THE UNIVERSITY OF SYDNEY

Economics Working Paper Series

2011 - 07

Likelihood-Ratio Based Confidence Sets for the Timing of Structural Breaks

Yunjong Eo & James Morley

February 2014

Likelihood-Ratio-Based Confidence Sets for the Timing of Structural Breaks*

Yunjong Eo†
University of Sydney

James Morley‡
University of New South Wales

February 20, 2014

Abstract

We propose the use of likelihood-ratio-based confidence sets for the timing of structural breaks in parameters from time series regression models. The confidence sets are valid for the broad setting of a system of multivariate linear regression equations under fairly general assumptions about the error and regressors and allowing for multiple breaks in mean and variance parameters. In our asymptotic analysis, we determine the critical values for a likelihood ratio test of a break date and the expected length of a confidence set constructed by inverting the likelihood ratio test. Notably, the likelihood-ratio-based confidence sets are more precise than other confidence sets considered in the literature. Monte Carlo analysis confirms their greater precision in finite samples, including in terms of maintaining accurate coverage even when the sample size or magnitude of a break is small. An application to postwar U.S. real GDP and consumption leads to a shorter 95% confidence set for the timing of the "Great Moderation" in the mid-1980s than previously found in the literature. Furthermore, when taking cointegration between output and consumption into account, confidence sets for structural break dates become even shorter and suggest a "productivity growth slowdown" in the early 1970s and an additional large, abrupt decline in long-run growth in the mid-1990s.

Keywords: Inverted Likelihood Ratio; Multiple Breaks; System of Equations; Great Moderation; Productivity Growth Slowdown

JEL classification: C22; C32; E20

*We thank Frank Schorfheide, Elie Tamer, and three anonymous referees for helpful editorial suggestions. Graham Elliott, Junsoo Lee, Alessio Moro, Charles Nelson, Irina Panovska, Jeremy Piger, Werner Ploberger, Barbara Rossi, Rami Tabri, Farshid Vahid, Tao Zha, and seminar and conference participants at Monash University, University of Alabama, Universitat Pompeu Fabra, University of Western Australia, University of Wollongong, the North American Summer Meeting of the Econometric Society, the NBER-NSF Time Series Conference, the Midwest Econometrics Group, and the Society for Nonlinear Dynamics and Econometrics Symposium also provided useful feedback. We acknowledge the use of GAUSS code by Zhongjun Qu and Pierre Perron and can provide compatible code upon request that calculates likelihood-ratio-based confidence sets for structural breaks within their modeling framework. The usual disclaimers apply.

†Yunjong Eo, School of Economics, University of Sydney, Sydney 2006, Australia; Tel: +612 9351 3075; Fax: + 612 9351 4341; Email: yunjong.eo@sydney.edu.au

‡*Corresponding Author:* James Morley, School of Economics, University of New South Wales, Sydney 2052, Australia; Tel: +612 9385 3366; Fax: +612 9313 6337; Email: james.morley@unsw.edu.au

1 Introduction

The exact timing of structural breaks in parameters from time series models is generally unknown *a priori*. Much of the literature on structural breaks has focused on accounting for uncertainty about this timing when testing for the existence of structural breaks (e.g., Andrews (1993)). However, there has also been considerable interest in how to make inference about the timing itself, with an important contribution made by Bai (1997). Employing asymptotic analysis for a slowly-shrinking magnitude of a break, Bai derives the distribution of a break date estimator in a linear time series regression model and uses a related statistic to construct a confidence interval for the timing of a break. One problem with Bai's approach highlighted in a number of studies (e.g., Bai and Perron (2006) and Elliott and Müller (2007)) is that the confidence interval tends to undercover in finite samples, even given a moderately-sized break. Elliott and Müller (2007) propose a different approach based on the inversion of a test for an additional break under the null hypothesis of a given break date and employing asymptotic analysis for a quickly-shrinking magnitude of break. Their approach produces a confidence set (not necessarily an interval) for the timing of a break that has very accurate coverage rates in finite samples, even given small breaks. However, it is only applicable for a single break and the confidence set tends to be quite wide, including when breaks are large.

In this paper, we propose the use of likelihood-ratio-based confidence sets for the timing of structural breaks in parameters from time series regression models. Employing asymptotic analysis for a slowly-shrinking magnitude of break, as in Bai (1997) and originally proposed by Picard (1985), we show that likelihood-ratio-based confidence sets are valid in Qu and Perron's (2007a) broad setting of quasi maximum likelihood estimation for a system of multivariate linear regression equations under fairly general assumptions about regressors and errors. Building on the literature on structural breaks, this setting allows for heteroskedasticity and autocorrelation in the errors, multiple breaks (e.g., Bai and Perron (1998)) in mean and variance parameters (e.g., Bai (2000)), and potentially produces more precise inferences as additional equations are added to the system (e.g., Bai, Lumsdaine, and Stock (1998)). Our asymptotic analysis provides critical values for a likelihood ratio test of a break date and an analytical expression for the expected length of a confidence set based on inverting the likelihood ratio test. Notably, the asymptotic expected length of a likelihood-ratio-based confidence set is generally much shorter than for the corresponding confidence intervals based on the break date estimator, such as Bai's for one break in mean and Qu and Perron's for their broader setting.[1]

Our proposed approach is motivated by Siegmund (1988), who considers confidence sets in the simpler context of a changepoint model of independent Normal observations with a one-time break in mean and assuming known parameters (other than the break date). In particular, we follow Siegmund's suggestion of constructing an inverted likelihood ratio (ILR) confidence set for the break date.[2] Also, our calculation of the asymptotic expected length of an ILR confidence set builds on his analysis in the simpler setting for which he also finds that the ILR confidence set is more precise than for a Wald-type approach along the lines of Bai (1997). Another related study is by Dümbgen (1991), who derives the asymptotic distribution of a break date estimator given independent, but not necessarily Normal observations and proposes inverting a bootstrap version of a likelihood ratio test to construct a confidence set for the break date. More recently, Hansen (2000) proposes the use of ILR confidence sets in the related context of a threshold regression model. However, he maintains the assumption of a stationary threshold variable, thus precluding the use of a deterministic time trend as a threshold variable in order to capture a structural break. Despite a somewhat different setup, our asymptotic analysis builds on Hansen's, in addition to the literature on structural breaks discussed above.

[1]Expected length is more difficult to determine for the confidence set proposed by Elliott and Müller (2007). However, if the asymptotic power for the test of an additional break is strictly less than one when the true break date is within some fixed fraction of the sample period away from the hypothesized break, the expected length of their confidence set will increase with the sample size. This pattern is confirmed in our Monte Carlo analysis, even for a large magnitude of break for which the power of a test for the existence of a break will be high regardless of its timing.

[2]Siegmund (1988) also suggests constructing a confidence set using what can be thought of as the marginal "fiducial distribution" of a break date. In particular, a marginal fiducial distribution of a break date is equivalent to a Bayesian marginal posterior distribution for the break date given a flat prior and integrating out other parameters over the likelihood. The motivation for using a fiducial distribution to construct a frequentist confidence set for a break date, which Siegmund (1988) attributes to Cobb (1978), ultimately comes from Fisher's (1930) idea of using fiducial inference to construct a confidence set for a location parameter. In practice, we find that both methods of constructing sets perform very similarly, but inverting a likelihood ratio test is far more computationally efficient. Thus, we focus on ILR confidence sets in this paper.

We consider a range of Monte Carlo experiments in order to evaluate the finite-sample performance of the competing methods for constructing confidence sets of structural break dates. We allow for both large and small breaks in mean and/or variance, including in the presence of serial correlation, multiple breaks, and a system of equations.[3] The Monte Carlo analysis supports the asymptotic results in the sense that the ILR confidence sets have the shortest average length even in large samples, while at the same time demonstrating accurate, if somewhat conservative, coverage in small samples. Bai's approach and the extension of it to a broader setting by Qu and Perron (2007a) produce confidence intervals that are longer on average, consistent with the asymptotic results, and they tend to undercover in small samples, even for moderately-sized breaks. Meanwhile, as emphasized by Elliott and Müller (2007), their approach always has very accurate coverage in finite samples. However, their confidence sets are much longer on average than for the ILR approach, including for small breaks and especially for larger sample sizes.

To demonstrate the empirical relevance of the shorter expected length of the ILR confidence sets, we apply the various methods to make inference about the timing of structural breaks in postwar U.S. real GDP and consumption. Consistent with the asymptotic and Monte Carlo results, we find the ILR confidence set for the timing of the so-called "Great Moderation" in quarterly output growth is about half the length as for Qu and Perron's approach. Indeed, the 95% ILR confidence set is similar to the 67% confidence interval reported in Stock and Watson (2002) based on Bai's approach.[4] The short length of the ILR confidence set supports the idea that the Great Moderation was an abrupt change in the mid-1980s rather than a gradual reduction in volatility, potentially providing insight into its possible sources (see, Morley (2009)). Meanwhile, when taking cointegration between out-

[3]Following Elliott and Müller (2007), we refer to 'large' breaks as those that can be detected with near certainty using a test for structural instability and 'small' breaks as those that cannot.

[4]Stock and Watson (2002) consider the four-quarter growth rate for U.S. real GDP, rather than the annualized quarterly growth rate, as considered here. They discuss that because they use Bai's approach by regressing the absolute value of residuals from an autoregression of real GDP growth on a constant and allowing a break in the constant from the auxiliary regression, the break estimator has a non-Normal and heavy-tailed distribution, and the 95% confidence interval would be very wide, hence their reporting of the 67% interval. Meanwhile, our ILR confidence sets are much more similar to the 95% credibility set for the timing of the Great Moderation found in Kim, Morley, and Piger (2008) based on the marginal posterior distribution of the break date given a flat/improper prior for the parameters of a linear time series regression model, which is computationally (but not conceptually) equivalent to the approach based on a marginal fiducial distribution suggested by Siegmund (1988).

put and consumption into account, confidence sets for structural break dates are even more precise, consistent with the findings in Bai, Lumsdaine, and Stock (1998). In addition to the Great Moderation, we find evidence of a large decline in the long-run growth rate of the U.S. economy in the early 1970s, corresponding to the "productivity growth slowdown", and another abrupt decline in long-run growth in the mid-1990s that has not, to our knowledge, been documented in the literature before.

The rest of this paper is organized as follows. Section 2 establishes the asymptotic properties of the likelihood-ratio-based confidence sets for the timing of structural breaks in parameters from time series regression models. Section 3 presents Monte Carlo analysis comparing the finite-sample performance of the likelihood-ratio-based approach to the widely-used methods developed by Bai (1997), Qu and Perron (2007a), and Elliott and Müller (2007). Section 4 provides an application to the timing of structural breaks in postwar U.S. real GDP and consumption. Section 5 concludes.

