0$ uniformly in s.

Proof

For $\delta>0$,

$$\Pr[\sup_s|T^{-1}\sum_{t=1}^{[Ts]}z_tw_t|>\delta] \leq \delta^{-4}E\max_{1\leq r\leq T}(T^{-1}\sum_{t=1}^{r}z_tw_t)^4$$

$$\leq \delta^{-4} E \sum_{r=1}^{T} (T^{-1} \sum_{t=1}^{r} z_t w_t)^4$$
$$\leq \delta^{-4} T^{-3} \max_{1 \leq r \leq T} E(\sum_{t=1}^{r} z_t w_t)^4$$
$$\leq \delta^{-4} T^{-3} \max_{1 \leq r \leq T} \sum_{t_1=1}^{r} \cdots \sum_{t_4=1}^{r} E(z_{t_1} \cdots z_{t_4}) E(w_{t_1} \cdots w_{t_4})$$
$$\leq \delta^{-4} \max_t E|z_t|^4 \max_t E|w_t|^4 [T^{-2}(6m+1)^3 + 3T^{-1}(2m+1)^2],$$

where the final inequality follows from noting that $E(Z_{t_1} \cdots Z_{t_4}) \leq \max_t E|z_t|^4 \{1(|t_1-t_2| \leq 3m)$ $1(|t_1-t_3| \leq 3m) 1(|t_1-t_4| \leq 3m) + 3 1(|t_1-t_2| \leq m) 1(|t_3-t_4| \leq m)$ because of the m-dependence of the conditional mean of z_t. Because the fourth moments are bounded by assumption, the final term in the expression $\to 0$, proving uniform consistency.

Proof of Theorem 1.

Let $a(L) = \sum_{i=0}^{p} a_i L^i$ with $a_0 = 1$, and let $X_t^{\dagger} = a(L) X_t$. An implication of assumptions A, B and C, or alternatively of assumptions A and D, is that

(A1) $$(\sigma_\epsilon^{-1} \Gamma^{-\frac{1}{2}} T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} X_t^{\dagger} \epsilon_t, \Omega^{-\frac{1}{2}} T^{-\frac{1}{2}} \sum_{t=1}^{[Ts]} \nu_t) => (W_1, W_2),$$

where W_1 and W_2 are independent k-dimensional standard Brownian motions.

We first prove the theorem under assumptions A, B and C.

(a) Let $\tilde{u}_t = \hat{a}(L) u_t$ and $\tilde{\omega}_t = -\sum_{j=1}^{p} \hat{a}_j X_{t-j}' \sum_{i=0}^{j-1} v_{t-i}$, so that $\tilde{y}_t = \beta_0' \tilde{X}_t + (\sum_{r=1}^{t} v_r)' \tilde{X}_t + \tilde{\omega}_t + \tilde{u}_t$. Accordingly,

(A2) $$\xi_T(s) = \xi_{1T}(s) + \lambda \xi_{2T}(s) + \xi_{3T}(s) - \kappa_T(s) \{\xi_{1T}(1) + \lambda \xi_{2T}(1) + \xi_{3T}(1)\}$$

where

$$\xi_{1T}(s)=T^{-1/2}\sum_{t=1}^{[Ts]}\tilde{X}_t\tilde{u}_t$$
$$\xi_{2T}(s)=T^{-3/2}\sum_{t=1}^{[Ts]}\tilde{X}_t\tilde{X}_t'\sum_{r=1}^{t}\nu_r$$
$$\xi_{3T}(s)=T^{-1/2}\sum_{t=1}^{[Ts]}\tilde{X}_t\tilde{\omega}_t$$
$$\kappa_T(s)=[T^{-1}\sum_{t=1}^{[Ts]}\tilde{X}_t\tilde{X}_t'][T^{-1}\sum_{t=1}^{T}\tilde{X}_t\tilde{X}_t']^{-1}.$$

Limits are obtained for these terms in turn. All limits are uniform in the index s.

(i) Write $\xi_{1T}(s)=\Delta_{1T}(s)+\Delta_{2T}(s)+\xi_{1T}^{\dagger}(s)$, where $\Delta_{1T}(s)=\sum_{j=0}^{p}\hat{a}_j\sum_{i=1}^{p}T^{1/2}(\hat{a}_i-a_i)$ $(T^{-1}\sum_{t=1}^{[Ts]}X_{t-j}u_{t-i})$, $\Delta_{2T}(s)=T^{-1/2}\sum_{t=1}^{[Ts]}(\tilde{X}_t-X_t^{\dagger})\epsilon_t$, and $\xi_{1T}^{\dagger}(s)=T^{-1/2}\sum_{t=1}^{[Ts]}X_t^{\dagger}\epsilon_t$. Assumptions A, B and C imply that $T^{-1}\sum_{t=1}^{[Ts]}X_{t-j}u_{t-i}$ satisfies the conditions of lemma 2 with $z_t=X_{t-j}u_{t-i}$ and $w_t=1$ so, because p is fixed, $\Delta_{1T}\xrightarrow{p}0$. By an analogous argument, $\Delta_{2T}\xrightarrow{p}0$. Using the limit in (A1), we have $\xi_{1T}(s) => \sigma_\epsilon\Gamma^{1/2}W_1(s)$.

