March 19, 1999

Unit Root Tests for Time Series with a Structural Break When the Break Point is Known*

by

Helmut Lütkepohl, Christian Müller and Pentti Saikkonen

Institut für Statistik und Ökonometrie
Wirtschaftswissenschaftliche Fakultät
Humboldt–University
Spandauer Str. 1
10178 Berlin
GERMANY
Tel.: +49-30-2093-5718
Fax.: +49-30-2093-5712

Department of Statistics
University of Helsinki
P.O. Box 54
SF-00014 University of Helsinki
FINLAND
+358-9-1918867
+358-9-1918872

Abstract

Unit root tests for time series with level shifts are considered. The level shift is assumed to occur at a known time point. In contrast to some other proposals the level shift is modeled as part of the intercept term of the stationary component of the data generation process which is separated from the unit root component. In this framework simple shift functions result in a smooth transition from one state to another both under the null and under the alternative hypothesis. In order to test for a unit root in this context the nuisance parameters are estimated in a first step and a standard unit root test e.g. of the Dickey-Fuller type is then applied to the residuals. The resulting test is shown to have a known asymptotic distribution under the null hypothesis of a unit root and nearly optimal asymptotic power under local alternatives. An empirical comparison with previous proposals is performed.

JEL classification: C22, C12

Keywords: Univariate time series, unit root, structural shift, autoregression

*The Deutsche Forschungsgemeinschaft, SFB 373, and the European Commission under the Training and Mobility of Researchers Programme (contract No. ERBFMRXCT980213) provided financial support. Part of this research was done while the third author was visiting the Institute of Statistics and Econometrics at the Humboldt University in Berlin.

1 Introduction

A number of studies consider testing for unit roots in univariate time series which have a level shift. Examples are Perron (1989, 1990), Perron & Vogelsang (1992), Banerjee, Lumsdaine & Stock (1992), Zivot & Andrews (1992), Amsler & Lee (1995), Leybourne, Newbold & Vougas (1998), Montañés & Reyes (1998) and Saikkonen & Lütkepohl (1999). These tests are important because the trending properties of a set of time series determine to some extent which model and statistical procedures are suitable for analyzing their relationship. In the aforementioned studies different models and assumptions for the structural shift are considered. In some of the studies the timing of the break point is assumed to be known whereas in others a shift in an unknown period is considered. There seems to be general consensus, however, that if the break point is known, this is useful information which should be taken into account in the subsequent analysis and in particular in testing for unit roots. Therefore we will focus on the latter case in the following. In practice, knowledge of a known break point is quite common. For instance, many German macroeconomic time series are known to have a shift in 1990 where the German reunification took place.

For the case of a known break point we will propose a framework which generalizes previously considered models. In this framework the shift is modeled as part of the intercept term of the stationary part of the data generation process (DGP) which is clearly separated from the unit root part. Our model has the convenient feature that even simple shift functions result in a smooth transition from one state to another both under the null of a unit root and under the alternative hypothesis of stationarity. Such a behaviour is often more realistic than an abrupt one-time shift. For instance, in some German macroeconomic time series such as GNP there is a clear shift in 1990 where the German reunification has occurred. However, the eastern part of the economy was in a quite different economic situation than West Germany at that time and entered into a long lasting adjustment process. Hence, a gradual adjustment after an initial shift may be a more realistic model in this case.

We will compare our new model to previously proposed models in an empirical comparison of different frameworks. A major advantage of the present approach relative to other approaches is that estimation of the nuisance parameters reduces to a fairly simple nonlinear least squares (LS) problem (see Amemiya (1983) for a review of nonlinear regression). In special cases estimation can even be done by linear LS although the shift from one regime to another is nonlinear.

The structure of this study is as follows. In the next section the general setup is presented and in Section 3 the tests are considered. Empirical examples are discussed in Section 4 and conclusions are given in Section 5. The proof of a theorem regarding the asymptotic properties of the test statistic is provided in the appendix.

The following notation is used. The lag and differencing operators are denoted by L and Δ, respectively, so that for a time series variable y_t, $Ly_t = y_{t-1}$ and $\Delta y_t = y_t - y_{t-1}$. Convergence in probability and in distribution are denoted by $\xrightarrow{p}$ and $\xrightarrow{d}$, respectively. Independently, identically distributed will be abbreviated as $iid(\cdot,\cdot)$, where the first and second moments are indicated in parentheses. Furthermore, $O(\cdot)$, $o(\cdot)$, $O_p(\cdot)$ and $o_p(\cdot)$ are the usual symbols for the order of convergence and convergence in probability, respectively, of a sequence. The symbol $\lambda_{min}(A)$ is reserved to denote the minimal eigenvalue of a matrix A. Moreover, $\|\cdot\|$ denotes the Euclidean norm. The abbreviations sup and inf are used as usual for supremum and infimum, respectively. The n-dimensional Euclidean space is signified as $\mathbf{R}^n$.

2 Models for Time Series with Level Shifts

Saikkonen & Lütkepohl (1999) (henceforth S&L) consider the following general model for a time series with a unit root and a level shift:

$$y_t = \mu_0 + \mu_1 t + f_t(\theta)'\gamma + x_t, \qquad t = 1, 2, \ldots, \tag{2.1}$$

where the scalars μ_0 and μ_1, the $(m \times 1)$ vector θ and the $(k \times 1)$ vector γ are unknown parameters and $f_t(\theta)$ is a $(k \times 1)$ vector of deterministic sequences depending on the parameters θ. The quantity x_t represents an unobservable stochastic error term which is assumed to have a finite order autoregressive (AR) representation,

$$a(L)x_t = \varepsilon_t, \tag{2.2}$$

where $a(L) = 1 - a_1 L - \cdots - a_{p+1}L^{p+1}$ is a polynomial in the lag operator and $\varepsilon_t \sim iid(0, \sigma^2)$. For simplicity, we assume that a suitable number of presample values of the observed series y_t is available. Obviously, if the DGP of x_t has a unit root, then the same is true for y_t. Therefore, S&L derive a test for a unit root in $a(L)$.

A simple version of a function $f_t(\theta)$ that has been considered in the literature is one which represents a single shift in the mean,

$$f_t(\theta) = d_{1t} := \begin{cases} 0, & t < T_1 \\ 1, & t \geq T_1 \end{cases} \tag{2.3}$$

that is, d_{1t} is a shift dummy variable which does not depend on any unknown parameters. In other words, the parameter vector θ does not appear in this case. The dummy d_{1t} represents a shift in the mean of the series in period T_1 which is assumed to be known. Smooth transitions from one level to another can also be accommodated in the above model by an appropriate definition of $f_t(\theta)$. An alternative way to generate smooth level shifts over a longer period of time is possible in a model of the form

$$b(L)y_t = \mu_0 + \mu_1 t + \gamma d_{1t} + v_t, \qquad t = 1, 2, \ldots, \tag{2.4}$$

where the operator $b(L) = 1 - b_1 L - \cdots - b_p L^p$ is assumed to have all its zeros outside the unit circle, the error term v_t is assumed to be an AR process of order 1,

$$v_t = \rho v_{t-1} + \varepsilon_t, \tag{2.5}$$

where again $\varepsilon_t \sim iid(0, \sigma^2)$ and $-1 < \rho \leq 1$ with $\rho = 1$ implying a unit root in y_t. In the model (2.4) the shift dummy variable generates a smooth transition to a new level via $\gamma b(L)^{-1} d_{1t}$. Defining $b(L)^{-1} = 1 + \sum_{i=1}^{\infty} \alpha_i L^i$, we get for $t > T_1$,

$$b(L)^{-1} d_{1t} = 1 + \sum_{i=1}^{t-T_1} \alpha_i.$$

Thus, in this model a smooth transition of the level of y_t is generated although just a shift dummy variable appears in the deterministic term. More flexibility of this kind of model can be obtained by replacing d_{1t} by a more general sequence $f_t(\theta)$ as in (2.1).

In this study we will consider models of the general type

$$b(L)y_t = \mu_0 + \mu_1 t + f_t(\theta)'\gamma + v_t, \qquad t = 1, 2, \ldots, \tag{2.6}$$

where all symbols are as defined in (2.1), (2.4) and (2.5). The parameters μ_0, μ_1 and γ in the model (2.6) are supposed to be unrestricted. Conditions required for the parameters θ and the sequence $f_t(\theta)$ are collected in the following set of assumptions from S&L.

Assumption 1

(a) The parameter space of θ, denoted by Θ, is a compact subset of $\mathbf{R}^m$.