2 Asymptotics

In this section, we make explicit some assumptions for which a likelihood-ratio-based confidence set of a structural break date is asymptotically valid. In particular, we consider Qu and Perron's (2007a) broad setting of a system of multivariate linear regression equations with possible multiple breaks in mean and variance parameters. However, it should be emphasized that this setting encompasses the simpler univariate and single-equation models that are often considered in structural break analysis (see, for example, Bai (1997) and Bai and Perron (1998, 2003)).

Our asymptotic analysis proceeds as follows: First, we present the general model and assumptions. Second, we discuss quasi maximum likelihood estimation of the model and establish results for the asymptotic distribution of the likelihood ratio test of a break date and a confidence set for the break date based on inverting the likelihood ratio test.

2.1 Model and Assumptions

We consider a multivariate regression model with multiple structural changes in the regression coefficients and/or the covariance matrix of the errors. The model is assumed to have n equations with $t = 1, ..., T$ observations for which there are m structural breaks (i.e. $m+1$ regimes) at break dates $\tau = (\tau_1, \dots, \tau_m)$.

Following the notation of Qu and Perron (2007a), the model in the jth regime for $j = 1, ..., m+1$ is given by

$$y_t = (I_n \otimes z_t')S\beta_j + u_t, \text{ for } \tau_{j-1} < t \leq \tau_j, \tag{1}$$

where y_t is a $n \times 1$ vector, $z_t = (z_{1t}, \dots, z_{qt})'$ is a $q \times 1$ vector of regressors, β_j is a $p \times 1$ vector of regression coefficients, and u_t is a $n \times 1$ vector of errors with mean 0 and covariance matrix Σ_j. The matrix S is a selection matrix for regressors z_t. It consists of 0 or 1 elements and has the dimension $nq \times p$ with full column rank.[5] Also, it is possible to impose a set of h cross- and within-equation restrictions across or within structural regimes in the general form of

$$g(\beta, vec(\Sigma)) = 0,$$

where $\beta = (\beta_1, \dots, \beta_{m+1})$, $\Sigma = (\Sigma_1, \dots, \Sigma_{m+1})$, and $g(\cdot)$ is an h-dimensional vector. For notational simplicity, we can rewrite (1) as

$$y_t = x_t'\beta_j + u_t, \tag{2}$$

where the $p \times n$ matrix x_t is defined by $x_t' = (I_n \otimes z_t')S$.

In developing our asymptotic results, we closely follow the assumptions in Bai (1997,

[5]For example, suppose there are two equations ($n = 2$) and three regressors ($q = 3$). If the first and second regressors are used in the first equation and the first and third regressors are used in the second equation, the selection matrix S would be specified as follows:

$$S = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Note that, if all the regressors are included in each equation, $nq = p$ and $S = I_p$.

2000) and Qu and Perron (2007a). Let $||X||_r = \left(\sum_i \sum_j E|X_{ij}|^r\right)^{1/r}$ for $r \geq 1$ denote the L_r norm of a random matrix X, $< \cdot >$ denote the usual inner product, $\lambda_{min}(\Sigma)$ and $\lambda_{max}(\Sigma)$ denote the smallest and largest eigenvalues of Σ respectively, and $[\cdot]$ denote the greatest integer function. Also, let the true values of the parameters be denoted with a superscript 0. Then, the assumptions are given as follows:

Assumption 1 *$\tau_j^0 = [T\lambda_j^0]$ for $j = 1, \ldots, m+1$ with $0 < \lambda_1^0 < \cdots < \lambda_m^0 < 1$.*

Assumption 2 *For each $j = 1, \ldots, m+1$ and $l_j \leq \tau_j^0 - \tau_{j-1}^0$, $(1/l_j) \times \sum_{t=\tau_{j-1}^0+1}^{\tau_{j-1}^0+l_j} x_t x_t' \xrightarrow{a.s.} H_j^0$ as $l_j \to \infty$ with H_j^0 a nonrandom positive definite matrix not necessarily the same for all j. In addition, for $\Delta\tau_j^0 = \tau_j^0 - \tau_{j-1}^0$, as $\Delta\tau_j^0 \to \infty$, uniformly in $s \in [0,1]$, $(1/\Delta\tau_j^0) \times \sum_{t=\tau_{j-1}^0+1}^{\tau_{j-1}^0+[s\Delta\tau_j^0]} x_t x_t' \underset{p}{\longrightarrow} sH_j^0$.*

Assumption 3 *There exists $l_0 > 0$ such that for all $l > l_0$, the matrices $(1/l) \times \sum_{t=\tau_j^0+1}^{\tau_j^0+l} x_t x_t'$ and $(1/l) \times \sum_{t=\tau_j^0-l}^{\tau_j^0} x_t x_t'$ have the minimum eigenvalues bounded away from zero for all $j = 1, \ldots, j$.*

Assumption 4 *The matrix $\sum_{t=k}^{l} x_t x_t'$ is invertible for $l - k \geq k_0$ for some $0 < k_0 < \infty$.*

Assumption 5 *If $x_t u_t$ is weakly stationary within each segment, then*

(a) $\{x_t u_t, \mathcal{F}_t\}$ form a strongly mixing ($\alpha - mixing$) sequence with size $-4r/(r-2)$ for some $r > 2$ for $\mathcal{F}_t = \sigma - fields\,\{\ldots, x_{t-1}, x_t, \ldots, u_{t-2}, u_{t-1}\}$,

(b) $E(x_t u_t) = 0$ and $||x_t u_t||_{2r+\delta} < M < \infty$ for some $\delta > 0$ and,

(c) letting $S_{k,j}(l) = \sum_{t=\tau_{j-1}^0+l+1}^{\tau_{j-1}^0+l+k} x_t u_t, j = 1, \ldots, m+1$, for each $e \in R^n$ of length 1, var($< e, S_{k,l}(0) >) \geq v(k)$ for some function $v(k) \to \infty$ as $k \to \infty$.

Or, if $x_t u_t$ is not weakly stationary within each segment, assume (a)-(c) holds and, in addition, there exists a positive definite matrix $\Omega = [w_{i,s}]$ such that, for any $i, s = 1, \ldots, p$, we have, uniformly in l, that $|k^{-1}E((S_{k,j}(l)_i S_{k,j}(l)_s) - w_{i,s}| \leq C_2 k^{-\psi}$ for some C_2 and $\psi > 0$.

Assumption 6 *Assumption 5 holds with $x_t u_t$ replaced by u_t or $u_t u_t' - \Sigma_j^0$ for $\tau_{j-1} < t \leq \tau_j$ $(j = 1, \ldots, m+1)$.*

Assumption 7 *The magnitudes of the shifts satisfy* $\Delta\beta_{T,j} = \beta^0_{T,j+1} - \beta^0_{T,j} = \upsilon_T\delta_j$, *and* $\Delta\Sigma_{T,j} = \Sigma^0_{T,j+1} - \Sigma^0_{T,j} = \upsilon_T\Phi_j$ *where* $(\delta_j, \Phi_j) \neq 0$ *and they are independent of T. Moreover,* υ_T *is a sequence of positive numbers that satisfy* $\upsilon_T \to 0$ *and* $T^{1/2}\upsilon_T/(logT)^2 \to \infty$. *(Note that, for simplicity, we use* β^0_j *and* Σ^0_j *from now on, suppressing the subscript* T.*)*

Assumption 8 $(\beta^0, \Sigma^0) \in \overline{\Theta}$ *with* $\overline{\Theta} = \{(\beta, \Sigma) : ||\beta|| \leq c_1, \lambda_{min}(\Sigma) \geq c_2, \lambda_{max}(\Sigma) \leq c_3\}$ *for some* $c_1 \leq \infty$, $0 < c_2 \leq c_3 < \infty$.

While building off of earlier work by Bai (1997, 2000), this particular formulation of assumptions is drawn directly from Qu and Perron (2007a) and is discussed in detail in their paper. However, we provide a brief explanation here. Assumption 1 restricts the break dates to be asymptotically distinct. Assumption 2 is used for the central limit theorem and allows the regressors to have different distributions across regimes, although it excludes unit root regressors and trending regressors. Assumption 3 requires that there is no local collinearity in the regressors near the break dates. Assumption 4 is a standard invertibility condition to ensure well-defined estimates. Assumptions 5 and 6 determine the structure of the $x_t u_t$ and u_t processes and imply short memory for $x_t u_t$ and $u_t u_t'$ with bounded fourth moments. These assumptions guarantee strongly consistent estimates and a well-behaved likelihood function while, at the same time, they are mild in the sense of allowing for substantial heteroskedasticity and autocorrelation and encompassing a wide range of econometric models. Assumption 7 implies that, although the magnitude of structural change shrinks as the sample size increases, it is large enough that we can derive limiting distributions for estimators of break dates that are independent of the exact distributions of regressors and errors. This assumption follows from Picard (1985) and Bai (1997), among many others, although Elliott and Müller (2007) make the assumption that υ_T shrinks at a faster rate in their analysis in order to consider a smaller magnitude of break. Finally, Assumption 8 implies that the data are generated by innovations with a nondegenerate covariance matrix and a finite conditional mean.

2.2 Estimation, Likelihood Ratio, and Likelihood-Ratio-Based Confidence Set

As discussed in Qu and Perron (2007a) and building on the results in Bai, Lumsdaine, and Stock (1998) and Bai (2000), the parameters for the model in (2) can be consistently estimated by restricted quasi maximum likelihood estimation with the likelihood constructed based on the (potentially false) assumption of serially-uncorrelated Normal errors. Specifically, the quasi-likelihood function is

$$L_T(\tau, \beta, \Sigma) = \prod_{j=1}^{m+1} \prod_{t=\tau_{j-1}+1}^{\tau_j} f\left(y_t | x_t; \beta_j, \Sigma_j\right),$$

where

$$f\left(y_t | x_t; \beta_j, \Sigma_j\right) = \frac{1}{(2\pi)^{n/2}|\Sigma_j|^{1/2}} \exp\left\{-\frac{1}{2}(y_t - x_t'\beta_j)\Sigma_j^{-1}(y_t - x_t'\beta_j)\right\}.$$

Let $l_T(\tau, \beta, \Sigma)$ be the natural logarithm of the quasi-likelihood function $L_T(\tau, \beta, \Sigma)$:

$$l_T(\tau, \beta, \Sigma) = \log L_T(\tau, \beta, \Sigma) = \sum_{j=1}^{m+1} \sum_{t=\tau_{j-1}+1}^{\tau_j} \left\{-\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_j| - \frac{1}{2}(y_t - x_t'\beta_j)\Sigma_j^{-1}(y_t - x_t'\beta_j)\right\}.$$

The estimates for (τ, β, Σ) are found by maximizing the quasi-log-likelihood function subject to the restrictions $g(\beta, vec(\Sigma)) = 0$:

$$(\hat{\tau}, \hat{\beta}, \hat{\Sigma}) = \arg\max_{(\tau,\beta,\Sigma)} l_T^r(\tau, \beta, \Sigma), \tag{3}$$

where $l_T^r(\tau, \beta, \Sigma) = l_T(\tau, \beta, \Sigma) + \lambda' g(\beta, vec(\Sigma))$. We also assume that this maximization is taken over all partitions from a set of break dates $\tau = (\tau_1, \ldots, \tau_m) = (T\lambda_1, \ldots, T\lambda_m)$, where $(\lambda_1, \ldots, \lambda_m) \in \Lambda_\epsilon = \{(\lambda_1, \ldots, \lambda_m); |\lambda_{j+1} - \lambda_j| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_m \leq 1 - \epsilon\}$ and ϵ is a trimming fraction that imposes a minimal length for each regime.

