EFWW
7/17/2003
11:30am

# A Structural Break in U.S. GDP?

David N. DeJong*
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260

Roman Liesenfeld
Department of Economics
Universität Tübingen
72074 Tübingen, Germany

Jean-François Richard
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260

June 2003

**Abstract**

The volatility of growth in U.S. real GDP declined dramatically in the mid-1980s. Viewed through the lens of linear autoregressive models, this phenomenon appears to be the result of a structural break in the innovation process that drives GDP fluctuations. We present an alternative model that is structurally stable over the entire post-war period. This model characterizes growth as following non-linear trajectories that fluctuate stochastically between alternative periods of general acceleration and deceleration. The specific trajectories can differ across regimes due to randomness in the parameters that govern their behavior; but the underlying distribution from which these parameters are realized is stable. The model also allows the variance of growth-rates innovations to differ across regimes, as a function of their corresponding regime-specific trajectories. We find no evidence of a structural break when viewing GDP growth through the lens of this model. Moreover, using an encompassing exercise, we show that this model can account for the evidence favoring a structural-break interpretation of the volatility reduction under linear autoregressive processes. The model also encompasses general patterns of business-cycle activity.

**Keywords:** business cycles; encompassing; regime switching; error correction; conditional heteroscedasticity
**JEL Codes:** C22, C51, C52

*Corresponding author: DeJong (email: dejong@pitt.edu). For helpful comments, we thank conference participants in "On the Wealth of Nations - Extending the Tinbergen Heritage" held in Rotterdam, April 2003. We are also grateful to Bruce Hansen for providing, via his website, GAUSS programs used to conduct tests for structural stability. The usual disclaimer applies.

"*In economic life there is an elusive mixture of relatively systematic interrelations, for which one can find a more or less regular repetitive pattern and historically unique events and disruptions.*" (From the Nobel Presentation Speech in honor of Ragnar Frisch and Jan Tinbergen, 1968, by Erik Lundberg of Sweden's Royal Academy of Sciences.)

# 1 Introduction

Evidence regarding the dramatic volatility reduction in post-war real GDP growth observed in the mid-1980s has become so extensive that this phenomenon now rates as a stylized fact. Viewed through the lens of linear autoregressive (AR) time-series representations, this phenomenon appears to be the result of a structural break in the volatility of the fundamental shocks of the process. Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) first presented evidence of this nature; see Stock and Watson (2002) for a thorough investigation of a broad range of aggregate time series and a comprehensive literature review. In a previous paper (DeJong, Liesenfeld and Richard, DLR, 2003), we developed a non-linear model of GDP growth that we found to be effective in forecasting business-cycle turning points. The purpose of the present paper is to address the structural-break issue viewed from the perspective of this alternative to the linear AR framework.

The model characterizes growth as following non-linear trajectories that fluctuate stochastically between alternative periods of general acceleration and deceleration. Regime changes occur stochastically, with probabilities determined by an observed indicator variable via a logistic link function. We refer to the observed indicator variable as a "tension index", which is constructed as a geometric sum of past deviations of actual GDP growth from a corresponding "sustainable rate" (interpretable as the growth rate of potential GDP).

The model is specifically designed to capture two key forms of heterogeneity observed in the post-war behavior of GDP growth. The first involves the decline in volatility noted above. The second involves growth patterns observed across cycles: no two business cycles are alike. In order to account for the latter, we model the parameters that dictate growth trajectories as random variables that vary across regimes. At the beginning of each regime a new set of parameters is drawn from a fixed distribution. In order to account for the former, we employ two alternative specifications for the variance of growth-rate innovations. These alternatives reflect alternative perspectives regarding the nature of the volatility reduction that have emerged in the literature. The first perspective, put forth by Simon and Blanchard (2001), views the reduction as part of an ongoing trend. The second, put forth, e.g., by McConnell and Perez-Quiros (2000) and Stock and Watson (2002), views the reduction as a discrete jump. To capture ongoing-trend behavior, we characterize innovation-variance heterogeneity using a GARCH (generalized autoregressive conditional heteroscedasticity) type specification (DLR-GARCH hereafter). To capture discrete-jump behavior, we model the variance of growth-rate innovations as differing across regimes as a function of their regime-drift trajectories (DLR-$b$ hereafter). We subject both specifications to a battery of structural-stability tests, and find clear evidence supporting the discrete-jump specification.

The structural-stability tests we use have been employed extensively in characterizing the volatility reduction in GDP growth. These tests involve the adaption of Quandt's (1960) testing procedure to the case in which the timing of a potential structural break is unknown. Implementation of the tests is facilitated by the work of Andrews (1993), Andrews and Ploberger (1994), and Hansen (1997). The tests provide evidence of structural instability when applied to the DLR-GARCH specification, but not the DLR-$b$ specification. Thus the latter model provides one possible structural account of the volatility reduction observed in GDP growth.

Next, we conduct two sets of encompassing exercises using the DLR-$b$ model (for a discussion of encompassing as a model-validation concept see, e.g., Hendry and Richard, 1990). The first involves the generation of artificial growth realizations that track the specific set of regime-change dates and regime-drift trajectories estimated from the data. Applying structural-stability tests to linear AR representations estimated using these conditional artificial realizations, we obtain a pattern of results that closely mimics that obtained in applications of these tests to the actual data. Thus we find that the model encompasses the structural-break interpretation of the volatility reduction in GDP growth that emerges when viewing the data through an AR lens.

The second exercise involves the generation of unconditional growth realizations from the DLR-$b$ model: that is, unconditional on the set of regime-change dates and regime-drift trajectories inferred from the actual data. Here, rather than focusing on a specific episode, we seek to determine whether the model is capable of accounting for the general pattern of business-cycle behavior observed in the post-war U.S. We examine five summary statistics in this exercise: the number of recessions observed over a 222 quarter period (the number of periods in the actual sample); the average length of expansions; the standard deviation of this length; the average length of recessions; and the standard deviation of this length. Once again, we find that our model encompasses these critical features of the data.

These findings convey a cautionary message. Under the structural-break interpretation of the volatility reduction, one might be tempted to conclude that this reduction has ushered in an era of economic tranquility that carries with it a certain air of permanence. But in the context of our model, the volatility reduction is associated merely with the realization of a particular regime sequence with tenuous longevity. Indeed, such a regime sequence is not unprecedented: the period 1961 to 1971 also clearly emerges as a transient low-volatility episode according to our model.

## 2 Data

We begin by characterizing evidence favoring a structural-break interpretation of the behavior of GDP growth, denoted by $g_t$. We compute $g_t$ as logged differences in quarterly GDP measured in chain-weighted 1996 prices, annualized by multiplying by 400, spanning 1947:II through 2002:III.

This series is illustrated in Figure 1, along with the residuals associated with an AR(4) specification including a constant (more parsimonious specifications yield similar results). Summary statistics of these series are reported in Table 1. The volatility reduction mentioned above is clearly evident in both the raw series and the AR(4) residuals. Prior to 1984:I (the break date identified using the tests discussed below), the standard deviations of these series were (4.73, 4.42), in comparison with measures of (4.05, 3.74) over the full sample, and (2.13, 1.99) in the latter half of the sample.

To analyze whether this behavior is consistent with a structural-break interpretation, we use a testing procedure based on the Wald form of the Quandt (1960) statistic with a heteroskedasticity-consistent covariance matrix, following the strategy employed, e.g., by McConnell and Perez-Quiros (2000). This involves sequential tests for a structural break of unknown timing in the variance and mean parameters of an AR($p$) specification:

$$g_t = \mu + \phi_1 g_{t-1} + \cdots + \phi_p g_{t-p} + \epsilon_t \ , \qquad \mathrm{Var}(\epsilon_t) = \sigma^2 \ , \qquad t : 1 \rightarrow T \ . \tag{1}$$

First, we test the stability of $\sigma^2$ using the sup-Wald statistic, which is given by the largest Wald statistic for a structural break in $\sigma^2$ over all potential breakdates $\ell$ between dates $T_1$ and $T_2$. (Following McConnell and Perez-Quiros (2000), we use $T_1 = [0.15 \cdot T]$ and $T_2 = [0.85 \cdot T]$.) Asymptotic critical values for the sup-Wald statistic are provided by Andrews (1993), and asymptotic $p$-values by Hansen (1997) [1]. In addition to the sup-Wald statistic, we also use the exp-Wald statistic proposed by Andrews and Ploberger (1994):

$$\mathrm{exp}W = \ln[1/(T_2 - T_1 + 1)] \sum_{\ell=T_1}^{T_2} \exp\{W(\ell)/2\} \ , \tag{2}$$

where $W(\ell)$ is the Wald test statistic obtained for a candidate break at date $\ell$. Second, we test for stability of the AR-parameters ($\mu$, $\phi_1$,...,$\phi_p$), again using the sup- and exp-Wald statistic. Results of the full set of tests are given in Table 1; and the sequence of Wald statistics obtained for the variance-stability test is plotted in Figure 1.