(b) For each $t = 1, 2, \ldots$, $f_t(\theta)$ is a continuous function of θ and

$$\sup_T \sum_{t=1}^{T} \sup_{\theta \in \Theta} \|\Delta f_t(\theta)\| < \infty$$

where $f_0(\theta) = 0$.

(c) Defining $g_t(\theta) = [1 : f_t(\theta)']'$ for $t = 1, 2, \ldots$, and $\Delta g_1(\theta) = [1 : f_1(\theta)']'$, there exists a real number $\epsilon > 0$ and an integer T_* such that, for all $T \geq T_*$,

$$\inf_{\theta \in \Theta} \lambda_{min} \left\{ \sum_{t=1}^{T} \Delta g_t(\theta) \Delta g_t(\theta)' \right\} \geq \epsilon.$$

□

This assumption is discussed in more detail in S&L. It is not very restrictive for our purposes because it is satisfied by the sequences $f_t(\theta)$ we will consider in the following. For instance, it is easy to check that the assumption is satisfied if $f_t = d_{1t}$. The assumption guarantees estimators of the parameters with suitable properties.

We will present unit root tests within this model framework. More precisely, we will present a test of the pair of hypotheses

$$H_0 : \rho = 1 \qquad \text{vs.} \qquad H_1 : \rho < 1$$

in the next section.

It is not clear a priori which one of the two general models (2.1) or (2.6) is best suited for testing for unit roots in time series with level shifts. In fact, this is an empirical question and therefore we will use both model types in Section 4 to analyze the unit root properties of a number of time series with a level shift at a known point in time. We will compare the resulting tests for some real life macroeconomic time series.

For completeness we mention that seasonal dummies may be added to both models (2.1) and (2.6) without changing the theoretical analysis in any substantive way. For instance, in that case model (2.6) becomes

$$b(L)y_t = \mu_0 + \mu_1 t + \sum_{i=1}^{q} \nu_i D_{it} + f_t(\theta)'\gamma + v_t, \qquad t = 1, 2, \ldots,$$

where the ν_i are scalar parameters and the D_{it} ($i = 1, \ldots, q$) represent seasonal dummy variables. For example, for quarterly data, D_{it} assumes the value 1 if t is associated with the ith quarter and zero otherwise. For quarterly data $q = 3$ is used because an intercept term is included in the model. This modification does not affect the asymptotic properties of the subsequently considered test. Therefore we do not include seasonal dummies at this stage to avoid notational complications. They will be used in the empirical analysis in Section 4.

3 A Unit Root Test

The basic idea underlying our test procedure is to estimate the nuisance parameters in (2.6) first and then apply a Dickey-Fuller type test to the residuals $\tilde{v}_t$. Our approach for estimating the nuisance parameters μ_0, μ_1, θ, γ and $b_1, \ldots, b_p$ is similar to that in Elliott, Rothenberg & Stock (1996) and Hwang & Schmidt (1996). These authors use a generalized LS procedure which does not necessarily assume validity of the null hypothesis but is based on appropriate local alternatives to be specified by the analyst. Thus, suppose that the error process v_t specified in (2.5) is near integrated so that

$$\rho = \rho_T = 1 + \frac{c}{T}, \tag{3.1}$$

where $c \leq 0$ is a fixed real number. Then the generating process of v_t can be written as

$$v_t = \rho_T v_{t-1} + \varepsilon_t, \qquad t = 1, 2, \ldots \tag{3.2}$$

For simplicity we make the initial value assumption $v_0 = 0$ although our asymptotic results also hold under more general conditions (cf. Elliott et al. (1996) for a discussion of the implications of initial value assumptions). It follows from the stated assumptions that

$$T^{-1/2} v_{[sT]} \xrightarrow{d} \sigma B_c(s), \tag{3.3}$$

where $B_c(s) = \int_0^s \exp\{c(s-u)\} dB_0(u)$ with $B_0(u)$ a standard Brownian motion (cf. Elliott et al. (1996)).

Our estimation procedure employs an empirical counterpart of the parameter c. This means that we shall replace c by a chosen value $\bar{c}$ and pretend that $\bar{c} = c$ although we do not assume that this presumption is actually true. The choice of $\bar{c}$ will be discussed later.

Now, if $\bar{\rho}_T = 1 + \frac{\bar{c}}{T}$, the idea is to first transform the variables in (2.6) by the filter $1 - \bar{\rho}_T L$. For convenience we will use matrix notation and define

$$Y = [y_1 : (y_2 - \bar{\rho}_T y_1) : \cdots : (y_T - \bar{\rho}_T y_{T-1})]',$$

$$Z_1 = [1 : (2 - \bar{\rho}_T) : \cdots : (T - \bar{\rho}_T(T-1))]'$$

and

$$Z_2(\theta) = \begin{bmatrix} 1 & 1 - \bar{\rho}_T & \cdots & 1 - \bar{\rho}_T \\ f_1(\theta) & f_2(\theta) - \bar{\rho}_T f_1(\theta) & \cdots & f_T(\theta) - \bar{\rho}_T f_{T-1}(\theta) \end{bmatrix}'.$$

Here, for simplicity, the notation ignores the dependence of the quantities on the chosen value $\bar{c}$. Using this notation, the transformed form of (2.6) can be written as

$$Y = W(\theta)\beta + \mathcal{E}, \tag{3.4}$$

where $W(\theta) = [V : Z(\theta)]$ with $Z(\theta) = [Z_1 : Z_2(\theta)]$ and V the $(T \times p)$ matrix containing lagged values of the regressand. Furthermore, $\beta = [b' : \mu_1 : \mu_0 : \gamma']'$ and $\mathcal{E} = [e_1 : \cdots : e_T]'$ is an error term such that $e_t = v_t - \bar{\rho}_T v_{t-1}$. It follows from the definitions that

$$e_t = \varepsilon_t + T^{-1}(c - \bar{c})v_{t-1}. \tag{3.5}$$

The second term on the r.h.s. of this equation is asymptotically negligible because, as a consequence of (3.3), $T^{-1}\max_{1\leq t\leq T}|v_t| = O_p(T^{-1/2})$. Thus, we shall consider a nonlinear LS estimation of (3.4) by proceeding in the same way as in the case $c = 0$, that is, $e_t = \varepsilon_t$ or under the null hypothesis. The reason why we still do not assume $\bar{c} = 0$ is that choosing $\bar{c} < 0$ yields more powerful tests (see Elliott et al. (1996)). Our estimators are thus obtained by minimizing the sum of squares function

$$S_T(\theta, \beta) = (Y - W(\theta)\beta)'(Y - W(\theta)\beta). \tag{3.6}$$

Assuming that the matrix $W(\theta)$ is of full column rank for all values of $\theta \in \Theta$ one can repeat the argument used by S&L for Model (2.1), and conclude that a minimizer of $S_T(\theta, \beta)$, denoted by $[\tilde{\theta}' : \tilde{\beta}']'$, exists when Assumption 1 holds. It is seen in the appendix that this is the case for all T large enough.

The estimator of β can be written as

$$\tilde{\beta} = (W(\tilde{\theta})'W(\tilde{\theta}))^{-1}W(\tilde{\theta})'Y. \tag{3.7}$$

Of course, the computation of $\tilde{\beta}$ requires iterative methods if a parameter θ actually appears in the model. However, if preliminary estimators of θ are available they can be used on the r.h.s. of (3.7) in place of θ to yield an LS estimator of β conditional on the given θ. If computationally simple alternatives to a full minimization of $S_T(\theta, \beta)$ are desired conventional two-step estimators may be considered. The asymptotic properties of our test procedures are the same even if these estimators are employed. However, in finite samples it may be worthwhile to use proper (nonlinear) LS estimation which is still very simple. Obviously, if $W(\theta)$ is independent of θ, like in (2.4), the above estimation procedure reduces to linear regression. When $W(\theta)$ is not independent of θ, a grid search over the values of θ may provide a convenient estimation procedure if θ is scalar or possibly even if it is two-dimensional but takes values in a reasonably small set. Alternatively, one of the available nonlinear estimation algorithms may be applied (see, e.g., Amemiya (1983, 1985, Section 4.4), Judge et al. (1985, Appendix B) or Seber & Wild (1989, Chapters 13 and 14)). Asymptotic properties of the nonlinear LS estimators are given in the appendix.