Qu and Perron (2007a) establish the same rate of convergence for parameter estimates in this setting as is found in much of the previous literature on structural breaks (e.g., Bai (1997, 2000), Bai and Perron (1998), Bai, Lumsdaine, and Stock (1998)). Specifically, for $j = 1, \ldots, m$, $v_T^2(\hat{\tau}_j - \tau_j^0) = O_p(1)$ and, for $j = 1, \ldots, m + 1$, $\sqrt{T}(\hat{\beta}_j - \beta_j^0) = O_p(1)$ and

$\sqrt{T}(\hat{\Sigma}_j - \Sigma_j^0) = O_p(1)$ (see the proof of Lemma 1 in Qu and Perron (2007b) for more details). Based on this result, we study the limiting distributions by using the restricted log-likelihood function in a compact set of the parameter space in the neighborhood of the true parameter values. In particular, we take the *arg max* of $l_T^r(\tau, \beta, \Sigma)$ over the compact set C_M, where

$$\begin{aligned} C_M &= \{(\tau, \beta, \Sigma) : v_T^2 |\tau_j - \tau_j^0| \leq M \text{ for } j = 1, \ldots, m, \\ & \quad |\sqrt{T}(\beta_j - \beta_j^0)| \leq M, |\sqrt{T}(\Sigma_j - \Sigma_j^0)| \leq M \text{ for } j = 1, \ldots, m+1\} \end{aligned}$$

and M is a fixed positive number that is large enough to be equivalent to taking the *arg max* in an unrestricted set because the estimates will fall in C_M with probability arbitrarily close to 1 (see also Lemma 1 in Qu and Perron (2007b)).

Motivated by Siegmund (1988), we propose confidence sets for the break dates $(\tau_1, \ldots, \tau_m)$ based on inverting likelihood ratio tests. Let $l_j^r(\tau_j)$ denote the natural logarithm of the profile likelihood function for the jth break date subject to the restrictions $g(\beta, vec(\Sigma)) = 0$:

$$\begin{aligned} l_j^r(\tau_j) &= l_j^r(\tau_j, \hat{\beta}_j(\tau_j), \hat{\Sigma}_j(\tau_j), \hat{\beta}_{j+1}(\tau_j), \hat{\Sigma}_{j+1}(\tau_j)) \\ &= \max_{(\beta_j, \Sigma_j, \beta_{j+1}, \Sigma_{j+1})} \sum_{t=\hat{\tau}_{j-1}+1}^{\tau_j} \left\{ -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_j| - \frac{1}{2}(y_t - x_t'\beta_j)\Sigma_j^{-1}(y_t - x_t'\beta_j) \right\} \\ &+ \sum_{t=\tau_j+1}^{\hat{\tau}_{j+1}} \left\{ -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_{j+1}| - \frac{1}{2}(y_t - x_t'\beta_{j+1})\Sigma_{j+1}^{-1}(y_t - x_t'\beta_{j+1}) \right\} + \lambda' g(\beta, vec(\Sigma)), \end{aligned}$$

Given this profile likelihood, we construct a $1 - \alpha$ confidence set for the jth break date by inverting the following α-level likelihood ratio test of $H_0 : \tau_j = \tau_j^0$ sequentially for different values of τ_j:

$$LR_j(\tau_j) = -2\left[l_j^r(\tau_j) - l_j^r(\hat{\tau}_j)\right], \tag{4}$$

where $l_j^r(\hat{\tau}_j) = \max_{\tau_j} \; l_j^r(\tau_j)$ and $\hat{\tau}_j = \arg\max_{\tau_j} \; l_j^r(\tau_j)$.

In practice, in order to construct confidence sets for break dates by inverting likelihood ratio tests, we first need consistent estimates of the number of breaks, $\hat{m}$, the break dates $(\hat{\tau}_1, ..., \hat{\tau}_j, ..., \hat{\tau}_m)$, and the regression parameters under the alternative. We obtain these based on the procedures in Qu and Perron (2007a). Given these estimates, we then proceed as follows for each break $j = 1, ..., \hat{m}$:

Step 1: Calculate the critical value, $\kappa_{\alpha,j}$, for an α-level likelihood ratio test of a break date (see Proposition 1 below).

Step 2: Invert a sequence of tests for all hypothesized dates τ_j^h within the trimmed subsample $(\hat{\tau}_{j-1} + \epsilon T, \hat{\tau}_{j+1} - \epsilon T)$ by computing the likelihood ratio value $LR_j(\tau_j^h)$ in (4) where ϵ is the same trimming fraction used in estimation and the relevant regression parameters are re-estimated for each hypothesized date.[6]

Step 3: Include the hypothesized date τ_j^h in the level $1 - \alpha$ confidence set for the jth break date if $LR_j(\tau_j^h) \leq \kappa_{\alpha,j}$ and exclude it otherwise.

For this procedure, we establish some asymptotic results relating to the distribution of the likelihood ratio statistic and the expected length of the likelihood-ratio-based confidence set. Letting $\eta_t = (\eta_{1,t}, \ldots, \eta_{n,t}) = (\Sigma_j^0)^{-1/2} u_t$ for $t \in [\tau_{j-1}^0 + 1, \tau_j^0]$ and $j = 1, \ldots, m$ and assuming that $E[\eta_{k,t}\eta_{l,t}\eta_{h,t}] = 0$ for all k, l, h and for every t, we define the following terms and then present two propositions:

[6] When computing the likelihood ratio for the jth break date, the estimates of the regression parameters and the break dates for the breaks $i \neq j$ are fixed.

$$
\begin{aligned}
B_{1,j} &= (\Sigma_j^0)^{1/2}(\Sigma_{j+1}^0)^{-1}\Delta\Sigma_j(\Sigma_j^0)^{-1/2}, \\
B_{2,j} &= (\Sigma_{j+1}^0)^{1/2}(\Sigma_j^0)^{-1}\Delta\Sigma_j(\Sigma_{j+1}^0)^{-1/2}, \\
Q_{1,j} &= \underset{T\to\infty}{plim}(\tau_j^0-\tau_{j-1}^0)^{-1}\sum_{t=\tau_{j-1}^0+1}^{\tau_j^0} x_t(\Sigma_{j+1}^0)^{-1}x_t', \\
Q_{2,j} &= \underset{T\to\infty}{plim}(\tau_{j+1}^0-\tau_j^0)^{-1}\sum_{t=\tau_j^0+1}^{\tau_{j+1}^0} x_t(\Sigma_j^0)^{-1}x_t', \\
\Pi_{1,j} &= \lim_{T\to\infty} var\left\{(\tau_j^0-\tau_{j-1}^0)^{-1/2}\left[\sum_{t=\tau_{j-1}^0+1}^{\tau_j^0} x_t(\Sigma_{j+1}^0)^{-1}(\Sigma_j^0)^{1/2}\eta_t\right]\right\}, \\
\Pi_{2,j} &= \lim_{T\to\infty} var\left\{(\tau_{j+1}^0-\tau_j^0)^{-1/2}\left[\sum_{t=\tau_j^0+1}^{\tau_{j+1}^0} x_t(\Sigma_j^0)^{-1}(\Sigma_{j+1}^0)^{1/2}\eta_t\right]\right\}, \\
\Omega_{1,j} &= \lim_{T\to\infty} var\left\{vec\left[(\tau_j^0-\tau_{j-1}^0)^{-1/2}\sum_{t=\tau_{j-1}^0+1}^{\tau_j^0}(\eta_t\eta_t'-I)\right]\right\}, \\
\Omega_{2,j} &= \lim_{T\to\infty} var\left\{vec\left[(\tau_{j+1}^0-\tau_j^0)^{-1/2}\sum_{t=\tau_j^0+1}^{\tau_{j+1}^0}(\eta_t\eta_t'-I)\right]\right\}, \\
\Gamma_{1,j} &= \left(\frac{1}{4}vec(B_{1,j})'\Omega_{1,j}^0 vec(B_{1,j})+\Delta\beta_j'\Pi_{1,j}\Delta\beta_j\right)^{1/2}, \\
\Gamma_{2,j} &= \left(\frac{1}{4}vec(B_{2,j})'\Omega_{2,j}^0 vec(B_{2,j})+\Delta\beta_j'\Pi_{2,j}\Delta\beta_j\right)^{1/2}, \\
\Psi_{1,j} &= \left(\frac{1}{2}tr(B_{1,j}^2)+\Delta\beta_j'Q_{1,j}\Delta\beta_j\right), \\
\Psi_{2,j} &= \left(\frac{1}{2}tr(B_{2,j}^2)+\Delta\beta_j'Q_{2,j}\Delta\beta_j\right).
\end{aligned}
$$

Proposition 1 *Under Assumptions 1-8 with $\Rightarrow$ denoting weak convergence under the Skorohod topology, the likelihood ratio statistic for the jth break date*

$$
LR_j(\tau_j^0) \Rightarrow \xi = \max_v \begin{cases} \omega_{1,j}\left(-|v|+2W_j(v)\right) & for \quad v\in(-\infty,0] \\ \omega_{2,j}\left(-|v|+2W_j(v)\right) & for \quad v\in(0,\infty) \end{cases}, \tag{5}
$$

where $W_j(v)$ is a standard Wiener processes defined on the real line,

$$\omega_{1,j} = \frac{\Gamma_{1,j}^2}{\Psi_{1,j}}, \text{ and } \omega_{2,j} = \frac{\Gamma_{2,j}^2}{\Psi_{2,j}}.$$