(ii) Write $\xi_{2T}(s)=\Delta_{3T}(s)+\Delta_{4T}(s)+\xi_{2T}^{\dagger}(s)$, where $\Delta_{3T}(s)=T^{-3/2}\sum_{t=1}^{[Ts]}(X_t^{\dagger}X_t^{\dagger\prime}-\Gamma)\sum_{r=1}^{t}\nu_r$, $\Delta_{4T}(s)=T^{-3/2}\sum_{t=1}^{[Ts]}(\tilde{X}_t\tilde{X}_t'-X_t^{\dagger}X_t^{\dagger\prime})\sum_{r=1}^{t}\nu_r$, and $\xi_{2T}^{\dagger}(s)=\Gamma T^{-3/2}\sum_{t=1}^{[Ts]}\sum_{r=1}^{t}\nu_r$. To show $\Delta_{3T}\xrightarrow{p}0$ and $\Delta_{4T}\xrightarrow{p}0$, consider for notational simplicity the case $k=1$ (the argument for $k>1$ is similar). Note that $T^{-3/2}\max_{t_1\cdots t_4}E|\sum_{r_1=1}^{t_1}\nu_{r_2}\cdots\sum_{r_4=1}^{t_4}\nu_{r_4}|\rightarrow$ $\sup_{s_1\cdots s_4}E|W_2(s_1)\cdots W_2(s_4)|/\Omega^2<\infty$. Because X_t has eight moments and is m-dependent, $X_t^{\dagger 2}-\Gamma$ has four moments, has mean zero, and is m-dependent. Thus Δ_{3T} satisfies the conditions of lemma 2 with $z_t=X_t^{\dagger 2}-\Gamma$ and $w_t=T^{-1/2}\sum_{r=1}^{t}\nu_r$, so $\Delta_{3T}\xrightarrow{p}0$. Turning to Δ_{4T}, $\Delta_{4T} = \sum_{j=0}^{p}\sum_{i=0}^{p}(\hat{a}_i+a_i)$ $T^{1/2}(\hat{a}_j-a_j)\Delta_{4T,ij}(s)$, where $\Delta_{4T,ij}(s)=[T^{-3/2}\sum_{t=1}^{[Ts]}X_{t-j}X_{t-i}T^{-1/2}\sum_{r=1}^{t}\nu_r]$. An argument analogous to that used for Δ_{3T} shows that $\Delta_{4T,ij}\xrightarrow{p}0$ and, because p is finite, $\Delta_{4T}\xrightarrow{p}0$. The limit of $\xi_{2T}^{\dagger}$ follows from (A1). Thus $\xi_{2T}(s) => \Gamma\,\Omega^{1/2}\int_0^s W_2(r)dr$.

(iii) Write $\xi_{3T}(s) = -\lambda \sum_{j=0}^{p} \sum_{\ell=0}^{p} \sum_{i=0}^{j-1} \xi_{3T,i\ell j}(s)$, where $\xi_{3T,i\ell j}(s) =$ $(T^{-3/2} \sum_{t=1}^{[Ts]} X_{t-\ell} X'_{t-j} v_{t-i})$. As before, consider the case k=1. Now, $T^{1/2}\xi_{3T,i\ell j}(s)$ satisfies lemma 1 with $z_t = X_{t-\ell}X'_{t-j}v_{t-i}$ and $w_t = 1$; thus $\xi_{3T} \xrightarrow{p} 0$.

(iv) Let $\Delta_{5T}(s) = T^{-1}\sum_{t=1}^{[Ts]}(\tilde{X}_t\tilde{X}_t' - X_t^\dagger X_t^{\dagger\prime})$ and $\Delta_{6T}(s) = T^{-1}\sum_{t=1}^{[Ts]}(X_t^\dagger X_t^{\dagger\prime} - \Gamma)$, and let $\Delta_{7T} = \Delta_{5T} + \Delta_{6T} = T^{-1}\sum_{t=1}^{[Ts]}(\tilde{X}_t\tilde{X}_t' - \Gamma)$. The argument that $\Delta_{5T} \xrightarrow{p} 0$ follows the argument that $\Delta_{3T} \xrightarrow{p} 0$ with $T^{-1/2}\sum_{r=1}^{t}\nu_r$ replaced by 1, and the argument that $\Delta_{6T} \xrightarrow{p} 0$ follows the argument that $\Delta_{4T} \xrightarrow{p} 0$ with the same replacement. Thus $\Delta_{7T} \xrightarrow{p} 0$, so $\kappa_T(s) \xrightarrow{p} sI_k$.