Consistent with the findings of McConnell and Perez-Quiros (2000), these tests yield no evidence of a break in the AR parameters estimated for $g_t$, but sharp evidence of a reduction in $\sigma^2$. Regarding AR parameters, the lowest $p$-value we obtain is 0.34, for $\phi_4$; in contrast, the $p$-value associated with $\sigma^2$ is 0 for both tests. An estimate of the date for a break in $\sigma^2$ using the largest Wald statistic is 1984:I, but note from Figure 1 that there is considerable uncertainty associated with this specific date. In particular, Wald statistics become significant at the 5% level as early as 1978, and the sequence of Wald statistics is fairly flat between 1982 and 1995. So there is uncertainty regarding the specific date of the break, but not about its occurrence: viewed through the lens of a linear AR framework, the dramatic decline in the volatility of GDP growth appears to be the result of a structural break. We now turn to an alternative, structurally coherent, interpretation of this phenomenon.

[1] To perform these tests, we used the software provided by Hansen at his website: `http://www.ssc.wisc.edu/~bhansen/`.

## 3 The Model

Our analysis is based on two versions of the non-linear forecasting model we developed previously (DLR, 2003). The mean specification is common to both versions, and has two key features: an error correction mechanism (ECM) and regime-switching behavior. Regime switches are manifested in the behavior of a stochastic drift component that fluctuates between periods of general acceleration and deceleration. Regime changes are triggered by a tension index, constructed as a geometric sum of past deviations of actual GDP growth $g_t$ from a corresponding "sustainable" growth rate (interpreted as the growth rate of potential GDP). Letting $g_t^*$ denote the sustainable rate, and $y_t$ the deviations $y_t = g_t - g_t^*$, the tension index $G_t$ is given by

$$G_t = \sum_{i=0}^{\infty} \delta^i y_{t-i} , \tag{3}$$

where $0 < \delta < 1$ measures the persistence of past deviations on current $G_t$. Since $g_t^*$ is not observable, we estimated $g_t^*$ by fitting a third-order polynomial in time to $g_t$, which reflects that $g_t^*$ - based on its economic interpretation - is expected to move slowly and smoothly over time; the results presented below are robust to the use of alternative representations of $g_t^*$.

The behavior of $G_t$ is illustrated in the top panel of Figure 2. (To produce the $G_t$ series, $\delta$ was set at 0.65, which results from optimizing a criterion function specified below.) Note that the series tends to pass between phases of general expansionary and contractionary behavior. Note also that peaks in $G_t$ tend to precede NBER-dated business-cycle peaks, and troughs tend to precede or coincide with business-cycle troughs. By implication, $g_t$ tends to pass between phases during which it alternately tends to outstrip and fall short of $g_t^*$. Under the interpretation of our model, neither phase is sustainable: both produce tension buildups (captured by increases in the absolute value of $G_t$) that ultimately lead to regime changes. We model regime-change probabilities using the following logit specification:

$$\pi_t = P(s_{t+1} = -\bar{s}_t | s_t = \bar{s}_t, G_t) = \frac{1}{1 + \exp\{\beta_0 - \beta_1 \bar{s}_t G_t\}} , \tag{4}$$

where the variable $s_t$ indicates the regime prevailing in period $t$, being 1 if $G_t$ is in an expansionary regime and -1 otherwise. Thus as the absolute value of $G_t$ increases, so too does the probability of a regime change.

The structural model for GDP growth, in terms of its deviations from $g_t^*$, is given by

$$y_t = m_t - \nu G_{t-2} + \gamma y_{t-2} + \epsilon_t , \quad \epsilon_t | \sigma_t \sim \mathrm{N}(0, \sigma_t^2) , \tag{5}$$

where $m_t$ represents a stochastic latent regime drift, and $\nu G_{t-2}$ and $\gamma y_{t-2}$ direct the ECM. (Alternative specifications of the conditional variance $\sigma_t^2$ are discussed below.) The specification for the regime drift, which allows for jumps at dates featuring a regime change and for different (possibly non-linear) slopes across the regimes, has the form

$$m_t = a_j + b_j s_t \sum_{\tau=1}^{t-t(j-1)-1} d^{\tau-1} , \tag{6}$$

where the index $j$ ($j : 1 \to J$) denotes the regime prevailing in period t, and $t(j)$ is the regime-change period from regime $j$ to regime $j+1$, with $t(0) \equiv 0$. The variable $a_j$ represents the value of the regime drift in the first period of regime $j$, and $b_j s_t \sum_\tau d^{\tau-1}$ directs the slope of the $m_t$ trajectory during regime $j$. Specifically, $s_t \sum_\tau d^{\tau-1}$ captures the general trend behavior of $m_t$, which is increasing in periods of acceleration and decreasing in periods of deceleration; the non-negative parameter $d$ governs the curvature of the trend, which is convergent for $d < 1$, divergent for $d > 1$, and linear for $d = 1$. The non-negative variable $b_j$, which technically represents the absolute value of the initial slope of $m_t$, can be interpreted as the loading factor of $m_t$ on the general trend component during regime $j$.

In order to allow for variation in the drift component $m_t$ across the $J$ regimes, accounting for the fact that no two business cycles are alike, it is assumed that $a_j$ and $b_j$ are latent random variables. While in DLR (2003) $a_j$ and $b_j$ are modelled as random variables following a Cauchy and a lognormal distribution, respectively, the estimations presented here are carried out using a Gaussian distribution for $a_j$ and an exponential for $b_j$. The parameter estimates under both sets of distributions are very similar, though the Gaussian-exponential specification produces significantly higher likelihood values than the Cauchy-lognormal combination. In particular, we specify

$$a_j \sim \mathrm{N}(\alpha_0 + \alpha_1 z_j, \sigma_a^2) , \quad b_j \sim \mathrm{Exponential}(\sigma_b) , \tag{7}$$

where $z_j$ is an indicator variable being 1 if regime $j$ is expansionary and 0 otherwise, and $\mathrm{E}(b_j) = \mathrm{Var}(b_j)^{1/2} = \sigma_b$. The dependence of the mean of $a_j$ on $z_j$ allows average starting values of $m_t$ to be different in expansionary and contractionary regimes. Finally, to capture observed heterogeneity in GDP volatility, in DLR (2003) the following GARCH-type specification was used for the conditional variance of growth-rate innovations:

$$\sigma_t^2 = \omega_0 + \omega_1 \sigma_{t-1}^2 + \omega_2 y_{t-1}^2 . \tag{8}$$

In the context of the model (3)-(8) (DLR-GARCH hereafter), an interpretation of the significant and persistent decline in GDP volatility in the mid-1980s is that the late-1980s and 1990s represent an episode featuring slowly evolving $m_t$ trajectories (estimated trajectories are presented in section 4). Such behavior produces low values of $y_t^2$, and thus generates a gradual and persistent decrease in the conditional variance $\sigma_t^2$. This interpretation is consistent with that provided by Simon and Blanchard (2001) who suggested that the volatility decline is part of an ongoing trend. However, it is at odds with the results of, e.g., McConnell and Perez-Quiros (2000) and Stock and Watson (2002), who suggested that the decline was due to a sharp break in the level of the volatility process and hence interpretable a discrete event. This latter interpretation implies the volatility process is best approximated by a step function based upon, e.g., a dummy variable reflecting regimes prevailing before and after the break. Here, we replace the GARCH specification (8) by introducing an alternative step function linking the conditional variance directly to the slope variable $b_j$. In particular, we use the following specification:

$$\sigma_t^2 = \omega_0 + \omega_3(1 - \exp\{-\omega_4 b_j\}) \ , \tag{9}$$

DLR-$b$ hereafter[2]. Note that in contrast to a variance specification featuring a regime dummy, the DLR-$b$ model remains structurally stable. This distinction is critical for forecasting, since under the dummy variable specification, an in-sample variance reduction automatically extends over the forecast horizon. In contrast, by allowing for stochastic regime changes, our model does not require that the volatility reduction be permanent: new regimes imply new variance levels.

Both models are estimated by maximizing their corresponding likelihood functions. In order to evaluate the likelihood function, the persistence parameter in the definition of the tension index $\delta$ and the set of regime-change dates $\{t(j)\}$ must be known. Using a grid search procedure, the parameter $\delta$ is chosen in order to maximize the ML estimate of $\beta_0$ in (4), which implies maximizing the $t$-statistic associated with a regime change. This yields a value of $\delta$ equal to 0.65, but the empirical results are fairly robust to a range of alternative values between 0.5 and 0.95. Regime-change dates are determined using an iterative back-induction procedure based upon the implied probability estimates on alternative sets of regime-change dates (see DLR, 2003 for details). This procedure yields a coherent set of dates, which are given in the descriptions of Table 2.