The distribution function of ξ is

$$P(\xi \le x) = \left(1 - \exp(-\frac{x}{2\omega_{1,j}})\right)\left(1 - \exp(-\frac{x}{2\omega_{2,j}})\right). \tag{6}$$

Then, using (6) to solve for the critical value $\kappa_{\alpha,j}$ of a α-level likelihood ratio test of a break date, a $1-\alpha$ likelihood-ratio-based confidence set for τ_j is given by

$$C_{j,1-\alpha} = \{\tau_j : LR_j(\tau_j) \le \kappa_{\alpha,j}\}.$$

Proposition 1 establishes the asymptotic distribution of the likelihood ratio test for a break date and shows how to calculate a confidence set based on inverting the likelihood ratio test. Note that the simpler distribution $\max_v -\frac{1}{2}|v| + W(v)$ was studied in Bhattacharya and Brockwell (1976), but the scaling factors $\omega_{1,j}$ and $\omega_{2,j}$ generally make the distribution of the likelihood ratio statistic asymmetric when allowing for different distributions of regressors and/or errors before and after the structural break. Note that $\omega_{1,j}$ and $\omega_{2,j}$ are replaced by consistent estimates from (3) in practice and the calculation of a critical value using (6) is straightforward (see proof for more details). Also, it should be noted that the likelihood-ratio-based confidence set in Proposition 1 is constructed under the assumption that the magnitude of the break $\Delta\beta_{T,j} \to 0$ and $\Delta\Sigma_{T,j} \to 0$ as $T \to \infty$, so the actual coverage should exceed the desired level $1-\alpha$ for a given fixed magnitude of break, at least for Normal errors (see Hansen (2000)).

Proposition 2 *Under Assumptions 1-8, the expected length of a $1-\alpha$ likelihood-ratio-based confidence set is*

$$\begin{aligned} & 2\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2\omega_{1,j}})\right)\left\{\frac{\kappa_{\alpha,j}}{2\omega_{1,j}} - \frac{1}{2}\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2\omega_{1,j}})\right)\right\} \\ + \;\; & 2\left(\Gamma_{2,j}^2/\Psi_{2,j}^2\right)\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2\omega_{2,j}})\right)\left\{\frac{\kappa_{\alpha,j}}{2\omega_{2,j}} - \frac{1}{2}\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2\omega_{2,j}})\right)\right\}. \end{aligned}$$

Proposition 2 establishes the expected length of a $1-\alpha$ likelihood-ratio-based confidence set. The length is calculated by measuring the expected size of the set of τ_j's such that $LR_j(\tau_j) \leq \kappa_{\alpha,j}$. Note that Siegmund (1986, 1988) considers a related calculation using Brownian motion with a break in drift as follows:

$$\begin{aligned} dX(t) &= \mu_1 dt + dW(t), \text{ if } t \leq \tau^0, \\ dX(t) &= \mu_2 dt + dW(t), \text{ if } t > \tau^0, \end{aligned}$$

where $\mu_1 \neq \mu_2$ and the Brownian motion is assumed to approximate the simple changepoint model of independent Normal observations with a one-time break in mean. For his analysis, the magnitude of the break is assumed to be fixed and known, while a variance of unity is also assumed to be known. In this case, Siegmund shows that the likelihood ratio statistic $LR_j(\tau^0)$ can be approximated by the distribution of $\max_r 2\left(-\frac{1}{2}|r| + W(r)\right)$. In our case, by contrast, we derive the asymptotic distribution of the likelihood ratio statistic for a more general setting with parameter and break date estimates that do not depend on the exact distributions of the regressors and the errors. Thus, a shrinking magnitude of break, as in Bai (1997), is required for the development of the limiting theory. Importantly, the distance between the break dates under the null and alternative hypotheses is scaled using a change in variables to obtain the distribution in (6). As a result, v in $\max_v -|v| + 2W(v)$ is not the distance between two break dates. Instead, we calculate expected length based on the distribution of $\max_v -|v| + 2W(v)$ for $v \geq 0$ and $v < 0$, respectively, and rescaled by $\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)$ for $\tau_j \leq \tau_j^0$ and $\left(\Gamma_{2,j}^2/\Psi_{2,j}^2\right)$ for $\tau_j > \tau_j^0$. Thus, the likelihood ratio statistic is invariant to the scales for the break dates, $\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)$ and $\left(\Gamma_{2,j}^2/\Psi_{2,j}^2\right)$, but the distance between break dates is not invariant to the transformations and should be taken into account for the calculation of length (see proof for more details).

In the following two corollaries, we consider simplified cases for either breaks in conditional mean or breaks in variance and solve for the simplified asymptotic distribution of the likelihood ratio statistic for a break date, critical values, and expressions for expected length:

Corollary 1 *Under Assumptions 1-8 and additionally if (i) there are only changes in conditional mean and (ii) the errors form a martingale difference sequence, then for the jth break* *date* $\omega_{1,j} = \omega_{2,j} = 1$ *and*

$$LR_j(\tau_j^0) \Rightarrow \max_v \ -|v| + 2W_j(v) \quad for \quad v \in (-\infty, \infty).$$

Also, the asymptotic critical value of a $1-\alpha$ *likelihood-ratio-based confidence set is*

$$\kappa_{\alpha,j} = -2\log\left(1-(1-\alpha)^{1/2}\right)$$

and the expected length of the confidence set is

$$\left(\frac{1}{\Delta\beta_j' Q_1 \Delta\beta_j} + \frac{1}{\Delta\beta_j' Q_2 \Delta\beta_j}\right) 2(1 - exp(-\frac{\kappa_{\alpha,j}}{2}))\left\{\frac{\kappa_{\alpha,j}}{2} - \frac{1}{2}(1 - exp(-\frac{\kappa_{\alpha,j}}{2}))\right\}$$

or, equivalently,

$$\left(\frac{1}{\Delta\beta_j' Q_1 \Delta\beta_j} + \frac{1}{\Delta\beta_j' Q_2 \Delta\beta_j}\right) 2(1-\alpha)^{1/2}\{-\log[1-(1-\alpha)^{1/2}] - \frac{1}{2}(1-\alpha)^{1/2}\}.$$

Remark 1 *If, in addition to Assumptions in Corollary 1, the distribution of the regressors is stable,* $Q = \Pi_{1,j} = Q_{1,j} = \Pi_{2,j} = Q_{2,j}$ *and* $\omega_{1,j} = \omega_{2,j} = 1$*. Thus, the expected length of the confidence set would further simplify to*

$$\left(\frac{1}{\Delta\beta_j' Q \Delta\beta_j}\right) 4\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2})\right)\left\{\frac{\kappa_{\alpha,j}}{2} - \frac{1}{2}\left(1 - exp(-\frac{\kappa_{\alpha,j}}{2})\right)\right\}$$

or, equivalently,

$$\left(\frac{1}{\Delta\beta_j' Q \Delta\beta_j}\right) 4(1-\alpha)^{1/2}\{-\log[1-(1-\alpha)^{1/2}] - \frac{1}{2}(1-\alpha)^{1/2}\}.$$

The asymptotic critical value is the same as in Corollary 1.

Remark 2 *If we replace the assumption of martingale difference errors in Remark 1 with the assumption that the errors are identically distributed,* $\Pi = lim_{T\to\infty}\, var\left\{T^{-1/2}\left[\sum_{t=1}^T x_t(\Sigma^0)^{-1/2}\eta_t\right]\right\}$, $Q = plim_{T\to\infty} T^{-1}\sum_{t=1}^T x_t(\Sigma^0)^{-1}x_t'$*, and* $\omega_{1,j} = \omega_{2,j} = \omega_j = \frac{\Delta\beta_j'\Pi\Delta\beta_j}{\Delta\beta_j' Q\Delta\beta_j}$*. Thus, the asymptotic* *critical value of a* $1-\alpha$ *likelihood-ratio-based confidence set is*

$$\kappa_{\alpha,j} = -2\omega_j \log\left(1-(1-\alpha)^{1/2}\right)$$

and the expected length of the confidence set is

$$\frac{\Delta\beta_j'\Pi\Delta\beta_j}{(\Delta\beta_j'Q\Delta\beta_j)^2}4\left(1-exp(-\frac{\kappa_{\alpha,j}}{2\omega_j})\right)\left\{\frac{\kappa_{\alpha,j}}{2\omega_j}-\frac{1}{2}\left(1-exp(-\frac{\kappa_{\alpha,j}}{2\omega_j})\right)\right\}$$

or, equivalently,

$$\frac{\Delta\beta_j'\Pi\Delta\beta_j}{(\Delta\beta_j'Q\Delta\beta_j)^2}4(1-\alpha)^{1/2}\{-\log[1-(1-\alpha)^{1/2}]-\frac{1}{2}(1-\alpha)^{1/2}\}.$$

Corollary 2 *Under Assumptions 1-8 and additionally if (i) there are only changes in variance and (ii) the errors are Normally distributed, then for the* j*th break date* $\omega_{1,j}=\omega_{2,j}=1$ *and*

$$LR_j(\tau_j^0) \Rightarrow \max_v \; -|v|+2W_j(v) \quad for \quad v\in(-\infty,\infty).$$

Also, the asymptotic critical value of a $1-\alpha$ *likelihood-ratio-based confidence set is*

$$\kappa_{\alpha,j} = -2\log\left(1-(1-\alpha)^{1/2}\right)$$

and the expected length of the confidence set is

$$\left(\frac{2}{tr(B_1^2)}+\frac{2}{tr(B_2^2)}\right)2(1-exp(-\frac{\kappa_{\alpha,j}}{2}))\left\{\frac{\kappa_{\alpha,j}}{2}-\frac{1}{2}(1-exp(-\frac{\kappa_{\alpha,j}}{2}))\right\}$$

or, equivalently,

$$\left(\frac{2}{tr(B_1^2)}+\frac{2}{tr(B_2^2)}\right)2(1-\alpha)^{1/2}\{-\log[1-(1-\alpha)^{1/2}]-\frac{1}{2}(1-\alpha)^{1/2}\}.$$

In the simplified cases of Corollaries 1 and 2, the critical values for the likelihood ratio test of a break date are the same as reported in Table 1 of Hansen (2000) for a likelihood ratio test of a threshold parameter. These values are 5.94, 7.35 and 10.59 at the 10%, 5%, and 1%

levels, respectively. Meanwhile, the simplified expected length expressions again make use of results in Siegmund (1986, 1988) and allow for easy comparison with the expected lengths of the confidence intervals in Bai (1997) and Qu and Perron (2007a), as is done throughout the next section.