Similar calculations imply that $\tilde{\sigma}_\epsilon^2 \xrightarrow{p} \sigma_\epsilon^2$ so $\tilde{V} \xrightarrow{p} \Gamma^{-1}\sigma_\epsilon^2$. By collecting terms and using (A2), it follows that $\tilde{V}^{-1/2}\xi_T(s) => h_\lambda(s) - sh_\lambda(1)$, where $h_\lambda(s) = W_1(s) + \lambda(\Gamma^{-1/2}\Omega^{1/2}/\sigma_\epsilon)\int_0^s W_2(r)dr$.

(b) This follows from the continuous mapping theorem.

(c) This follows by straightforward but tedious manipulations using the previous limiting results.

Next turn to the proof under assumptions A and D. Under assumption D, $a(L) = \hat{a}(L) = 1$ so $\tilde{X}_t = X_t^\dagger = X_t$ and $\tilde{e}_t = y_t - \hat{\beta}'X_t$ (where $\hat{\beta}$ remains the OLS estimator). The proof under these conditions follows the proof above but is simpler. In particular, (A2) now holds with $\xi_{1T}(s) = T^{-1/2}\sum_{t=1}^{[Ts]}X_t\epsilon_t$, $\xi_{2T}(s) = T^{-3/2}\sum_{t=1}^{[Ts]}X_tX_t'\sum_{r=1}^{t}\nu_r$, $\xi_{3T}(s) = 0$, and $\kappa_T(s) =$ $[T^{-1}\sum_{t=1}^{[Ts]}X_tX_t'][T^{-1}\sum_{t=1}^{T}X_tX_t']^{-1}$. The limit of ξ_{1T} follows from (A1). Write $\xi_{2T}(s)$ as, $\xi_{2T}(s) = T^{-1}\sum_{t=1}^{[Ts]}(X_tX_t' - \Gamma)(T^{-1/2}\sum_{r=1}^{t}\nu_r) + \Gamma T^{-3/2}\sum_{t=1}^{[Ts]}\sum_{r=1}^{t}\nu_r$. The first term in this expression $\xrightarrow{p} 0$ as a consequence of Lemma A2 and the independence of $\{\nu_t\}$ and $\{X_t\}$, as discussed above for the term Δ_{3T}. The limit of the second term follows from (A1) and the continuous mapping theorem. The argument given above for $\kappa_T(s) \xrightarrow{p} sI_k$ applies under these assumptions, and $\tilde{V} \xrightarrow{p} \sigma_\epsilon^2\Gamma$. This proves part (a) under assumptions A and D; parts (b) and (c) follow accordingly.

References

Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383-1414.

Andrews, D.W.K., I. Lee and W. Ploberger (1996), "Optimal Changepoint Tests for Normal Linear Regression," *Journal of Econometrics*, 70, 9-38.

Chow, G.C. (1984), "Random and Changing Coefficient Models," in Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics*, North Holland, Amsterdam, Chapter 21.

Cooley, T.F. and E.C. Prescott (1973a), "An Adaptive Regression Model," *International Economic Review*, 14, pp. 364-371.

Cooley, T.F. and E.C. Prescott (1973b), "Tests of an Adaptive Regression Model," *Review of Economics and Statistics*, 55, pp. 248-256.

Cooley, T.F. and E.C. Prescott (1976), "Estimation in the Presence of Stochastic Parameter Variation," *Econometrica*, 44, pp. 167-184.

Davis, R.A. and W.T.M. Dunsmuir (1996), "Maximum Likelihood Estimation for MA(1) Processes with a Root On or Near the Unit Circle," *Econometric Theory*, Vol. 12, No. 1, pp. 1-29.

Engle, R.F. and M.W. Watson (1987), "The Kalman Filter Model: Applications to Forecasting and Rational Expectations Models," in *Advances in Econometrics: Fifth World Congress of the Econometric Society*, T. Bewley (ed.), Cambridge University Press, Cambridge, U.K.

Franzini, L. and A.C. Harvey (1983), "Testing for Deterministic Trend and Seasonal Components in Time Series Models," *Biometrika*, 70, 673-682.

Hansen, B.E. (1992), "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321-336.

Harvey, A.C. (1985), "Trends and Cycles in Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 3, 216-227.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models, and the Kalman Filter*, Cambridge: Cambridge University Press.

King, M.L. and G.H. Hillier (1985), "Locally Best Invariant Tests of the Error Covariance Matrix of the Linear Regression Model," *Journal of the Royal Statistical Society-B*, 47, 98-102.

Nabeya, S. and K. Tanaka (1988), "Asymptotic Theory of a Test for the Constancy of Regression Coefficients against the Random Walk Alternative," *Annals of Statistics*, 16, 218-35

Nichols, D.F. and A.R. Pagan, (1985), "Varying Coefficient Regression," in E.J. Hannan, P.R. Krishnaiah and M.M. Rao, eds., *Handbook of Statistics*, Vol. 5, North Holland, Amsterdam, pp. 413-449.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223-230.