Letting $x_t$ denote the observable variables of period $t$, i.e., $x_t = \{y_t, s_t\}$, the likelihood associated with both models can be written as

$$L(\theta) = \prod_{j=1}^{J} \int \int f(x_{t(j-1)+1}, \dots, x_{t(j)} | a_j, b_j; \theta) h(a_j, b_j; \theta) da_j db_j \ , \tag{10}$$

where $\theta$ is the unknown parameter vector, $f(\cdot|\cdot)$ denotes the conditional joint density of the observable variable of regime $j$ given the corresponding latent variables, and $h(\cdot)$ represents the joint density of the latent variables[3].

[2] We also explored alternative functional forms for the $\sigma_t^2$-$b_j$ relationship, but found that they did not deliver improved performance.

[3] The evaluation of the $J$ two-dimensional integrals are accomplished using a two-dimensional $20 \times 40$-point quadrature rule. In the Gaussian $a_j$-dimension we employ a 20-point Gauss-Hermite and in the exponential $b_j$-dimension a 40-point Gauss-Laguerre quadrature. ML estimates are obtained using the BFGS optimization algorithm implemented in GAUSS.

# 4 Empirical Results

## 4.1 Model Estimates and Structural Stability Tests

Estimation results of the DLR-GARCH model are reported in the two left columns of Table 2. Notably, the estimated $\gamma$ and $\nu$ imply a fairly strong error-correction effect with estimated ECM coefficients $(1-\gamma)$ and $\nu$ of 0.712 (s.e. 0.090) and 0.235 (0.068). While the parameters governing the initial values $a_j$ are estimated with considerable imprecision, $\sigma_b$ and $\ln d$, determining slopes of the $m_t$ trajectories, are estimated very precisely. The estimates of $\sigma_b$ and $d$ are 0.225 (0.096) and 1.116 (0.027), implying significant variation across regimes in the loading factors on a divergent trend component. Finally, the estimate of the GARCH parameter $\omega_1$ at 0.917 (0.034) reflects a high degree of persistence in both high- and low-volatility episodes.

To illustrate the interpretation of the persistent volatility decline in the mid-1980s in the context of the DLR-GARCH model, Figure 2 plots in the middle panel the estimated $m_t$ process together with estimated deviations of growth from the ECM component $y_t+\nu G_{t-1}-\gamma y_{t-1}$, and in the bottom panel the estimated conditional standard deviation $\sigma_t$. (The estimated $m_t$ process is obtained using smoothed estimates of the latent variables $a_j$ and $b_j$; i.e., using their conditional expectations given complete sample information.) It can be seen that the years 1984-1999 represent a period with a slowly moving $m_t$ trajectory, generated by two successive long regimes with low $b_j$s. Due to the high persistence in the estimated GARCH process, this translates into a locally decreasing trend in the conditional variance and leads to a persistent low-volatility episode. As noted above, this characterization of the volatility decline is consistent with the trend interpretation of Simon and Blanchard (2001). However, as shown, e.g., by Lamoureux and Lastrapes (1990), the appearance of high persistence in the estimated GARCH component of the DLR model might simply be the result of a sharp structural shift in the level of the volatility process, which accords with the structural-break interpretation of the volatility decline put forth by McConnell and Perez-Quiros (2000) and Stock and Watson (2002).

To address this issue, we subjected the DLR-GARCH model to two tests of structural stability in the volatility level, captured by the intercept of the GARCH-equation $\omega_0$. The first test is based on the standardized residuals, given by $u_t=(y_t-\hat{m}_t+\hat{\nu}G_{t-2}-\hat{\gamma}y_{t-2})/\hat{\sigma}_t$, where $\hat{\nu}$ and $\hat{\gamma}$ are the corresponding parameter estimates, $\hat{m}_t$ is the estimated regime drift based on the smoothed estimates of $a_j$ and $b_j$, and $\hat{\sigma}_t$ represents the estimated conditional standard deviation. Under the hypothesis of no structural break in the intercept $\omega_0$, the variance of $u_t$ should be stable over time. This stability is tested using the sup-Wald and exp-Wald statistic based upon an AR(1) model fitted to $u_t$ (as discussed in Section 2). The second test is based directly on the DLR model using the following generalized GARCH-equation

$$\sigma_t^2=\omega_0+\omega_1\sigma_{t-1}^2+\omega_2 y_{t-1}^2+\kappa I_{(\ell)}\ , \tag{11}$$

where $I_{(\ell)}$ is a binary variable that equals to 1 if $t\leq\ell$ and equals 0 otherwise. To test the stability in the volatility level, we apply – analogously to the sup-Wald statistic – the sup-likelihood-ratio (sup-LR) statistic associated with the hypothesis that $\kappa$ equals 0 over all potential break dates $\ell$ between $T_1 = [0.15 \cdot T]$ and $T_2 = [0.85 \cdot T]$. Results of the stability tests are given in Table 2; and standardized residuals, together with the sequence of the Wald and LR statistics, are plotted in Figure 3.

Consistent with the findings of DLR (2003), the tests for structural change in the residual variance yield no evidence of a break in variance of the DLR-GARCH model. In particular, the $p$-values associated with the sup-Wald and exp-Wald statistic for the residuals are 0.35 and 0.50. However, the sup-LR statistic indicates a strong rejection of the null hypothesis that there is no structural change in the GARCH-equation; regarding the break date, note from the bottom panel of Figure 3 that the LR-statistic reaches its largest value in 1982:IV. An explanation for this difference in test results is that the power of the residual-based test is significantly lower than that of the test comparing the variance specification under the null and the alternative directly. So these results seem to suggest that the volatility decline in the 1980s is better-characterized as a discrete event than as a gradual transition into a low-volatility period, consistent with the results of McConnell and Perez-Quiros (2000) and Stock and Watson (2002).

For comparison, estimation results of the DLR model using a variance specification with a dummy generating a break at the estimated break date, i.e. $\sigma_t^2 = \omega_0 + \kappa I_{(82:\mathrm{IV})}$, are reported in the two middle columns of Table 2. They show that the parameter associated with the regime dummy $I_{(82:\mathrm{IV})}$ is highly significant based on a conventional $t$-test. The additional inclusion of $\sigma_{t-1}^2$ and $y_{t-1}^2$ in this specification (not presented here) does not improve the fit of the model, suggesting that the persistence in the estimated GARCH equation indeed proxies for a discrete shift in the volatility process, as suggested by the results of Lamoureux and Lastrapes (1990). However, as mentioned above, a disadvantage of a variance specification featuring such a dummy variable is that its application to forecasting exercises is problematic.

The two rightmost columns of Table 2 contain estimation results for the DLR-$b$ model, which provides a structurally stable alternative to the dummy variable approach in characterizing changes in the volatility as discrete events. Note from Figure 2 that relating the conditional variance directly to $b_j$ represents an empirically sensible approach, as the low-volatility episode at the end of the sample coincides with regimes with low $b_j$s. This is confirmed by the fact that the parameter $\omega_3$, measuring the impact of $b_j$ on the variance, is statistically significant. Furthermore, the log-likelihood value indicates that the DLR-$b$ presents a significant improvement over the DLR-GARCH specification. Finally, note that estimates of the parameters governing the stochastic drift $m_t$ and the ECM component are, except for $\alpha_1$, very similar across the different variance specifications, indicating that the mean equation of the DLR models is fairly robust.

As with the DLR-GARCH model, we test for a structural break in the conditional variance of the DLR-$b$ model using the sup-Wald and exp-Wald statistics for the standardized residuals, as well as the sup-LR statistic based on the generalized variance equation

$$\sigma_t^2 = \omega_0 + \omega_3(1 - \exp\{-\omega_4 b_{j_t}\}) + \kappa I_{(\ell)} \ . \tag{12}$$

Corresponding test statistics are reported in Table 2, and the sequence of Wald and LR statistics are plotted in Figure 4. In contrast to the DLR-GARCH model, neither the residual-based tests nor the direct sup-LR test yields evidence of a structural break in the variance of the DLR-$b$ model. In particular, note that the $p$-value for the sup-LR test is 0.18. Hence these results suggest that the DLR-$b$ model provides a structurally coherent characterization of the sharp decline in the GDP volatility.