3 Monte Carlo Analysis

In this section, we present extensive Monte Carlo analysis of the finite-sample performance of competing methods for constructing confidence sets of structural break dates. In addition to the likelihood-ratio-based approach proposed in the previous section, we also consider the methods developed by Bai (1997), Qu and Perron (2007a), and Elliott and Müller (2007). For brevity, we omit many of the details of these widely-used methods and encourage interested readers to consult the original papers. However, we provide some background for these other approaches in the following subsection to help motivate our Monte Carlo experiments and facilitate interpretation of our results.

3.1 Widely-Used Methods for Constructing Confidence Sets of Structural Break Dates

Bai (1997) solves for the distribution of the least squares break date estimator using asymptotic analysis for a slowly-shrinking magnitude of break. In terms of the notation in the previous section, he assumes that $v_T \to 0$ and $v_T T^{\varepsilon} \to \infty$ for some $\varepsilon \in (0, 1/2)$ when $\Delta\beta = v_T \delta$, where the break subscript is dropped from $\Delta\beta_j$ and δ_j for convenience given the assumption of only one break. His confidence intervals are constructed based on the asymptotic distribution of this break date estimator. Bai's approach is designed for univariate analysis under fairly general assumptions about the error term and even allowing for the possibility of a deterministic time trend regressor. His approach has been generalized to more complicated settings of multiple breaks and multivariate models (see Bai, Lumsdaine, and Stock (1998), Bai and Perron (1998, 2003), Bai (2000), and Qu and Perron (2007a)).

Qu and Perron (2007a) consider a system of multivariate linear regression equations with potentially serially correlated errors and allow for multiple breaks in mean and variance parameters. In order to calculate a confidence interval for jth break date, they construct a Bai-type confidence interval based on the following statistic with a non-standard distribution:

$$\frac{\Psi_{1,j}^2}{\Gamma_{1,j}^2}(\hat{\tau} - \tau_0) \Rightarrow \arg\max_s Z(s), \tag{7}$$

where

$$Z(s) = \begin{cases} -\frac{1}{2}|s| + W_1(-s) & \text{if } s \leq 0 \\ -\frac{\varsigma}{2}|s| + \sqrt{\varphi} W_2(s) & \text{if } s > 0, \end{cases} \tag{8}$$

with $W_i(s)$, $i = 1, 2$ denoting two independent standard Wiener processes defined on $[0, \infty]$, $\varsigma = \frac{\Psi_{2,j}^2}{\Psi_{1,j}^2}$, $\varphi = \frac{\Gamma_{2,j}^2}{\Gamma_{1,j}^2}$, and $\Psi_{i,j}$ and $\Gamma_{i,j}$, $i = 1, 2$ are as defined in the previous section. The confidence intervals are then constructed using least squares estimates and equal-tailed quantile values:

$$\left[\hat{\tau} - \frac{\Gamma_{1,j}^2}{\Psi_{1,j}^2} \times q(1 - \alpha/2), \hat{\tau} - \frac{\Gamma_{1,j}^2}{\Psi_{1,j}^2} \times q(\alpha/2)\right],$$

where $q(\cdot)$ is the quantile function for the non-standard distribution in (7).[7]

Bai's confidence interval is a special case of Qu and Perron's confidence interval under the following assumptions: (i) no break in variance, (ii) a single break ($m = 1$), and (iii) single equation ($n = 1$) in (1). In this simplified case, $\frac{\Psi_{1,j}^2}{\Gamma_{1,j}^2} = \frac{(\Delta\beta' Q_1 \Delta\beta)^2}{\Delta\beta' \Pi_1 \Delta\beta}$, $\varphi = \frac{\Delta\beta' \Pi_2 \Delta\beta}{\Delta\beta' \Pi_1 \Delta\beta}$, and $\varsigma = \frac{\Delta\beta' Q_2 \Delta\beta}{\Delta\beta' Q_1 \Delta\beta}$.[8] Furthermore, when regressors and errors are stationary across regimes (i.e. $Q = Q_1 = Q_2$ and $\Pi = Q$), the asymptotic expected length of Bai's confidence interval is given by

$$2\frac{1}{(\Delta\beta' Q \Delta\beta)} \times q(1 - \alpha/2),$$

where the quantile function $q(\cdot)$ is determined by (8) under more simplifying conditions that $\varphi = 1$ and $\varsigma = 1$. For example, the asymptotic expected length of the confidence interval at 95% confidence level is approximately $22 \times \frac{1}{(\Delta\beta' Q \Delta\beta)}$. Notably, this is almost twice the

[7]The quantile function $q(\cdot)$ can obtained from the CDF $G(x)$ for $\arg\max_s Z(s)$ shown in Bai (1997). Note that this is different from the CDF for $\max_s Z(s)$ that we use to construct the likelihood-ratio-based confidence set.

[8]Note that Q_i and Π_i are normalized by the conditional variance, as in Qu and Perron (2007a), but different to Bai (1997).

asymptotic expected length of approximately $12 \times \frac{1}{(\Delta\beta' Q \Delta\beta)}$ for the equivalent 95% likelihood-ratio-based confidence set implied by Corollary 1 in the previous section. It is also worth noting from these expressions that the asymptotic expected length depends on the squared magnitude of the break relative to the variance of the errors, which implies that the expected length will increase in proportion to the sampling frequency (i.e., it will be three times the length in terms of monthly observations as for quarterly observations, corresponding to the same length of calendar time).

Elliott and Müller (2007) take a different approach than Bai (1997) and propose constructing a confidence set (not interval) for a break date based on the inversion of a sequence of tests for an additional break given a maintained break date. The validity of their approach is established using asymptotic analysis for a quickly-shrinking magnitude of break (i.e. $\Delta\beta = \delta T^{-1/2}$). They argue that Bai's approach has poor finite-sample performance due his asymptotic analysis based on a slowly-shrinking break being innappropriate for the moderately-sized breaks that appear to occur in practice. It should be noted, however, that the use of a slowly-shrinking break, originally proposed by Picard (1985), is common in the literature on structural breaks, including in Qu and Perron (2007a) and in our asymptotic analysis in the previous section as well. Meanwhile, it should also be noted that, because Elliott and Müller's approach is based on tests for an additional break, it is only suitable for a one-time break and cannot be generalized to multiple breaks unlike Bai and Perron (1998) for Bai's approach or the likelihood-ratio-based approach proposed in this paper.

3.2 Experiments

For our Monte Carlo experiments, we calculate the effective coverage rates and average lengths of confidence sets (or intervals) for break dates based on 1,000 replications given data generating processes involving structural breaks. We first consider a simple univariate model with one break in mean and/or variance. Then, we consider extended models with multiple breaks or a system of equations.

3.2.1 A simple univariate model with one break in mean and/or variance

For the experiments assuming one break in mean and/or variance, the general univariate model for our data generating process is given by

$$y_t = z_t'\beta_1 + z_t'\Delta\beta\mathbf{1}[t > \tau] + u_t, \tag{9}$$

where $u_t = \left(\sqrt{\Sigma_1 + (\Sigma_2 - \Sigma_1)\mathbf{1}[t > \tau]}\right) e_t$, $e_t \sim i.i.d.\mathcal{N}(0, 1)$, $\mathbf{1}[\cdot]$ is an indicator function, and $\tau = [rT]$ with r denoting the true break point fraction. Unless otherwise specified, we set $z_t = 1$, $\beta_1 = 0$, and $r = 0.5$.

Our first experiment considers large-sample coverage rates and lengths of confidence sets for a large slowly-shrinking break in mean. Given the large samples, this experiment is designed to verify our asymptotic analysis. Meanwhile, the break is "large" in Elliott and Müller's (2007) sense that its existence would be reliably detected with a test for the presence of a structural break. We parameterize a slowly-shrinking break in mean as $\Delta\beta = \delta/T^{1/4}$, with $\delta = 5$, which implies $\Delta\beta = 1.06$ for $T = 500$ and $\Delta\beta = 0.89$ for $T = 1,000$. Given fixed variance $\Sigma_1 = \Sigma_2 = 1$, a break magnitude of close to 1 is roughly calibrated to the estimated reduction in the long-run growth rate of the U.S. economy in the early 1970s when measured relative to the volatility of consumption growth in our application in the next section.[9] For each simulated sample, we estimate the parameters of a restricted version of the model in (9) with a fixed variance. Estimation is via maximum likelihood assuming one break with 15% trimming at the beginning and end of the sample period (i.e., the inner 70% of sample period provides the admissible set of break dates for estimation and calculation of the ILR confidence sets). In this experiment, Bai's approach and Qu and Perron's approach are equivalent and referred to by the label "Bai/QP" hereafter. For convenience, we also refer to Elliott and Müller's approach by the label "EM" hereafter.

Table 1 reports the results for the first experiment. Even with such large sample sizes,

[9]From Table 9, the implied reduction in 1972Q4 is 0.89 based on the estimated reduction in long-run growth and the conditional standard deviation of consumption growth. Ideally, we should standardize the magnitude of a break by the "long-run standard deviation" based on the spectral density at frequency zero. However, assuming consumption growth has little or no persistence, the conditional standard deviation provides a reasonable approximation.

the ILR confidence sets overcover at the 90% and 95% levels.[10] Bai/QP confidence intervals are also somewhat conservative, while the EM confidence sets have very accurate coverage. All three approaches have accurate coverage at the 99% level. Unlike with undercoverage, conservative confidence sets are not, in themselves, a problem as long as they are informative. So the key result in Table 1 is that, even though the ILR confidence sets overcover somewhat in large samples, they always have shortest average length and are, therefore, the most informative.

The average lengths in Table 1 correspond closely to the asymptotic expected lengths for both of the ILR and Bai/QP approaches.[11] In practice, the average lengths are slightly longer than the asymptotic lengths, which could be due in part to the overcoverage discussed above. But it is also related to the fact that the analytical expressions for the asymptotic expected length are for a continuous measure, while the average length captures the number of discrete periods in finite samples. By including an entire discrete time period instead of fractions of periods in a set, there is a natural rounding up in the average length relative to the asymptotic length. This rounding up will be more severe given disjointed sets, which occur for the ILR approach, but not for the Bai/QP approach. The main point, however, is that the ratio of average lengths for the competing methods converges to the ratio of asymptotic expected lengths as the sample size gets larger. The average lengths for the Bai/QP approach are much longer and sometimes twice as long as for the ILR approach. Meanwhile, even though we cannot calculate the asymptotic expected length for the EM approach, the average lengths for it are generally at least three times as long as for the ILR approach. Thus, the ILR confidence sets perform best and it is not just a small-sample issue.