Once the nuisance parameters in (2.6) have been estimated, the residual series $\tilde{v}_t = \tilde{b}(L)y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - f_t(\tilde{\theta})'\tilde{\gamma}$ may be used to obtain unit root tests. There are several possible choices. One possibility is to use Dickey-Fuller (DF) tests like, for instance, Elliott et al. (1996). In the following we shall also consider these tests.

Consider the auxiliary regression model

$$\tilde{v}_t = \rho \tilde{v}_{t-1} + e_t^*, \qquad t = 1, \ldots, T, \tag{3.8}$$

where $\tilde{v}_0 = 0$. If $\tilde{v}_t$ is replaced by v_t the error term in (3.8) becomes ε_t so that we can use simple LS to obtain a test statistic. Specifically, define the LS estimator

$$\tilde{\rho} = \left(\sum_{t=1}^{T} \tilde{v}_{t-1}^2 \right)^{-1} \sum_{t=1}^{T} \tilde{v}_{t-1} \tilde{v}_t, \tag{3.9}$$

with associated error variance estimator

$$\tilde{\sigma}^2 = T^{-1} \sum_{t=1}^{T} (\tilde{v}_t - \tilde{\rho} \tilde{v}_{t-1})^2 \tag{3.10}$$

and introduce the test statistic

$$\tau_{alt} = \left(\sum_{t=1}^{T} \tilde{v}_{t-1}^2 \right)^{1/2} (\tilde{\rho} - 1)/\tilde{\sigma}. \tag{3.11}$$

The notation τ_{alt} is used here to distinguish the statistic from the one given in S&L and to indicate that it is based on an *alt*ernative model. The statistic τ of S&L will be denoted by $\tau_{S\&L}$ in the following. The limiting distribution of the test statistic τ_{alt} is given in the next theorem.

Theorem 1. Suppose that Assumption 1 holds and that the matrix $W(\theta)$ is of full column rank for all $T \geq k+p+2$ and all $\theta \in \Theta$. Then,

$$\tau_{alt} \xrightarrow{d} \frac{1}{2}\left(\int_0^1 G_c(s;\bar{c})ds\right)^{-1/2}(G_c(1;\bar{c})^2-1),$$

where

$$G_c(s;\bar{c}) = B_c(s) - s\left(\lambda B_c(1) - 3(1-\lambda)\int_0^1 sB_c(s)ds\right)$$

and $\lambda = (1-\bar{c})/(1-\bar{c}+\bar{c}^2/3)$. □

We have included the condition for the rank of the matrix $W(\theta)$ in the theorem because it is plausible and simplifies the exposition. It is seen in the proof given in the appendix that, as a consequence of Assumption 1(c), this condition always holds for T large enough. The limiting distribution in Theorem 1 is the same as that obtained by S&L for their test statistic $\tau_{S\&L}$ and also the one Elliott et al. (1996) obtained for their t-statistic in a model whose deterministic part only contained a mean and linear trend term. The limiting null distribution, obtained by setting $c=0$, is free of unknown nuisance parameters but depends on the quantity $\bar{c}$. Elliott et al. (1996) suggest using $\bar{c}=-13.5$ and give some critical values for this choice. They show that with this choice of $\bar{c}$ the asymptotic local power of their t-test is nearly optimal for all values of c. From their results and Theorem 1 we can conclude that this is also the case for our test. Since our alternative is a stationary process v_t (i.e., $|\rho|<1$), small values of τ_{alt} are critical. It is shown in the appendix that the limiting distribution of τ_{alt} is unaffected by including seasonal dummies in the model.

In the same way as in Elliott et al. (1996) we could derive point optimal tests. These tests would be based on the statistics $\tilde{\sigma}^2(1)$ and $\tilde{\sigma}^2(\bar{\rho}_T)$ defined by replacing $\tilde{\rho}$ in (3.10) by unity and $\bar{\rho}_T$, respectively. According to the simulation results of Elliott et al. (1996) the overall properties of their DF t-statistic appeared somewhat better than those of the point optimal tests. Therefore we use the DF test version τ_{alt} in the following. Finally, note that if we have the a priori restriction that there is no linear trend term so that $\mu_1 = 0$, the above test remains essentially the same except for the limiting distribution which is then the same as in a model without any deterministic terms. Furthermore Elliott et al. (1996) recommend $\bar{c} = -7$ in this case.

4 Empirical Comparison of Tests

As mentioned earlier, which model to use for a time series with a shift in mean is primarily an empirical question because it is usually not clear a priori what kind of adjustment is required to capture the level shift in an adequate way. Therefore we have applied the different tests to some economic time series. In particular, we use a set of German macroeconomic series which was also used by S&L consisting of quarterly, seasonally unadjusted log GNP (1975(1) - 1996(4)), money stock M1 (1960(1) - 1997(1)) and M3 (1972(1) - 1996(4)). In addition we use Polish log Industrial Production (IP) (1982(1) - 1995(4)).* S&L used the $\tau_{S\&L}$ test based on the following three shift functions:

$$f_t^{(1)}(\theta) = d_{1t},$$

$$f_t^{(2)}(\theta) = \begin{cases} 0, & t < T_1 \\ 1 - \exp\{-\theta(t - T_1)\}, & t \geq T_1 \end{cases}$$

and

$$f_t^{(3)}(\theta) = \left[\frac{d_{1,t}}{1-\theta L} : \frac{d_{1,t-1}}{1-\theta L} \right]'.$$

The first one of these shift functions in conjunction with model (2.1) corresponds to an abrupt shift whereas $f_t^{(2)}(\theta)$ and $f_t^{(3)}(\theta)$ allow for a smooth transition to a new level. All three functions result in a nonlinear optimization problem in computing the $\tau_{S\&L}$ statistic. In contrast, even $f_t^{(1)}(\theta)$ can generate a smooth adjustment to a new level if the framework

*The data sources are: GNP – quarterly, seasonally unadjusted data, 1975(1) - 1990(2) West Germany, 1990(3) - 1996(4) all of Germany, Deutsches Institut für Wirtschaftsforschung, Volkswirtschaftliche Gesamtrechnung.

M1 – quarterly, seasonally unadjusted data, 1960(1) - 1990(1) West Germany, 1990(2) - 1997(1) all of Germany, OECD.

M3 – quarterly, seasonally unadjusted data, 1972(1) - 1990(2) West Germany, 1990(3) - 1996(4) all of Germany, Monatsbericht der Deutschen Bundesbank.

IP – quarterly, seasonally unadjusted data from Poland 1982(1) - 1995(4), International Monetary Fund.

of the τ_{alt} statistic is used. Moreover, in computing τ_{alt} for $f_t^{(1)}(\theta)$ linear regression is needed only. Since $f_t^{(2)}(\theta)$ and $f_t^{(3)}(\theta)$ involve just a single parameter θ, the nonlinear LS estimators for these functions are conveniently obtained by a grid search.

In Figures 1 - 4 the series and the estimated shift functions $\hat{b}(L)^{-1}f_t^{(i)}(\hat{\theta})'\hat{\gamma}$ $(i = 1, 2, 3)$ are plotted. All four series have obvious shifts. In the three German series it occurs in 1990 and is due to the German unification. Before the unification the series refer to West Germany only and after the unification they are defined for all of Germany. Hence, the shift is due to a change in the definition of the series. In Poland the introduction of a market economy in the first quarter of 1989 had a substantial and quite visible impact on the IP series. Whereas the change in the definition of the German series resulted in a quite abrupt shift, the shift in Polish IP is spread out over a number of periods. Thus, one would expect that the model in (2.1) may be better suited for capturing the shift in the German series whereas it may be necessary to allow for a gradual adjustment in the Polish series and, hence, the model (2.6) may be advantageous for this series.

The expectation with respect to the German series is supported by the abrupt shifts found by S&L in these series by fitting model (2.1) with the three shift functions mentioned earlier. They were in fact quite similar to the shifts depicted in Figures 1 - 3 for these series. It turns out that the estimated parameters in $b(L)$ are all very close to zero and, hence, $\hat{b}(L)^{-1}f_t^{(i)}(\hat{\theta})'\hat{\gamma}$ is very similar to $f_t^{(i)}(\hat{\theta})'\hat{\gamma}$ $(i = 1, 2, 3)$. It may be worth noting that there remains some autocorrelation in the residuals of the estimated model (2.6) although quite large orders of $b(L)$ are considered and using the same orders for $a(L)$ in (2.1) largely removes the residual autocorrelation. This observation indicates that the steep shift dominates the series to such an extent that even the parameter estimates in $\hat{b}(L)$ from (2.6) are distorted. They are shrunk towards zero to enforce an abrupt shift in the model and, consequently, they cannot take care of the residual autocorrelation. In contrast the estimated shift in the Polish series in Figure 4 is more gradual and, hence, for this series the model in (2.6) may be more suitable.