Nyblom, J. and T. Mäkeläinen (1983), "Comparisons of Tests for the Presence of Random Walk Coefficients in a Simple Linear Model," *Journal of the American Statistical Association*, 78, 856-864.

Pagan, A. (1980), "Some Identification and Estimation Results for Regression Models with Stochastically Varying Coefficients," *Journal of Econometrics*, 12, pp. 341-363.

Pantula, S.G. (1991), "Asymptotic Distributions of the Unit-Root Tests when the Process is Nearly Stationary," *Journal of Business and Economic Statistics*, 9, 325-353.

Quandt, R.E. (1960), "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55, 324-330.

Reinsel, G.C. and W.-K. Tam, (1996), "Tests for Seasonal Moving Average Unit Root in ARIMA Models," manuscript, Department of Statistics, University of Wisconsin-Madison.

Rosenberg, B. (1972), "The Estimation of Stationary Stochastic Regression Parameters Re-Examined," *Journal of the American Statistical Association*, 67, pp. 650-654.

Rosenberg, B. (1973), "The Analysis of a Cross-Section of Time Series by Stochastically Convergent Parameter Regression," *Annals of Economic and Social Measurement*, 2, 461-484.

Saikkonen, P. and R. Luukonen (1993), "Point Optimal Test for Testing the Order of Differencing in ARIMA Models," *Econometric Theory*, Vol. 9, No. 3, pp. 343-62.

Sarris, A.H. (1973), "A Bayesian Approach to Estimation of Time Varying Regression Coefficients," *Annals of Economic and Social Measurement*, 2, 501-523.

Sargan, J.D. and A. Bhargava (1983), "Maximum Likelihood Estimation of Regression Models with First Order Moving Average Errors when the Root Lies on the Unit Circle," *Econometrica*, 51, 799-820.

Schwert, G.W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation," *Journal of Business and Economic Statistics* 7, 147-160.

Shephard, N.H. (1993), "Distribution of the ML Estimator of an MA(1) and a Local Level Model," *Econometric Theory* 9, 377-401.

Shephard, N.G. and A.C. Harvey (1990), "On the Probability of Estimating a Deterministic Component in the Local Level Model," *Journal of Time Series Analysis*, Vol. 11, No. 4, pp. 339-347.

Shively, T.S. (1988), "An Exact Test of a Stochastic Coefficient in a Time Series Regression Model," *Journal of Time Series Analysis*, Vol. 9, No. 1, 81-88.

Stock, J.H. (1991), "Confidence Intervals for the Largest Autoregressive Root in U.S. Economic Time Series," *Journal of Monetary Economics*, 28, 435-460.

Stock, J.H. (1994), "Unit Roots, Structural Breaks, and Trends," in R.F. Engle and D. McFadden (eds), *The Handbook of Econometrics, Vol. IV.* Amsterdam: Elsevier.

Stock, J.H. and M.W. Watson (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business and Economic Statistics*, Vol. 14, No. 1, pp. 11-30.

Tanaka, K. (1990), "Testing for a Moving Average Unit Root," *Econometric Theory*, Vol. 6, No. 4, pp. 433-44.

Table 1
Pile-up probability that $\hat{\lambda}=0$ for MLE's and various median-unbiased estimators

λ	MPLE	MMLE	L	MW	EW	QLR	POI(7)	POI(17)
0	0.96	0.66	0.50	0.50	0.50	0.50	0.50	0.50
1	0.91	0.60	0.47	0.47	0.47	0.46	0.47	0.47
2	0.88	0.57	0.42	0.42	0.42	0.43	0.44	0.43
3	0.81	0.47	0.34	0.34	0.34	0.35	0.35	0.37
4	0.72	0.40	0.28	0.28	0.29	0.29	0.29	0.30
5	0.65	0.35	0.24	0.24	0.24	0.24	0.24	0.26
6	0.56	0.28	0.19	0.19	0.19	0.19	0.18	0.20
7	0.48	0.24	0.15	0.16	0.16	0.16	0.14	0.15
8	0.42	0.19	0.13	0.13	0.13	0.13	0.12	0.13
9	0.37	0.17	0.11	0.12	0.12	0.12	0.09	0.10
10	0.30	0.13	0.09	0.09	0.09	0.09	0.07	0.07
12	0.24	0.09	0.06	0.07	0.07	0.06	0.05	0.05
14	0.15	0.06	0.03	0.04	0.04	0.04	0.03	0.02
16	0.13	0.04	0.03	0.03	0.03	0.03	0.01	0.01
18	0.09	0.03	0.02	0.03	0.02	0.02	0.01	0.01
20	0.07	0.02	0.01	0.01	0.01	0.01	0.01	0.01
25	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.01
30	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Notes: Enties for MPLE and MMLE for $\lambda=0$ are from Shepard and Harvey (1990). Entries for other values of λ are estimated using 5000 replications with $T=500$. To facilitate the computations, the likelihoods were computed on a discrete grid of 240 equally spaced values of $0\le\lambda\le60$ and the MLE's were computed by a search over this grid.