A natural question is why the ILR approach is so much better than the Bai/QP approach asymptotically. Both methods are based on inverting a test of a hypothesized break

[10]Although we are considering a slowly-shrinking magnitude of break across the different sample sizes in this Monte Carlo experiment, the break for any given sample size is, of course, of fixed magnitude. Thus, if the asymptotic distribution of the likelihood ratio statistic for a slowly-shrinking magnitude of break provides an upper bound on the distribution of a fixed magnitude of break, as it should according to Hansen (2000) under Normal errors, we would expect the coverage to be conservative for any given experiment.

[11]It should be noted that the asymptotic expected length calculations depend on the fixed magnitude of the change in mean for the ILR and Bai/QP approaches. Therefore, under a shrinking break, the asymptotic lengths get longer with the sample size. For the EM approach, expected length always increases with sample size, including given a fixed magnitude of the change in mean. So we cannot calculate an asymptotic expected length for the EM confidence sets.

Table 1: Large-Sample Coverage Rates and Lengths of Confidence Sets for a Large Slowly-Shrinking Break in Mean

(a) 90% Confidence Level

	T=500			T=1,000		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.95	11.09	8.47	0.94	14.36	11.98
Bai/QP	0.93	16.68	14.00	0.91	22.33	20.00
EM	0.90	34.83	-	0.90	56.50	-

(b) 95% Confidence Level

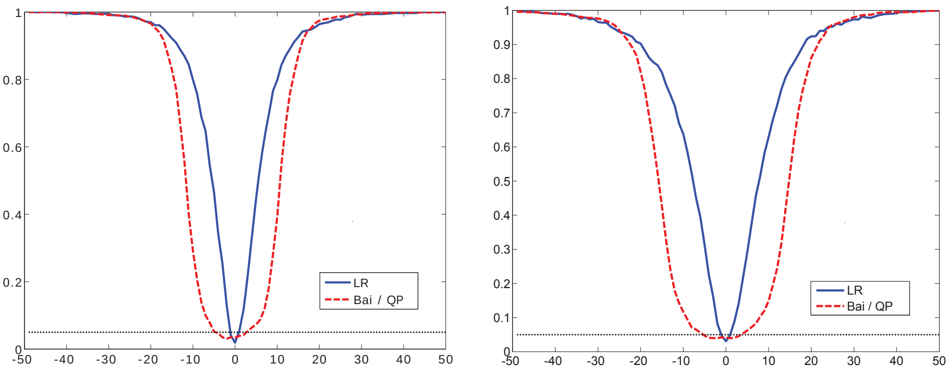
	T=500			T=1,000		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.97	13.89	11.12	0.97	18.20	15.73
Bai/QP	0.97	22.69	20.00	0.95	30.77	28.00
EM	0.95	43.06	-	0.95	69.34	-

(c) 99% Confidence Level

	T=500			T=1,000		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.99	20.28	17.08	0.99	27.04	24.16
Bai/QP	0.99	38.25	36.00	0.99	52.71	51.00
EM	0.99	58.79	-	0.99	93.59	-

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming one break, where coverage refers to the inclusion of the true break date in a confidence set. Asymptotic length is based on analytical results for expected length discussed in Sections 2.2 and 3.1. ILR refers to inverted likelihood ratio, Bai refers to Bai (1997), QP refers to Qu and Perron (2007a), and EM refers to Elliott and Müller (2007). In terms of the model in (9), the break in mean is set at $\Delta\beta = \delta/T^{1/4}$ with $\delta = 5$, which implies $\Delta\beta = 1.06$ for $T = 500$ and $\Delta\beta = 0.89$ for $T = 1,000$ and the variance fixed at $\Sigma_1 = \Sigma_2 = 1$.

Figure 1: Empirical Power Functions for 5% Tests of a Break Date Given a Large Slowly-Shrinking Break in Mean



(a) $T = 500$ (b) $T = 1,000$

Notes: Empirical rejection rates for tests of a hypothesized break date up to 50 periods away from the true break date based on 1,000 Monte Carlo replications. LR refers to likelihood ratio, Bai refers to Bai (1997), and QP refers to Qu and Perron (2007a). In terms of the model in (9), the break in mean is set at $\Delta\beta = \delta/T^{1/4}$ with $\delta = 5$, which implies $\Delta\beta = 1.06$ for $T = 500$ and $\Delta\beta = 0.89$ for $T = 1,000$ and the variance fixed at $\Sigma_1 = \Sigma_2 = 1$.

date. However, in the nonstandard environment of a test for a break date, the likelihood ratio (LR) test turns out to be more powerful than the Wald-type test used to construct the Bai/QP confidence intervals. Figure 1 displays empirical power functions for 5% tests of a hypothesized break date given the same data generating process and sample sizes considered in Table 1. The LR test clearly has a lot more power than the Bai/QP test against hypothesized break dates that are close to the true break date.[12] Thus, the ILR confidence sets can exclude a lot more break dates than the Bai/QP confidence intervals. Hence, their shorter average lengths.

Next, we consider small-sample coverage rates and lengths of 95% confidence sets for large fixed-magnitude breaks in mean and/or variance. This experiment is designed to capture how different methods would perform given empirically-relevant sample sizes of $T = 100, 200, 300$ and when the magnitude of break is a fixed quantity, as it would be in reality. In terms of the model in (9), the break in mean only is set at $\Delta\beta = 1$, with the variance fixed at $\Sigma_1 = \Sigma_2 = 1$. Again, this is roughly calibrated to the estimated productivity growth slowdown in the U.S.

[12]The LR test is also slightly undersized at the true break date, corresponding to the overcoverage of the ILR confidence sets reported in Table 1.

economy relative to the volatility of consumption growth. The break in variance only is set at $\Sigma_1 = 1.5, \Sigma_2 = 0.5$ with $\Delta\beta = 0$. This corresponds to a 40% reduction in the standard deviation, which is roughly calibrated to (albeit a bit smaller than) the estimated reduction of 50% or more in output growth volatility in the mid-1980s reported in many studies and also found in our application in the next section. The break in both mean and variance is set at $\Delta\beta = 0.5$ and $\Sigma_1 = 1, \Sigma_2 = 0.5$, which corresponds to somewhat smaller individual breaks, but their combined effect is such that the asymptotic expected lengths are similar to the other two cases. For each simulated sample, we again estimate the parameters of the model in (9) imposing a fixed variance in the case of a break in mean only and a fixed mean in the case of a break in variance only. As before, estimation is via maximum likelihood assuming one break with 15% trimming.

Table 2 reports the small-sample results for the breaks in mean and/or variance. As in the large-sample experiment, the 95% ILR confidence sets for a break in mean overcover and have the shortest expected length for all three sample sizes $T = 100, 200, 300$. Despite having a longer average length than the ILR approach, the Bai/QP confidence intervals undercover for a break in mean, especially for the smaller sample sizes, while the EM confidence sets have accurate coverage, but much longer average lengths. Bai's approach and the EM approach were both designed for a break in mean only. So for a break in variance and break in mean and variance, we compare the ILR approach to the QP approach only. Again, the QP confidence intervals undercover the true break date, especially for the break in variance. Even given the small sample sizes, the average lengths for the ILR and QP methods are generally in line with their corresponding asymptotic expected lengths, with a slight upward bias for the rounding reason discussed above, especially for the ILR confidence sets in the case of a break in mean and variance. Indeed, the QP confidence intervals have slightly shorter average lengths for the smaller sample sizes in this case, although their undercoverage raises serious concerns about their usefulness in practice given small samples.

Our third experiment considers small-sample coverage rates and lengths of 95% confidence sets for a small quickly-shrinking break in mean or no break. This experiment is designed to determine how well the ILR approach performs in the setting that the EM approach was designed for and for which Elliott and Müller (2007) show Bai's approach performs

Table 2: Small-Sample Coverage Rates and Lengths of 95% Confidence Sets for Large Fixed-Magnitude Breaks in Mean and/or Variance

(a) Break in Mean

	T=100		T=200		T=300		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.96	18.35	0.96	16.53	0.98	15.74	12.43
Bai/QP	0.90	22.66	0.92	22.92	0.94	22.90	24.00
EM	0.94	30.05	0.94	33.80	0.95	38.75	-

(b) Break in Variance

	T=100		T=200		T=300		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.96	34.29	0.97	30.69	0.96	29.23	31.08
QP	0.73	58.45	0.74	56.68	0.75	57.23	56.00

(c) Break in Mean and Variance

	T=100		T=200		T=300		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.97	35.99	0.98	37.38	0.98	35.48	29.59
QP	0.84	32.22	0.89	35.41	0.91	36.00	35.00

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming one break, where coverage refers to the inclusion of the true break date in a confidence set. ILR refers to inverted likelihood ratio, Bai refers to Bai (1997), QP refers to Qu and Perron (2007a), and EM refers to Elliott and Müller (2007). In terms of the model in (9), the break in mean only is set at $\Delta\beta = 1$ with the variance fixed at $\Sigma_1 = \Sigma_2 = 1$, the break in variance only is set at $\Sigma_1 = 1.5, \Sigma_2 = 0.5$ with $\Delta\beta = 0$, and the break in mean and variance is set at $\Delta\beta = 0.5$ and $\Sigma_1 = 1, \Sigma_2 = 0.5$.

particularly poorly. Given that the asymptotic validity of our approach is based on the same assumption as Bai's approach of a slowly-shrinking break, it is a reasonable concern that our approach might also perform poorly in the presence of a small break. Meanwhile, an extremely small break is essentially the same as no break at all, so it is an interesting question as to how different methods perform when there is actually no break (i.e., the true number of breaks is misspecified in the estimated model). We parameterize a quickly-shrinking break in mean as $\Delta\beta = \delta/T^{1/2}$, with $\delta = 4$. This corresponds to the smallest magnitude of break considered in Elliott and Müller's (2007) Monte Carlo analysis and implies $\Delta\beta = 0.40$ for $T = 100$, $\Delta\beta = 0.28$ for $T = 200$, and $\Delta\beta = 0.23$ for $T = 300$. Given fixed variance $\Sigma_1 = \Sigma_2 = 1$, the magnitude of the break is similar to the size of the estimated reduction in the long-run growth rate of the U.S. economy in the early 1970s when measured relative to the volatility of real GDP growth in our application in the next section.[13] As in the first experiment, we estimate the parameters of a restricted version of the model in (9) with a fixed variance. Again, estimation is via maximum likelihood assuming one break with 15% trimming.