In Table 1 the results for unit root tests for all four series are given. In addition to the $\tau_{S\&L}$ and τ_{alt} tests we also show the results of ordinary ADF tests which allow for a deterministic trend and do not include a shift term. For all series the results for AR order $p = 4$ are given which is a reasonable order for quarterly data. However, S&L also use

Table 1. Unit Root Tests

Variable	AR order	ADF test[a]	$\tau_{S\&L}$ test[b] $f_t^{(1)}(\theta)$	$f_t^{(2)}(\theta)$	$f_t^{(3)}(\theta)$	τ_{alt} test[b] $f_t^{(1)}(\theta)$	$f_t^{(2)}(\theta)$	$f_t^{(3)}(\theta)$
log GNP	4	−2.28	−1.59	−1.59	−1.79	−2.33	−2.33	−2.51
	5	−2.52	−1.80	−1.80	−2.19	−2.44	−2.44	−2.60
log M1	4	−3.19	−2.37	−2.53	−2.70	−2.36	−2.22	−2.22
	6	−1.82	−2.61	−2.36	−2.43	−2.37	−2.24	−2.24
log M3	4	−2.40	−0.80	−0.80	−1.15	−1.17	−1.17	−1.38
	6	−2.18	−0.80	−0.80	−1.15	−1.23	−1.23	−1.35
log IP	2	−1.68	−2.02	−3.08	−2.88	−1.68	−3.41	−3.41
	4	−1.80	−2.07	−2.26	−2.07	−1.60	−3.12	−3.10

[a]Critical values: −3.96 (1%), −3.41 (5%), −3.12 (10%)
(see Fuller (1976), Table 8.5.2, $\hat{\tau}_\tau$, $n = \infty$).

[b]Critical values: −3.48 (1%), −2.89 (5%), −2.57 (10%)
(see Elliott et al. (1996), Table I.C, $T = \infty$).

different orders in their study which eliminates residual autocorrelation in the model (2.1). These orders are also used in the table for the German series. Similarly, $p = 2$ is sufficient in model (2.1) for log IP to remove the residual autocorrelation and therefore results for that order are also given in Table 1 for the Polish series.

S&L found clear support for a unit root in log GNP and log M3 and weak evidence against a unit root in log M1. In Table 1 it can be seen that similar results are obtained based on τ_{alt} although the evidence against a unit root in log GNP is somewhat stronger when the latter test is used. Still, for log GNP none of the statistics is significant at the 5% level.

A different picture emerges for log IP. Again the ADF test which does not allow for a shift is not significant at any reasonable level. Hence, an uncritical application of this test leads to the conclusion that there is a unit root in log IP. Taking into account the shift by applying $\tau_{S\&L}$ or τ_{alt} tests based on the shift functions $f_t^{(i)}(\hat{\theta})$ $(i = 2, 3)$ the unit root null hypothesis is clearly rejected at the 10% level and in most cases also at the 5% level if a model with order $p = 2$ is used. The fact that $\tau_{S\&L}$ is not significant for $p = 4$ may be a reflection of the potential power loss due to overfitting the model. Also it may indicate that the model (2.1) is not as adequate as model (2.6) for this series. Moreover, the insignificant test values based on $f_t^{(1)}$ suggest that this shift function may be too rigid for the presently considered series. The overall conclusion from these examples is that both model types may be valuable tools for unit root analysis when series have level shifts.

5 Conclusions

Many economic time series exhibit level shifts in some known time period due to a special event. It is important to take such shifts into account in unit root tests because the standard tests for this purpose are distorted and may have low power if such shifts are ignored. A quite general class of tests has been proposed in this paper for taking care of deterministic level shifts. They have the convenient feature that they allow for smooth transitions from one level to some other level over an extended period of time. Such a smooth transition to a new state is often more realistic than assuming an abrupt shift to a new level. Although there are other unit root tests which can accommodate smooth shifts in the level of a time series the tests proposed here have the advantage that the corresponding test statistics are very easy to compute for quite general shift functions. Moreover, their asymptotic null distribution is known from the unit root literature and tables with critical values exist. The tests have been applied to economic time series to illustrate how they work in practice.

In empirical work it is quite common that the timing of the level shift is known as in the examples considered in the foregoing. However, there are also occasions where the exact time of the shift is unknown. We intend to investigate extensions of the tests to this case in future research.

Appendix. Proofs

We will first present some asymptotic properties of the estimators of the nuisance parameters and then prove Theorem 1.

A.1 Properties of Estimators

Some properties of the nonlinear LS estimators obtained via (3.4) are given in the following lemma. The lemma assumes local alternatives as specified by (3.1) so that the null hypothesis is obtained by setting $c = 0$.

Lemma A.1. Suppose that the assumptions of Theorem 1 hold. Then,

$$\tilde{b} \xrightarrow{p} b, \tag{A.1}$$

$$\tilde{\theta} = \theta + O_p(1), \tag{A.2}$$

$$\tilde{\gamma} = \gamma + O_p(1), \tag{A.3}$$

$$\tilde{\mu}_0 = \mu_0 + O_p(1) \tag{A.4}$$

and

$$T^{1/2}(\tilde{\mu}_1 - \tilde{b}(1)\mu_1/b(1)) \xrightarrow{d} \sigma\left(\lambda B_c(1) - 3(1-\lambda)\int_0^1 sB_c(s)ds\right), \tag{A.5}$$

where $\lambda = (1-\bar{c})/(1-\bar{c}+\bar{c}^2/3)$. □

Lemma A.1 shows that the estimators $\tilde{b}$ and $\tilde{\mu}_1$ are consistent whereas $\tilde{\mu}_0$, $\tilde{\theta}$ and $\tilde{\gamma}$ are generally not. These latter estimators are only bounded in probability in general. For $\tilde{\theta}$

boundedness is, of course, trivial because the parameter space of θ is compact by assumption. However, for $\tilde{\mu}_0$ and $\tilde{\gamma}$ the situation is different because the parameter space of μ_0 and γ is unrestricted. The situation is similar to Lemma 1 of S&L except that the result for $\tilde{\mu}_1$ now involves the quantities $\tilde{b}(1)$ and $b(1)$. However, this result is precisely the one we need in the proof of Theorem 1. It is also possible to obtain the limiting distribution of $\tilde{\mu}_1$. That distribution is not needed in subsequent derivations, however, and it is therefore omitted. Since Assumption 1(b) implies that $f_t(\theta) - \bar{\rho}_T f_{t-1}(\theta) \approx \Delta f_t(\theta) \to 0$ as $t \to \infty$ the inconsistency of the estimators $\tilde{\gamma}$ and $\tilde{\theta}$ is expected and a similar argument can be given for $\tilde{\mu}_0$.

The proof uses the same techniques as the proof of Lemma 1 in S&L. The following results from that proof are used here as well:

$$T^{-1}Z_1'Z_1 = h(\bar{c}) + O(T^{-1}), \tag{A.6}$$

where $h(\bar{c}) = 1 - \bar{c} + \bar{c}^2/3$,

$$T^{-1/2}Z_1'Z_2(\theta) = O(T^{-1/2}), \tag{A.7}$$

$$Z_2(\theta)'Z_2(\theta) = \sum_{t=1}^{T} \Delta g_t(\theta)\Delta g_t(\theta)' + O(T^{-1}) \tag{A.8}$$

uniformly in θ, where $g_t(\theta) = [1 : f_t(\theta)']'$ and $g_0(\theta) = 0$. Furthermore, defining $D_{1T} = \text{diag}[T^{1/2} : I_k]$ and $M_T(\theta) = \text{diag}\left[h(\bar{c}) : \sum_{t=1}^{T} \Delta g_t(\theta)\Delta g_t(\theta)'\right]$,

$$D_{1T}^{-1}Z(\theta)'Z(\theta)D_{1T}^{-1} = M_T(\theta) + O(T^{-1/2}) \tag{A.9}$$

uniformly in θ. We note in passing that $(A.9)$ implies that the matrix $Z(\theta)$ is of full column rank for all θ and all T large enough because, by Assumption 1(c), the matrix $M_T(\theta)$ is positive definite for all θ and all T large enough.