To illustrate the interpretation of the volatility decline in the context of the DLR-$b$ model, Figure 5 plots its (non-standardized) residuals $e_t = y_t - \hat{m}_t + \hat{\nu} G_{t-2} - \hat{\gamma} y_{t-2}$ together with its estimated conditional standard deviation. For comparison, it also plots the standard deviation under the DLR-GARCH and under the DLR model with a variance specification with only the dummy variable $I_{(82:\text{IV})}$ generating a break at the estimated break date. Like the pure-break specification, the DLR-$b$ model predicts the lowest volatility in the post-war period for the period starting in mid 1980s, even though it starts few quarters later than under the pure-break model. Note also that, in contrast to the pure-break specification, the DLR-$b$ model predicts a comparably low volatility during the 1960s (from 1960:I until 1970:IV). This prediction is consistent with the fact that the series of the DLR-$b$ residuals $e_t$ is comparably smooth during this period and strengthens the notion that the stochastic behavior of GPD in the post-war period was structurally stable. In order to assess the importance of accounting for this low-volatility period in the 1960s, we calculated the contributions of the two regimes spanning 1960:I through 1970:IV to the log-likelihood under the DLR-$b$ model and under the pure-break specification. Under the DLR-$b$ the log-likelihood contributions of the two regimes are -48.82 and -46.53, and under the pure break model -51.61 and -47.89, respectively, indicating that the improvement obtained by allowing for this earlier low-volatility episode is significant.

## 4.2 Encompassing Exercises

For further empirical validation of the DLR-$b$ model, we now turn to two sets of encompassing exercises. The first is designed to assess the model's capability to account for the structural break found under AR specifications of GDP growth. The second is designed to assess the model's ability to account for the general pattern of business-cycle behavior observed during the post-war period.

The implementation of these exercises follows standard practice, as outlined, e.g., by Hendry and Richard (1990). Treating the DLR-$b$ model as a data generating process (DGP), we use it to produce artificial realizations of GDP growth, which we then use to calculate sets of statistics. The degree of correspondence between the statistics obtained artificially with counterparts obtained using the actual data provides an indication of the model's empirical validity.

There are two distinct natural ways our model can be operationalized as a DGP. Under one approach, the latent variables $a_j$ and $b_j$ can remain fixed at a specific set of values throughout the exercise, and artificial data can be generated repeatedly using these fixed values. In this case, the results of the exercise are conditional on the chosen values. Alternatively, a new set of values for the variables can be obtained (as realizations from their corresponding distributions) for each artificial realization of the data. In this case, the results of the exercise are unconditional on the latent variables.

The goal of the first exercise is to determine whether the model is capable of encompassing the specific evidence regarding an apparent structural break presented in Section 2. Since part of what is at issue here is the timing of the apparent break, we generated artificial realizations of GDP growth from the model conditional on the specific set of regime-change dates, regime-drift trajectories and conditional variances estimated from the actual data. Conditional on these specifications, we generated 10,000 artificial realizations of $\{g_t\}$ from the model (each of length 222, as in the actual sample), and applied the battery of tests conducted in Section 2 using the actual data. Specifically, for each realization of $g_t$, we estimated the AR(4) model, and tested for the stability of the AR parameters and the variance of the innovations $\sigma^2$. Rejection probabilities associated with this exercise are presented in Table 3, and the distribution of Wald statistics obtained for the structural stability tests for $\sigma^2$ is illustrated in Figure 6.

The results of this exercise indicate that the DLR-$b$ model is indeed capable of encompassing the evidence of a break that emerges when working with AR representations for GDP growth. Just as in the actual data, AR parameters are typically stable in the artificial samples, while the variance of innovations is not. Regarding AR parameters, the highest rejection frequency for the no-break null hypothesis we obtained for any parameter is 6.3%, for $\phi_1$ using a 10% critical value. In contrast, the rejection frequency obtained for $\sigma^2$ is 99.9%. Regarding the timing of a break in $\sigma^2$, note from Figure 6 that the median of the distribution of Wald-test statistics generated in this experiment closely tracks the statistics obtained using the actual data over time. Both cross the asymptotic 5% critical value at the same date (1978:II); both have a local peak in 1984:I; and both have a second local peak in the early 1990s (1990:II in the actual data, 1991:II in the median generated in the experiment). Thus the performance of the model is solid along this dimension.

The goal of the second exercise is to determine whether the DLR-$b$ model is capable of encompassing the general pattern of business-cycle activity observed in the actual data. One challenge here is that in the U.S., the NBER provides the definitive characterization of business-cycle activity, yet their procedure for defining expansionary and contractionary periods is not a deterministic function of GDP growth. In order to conduct a simulated encompassing exercise we need to design an automated NBER-type dating mechanism based solely on GDP figures. Heuristically, the procedure we used represents our best approximation of how the NBER would map a given sequence of growth-rate observations into a sequence of expansions and recessions. (Hereafter, we refer to a given approximation as a set of DLR dates.) The specifics of the procedure are as follows. First, we represented the NBER's post-war characterization of expansions and recessions as a sequence of zeros and ones (ones representing recession dates). Next, we fit this sequence of zeros and ones to observations of actual post-war growth using a probit model. The specific model we used included a constant, the contemporaneous realization of growth, and one lead and lag of growth (additional leads and lags were insignificant in this specification). Combining the parameter estimates of this model with a sequence of growth realizations yields a corresponding sequence of estimated probabilities of the occurrence of an NBER recession date. We then used the following algorithm to map these probabilities into a sequence of zeros and ones:

- Assign a value of 1 to any date with an associated probability greater than $p$.
- If any date initially assigned a value of 1 fails to have at least one neighboring date that is also 1, reassign this date a value of zero; otherwise leave the assigned value at 1.

Motivation for this algorithm is provided in Figure 7, which illustrates NBER dates as shaded bars in the upper diagram, and corresponding probit probabilities obtained using the actual growth data in the lower diagram. (We assume the latest recession to have ended in 2001:IV, although this designation has not yet been announced officially.) Note that for each of the 10 NBER-defined recessions, corresponding probit probabilities reach virtually one at least once during the recession. Also, probabilities exceeding 50% are rarely observed during expansions, and probabilities exceeding 75% are even more rare. Finally, recession dates never occur in isolation. Thus using probabilities $p$ of 50% and greater, and eliminating isolated occurrences of large probabilities, we obtain close approximations of actual NBER dates. We used two values of $p$ in our encompassing exercises, 50% and 75%, and employed five summary statistics to quantify business-cycle behavior: the number of recessions observed over a 222-quarter period; the average length of expansions; the standard deviation of this length; the average length of recessions; and the standard deviation of this length. In calculating length statistics, we used only complete episodes. Thus, e.g., if a recession was observed to be ongoing at the end of an artificial sample, it was added to the total number of recessions observed in the sample period, but was not used in calculating length statistics. The top panel of Table 4 reports these statistics corresponding to NBER dates, and to DLR dates obtained using the actual growth data. Note that the number of recessions is 10 in all cases, and DLR summary statistics are all within one period of NBER summary statistics.

Unlike in the structural-break exercise, here the specific timing associated with the occurrence of recessions and expansions is not part of the focus of the analysis. Thus in this case we generated artificial data (again each of length 222) both conditional on the specific set of estimated regime-change dates and $a_j$ and $b_j$ values employed above, as well as unconditionally. Results obtained in both cases (means and standard deviations calculated using 10,000 replications obtained both conditionally and unconditionally) are presented in the lower panels of Table 4.

A priori, one might expect the performance of the conditional model to be superior to that of the unconditional model, since the business-cycle characteristics of the actual data weigh heavily in determining the break dates and $a_j$ and $b_j$ values that remain fixed throughout the exercise in this case. This turns out not to be true ex post: the performance of the unconditional model is on par with that of the conditional model. In both cases the performance is again solid. In fact, in both sets of simulations, the means obtained for all five summary statistics using both values of $p$ are all within one standard deviation of their empirical counterpart. Moreover, this result is not due to a tendency for the standard deviations to be particularly large. For example, using $p = 75\%$, the mean (standard deviation) number of recessions obtained conditionally is 9.21 (1.73), and unconditionally is 9.38 (2.31). The largest standard deviations we obtain are for the statistic "average expansion length" (4.78 quarters conditionally, 6.9 unconditionally, using $p = 75\%$), but in this case the means (19.59, 19.99) are particularly close to their empirical counterpart (19.89). Thus once again, the DLR-$b$ model is capable of closely encompassing patterns of behavior observed in the actual data.

Of course, these results do not establish the DLR-$b$ model as providing a definitive characterization of the data. Rather, it represents one potential structural characterization of the volatility reduction observed in GDP growth that, in addition, is capable of accounting for a the general pattern of cyclical activity. According to this characterization, the recent extended period of relative economic tranquility we have enjoyed does not represent a permanent break from the relatively volatile past. Rather, it represents the realization of a particular regime sequence with tenuous longevity and slowly moving regime-drift component. Only time will tell whether this interpretation is ultimately borne out.