Table 2
Asymptotic relative efficiencies of median-unbiased estimators relative to the MMLE

λ	MPLE	L	MW	EW	QLR	POI(7)	POI(17)
1	0.13	1.00	1.00	1.00	1.00	1.07	1.07
2	0.19	1.09	1.07	1.07	0.96	1.07	1.02
3	0.52	1.08	1.10	1.08	1.04	1.02	0.94
4	0.62	1.06	1.06	1.06	1.02	1.06	1.00
5	0.65	0.93	0.93	0.98	0.97	1.11	1.14
6	0.71	0.94	0.92	0.99	1.00	1.03	1.08
7	0.76	0.79	0.79	0.85	0.88	0.96	1.04
8	0.77	0.77	0.77	0.85	0.85	0.91	0.98
9	0.75	0.69	0.69	0.75	0.77	0.86	0.89
10	0.80	0.65	0.65	0.71	0.74	0.80	0.80
12	0.67	0.56	0.56	0.65	0.67	0.67	0.67
14	0.57	0.50	0.49	0.57	0.57	0.57	0.57
16	0.50	0.42	0.42	0.49	0.50	0.50	0.50
18	0.44	0.38	0.38	0.44	0.44	0.44	0.44
20	0.40	0.33	0.33	0.40	0.40	0.40	0.40
25	0.32	0.28	0.28	0.32	0.32	0.32	0.32
30	0.27	0.22	0.23	0.27	0.27	0.27	0.27

Notes: The reported ARE's are the limiting ratio of the number of observations necessary for the MLE to achieve the same probability of being in the region $\tau \pm 0.5\tau$ as the candidate estimator, as a function of $\lambda = \tau/T$, as described in the text. ARE's exceeding 1 indicate greater efficiency than the MLE. Entries are estimates based on interpolating probabilities from the values of λ shown in column 1. These probabilities were estimated using 5000 replications and T=500 for each value of λ.

Table 3
Lookup table for constructing median-unbiased estimator of λ for various test statistics when X_t=1 and D=1

λ	L	MW	EW	QLR	POI7	POI17
0	0.118	0.689	0.426	3.198	2.693	7.757
1	0.127	0.757	0.476	3.416	2.740	7.825
2	0.137	0.806	0.516	3.594	2.957	8.218
3	0.169	1.015	0.661	4.106	3.301	8.713
4	0.205	1.234	0.826	4.848	3.786	9.473
5	0.266	1.632	1.111	5.689	4.426	10.354
6	0.327	2.018	1.419	6.682	4.961	11.196
7	0.387	2.390	1.762	7.626	5.951	12.650
8	0.490	3.081	2.355	9.160	6.689	13.839
9	0.593	3.699	2.910	10.660	7.699	15.335
10	0.670	4.222	3.413	11.841	8.849	16.920
11	0.768	4.776	3.868	13.098	10.487	19.201
12	0.908	5.767	4.925	15.451	11.598	20.570
13	1.036	6.586	5.684	17.094	13.007	22.944
14	1.214	7.703	6.670	19.423	14.554	24.962
15	1.360	8.683	7.690	21.682	16.153	27.135
16	1.471	9.467	8.477	23.342	18.073	30.030
17	1.576	10.101	9.191	24.920	19.563	32.209
18	1.799	11.639	10.693	28.174	21.662	35.426
19	2.016	13.039	12.024	30.736	24.160	38.465
20	2.127	13.900	13.089	33.313	25.479	40.583
21	2.327	15.214	14.440	36.109	27.687	44.104
22	2.569	16.806	16.191	39.673	30.260	47.239
23	2.785	18.330	17.332	41.955	32.645	50.881
24	2.899	19.020	18.699	45.056	35.011	54.426
25	3.108	20.562	20.464	48.647	37.481	58.172
26	3.278	21.837	21.667	50.983	39.907	60.842
27	3.652	24.350	23.851	55.514	41.146	63.561
28	3.910	26.248	25.538	59.278	43.212	66.782
29	4.015	27.089	26.762	61.311	47.135	71.577
30	4.120	27.758	27.874	64.016	50.134	76.343

Notes: Entries are the value of the test statistic, for which the value of λ given in the first column is the median-unbiased estimator. Care must be taken to impose the normalization D=1 when using these estimates of λ. Estimates of τ are computed as λ/T. If the test statistic takes on a value smaller than that in the first row, the median-unbiased estimate is zero. Estimates for other values of the test statistics can be obtained by interpolation. For example, suppose QLR=5.0 is obtained empirically; using linear interpolation, the median unbiased estimator of λ is 4+(5.0-4.848)/(5.689-4.848). A more accurate computer-based version of this (which handles general X_t for the case $D=I_k$) is available from the authors by request. All entries in the table were estimated using 5000 replications and T=500.