Table 3 reports the results for a small break in mean or no break. In the case of a small break, the undercoverage of the Bai/QP confidence intervals highlighted in Elliott and Müller (2007) is confirmed, while the EM confidence sets again have very accurate coverage.[14] However, despite our asymptotic analysis being based on the same assumption as Bai's approach of a slowly-shrinking break, the ILR confidence sets retain their relatively accurate small-sample coverage properties, with some remaining overcoverage. Again, conservative confidence sets are not a problem in this case because the average length is lowest for the ILR approach. Meanwhile, the fact that the average lengths for the ILR and Bai/QP approaches are less than the respective asymptotic expected lengths is likely due to trimming for the ILR

[13] Again from Table 9, the implied reduction in 1972Q4 is 0.46 based on the estimated reduction in long-run growth and the conditional standard deviation of output growth. Furthermore, because output growth appears to be somewhat persistent even when accounting for structural breaks in mean, using the conditional standard deviation clearly overstates what the reduction would be relative to the long-run standard deviation.

[14] The experiment for $T = 100$ is essentially the same as the first experiment reported in Table 3 of Elliott and Müller (2007). However, Bai's approach does not perform quite as poorly here. One possible reason is that we trim the possible break dates to exclude the first and last 15% of the sample period, as is standard in the structural break literature, while Elliott and Müller only trim the first and last few observations when applying Bai's approach.

confidence sets and the small-sample undercoverage for the Bai/QP confidence intervals. In the case of no break, we report the coverage rate in terms of how often the entire admissible set of break dates based on 15% trimming is included in a confidence set. The idea is that, as the magnitude of a break gets smaller and smaller, a confidence set should get wider and wider until it almost always includes the entire admissible set. The results in the table confirm that confidence sets do, indeed, get wider, with the average length for the ILR approach close to the length of the entire admissible set, while the average lengths for the Bai/QP and EM approaches are even longer. Notably, the ILR confidence sets include the entire admissible over 90% of the time. By comparison, despite their longer average lengths, the Bai/QP and EM confidence sets include the entire admissible set only about 60% of the time.

The last experiment that we consider for the simple univariate model allows for serial correlation. For simplicity, we only consider a break in mean and the sample size of $T = 300$, which is roughly similar to the number of postwar quarterly observations for the U.S. macroeconomic time series considered in the application in the next section. In the first case, we modify the error process to have first-order serial correlation as follows:

$$u_t = \rho u_{t-1} + (1-\rho)e_t, \quad e_t \sim i.i.d.\mathcal{N}(0,1).$$

The break in mean is set at $\Delta\beta = 1$, which corresponds to a magnitude of 1 relative to the long-run standard deviation. This is the same as for the break in mean only in Table 2, which is why we have the same asymptotic lengths as before. We consider low or high persistence by setting $\rho = 0.3$ (as in Table 5 of Elliott and Müller (2007)) or $\rho = 0.6$. In this case, estimation of the parameters of the model in (9) is via quasi maximum likelihood assuming one break in mean with 15% trimming and we employ a HAC estimator of the long-run variance of u_t in order to calculate scaled test statistics with asymptotically pivotal distributions for the purposes of constructing confidence sets. Following Elliott and Müller (2007) and Qu and Perron (2007a), we consider the HAC estimator due to Andrews and Monahan (1992), which would also address heteroskedasticity if it were present, although we focus on the problem of serially-correlated errors in this Monte Carlo experiment. In the

Table 3: Small-Sample Coverage Rates and Lengths of 95% Confidence Sets for a Small Quickly-Shrinking Break in Mean or No Break

(a) Break in Mean

	T=100			T=200			T=300		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.96	57.95	77.70	0.95	116.01	155.40	0.96	169.07	358.00
Bai / QP	0.87	93.85	138.00	0.88	186.95	276.00	0.88	261.21	742.00
EM	0.95	76.66	-	0.95	157.34	-	0.96	233.63	-

(b) No Break in Mean

	T=100			T=200			T=300		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.92	67.99	-	0.92	135.39	-	0.91	201.98	-
Bai / QP	0.57	165.71	-	0.60	331.34	-	0.59	491.47	-
EM	0.60	90.11	-	0.59	184.54	-	0.57	280.35	-

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming one break, where coverage refers to the inclusion of the true break date in a confidence set when there is a break in mean and to the inclusion of the entire admissible set of break dates (i.e., inner 70% of the sample period given 15% trimming) when there is no break. ILR refers to inverted likelihood ratio, Bai refers to Bai (1997), QP refers to Qu and Perron (2007a), and EM refers to Elliott and Müller (2007). In terms of the model in (9), the break in mean is set at $\Delta\beta = \delta/T^{1/2}$ with $\delta = 4$, which implies $\Delta\beta = 0.40$ for $T = 100$, $\Delta\beta = 0.28$ for $T = 200$, and $\Delta\beta = 0.23$ for $T = 300$, and the variance fixed at $\Sigma_1 = \Sigma_2 = 1$.

second case, we capture serial correlation as a time-varying conditional mean for the model in (9) by considering an AR(1) process with a break in mean set at $\Delta\beta = ((1-\rho)\ 0)'$ given $z_t = (1\ y_{t-1})$, $\beta_1 = (0\ \rho)'$, and the variance fixed at $\Sigma_1 = \Sigma_2 = (1-\rho)^2$. In terms of the unconditional mean, this break corresponds to a magnitude of 1 relative to the long-run standard deviation, again implying the same asymptotic lengths as for the break in mean in Table 2, and we consider low or high persistence by setting $\rho = 0.3$ or $\rho = 0.6$. In this case, estimation of parameters in (9) is via conditional maximum likelihood assuming one break in mean with 15% trimming.

Table 4 reports the results for a break in mean in the presence of serial correlation. Again, the ILR approach performs best with the shortest average lengths despite conservative coverage. Serial correlation generally increases the average lengths compared to the corresponding results in Table 2, with higher serially correlation generally having a larger effect. For the ILR approach, the average lengths are shorter when the serial correlation is captured in the AR(1) model rather than allowed for in the errors, likely reflecting more efficient estimation and less overcoverage. For the Bai/QP approach, there appears to be slight undercoverage for the AR(1) model that worsens given higher persistence, although the coverage is more accurate when serial correlation is allowed for in the errors. As always, the EM approach has accurate coverage, although its average lengths increase a lot when there is high persistence.

To summarize the results from the first four experiments, the ILR approach always provides the most precise inferences about the timing of structural breaks. It tends to have conservative coverage, consistent with the analysis in Hansen (2000) for ILR confidence sets of threshold parameters under Normal errors, as are assumed in our Monte Carlo experiments. The EM approach does remarkably well in terms of coverage rates in finite samples, but it produces much less precise confidence sets than the ILR approach. The Bai/QP approach undercovers in smaller samples and typically has wide confidence intervals. In general, confidence sets become wider as the magnitude of the break gets smaller. However, in the case of misspecification of the number of breaks, the ILR confidence sets tend to include the entire admissible set while other methods often produce misleadingly narrow confidence sets in particular samples even though their average lengths are larger than for the ILR approach.

Table 4: Coverage Rates and Lengths of 95% Confidence Sets for $T = 300$ and a Large Break in Mean in the Presence of Serial Correlation

(a) Serially-Correlated Errors

	ρ=0.3		ρ=0.6		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.98	17.07	0.99	18.16	12.43
Bai / QP	0.95	23.28	0.96	23.00	24.00
EM	0.95	28.20	0.97	69.06	-

(b) AR(1) Model

	ρ=0.3		ρ=0.6		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.96	16.18	0.97	15.85	12.43
Bai / QP	0.94	23.49	0.93	22.71	24.00
EM	0.95	45.56	0.95	57.11	-

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming one break, where coverage refers to the inclusion of the true break date in a confidence set. ILR refers to inverted likelihood ratio, Bai refers to Bai (1997), QP refers to Qu and Perron (2007a), and EM refers to Elliott and Müller (2007). In terms of the model in (9), the break in mean is set at $\Delta\beta = 1$ and the error process modified to $u_t = \rho u_{t-1} + (1-\rho)e_t,\ e_t \sim i.i.d.\mathcal{N}(0,1)$, in the case of serially-correlated errors and the break in mean is set at $\Delta\beta = ((1-\rho)\ 0)'$ for $z_t = (1\ y_{t-1})$ and $\beta_1 = (0\ \rho)'$ with the variance fixed at $\Sigma_1 = \Sigma_2 = (1-\rho)^2$ in the case of the AR(1) model.

3.2.2 Extended models with multiple breaks or a system of equations

In this subsection, we consider extended models that allow for multiple breaks or a system of equations. For the experiment that allows for multiple breaks in mean and/or variance, the univariate model for our data generating process is now given by

$$y_t = \beta_1 + \Delta\beta_1 \mathbf{1}[t > \tau_1] + \Delta\beta_2 \mathbf{1}[t > \tau_2] + u_t, \tag{10}$$

where $u_t = \left(\sqrt{\Sigma_1 + (\Sigma_2 - \Sigma_1)\mathbf{1}[t > \tau_1] + (\Sigma_3 - \Sigma_2)\mathbf{1}[t > \tau_2]}\right) e_t$, $e_t \sim i.i.d.\mathcal{N}(0, 1)$, and $\tau_j = [r_j T]$ with r_j, $j = 1, 2$, denoting the true break point fractions. As in most of the previous subsection, we set $z_t = 1$ and $\beta_1 = 0$. For the timing of the two breaks, we set $r_1 = 0.3$ and $r_2 = 0.7$. Again, for simplicity, we only consider the sample size of $T = 300$. In one case, we consider two breaks in mean only by setting $\Delta\beta_1 = 1, \Delta\beta_2 = -1$ with the variance fixed at $\Sigma_1 = \Sigma_2 = \Sigma_3 = 1$. Thus, these breaks are the same magnitude as the break in mean in Tables 2 and 4. In a second case, we consider two breaks in variance only by setting $\Sigma_1 = 1.5, \Sigma_2 = 0.5, \Sigma_3 = 1.5$ with $\Delta\beta_1 = 0, \Delta\beta_2 = 0$. Thus, these breaks are the same magnitude as the breaks in variance in Table 2. In a third case, we consider different breaks in mean and variance by setting $\Delta\beta_1 = 1, \Delta\beta_2 = 0$ and $\Sigma_1 = \Sigma_2 = 1.5, \Sigma_3 = 0.5$. The first break in mean is smaller than in Table 2 because of the higher variance in the first two regimes, while the break in variance is the same magnitude as is Table 2. Estimation of the parameters in (10) is via maximum likelihood assuming two breaks with 15% trimming at the beginning and end of the sample period and between break dates. For simplicity, the true structure of the breaks in terms of whether they are in mean or variance is assumed to be known and imposed in estimation.