## 5 Conclusion

We have presented a non-linear regime-switching model of GDP growth that, in contrast to AR specifications, is structurally stable over the entire post-war period. In addition, the model is capable of encompassing the pattern of structural instability observed in AR specifications of growth, along with general patterns of business-cycle activity. These results convey a cautionary message. The structural instability of AR specifications implicitly associates an air of permanence with the period of relatively tranquil economic activity we have enjoyed over the past two decades. In contrast, viewed through the lens of our model, this tranquility is associated merely with the realization of a particular regime sequence with tenuous longevity, similar to the earlier low-volatility episode of the 1960s.

## References


Andrews, D.W.K. (1993). "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica,* vol. 61: 821-856.

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

Blanchard, O. and Simon, J. (2001). "The Long and Large Decline in U.S. Output Volatility." *Brooking Papers on Economic Activity, 2001:1*: 135-164.

DeJong, D., Liesenfeld, R. and Richard, J.-F. (2003). "A Non-Linear Forecasting Model of GDP Growth." University of Pittsburgh, Department of Economics, manuscript.

Hansen, B.E. (1997). "Approximate Asymptotic p-Values for Structural Change Tests." *Journal of Business and Economic Statistics,* vol. 15: 60-67.

Hendry, D.F. and Richard, J.-F. (1990). "Recent Developments in the Theory of Encompassing." in Cornet, B. and Tulkens, H. (eds), *Contributions to Operation Research and Economics.* MIT Press, Chapter 12.

Kim, C.J. and Nelson, C.R.(1999). "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of Business Cycle." *Review of Economics and Statistics,* vol. 81: 1-10.

Lamoureux, C.G. and W. Lastrapes (1990). "Persistence in the Variance, Structural Change, and the GARCH Model." *Journal of Business and Economic Statistics,* vol. 8: 225-234.

McConnell, M.M. and Perez-Quiros, G. (2000). "Output Fluctuations in the United States: What has Changed Since the Early 1980's?" *American Economic Review,* vol. 90: 1464-1476.

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

Stock, J.H. and Watson, M.W. (2002). "Has the Business Cycle Changed and Why?" Havard University, Department of Economics, manuscript.

*Table 1. Summary Statistics*

| | Mean | Std. Dev. | |
|---|---|---|---|
| Real GDP Growth | | | |
| Full sample | 3.35 | 4.05 | |
| 1947:II–1984:I | 3.50 | 4.73 | |
| 1984:II–2002:III | 3.04 | 2.13 | |
| AR(4) Residuals | | | |
| Full sample | 0.00 | 3.74 | |
| 1948:II–1984:I | 0.11 | 4.42 | |
| 1984:II–2002:III | −0.21 | 1.99 | |

Structural break tests in $g_t = \mu + \phi_1 g_{t-1} + \phi_2 g_{t-2} + \phi_3 g_{t-3} + \phi_4 g_{t-4} + \epsilon_t, \quad \mathrm{E}(\epsilon_t^2) = \sigma^2$

| Null | sup-Wald | exp-Wald | Break Date |
|---|---|---|---|
| constant $\sigma^2$ | 33.09 | 14.04 | 1984:I |
| | (0.00) | (0.00) | |
| constant $\mu$ | 2.22 | 0.44 | none |
| | (0.73) | (0.49) | |
| constant $\phi_1$ | 2.55 | 0.38 | none |
| | (0.65) | (0.55) | |
| constant $\phi_2$ | 1.04 | 0.14 | none |
| | (0.99) | (0.93) | |
| constant $\phi_3$ | 1.41 | 0.17 | none |
| | (0.93) | (0.84) | |
| constant $\phi_4$ | 4.09 | 0.65 | none |
| | (0.36) | (0.34) | |

*NOTE:* Asymptotic $p$-values are in parentheses.

*Table 2. ML Parameter Estimates and Break Tests*

| Parameter | Specification for the Variance $\sigma_t^2$: $\omega_0+\omega_1\sigma_{t-1}^2+\omega_2 y_{t-1}^2$ | | $\omega_0+\kappa I_{(82:\text{IV})}$ | | $\omega_0+\omega_3(1-e^{-\omega_4 b_j})$ | |
|---|---|---|---|---|---|---|
| $\nu$ | 0.235 | (0.068) | 0.221 | (0.061) | 0.235 | (0.057) |
| $\gamma$ | 0.288 | (0.090) | 0.320 | (0.088) | 0.314 | (0.078) |
| $\alpha_0$ | 0.579 | (0.836) | 0.396 | (0.588) | 0.476 | (0.636) |
| $\alpha_1$ | 0.001 | (0.960) | 0.217 | (0.759) | 0.105 | (0.802) |
| $\sigma_a$ | 1.262 | (0.698) | 1.692 | (0.428) | 1.796 | (0.551) |
| $\sigma_b$ | 0.225 | (0.096) | 0.203 | (0.092) | 0.204 | (0.087) |
| $\ln d$ | 0.110 | (0.024) | 0.113 | (0.024) | 0.117 | (0.021) |
| $\beta_0$ | 6.585 | (1.199) | 6.585 | (1.227) | 6.585 | (1.231) |
| $\beta_1$ | 0.546 | (0.112) | 0.546 | (0.117) | 0.546 | (0.116) |
| $\omega_0$ | 0.101 | (0.110) | 2.860 | (0.512) | 2.088 | (0.652) |
| $\omega_1$ | 0.917 | (0.034) | | | | |
| $\omega_2$ | 0.039 | (0.019) | | | | |
| $\omega_3$ | | | | | 16.178 | (7.707) |
| $\omega_4$ | | | | | 4.300 | (3.576) |
| $\kappa$ | | | 9.896 | (1.820) | | |
| Log-Lik. | −589.80 | | −580.05 | | −582.93 | |
| sup-Wald | 4.17 | [0.350] | | | 1.76 | [0.845] |
| exp-Wald | 0.43 | [0.497] | | | 0.13 | [0.942] |
| sup-LR | 19.49 | [0.000] | | | 5.81 | [0.175] |

*NOTE*: Estimates obtained using the tension index $G_t = 0.65 \cdot G_{t-1} + g_t - g_t^*$, where $g_t^*$ is obtained from a cubic time trend. The regime-change dates used are: 1949:IV, 1950:III, 1954:I, 1955:I, 1958:I, 1959:II, 1960:IV, 1966:I, 1970:IV, 1973:I, 1975:I, 1978:II, 1982:I, 1984:I, 1991:I, 1999:IV. Asymptotic standard errors obtained from a numerical approximation to the Hessian are given in parentheses. Asymptotic $p$-values are given in brackets.

*Table 3. Sup-Wald Structural Break Tests in an AR(4) Under the DLR-b Model*

Rejection frequencies (%) for structural break tests in
$g_t = \mu + \phi_1 g_{t-1} + \phi_2 g_{t-2} + \phi_3 g_{t-3} + \phi_4 g_{t-4} + \epsilon_t, \quad \mathrm{E}(\epsilon_t^2) = \sigma^2$

| Null | | 10% Significance level | 5% Significance level |
|---|---|---|---|
| constant | $\sigma^2$ | 99.9 | 99.9 |
| constant | $\mu$ | 0.5 | 0.1 |
| constant | $\phi_1$ | 6.3 | 2.5 |
| constant | $\phi_2$ | 5.8 | 2.6 |
| constant | $\phi_3$ | 5.3 | 2.5 |
| constant | $\phi_4$ | 4.7 | 2.1 |

*NOTE*: Rejection frequencies are calculated using 10,000 artificial trajectories of length 222 from the DLR-$b$ model, conditional on the $a_j$- and $b_j$-processes estimated from the data.