Table 4
Postwar U.S. GDP Growth, 1947:II-1995:IV:
Tests of $\tau=0$, median-unbiased estimates,
and 90% confidence intervals

Test	Statistic	(p-value)	$\hat{\lambda}$	(90% CI)	$\hat{\tau}$	(90% CI)
L	0.21	(0.25)	4.1	(0.00,19.4)	0.13	(0.00,0.62)
MW	1.16	(0.29)	3.4	(0.00,18.8)	0.11	(0.00,0.60)
EW	0.68	(0.32)	3.1	(0.00,17.0)	0.10	(0.00,0.55)
QLR	3.31	(0.48)	0.8	(0.00,13.3)	0.03	(0.00,0.41)
POI(7)	2.90	(0.45)	1.7	(0.00,12.9)	0.05	(0.00,0.37)
POI(17)	7.52	(0.54)	0.0	(0.00,11.3)	0.00	(0.00,0.36)

Notes: $\hat{\tau}$ is the estimate of the standard deviation of $\Delta\beta_t$ in (11), that is, $\hat{\tau}=T^{-1}\hat{\lambda}\hat{\sigma}_\epsilon/\hat{a}(1)$.

Table 5
Estimates of parameters in (10)-(12) for various values of λ and implied subsample trend growth rates

A. Parameter Estimates

Parameter	MPLE		MMLE		-- Estimates with fixed λ --			
τ	0.00	--	0.04	--	0.13	--	0.62	--
σ_ϵ	3.85	(0.20)	3.86	(0.20)	3.85	(0.20)	3.78	(0.20)
ρ_1	0.33	(0.24)	0.34	(0.07)	0.34	(0.07)	0.32	(0.08)
ρ_2	0.13	(0.19)	0.13	(0.08)	0.13	(0.08)	0.12	(0.08)
ρ_3	-0.01	(0.10)	-0.01	(0.02)	-0.01	(0.11)	-0.01	(0.08)
ρ_4	-0.09	(0.07)	-0.08	(0.07)	-0.09	(0.08)	-0.09	(0.07)
β_0	1.80	(0.44)	---		2.44	(0.90)	2.67	(1.89)

B. Estimated average trend growths

Date	$\overline{GY}$	MPLE	MMLE	τ=.13	τ=.62
1947-95	1.80	1.80	1.80	1.80	1.80
1947-70	2.46	1.80	1.89	2.16	2.43
1970-95	1.22	1.80	1.71	1.47	1.23
1950-60	2.75	1.80	1.91	2.25	2.27
1960-70	2.39	1.80	1.84	1.98	2.39
1970-80	1.20	1.80	1.75	1.56	1.07
1980-90	1.58	1.80	1.70	1.45	1.50
1990-95	0.62	1.80	1.68	1.36	1.04

Notes: Estimates were computed by maximum likelihood, with numerical standard errors computed from the inverse of the Hessian. Unrestricted MLE's (standard errors in parentheses) are reported in the first two columns. (Because of the nonnormal distribution of the MLE of λ, the standard error for τ is not reported.) The final two sets of columns report estimates by restricted MLE, with λ fixed to the indicated values. The column labeled $\overline{GY}$ in panel B is the sample mean of GY; the other entries are average values of $\beta_{t|T}$ over the indicated subsample for the indicated model, where $\beta_{t|T}$ are the estimates of β_t obtained from the Kalman smoother.