Next we shall obtain an expression for the observed series by solving the difference equation defined by (2.6). This yields

$$y_t = s_{1t} + b(1)^{-1}\mu_0 + b(L)^{-1}t\mu_1 + b(L)^{-1}f_t(\theta)'\gamma + x_t, \quad t = 1, 2, \ldots,$$

where $x_t = b(L)^{-1}v_t$ with $v_t = 0$ for $t \leq 0$, the trend term and the sequence $f_t(\theta)$ are defined as zero for $t \leq 0$ and the sequence s_{1t} contains transient effects due to the presample values of y_t. As is well known, s_{1t} converges to zero exponentially as $t \to \infty$. Using the decomposition $b(L)^{-1} = b(1)^{-1} + b^*(L)\Delta$ we can write $b(L)^{-1}t\mu_1 = b(1)^{-1}\mu_1 t + s_{2t}$, where s_{2t}

contains transient effects similar to those in s_{1t} except that now these effects do not converge to zero but are bounded. Thus, we have

$$y_t = \mu_* t + k_t + x_t, \quad t = 1, 2, \ldots, \tag{A.10}$$

where $\mu_* = b(1)^{-1}\mu_1$ and $k_t = s_{1t} + s_{2t} + b(1)^{-1}\mu_0 + b(L)^{-1}f_t(\theta)'\gamma$. Notice that here the parameters indicate true values and that the properties of the sequence k_t are similar to those of $b(1)^{-1}\mu_0 + f_t(\theta)'\gamma$. In particular, when Assumption 1(b) holds, the sequence Δk_t is absolutely summable. Transforming $(A.10)$ by the filter $1 - \bar{\rho}_T L$ yields

$$y_t - \bar{\rho}_T y_{t-1} = Z_{1t}\mu_* + k_t - \bar{\rho}_T k_{t-1} + u_t, \qquad t = 1, 2, \ldots, \tag{A.11}$$

where $k_0 = 0$ and

$$\begin{aligned} u_t &= x_t - \bar{\rho}_T x_{t-1} = x_t - \rho_T x_{t-1} + (\rho_T - \bar{\rho}_T)x_{t-1} \\ &= b(L)^{-1}(v_t - \rho_T v_{t-1}) + T^{-1}(c - \bar{c})x_{t-1} \\ &= b(L)^{-1}\varepsilon_t + T^{-1}(c - \bar{c})x_{t-1} \\ &\stackrel{def}{=} u_t^{(0)} + T^{-1}(c - \bar{c})x_{t-1}. \end{aligned} \tag{A.12}$$

Note that the above remark made for the sequence Δk_t implies that the second sample moments between Z_{1t} and $k_t - \bar{\rho}_T k_{t-1}$ as well as between u_t and $k_t - \bar{\rho}_T k_{t-1}$ converge to zero in probability (for the latter, see the justification of $(A.13)$ of S&L).

We shall demonstrate next that, uniformly in θ,

$$\begin{aligned} D_T^{-1}W(\theta)'W(\theta)D_T^{-1} &= \text{diag}[R_{11} : \textstyle\sum_{t=1}^T \Delta g_t(\theta)\Delta g_t(\theta)'] + o_p(1) \\ &\stackrel{def}{=} R_T(\theta) + o_p(1), \end{aligned} \tag{A.13}$$

where $D_T = \text{diag}[T^{1/2}I_{p+1} : I_{k+1}]$ and

$$R_{11} = \begin{bmatrix} \sigma^2\Sigma(b) + \mu_*^2 h(\bar{c})\mathbf{1}_p\mathbf{1}_p' & \mu_* h(\bar{c})\mathbf{1}_p \\ \mu_* h(\bar{c})\mathbf{1}_p' & h(\bar{c}) \end{bmatrix}.$$

Here $\Sigma(b) = \sigma^{-2}\text{Cov}(u_1^{(0)}, \ldots, u_p^{(0)})$, $\mathbf{1}_p = [1 : \cdots : 1]'$ $(p \times 1)$ and the other notation is as before. To justify $(A.13)$, recall that $W(\theta) = [V : Z(\theta)]$ where the ith column of the matrix V consists of $y_{t-i} - \bar{\rho}_T y_{t-i-1}$ $(i = 1, \ldots, p,\ t = 1, \ldots, T)$. Thus, a typical element of the matrix $T^{-1}V'V$ is

$$\begin{aligned} T^{-1}&\sum_{t=1}^{T}[y_{t-i}-\bar{\rho}_T y_{t-i-1}][y_{t-j}-\bar{\rho} y_{t-j-1}] \\ &= T^{-1}\sum_{t=1}^{T}[Z_{1,t-i}\mu_* + u_{t-i}][Z_{1,t-j}\mu_* + u_{t-j}] + o_p(1) \\ &= \mu_*^2 T^{-1}Z_1'Z_1 + T^{-1}\sum_{t=1}^{T} u_{t-i}u_{t-j} + o_p(1) \\ &= \mu_*^2 T^{-1}Z_1'Z_1 + T^{-1}\sum_{t=1}^{T} u_{t-i}^{(0)}u_{t-j}^{(0)} + o_p(1) \\ &\xrightarrow{p} \mu_*^2 h(\bar{c}) + \mathrm{Cov}(u_{t-i}^{(0)}, u_{t-j}^{(0)}). \end{aligned} \tag{A.14}$$

The first equality in $(A.14)$ follows from $(A.11)$ and the remark made after it. The second equality is based on the fact that the sample mean of u_t is of order $o_p(1)$. This can be established by using $(A.12)$ and well-known properties of stationary and near-integrated processes, which also imply the third equality. Finally, the stated convergence in probability is justified by $(A.6)$ and a weak law of large numbers.

Next recall the definition $Z(\theta) = [Z_1 : Z_2(\theta)]$ and notice that a typical component of the vector $T^{-1}V'Z_1$ is

$$\begin{aligned} T^{-1}\textstyle\sum_{t=1}^{T} Z_{1t}[y_{t-i}-\bar{\rho}_T y_{t-i-1}] &= T^{-1}\textstyle\sum_{t=1}^{T} Z_{1t}[Z_{1,t-i}\mu_* + u_{t-i-1}] + o_p(1) \\ &= T^{-1}Z_1'Z_1\mu_* + o_p(1) \\ &\xrightarrow{p} \mu_* h(\bar{c}). \end{aligned} \tag{A.15}$$

Here the stated conclusions can be justified in the same way as in $(A.14)$. Now, from $(A.14)$, $(A.15)$ and $(A.6)$ we can conclude that, as far as the $((p+1)\times(p+1))$ upper left hand corner is concerned, the result stated in $(A.13)$ holds. It follows from $(A.8)$ that the same is true for the $((k+1)\times(k+1))$ lower right hand corner of the matrix in $(A.13)$. Thus, it remains to show that $T^{-1/2}Z_2(\theta)'V = o_p(1)$. A typical column of this matrix is

$$T^{-1/2}\sum_{t=1}^{T} Z_{2t}(\theta)[y_{t-i}-\bar{\rho}_T y_{t-i-1}] = T^{-1/2}\sum_{t=1}^{T}\left[\Delta g_t(\theta) - \frac{\bar{c}}{T}g_{t-1}(\theta)\right][y_{t-i}-\bar{\rho}_T y_{t-i-1}] = o_p(1)$$

uniformly in θ. Here, the first equality follows from the definition of $Z_2(\theta)$ (see S&L). The second one is based on $(A.11)$ and remarks made regarding the properties of the sequences therein. Thus, we have established $(A.13)$.