*Table 4. Business-Cycle Characteristics Under the DLR-b Model*

Actual data

| Dating scheme | Number of recessions | Average expansion length | Std. dev. of expansion length | Average recession length | Std. dev. of recession length |
|---|---|---|---|---|---|
| NBER dates | 10.00 | 18.78 | 12.67 | 4.30 | 1.16 |
| DLR-Dates ($p = 0.50$) | 10.00 | 19.22 | 11.87 | 3.90 | 1.20 |
| DLR-Dates ($p = 0.75$) | 10.00 | 19.89 | 12.13 | 3.30 | 1.06 |

Under the DLR-$b$ model

| Simulation model | Number of recessions | Average expansion length | Std. dev. of expansion length | Average recession length | Std. dev. of recession length |
|---|---|---|---|---|---|
| | $p = 0.50$ | | | | |
| conditional DLR-$b$ | 10.80 | 16.76 | 13.61 | 3.60 | 1.66 |
| | (1.77) | (3.47) | (2.89) | (0.50) | (0.49) |
| unconditional DLR-$b$ | 10.92 | 16.49 | 14.04 | 3.52 | 1.56 |
| | (2.47) | (5.30) | (5.39) | (0.52) | (0.58) |
| | $p = 0.75$ | | | | |
| conditional DLR-$b$ | 9.21 | 19.59 | 15.32 | 3.37 | 1.47 |
| | (1.73) | (4.78) | (4.23) | (0.50) | (0.50) |
| unconditional DLR-$b$ | 9.38 | 19.99 | 16.05 | 3.22 | 1.29 |
| | (2.31) | (6.90) | (6.68) | (0.47) | (0.50) |

*NOTE*: DLR dates approximate NBER dates as follows: approximate NBER recession probabilities by combining growth observations with estimates from a probit specification; define a recession at each date for which the corresponding probit probability is greater that $p$, so long as this is also the case for an immediate neighbor. The probit specification was obtained by fitting NBER-defined recessions/expansions (1s and 0s) to a constant, contemporaneous growth, and one lead and lag of growth (actual growth data). The statistics under the DLR-$b$ model are the means calculated using 10,000 artificial trajectories of length 222 from the DLR-$b$ model, conditional on the $a_j$- and $b_j$-processes estimated from the data (conditional model) and unconditional on specific $a_j$- and $b_j$-processes (unconditional model). Standard deviations of the statistics calculated over the 10,000 trajectories are given in parentheses.

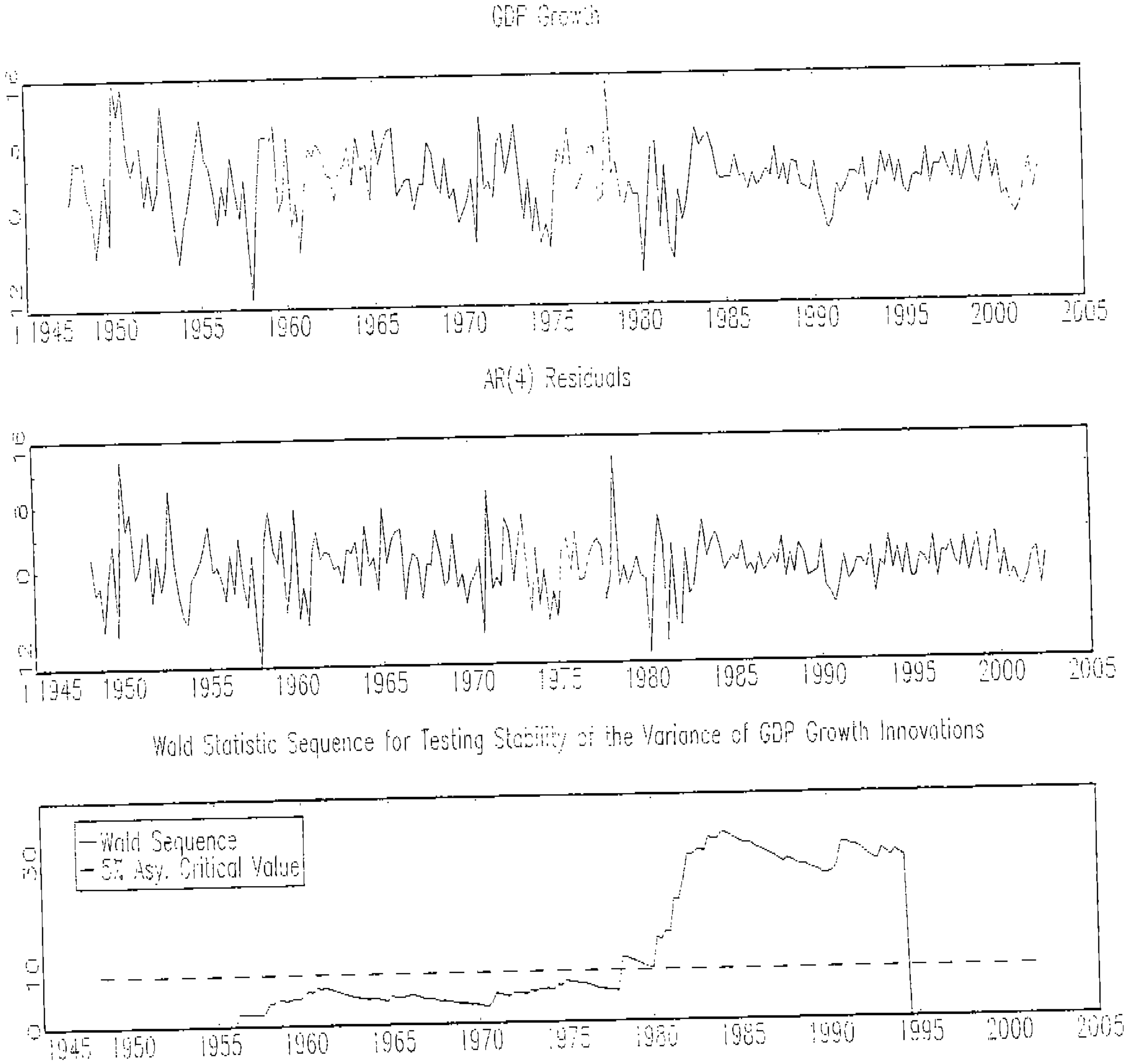


*Figure 1.* GDP growth measured as logged differences in quarterly GDP, measured in chain-weighted 1996 prices, annualized by multiplying by 400, spanning 1947:I through 2002:III (upper panel); residuals from a AR(4) with a constant for GDP growth (middle panel); Wald statistic sequence for testing for a structural break in the variance of the AR(4) fitted to the GDP growth rate (bottom panel).

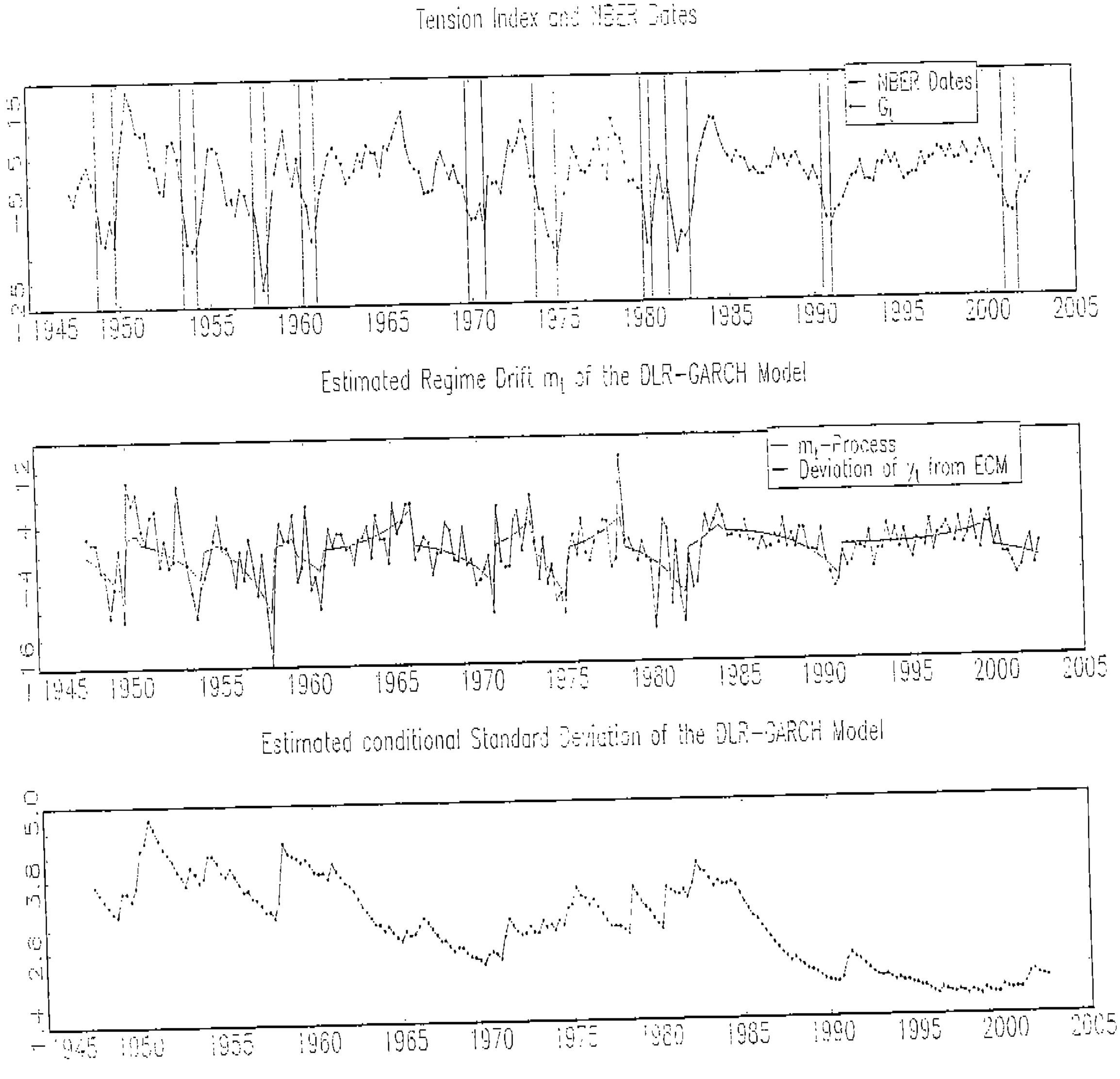


*Figure 2.* Tension index $G_t$ generated by $\delta = 0.65$ and the dates where NBER-recessions begin and end (upper panel); smoothed estimates of the stochastic regime drift $m_t$ and deviations of growth from the estimated ECM component $y_t + 0.235G_{t-2} - 0.288y_{t-2}$ under the DLR-GARCH model (middle panel); estimated conditional standard deviation $\sigma_t$ under the DLR-GARCH model (bottom panel).