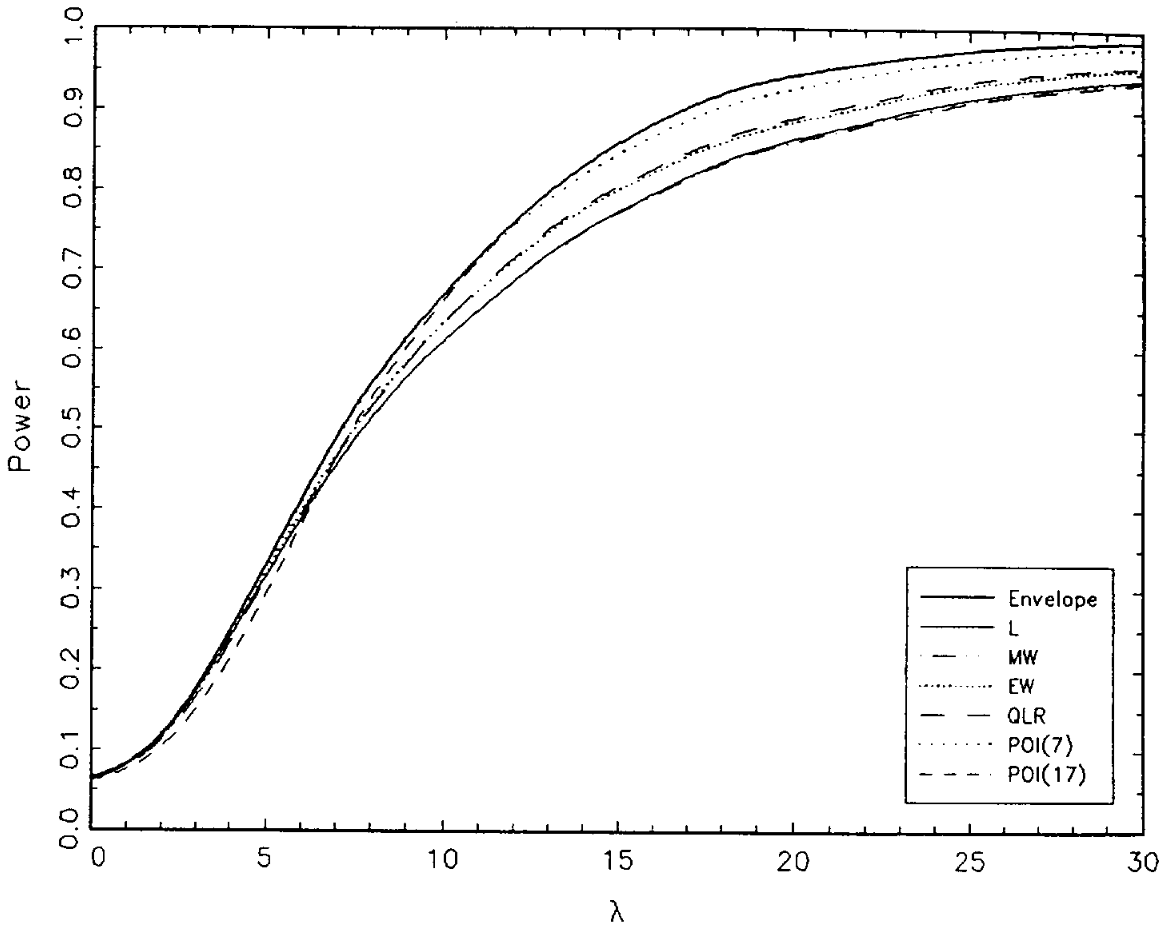

Figure 1.
Asymptotic power functions of 5% tests of $\tau=0$ against alternatives $\tau=\lambda/T$