Next note that from (3.7) and (3.4) it follows that

$$\begin{aligned} D_T(\tilde{\beta}-\beta) &= (D_T^{-1}W(\tilde{\theta})'W(\tilde{\theta})D_T^{-1})^{-1}D_T^{-1}W(\tilde{\theta})'\mathcal{E} \\ &\quad +(D_T^{-1}W(\tilde{\theta})'W(\tilde{\theta})D_T^{-1})^{-1}D_T^{-1}W(\tilde{\theta})'\zeta, \end{aligned} \tag{A.16}$$

where $\zeta = (Z_2(\theta) - Z_2(\tilde{\theta}))[\mu_0 : \gamma']'$. In the latter term on the r.h.s. of $(A.16)$ we have

$$D_T^{-1} W(\tilde{\theta})'\zeta = \begin{bmatrix} T^{-1/2}V'\zeta \\ T^{-1/2}Z_1'\zeta \\ Z_2(\tilde{\theta})'\zeta \end{bmatrix} = \begin{bmatrix} o_p(1) \\ o_p(1) \\ O_p(1) \end{bmatrix}. \tag{A.17}$$

Here the last two results can be established in exactly the same way as $(A.10)$ of S&L. As to the first one, note that the components of ζ define an absolutely summable sequence (see the discussion leading to $(A.10)$ of S&L). Then, the desired result is obtained by using $(A.11)$ and the remarks made below it. It is straightforward to check that the matrix R_{11} in $(A.13)$ is positive definite. From this property and Assumption 1(c) it further follows that the matrix $R_T(\theta)$ is positive definite for all θ and all T large enough. Thus, $(A.13)$ and Lemma A2 of Saikkonen & Lütkepohl (1996) yield

$$(D_T^{-1} W(\tilde{\theta})' W(\tilde{\theta}) D_T^{-1})^{-1} = R_T^{-1}(\tilde{\theta}) + o_p(1). \tag{A.18}$$

Using $(A.11)$ and (3.5) it is straightforward to check that $T^{-1/2}V'\mathcal{E} = O_p(1)$ while $T^{-1/2}Z_1'\mathcal{E} = O_p(1)$ and $Z_2(\tilde{\theta})'\mathcal{E} = O_p(1)$ can be established in the same way as $(A.12)$ and $(A.13)$ of S&L. Thus, combining $(A.18)$ with $(A.16)$ and $(A.17)$ shows that

$$\begin{bmatrix} T^{1/2}(\tilde{b} - b) \\ T^{1/2}(\tilde{\mu}_1 - \mu_1) \end{bmatrix} = R_{11}^{-1} \begin{bmatrix} T^{-1/2}V'\mathcal{E} \\ T^{-1/2}Z_1'\mathcal{E} \end{bmatrix} + o_p(1). \tag{A.19}$$

The definition of the matrix R_{11} shows that a premultiplication of $(A.19)$ by $[\mu_* \mathbf{1}_p' : 1]$ gives

$$T^{1/2}(\mu_* \mathbf{1}_p'(\tilde{b} - b) + \tilde{\mu}_1 - \mu_1) = h(\bar{c})^{-1} T^{-1/2} Z_1'\mathcal{E} + o_p(1). \tag{A.20}$$

On the l.h.s. we have

$$\begin{aligned} \mu_* \mathbf{1}_p'(\tilde{b} - b) &= b(1)^{-1}\mu_1(\textstyle\sum_{j=1}^p \tilde{b}_i - \sum_{j=1}^p b_j) \\ &= -b(1)^{-1}\mu_1(\tilde{b}(1) - b(1)) \\ &= -\tilde{b}(1)b(1)^{-1}\mu_1 + \mu_1. \end{aligned}$$

Thus, the l.h.s. of $(A.20)$ equals the l.h.s. of $(A.5)$. Moreover, the r.h.s. of $(A.20)$ converges in distribution to the r.h.s. of $(A.5)$ by arguments given in Elliott et al. (1996) and in the proof of Lemma 1 of S&L. Thus, we have proved $(A.5)$ while $(A.1)$ is a straightforward consequence of $(A.19)$. Finally, $(A.2)$ is trivial while $(A.3)$ and $(A.4)$ are obtained from

$(A.16)$ - $(A.18)$ and the fact that the smallest eigenvalue of $R_T(\tilde{\theta})$ is bounded away from zero by $(A.13)$ and Assumption 1(c). This completes the proof of Lemma A.1.

If seasonal dummies are included in the model the matrix $W(\theta)$ is defined as $W(\theta) = [V : Z_1 : Z_3 : Z_2(\theta)]$, where Z_3 is the matrix containing the values of the seasonal dummies corresponding to $y_1, \ldots, y_T$ transformed by the filter $1 - \bar{\rho}_T L$. Redefining the dimension of the identity matrix in the definition of D_T, we still have $(A.13)$ except that R_{11} is changed to

$$R_{11} = \begin{bmatrix} \sigma^2\Sigma(b) + \mu_*^2 \mathbf{1}_p \mathbf{1}_p' & \mu_* h(\bar{c})\mathbf{1}_p & A \\ \mu_* h(\bar{c})\mathbf{1}_p' & h(\bar{c}) & 0 \\ A' & 0 & B \end{bmatrix},$$

where $A = \text{plim } T^{-1}V'Z_3$, $B = \text{plim } T^{-1}Z_3'Z_3$ and the zero matrix results because plim $T^{-1}V'Z_3 = 0$, as shown in S&L. It is not difficult to check that $R_{11} > 0$. Since it is also straightforward to verify that $T^{-1/2}Z_3'\mathcal{E} = O_p(1)$ and $T^{-1/2}Z_3'\zeta = o_p(1)$ we have an analog of $(A.19)$ with the two vectors in the brackets augmented to allow for the estimation of the coefficients of the seasonal dummies. Premultiplying this new version of $(A.19)$ by the vector $[\mu_*\mathbf{1}_p' : 1 : 0]$ shows that we still have $(A.20)$ so that $(A.5)$ holds even when seasonal dummies are included in the model. According to what was said above it is clear that $(A.1) - (A.4)$ also hold and that the coefficient estimators related to the seasonal dummies are consistent.

A.2 Proof of Theorem 1

By the definition of $\tilde{v}_t$ and $(A.10)$ one obtains

$$\begin{aligned} \tilde{v}_t &= \tilde{b}(L)\mu_* t + \tilde{b}(L)k_t + \tilde{b}(L)x_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - f_t(\tilde{\theta})'\tilde{\gamma} \\ &= v_t + \tilde{b}(L)\mu_* t + \tilde{b}(L)k_t + (\tilde{b}(L) - b(L))x_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - f_t(\tilde{\theta})'\tilde{\gamma}, \end{aligned}$$

where $\tilde{b}(L)$ is defined in an obvious way and the latter equality is based on the identity $b(L)x_t = v_t$. Now recall that $\mu_* = b(1)^{-1}\mu_1$ and use the representation $b(L) = b(1) + b_*(L)\Delta$ to obtain from the above

$$\begin{aligned} \tilde{v}_t &= v_t - (\tilde{\mu}_1 - \tilde{b}(1)b(1)^{-1}\mu_1)t + \tilde{b}_*(1)\mu_* + \tilde{b}(L)k_t \\ &\quad + (\tilde{b}(1) - b(1))x_t + (\tilde{b}_*(L) - b_*(L))\Delta x_t - \tilde{\mu}_0 - f_t(\tilde{\theta})'\tilde{\gamma}. \end{aligned}$$

From (3.3) we have $T^{-1/2}v_{[Ts]} \stackrel{d}{\to} \sigma B_c(s)$. Thus, from the above equality, Lemma A.1 and arguments similar to those used in the proof of Theorem 1 in S&L it follows that

$$\begin{aligned} T^{-1/2}\tilde{v}_{[Ts]} &= T^{-1/2}v_{[Ts]} - T^{1/2}(\tilde{\mu}_1 - \tilde{b}(1)b(1)^{-1}\mu_1)\tfrac{[Ts]}{T} + o_p(1) \\ &\stackrel{d}{\to} \sigma G_c(s;\bar{c}). \end{aligned}$$

Proceeding in the same way as in the proof of Theorem 1 of S&L, it is straightforward to use the above result to obtain the limiting distribution of the test statistic τ_{alt}. Details are omitted.

Now suppose that seasonal dummies are included in the model. Then, according to what was said above about parameter estimation in this context it is clear that the counterpart of the residual series $\tilde{v}_t$ obtained in this case satisfies $T^{-1/2}\tilde{v}_{[Ts]} \stackrel{d}{\to} \sigma G_c(s;\bar{c})$ so that the resulting test statistic has the same limiting distribution as in the model where no seasonal dummies are included.

References

Amemiya, T. (1983), Non-linear regression models, in Z. Griliches & M.D. Intriligator (eds.), *Handbook of Econometrics, Volume I*, Amsterdam: North-Holland, pp. 333 - 389.

Amemiya, T. (1985), *Advanced Econometrics*, Cambridge, Ma.: Harvard University Press.

Amsler, C. & J. Lee (1995), An LM test for a unit root in the presence of a structural change, *Econometric Theory*, 11, 359 - 368.

Banerjee, A., R.L. Lumsdaine & J.H. Stock (1992), Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business & Economic Statistics*, 10, 271 - 287.

Elliott, G., T.J. Rothenberg & J.H. Stock (1996), Efficient tests for an autoregressive unit root, *Econometrica*, 64, 813 - 836.