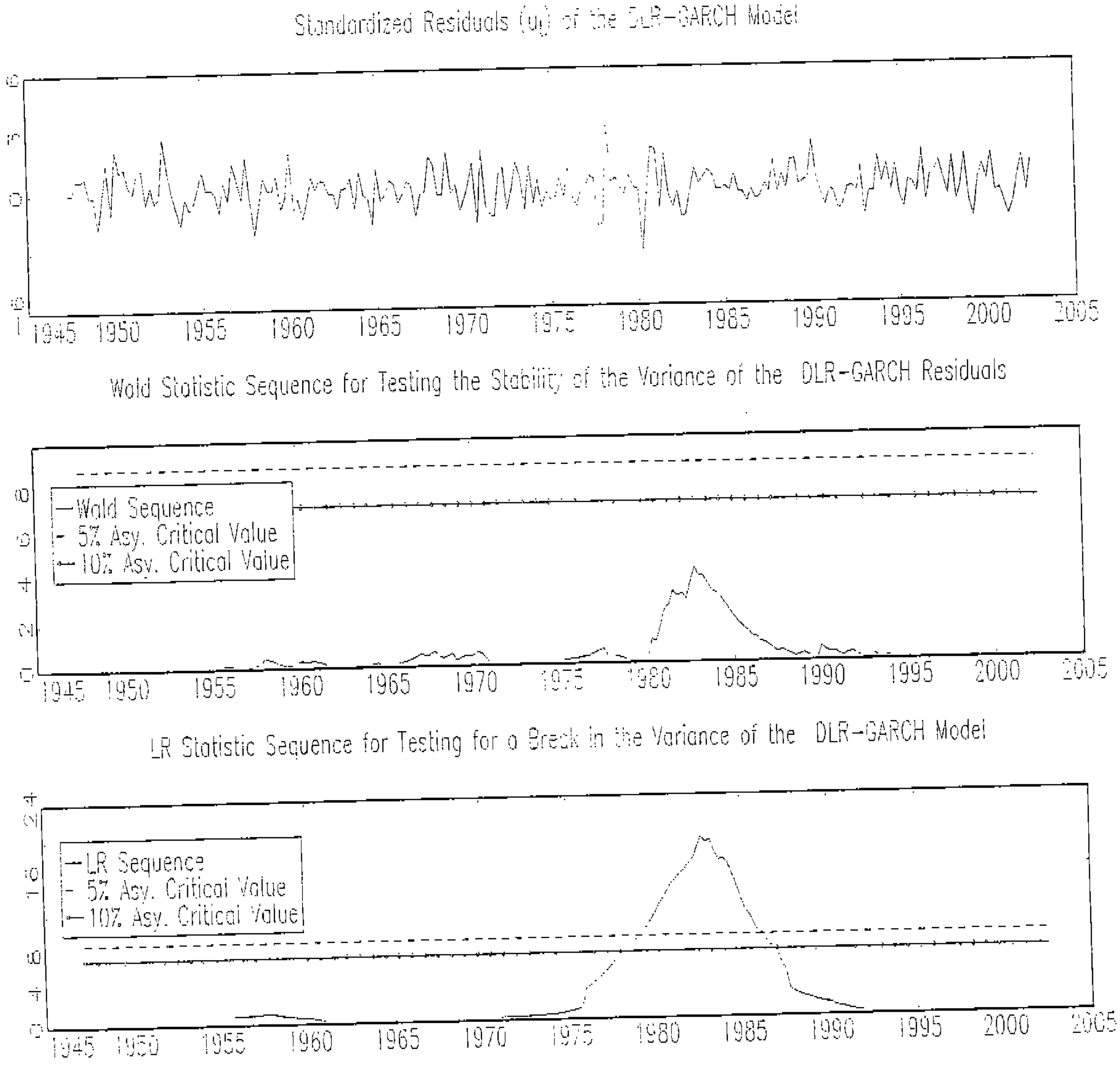


*Figure 3.* Smoothed standardized residuals $u_t$ of the DLR-GARCH model (upper panel); Wald statistic sequence for testing for a structural break in the variance of the AR(1) fitted to the smoothed standardized residuals $u_t$ of the DLR-GARCH model as a function of break date (middle panel); LR statistic sequence for testing for a structural break in the variance equation of the DLR-GARCH model (bottom panel).

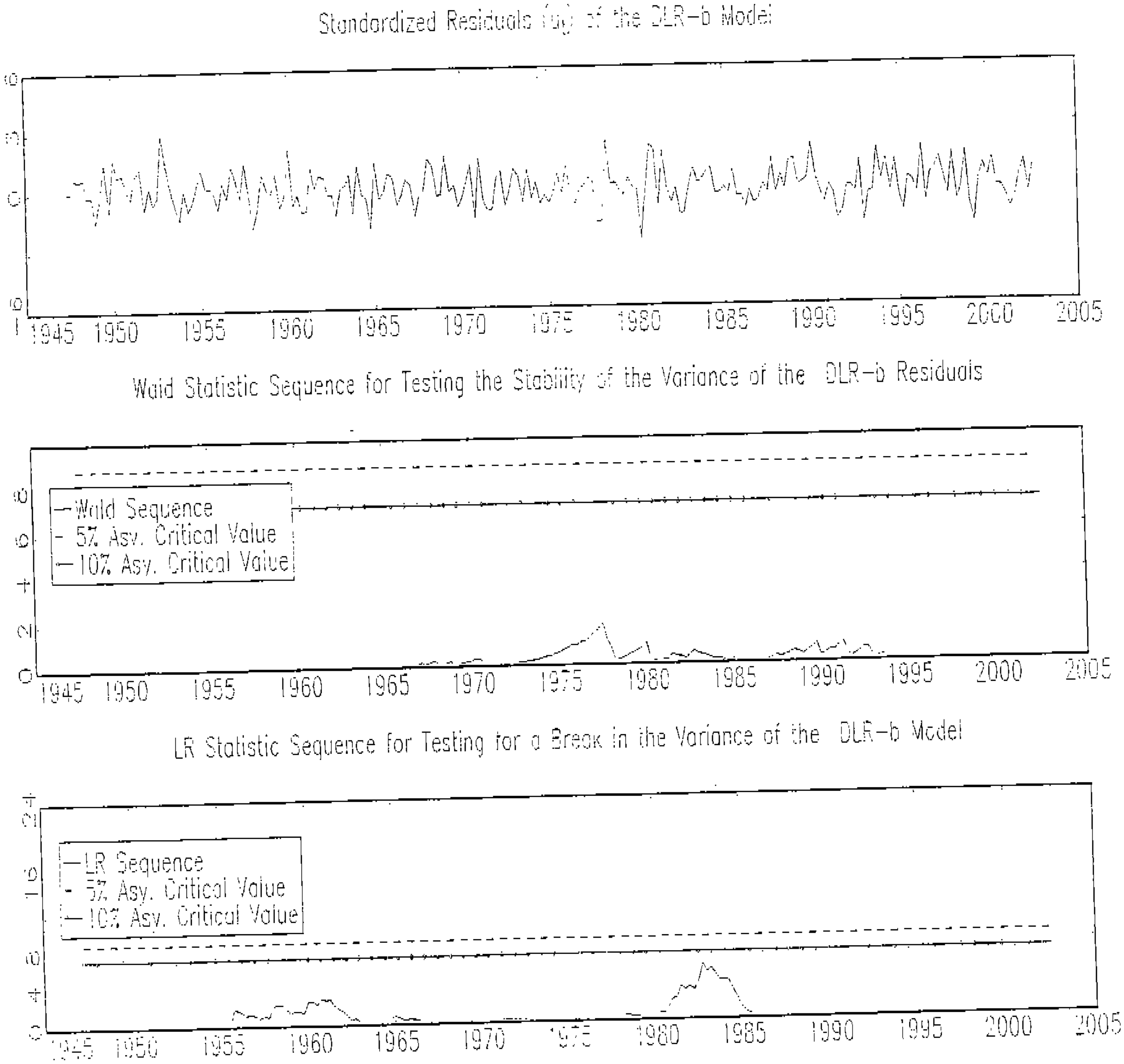


*Figure 4.* Smoothed standardized residuals $u_t$ of the DLR-$b$ model (upper panel); Wald statistic sequence for testing for a structural break in the variance of the AR(1) fitted to the smoothed standardized residuals $u_t$ of the DLR-$b$ model as a function of break date (middle panel); LR statistic sequence for testing for a structural break in the variance equation of the DLR-b model (bottom panel).