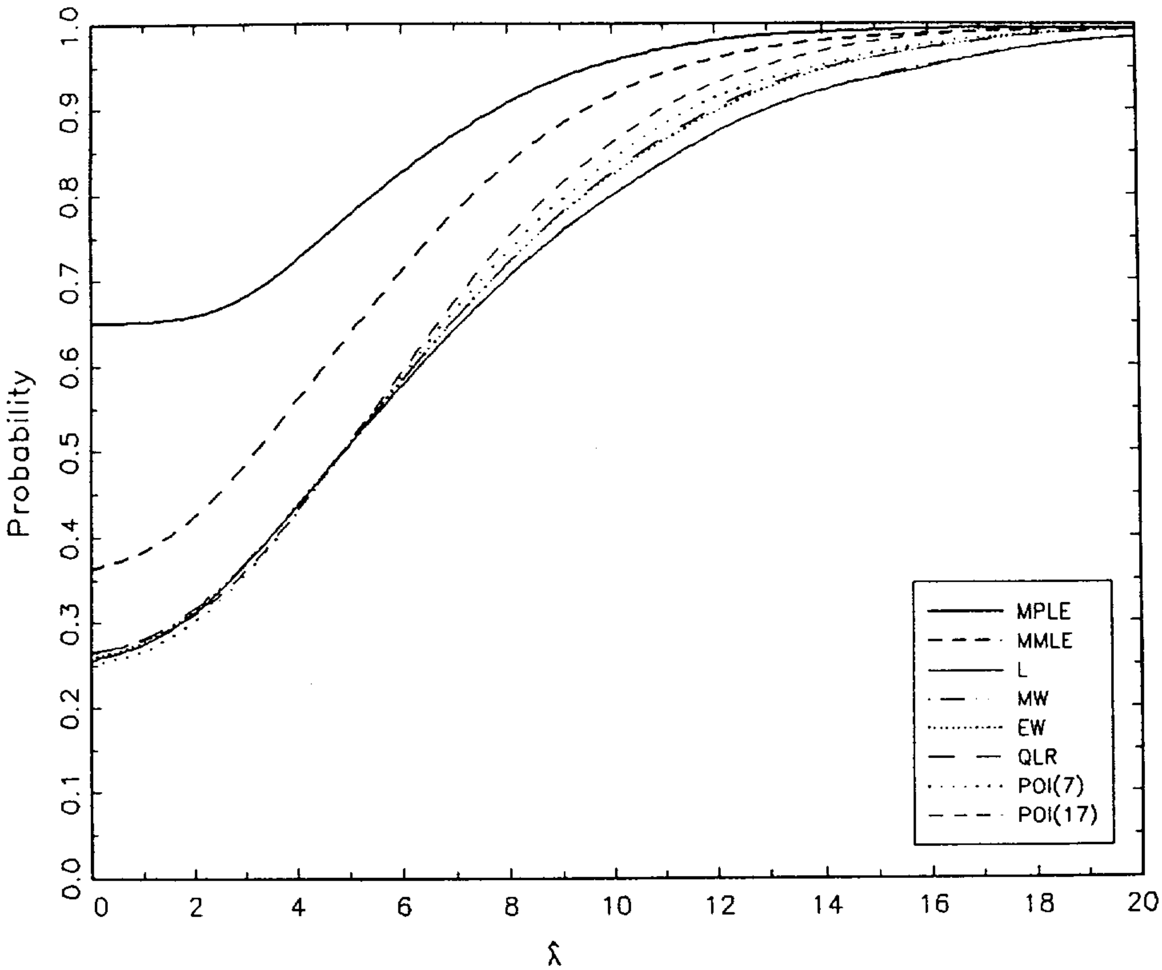

Figure 2.
Cumulative asymptotic distributions of the Gaussian MLEs
and six median-unbiased estimators of λ when λ=5

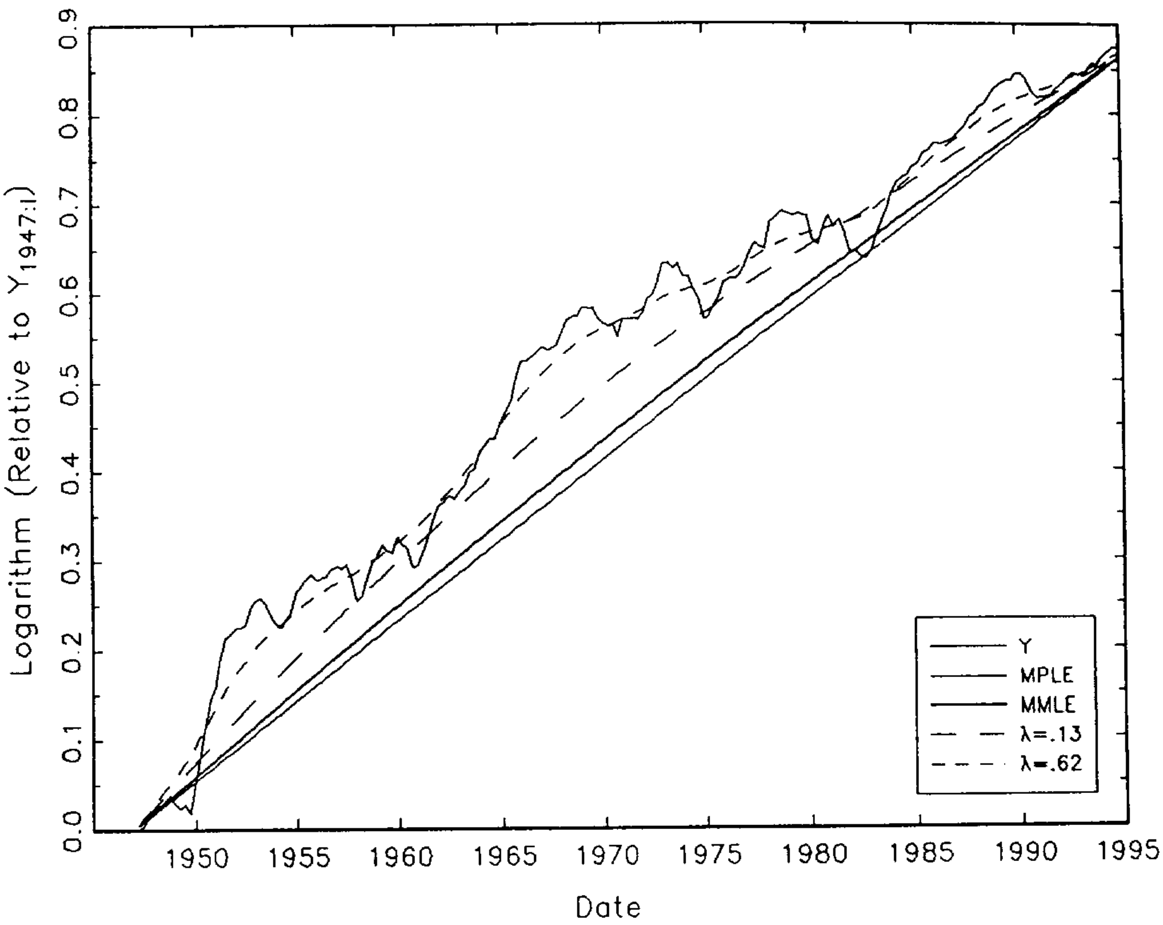

Figure 3.
U.S. real per capita GDP and estimated trends based on
the four models in table 5.