Hwang, J. & P. Schmidt (1996), Alternative methods of detrending and the power of unit root tests, *Journal of Econometrics*, 71, 227 - 248.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl & T.C. Lee (1985), *The Theory and Practice of Econometrics*, Second Edition, New York: John Wiley.

Leybourne, S., P. Newbold & D. Vougas (1998), Unit roots and smooth transitions, *Journal of Time Series Analysis*, 19, 83 - 97.

Montañés, A. & M. Reyes (1998), Effect of a shift in the trend function on Dickey-Fuller unit root tests, *Econometric Theory*, 14, 355 - 363.

Perron, P. (1989), The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361 - 1401.

Perron, P. (1990), Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics*, 8, 153 - 162.

Perron, P. & T.J. Vogelsang (1992), Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business & Economic Statistics*, 10, 301 - 320.

Saikkonen, P. & H. Lütkepohl (1996), Infinite-order cointegrated vector autoregressive processes: Estimation and inference, *Econometric Theory*, 12, 814 - 844.

Saikkonen, P. & H. Lütkepohl (1999), Testing for unit roots in time series with level shifts, Discussion Paper, SFB 373, Humboldt-Universität zu Berlin.

Seber, G.A.F. & C.J. Wild (1989), *Nonlinear Regression*, New York: John Wiley.

Zivot, E. & D.W.K. Andrews (1992), Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis, *Journal of Business & Economic Statistics*, 10, 251 - 270.

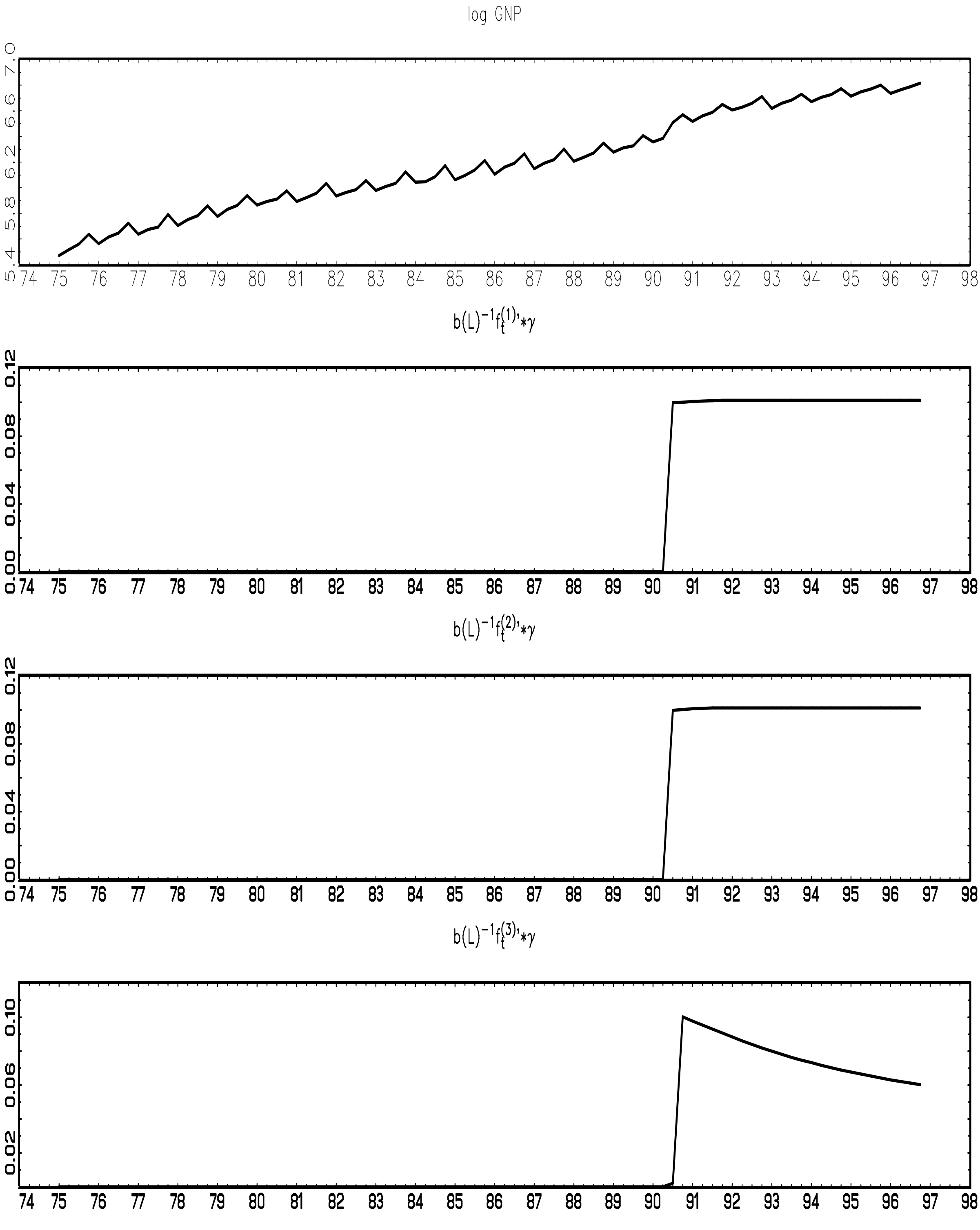

Figure 1: Plots of German log GNP and shift functions (based on AR order 4).

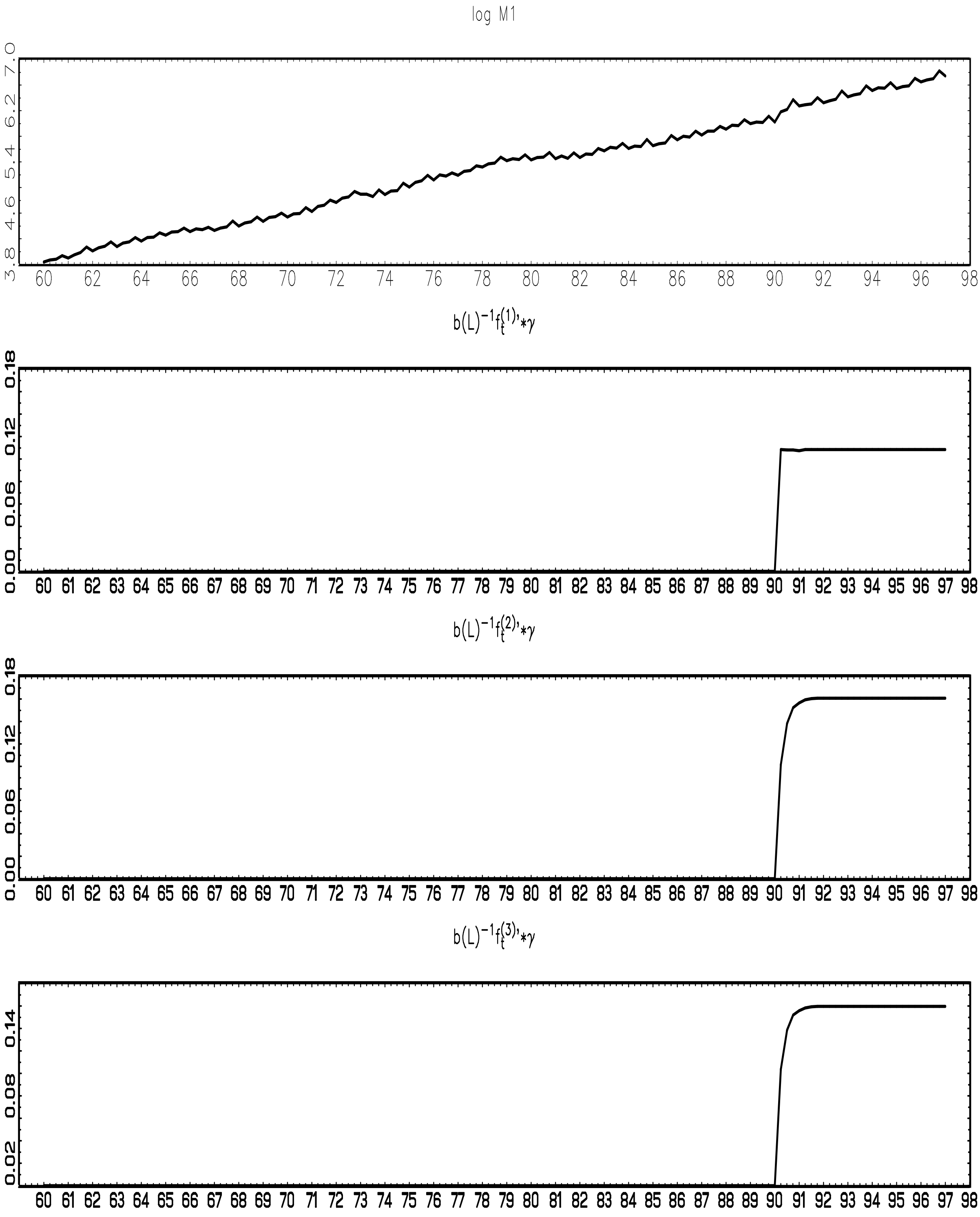

Figure 2: Plots of German log M1 and shift functions (based on AR order 4).