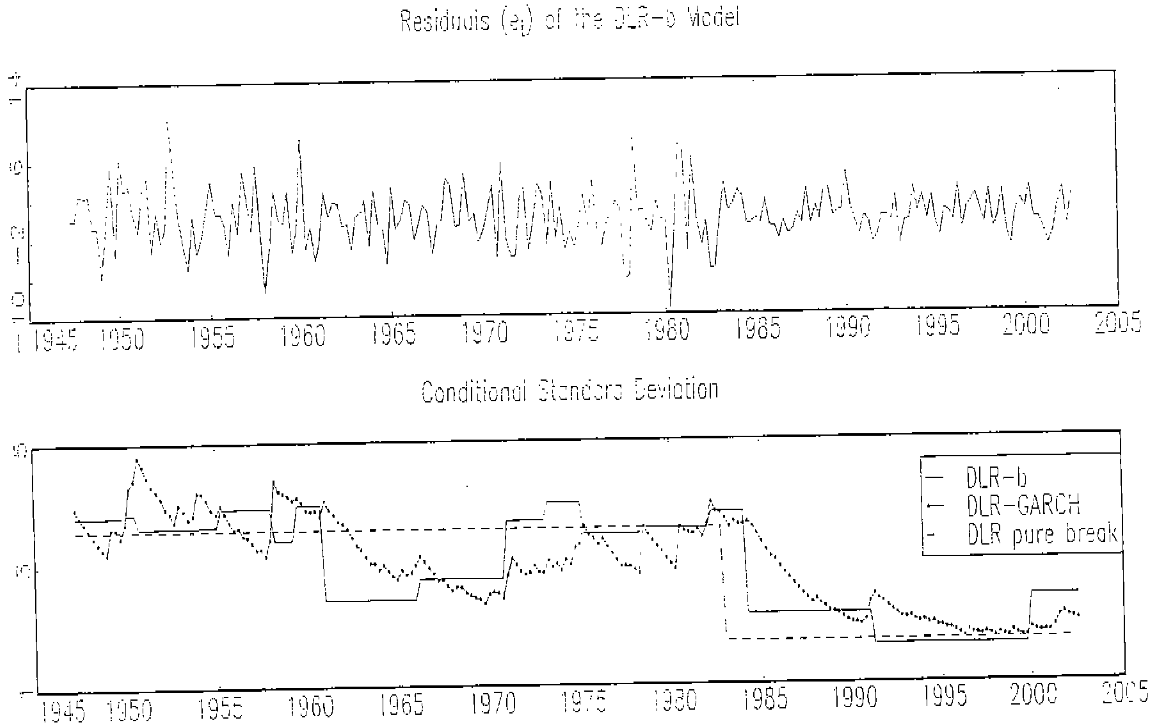


*Figure 5.* Smoothed residuals $e_t$ of the DLR-$b$ model (upper panel); estimated conditional standard deviation under the DLR-$b$, DLR-GARCH, and the DLR model with a pure break in the variance at 1982:IV (bottom panel).

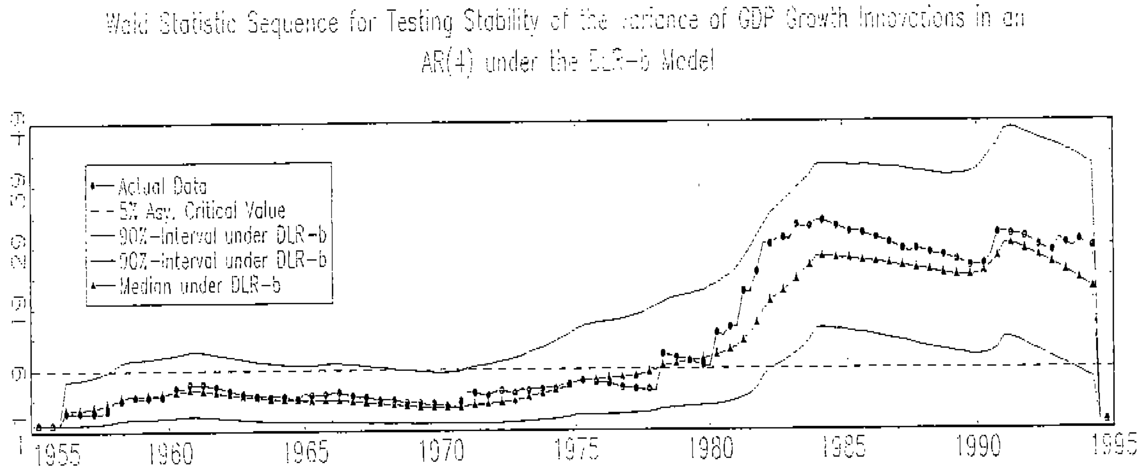


*Figure 6.* Wald statistic sequence for testing for a structural break in the variance of the AR(4) under the DLR-*b* model and for the real GDP growth rate. Median and the 90% interval under the DLR-*b* model are calculated using 10,000 artificial trajectories of length 222 from the DLR-*b* model, conditional on the $a_j$- and $b_j$-processes estimated from the data.

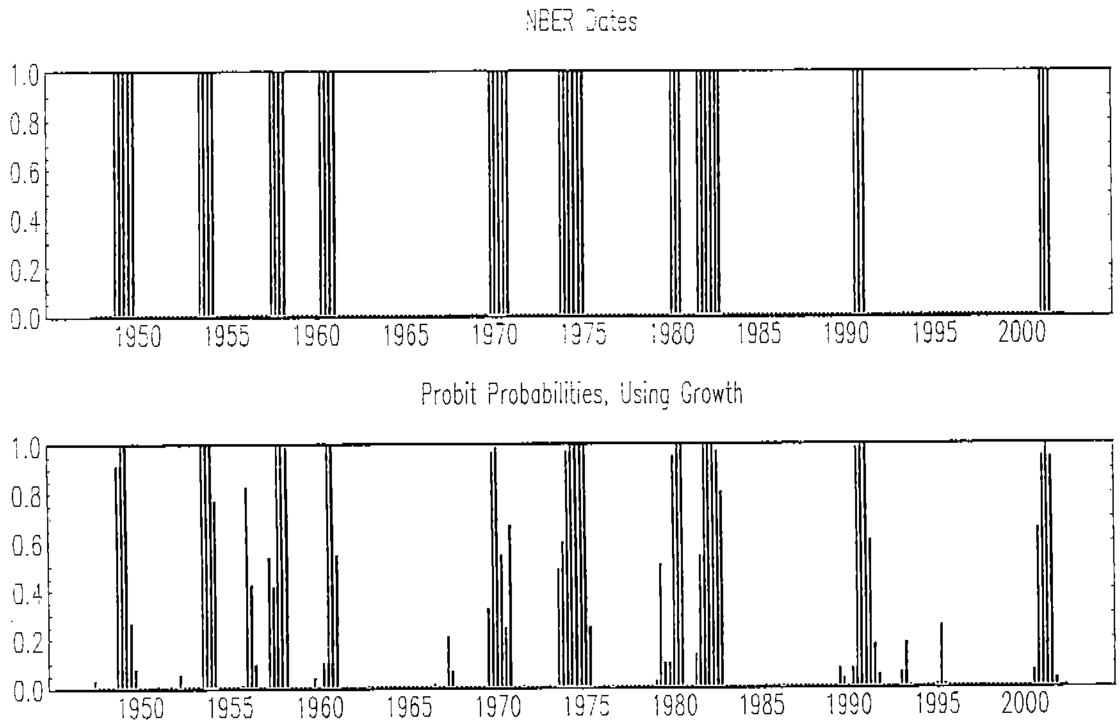


*Figure 7.* Dummy indicating NBER-defined recessions (upper panel); predicted probabilities of a NBER-recession from a fitted probit model based upon contemporaneous growth rate, and one lead and one lag of growth rate (bottom panel).