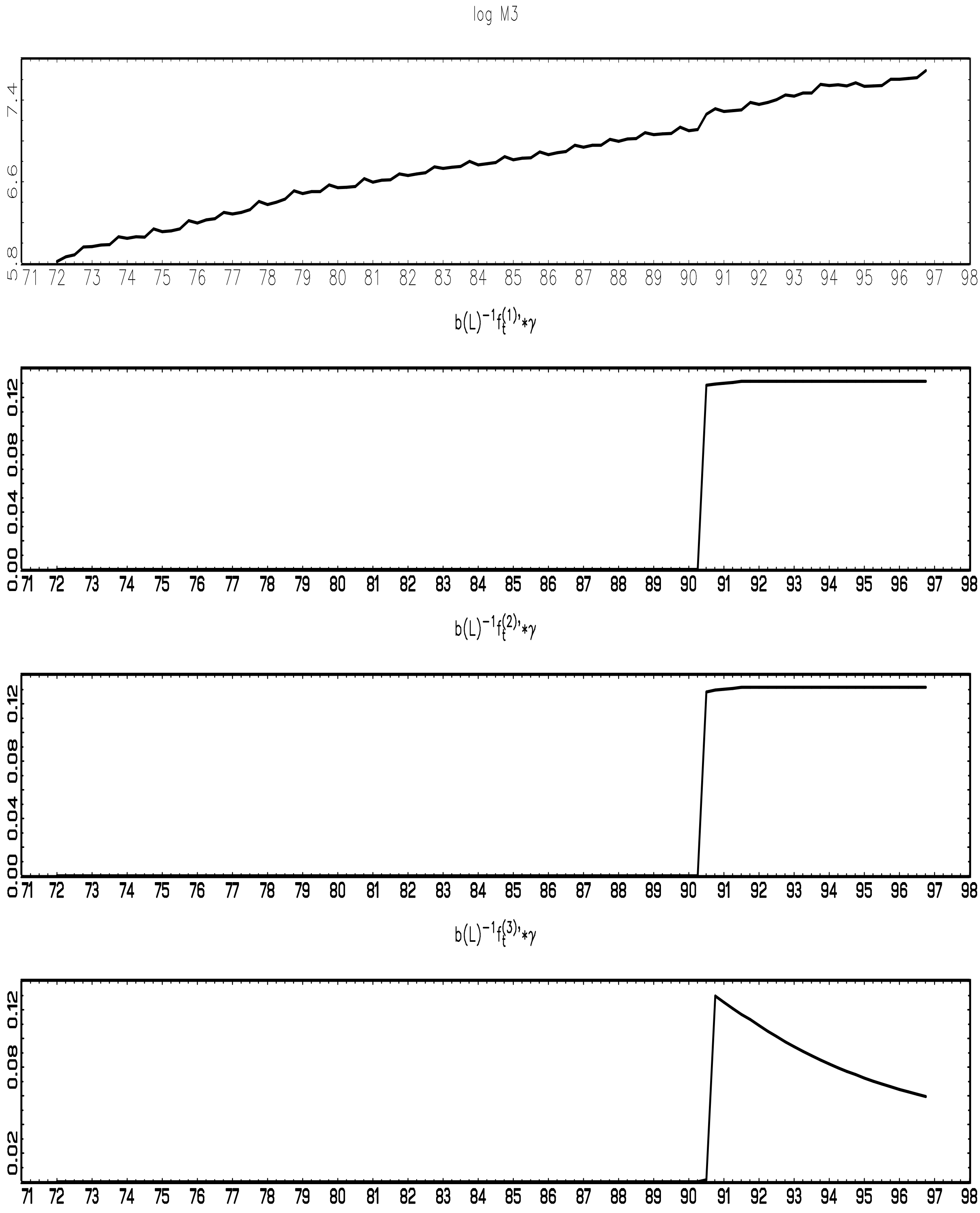

Figure 3: Plots of German log M3 and shift functions (based on AR order 4).

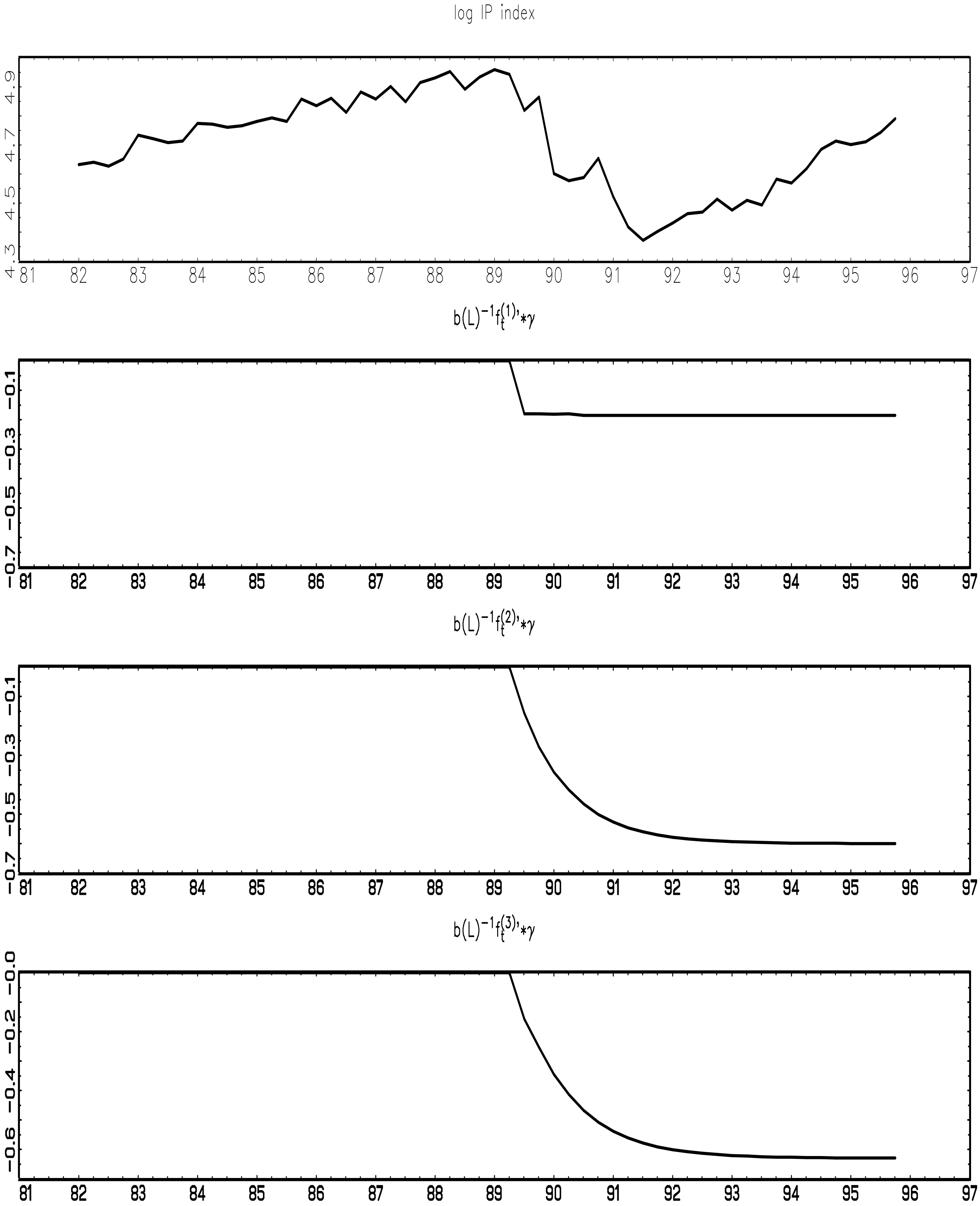

Figure 4: Plots of Polish log Industrial Production and shift functions (based on AR order 4).