Table 5 reports the results for multiple breaks in mean and/or variance. The ILR approach continues to perform better than the QP approach. For the case of two breaks in mean, the 95% ILR confidence sets are extremely accurate and have slightly shorter average lengths than in the corresponding case for one break reported in Table 2. This improvement in average length is likely due to more accurate coverage and the trimming given two breaks reducing the size of the admissible break sets. For the case of two breaks in variance, the ILR confidence sets are, again, slightly shorter than for the corresponding case in

Table 2, although they still overcover somewhat. The QP confidence sets again suffer from undercoverage, while still being almost twice the average length of the corresponding ILR confidence sets. For the case of different breaks in mean and variance, the ILR confidence sets overcover for the break in mean and the average length is relatively long compared to the asymptotic expected length. However, the ILR confidence sets remain more precise than the QP confidence intervals, which slightly undercover for the break in mean. For the break in variance, both methods perform similarly to the case of two breaks in variance, with a slight increase in the average lengths. The general point of this experiment is that the ILR confidence remain precise and better than the QP confidence sets when there are multiple breaks that occur closer to the beginning or end of the sample.

For the experiment that allows for a system of equations, the bivariate model for our data generating process is given by

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \beta_{1,1} + \Delta\beta_{1,1} 1[t > \tau_0] \\ \beta_{2,1} + \Delta\beta_{2,1} 1[t > \tau_0] \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}, \quad \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right), \tag{11}$$

where $\Delta\beta_{i,1} = \beta_{i,2} - \beta_{i,1}$, $i = 1, 2$. In our simulations, we set $\beta_{i,1} = 0$, $i = 1, 2$, and consider three different values for correlation between errors across equations of $\rho = -0.3, 0, 0.3$. Again, for simplicity, we only consider the sample size of $T = 300$. In one case, we consider a break in the mean for the first equation only by setting $\Delta\beta_{1,1} = 1, \Delta\beta_{2,1} = 0$. In a second case, we consider a break in the mean for both equations by setting $\Delta\beta_{1,1} = \Delta\beta_{2,1} = 1$. Estimation of the parameters in (11) is via maximum likelihood assuming one break with 15% trimming. Again, the true structure of the breaks in terms of which equations they occur in is assumed to be known and imposed in estimation.

Table 6 reports the results for a break in mean in a system of equations. Once again, the ILR confidence sets perform better than the QP approach, with shorter average lengths and correct coverage. Regardless of its sign, correlation in the errors across equations reduces the asymptotic expected length when there is a break in one equation only. The average lengths for both methods reflect this, while the average lengths when there is no correlation are slightly longer than in Table 2, presumably due to finite-sample uncertainty about the correlation when estimating the system of equations. In the case of a break in both variables,

Table 5: Coverage Rates and Lengths of 95% Confidence Sets for $T = 300$ and Multiple Large Breaks in Mean and/or Variance

(a) Two Breaks in Mean

	First Break		Second Break		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.95	14.01	0.95	14.46	12.43
QP	0.95	24.32	0.95	25.22	24.00

(b) Two Breaks in Variance

	First Break		Second Break		
	Coverage Rate	Average Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.96	28.89	0.96	28.36	31.08
QP	0.75	53.50	0.76	53.91	56.00

(c) Different Breaks in Mean and Variance

	First Break			Second Break		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.98	25.70	18.65	0.96	29.09	31.08
QP	0.92	33.55	34.00	0.77	55.42	56.00

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming two breaks, where coverage refers to the inclusion of the corresponding true break date in a confidence set. ILR refers to inverted likelihood ratio and QP refers to Qu and Perron (2007a). In terms of the model in (10), the two breaks in mean only are set at $\Delta\beta_1 = 1, \Delta\beta_2 = -1$ with the variance fixed at $\Sigma_1 = \Sigma_2 = \Sigma_3 = 1$, the two breaks in variance only are set at $\Sigma_1 = 1.5, \Sigma_2 = 0.5, \Sigma_3 = 1.5$ with $\Delta\beta_1 = 0, \Delta\beta_2 = 0$, and the different breaks in mean and variance are set at $\Delta\beta_1 = 1, \Delta\beta_2 = 0$ and $\Sigma_1 = \Sigma_2 = 1.5, \Sigma_3 = 0.5$.

the asymptotic expected length is reduced because a break in both variables is effectively a larger magnitude break. Indeed, when there is no correlation, the asymptotic lengths for both methods reduce by exactly 50%, as would happen with a doubling of the squared magnitude of a break in a univariate setting given the scale of the variance. Again, the average lengths drop roughly in proportion to the reduction in the asymptotic lengths. However, in this case, the asymptotic lengths depend on the sign of the correlation. A negative correlation for the errors better identifies the break date and reduces the asymptotic length because the structural break is assumed to correspond to a common positive movement in the two variables. A positive correlation increases the asymptotic length because it is harder to identify the break as distinct from a typical positive co-movement in the two variables due to the errors. The average lengths again reflect the asymptotic lengths. Overall, then, these results support the findings in Bai, Lumsdaine, and Stock (1998) and Qu and Perron (2007a) that adding equations to a multivariate model can produce more precise inferences, either given correlation in the errors when there is a break in one of the equations or given common breaks across equations, which can be helped or hindered by correlation in the errors depending on the common impact of the break.

4 Structural Breaks in Postwar U.S. Real GDP and Consumption

We apply our proposed likelihood-ratio-based method of constructing confidence sets to investigate structural breaks in postwar quarterly U.S. real GDP and consumption of nondurables and services. We first consider univariate models of the growth rates of output and consumption and then we consider a multivariate model that imposes balanced long-run growth between output and consumption. The data were obtained from the BEA website for the sample period of 1947Q1 to 2012Q1.[15] Annualized quarterly growth rates are calculated as 400 times the first differences of the natural logarithms of the levels data.

[15] The raw data are from the BEA Tables 1.1.5 and 1.1.6 for the vintage of April 27, 2012. We need both real and nominal measures for total consumption and consumption of durables in order to construct a chain-weighted measure of real consumption of nondurables and services based on Whelan's (2000) suggestion of a Tornqvist approximation to the ideal Fisher index.

Table 6: Coverage Rates and Lengths of 95% Confidence Sets for $T = 300$ and a Large Break in Mean for a Bivariate Model with Possibly Correlated Errors across Equations

(a) Break in One Equation

	ρ=-0.3			ρ=0			ρ=0.3		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.98	14.40	11.31	0.98	16.23	12.43	0.98	14.78	11.31
QP	0.96	21.69	22.00	0.96	24.00	24.00	0.94	22.07	22.00

(b) Break in Both Equations

	ρ=-0.3			ρ=0			ρ=0.3		
	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length	Coverage Rate	Average Length	Asymptotic Length
ILR	0.99	5.85	4.35	0.98	8.12	6.22	0.97	10.49	8.08
QP	0.97	9.69	8.00	0.95	12.90	12.00	0.94	15.95	16.00

Notes: Coverage rate and average length based on 1,000 Monte Carlo replications assuming one break, where coverage refers to the inclusion of the true break date in a confidence set. ILR refers to inverted likelihood ratio and QP refers to Qu and Perron (2007a). In terms of the model in (11), the break in mean is set at $\Delta\beta_{1,1} = 1, \Delta\beta_{2,1} = 0$ in the case of a break in one equation only and $\Delta\beta_{1,1} = \Delta\beta_{2,1} = 1$ in the case of a break in both equations.

4.1 Univariate Models

The typical approach to investigating structural breaks is to consider a univariate model. Although this can be less efficient than considering a multivariate model, as we found in our Monte Carlo analysis, it has the benefit of making the interpretation of estimated breaks straightforward. Thus, we begin our analysis with univariate models of output growth and consumption growth, respectively, as the results will help with understanding the results for the multivariate model presented below.

For the univariate analysis, we assume that log output has a stochastic trend with drift and a finite-order autoregressive representation. Specifically, our model for quarterly output growth is an AR(p) process:

$$\Delta y_t = \gamma_y + \sum_{j=1}^{p} \zeta_{y,j} \Delta y_{t-j} + e_{yt}, \quad e_{yt} \sim i.i.d.\mathcal{N}(0, \sigma_y^2). \tag{12}$$

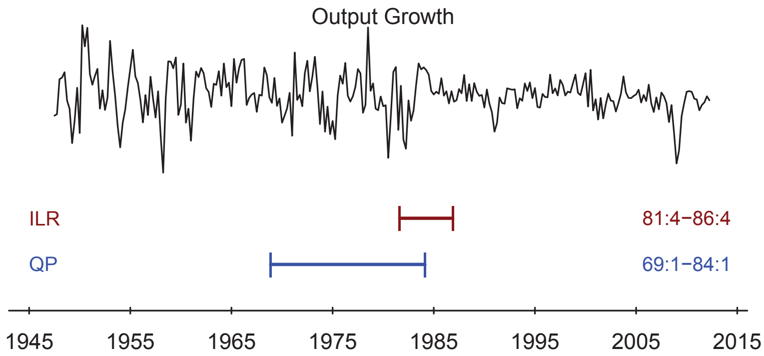
Similarly, we assume log consumption has a stochastic trend with drift and a finite-order autoregressive representation. Thus, our model for quarterly consumption growth is also an AR(p) process:

$$\Delta c_t = \gamma_c + \sum_{j=1}^{p} \zeta_{c,j} \Delta c_{t-j} + e_{ct}, \quad e_{ct} \sim i.i.d.\mathcal{N}(0, \sigma_c^2). \tag{13}$$

For lag selection, we employ Kurozumi and Tuvaandorj's (2011) modified BIC to account for the possibility of multiple structural breaks. Given an upper-bound of four lags and four breaks, with the common adjusted sample of 1948Q2 to 2012Q1, we find that the highest lag order selected is $p = 1$ for output growth and $p = 2$ for consumption growth.

Figure 2 plots the output growth series over the postwar period. Although the series clearly resembles the realization of a low-order autoregressive process with fairly low persistence, the parameters for this process may have changed over time. Applying Qu and Perron's (2007a) testing procedure to an AR(1) model estimated over the longest available sample period for conditional maximum likelihood of 1947Q3 to 2012Q1 with 15% trimming at the beginning and end of the sample period and between break dates, we find evidence of one break (the same as the number of breaks chosen by the modified BIC statistic mentioned above). The break is estimated to have occurred in 1983Q2, which corresponds closely to

Figure 2: U.S. Real GDP Growth and Confidence Sets for AR(1) Model



Notes: Sample period is 1947Q3 to 2012Q1. ILR refers to inverted likelihood ratio and QP refers to Qu and Perron (2007a).

the timing of the so-called "Great Moderation" widely reported in the literature (e.g., Kim and Nelson (1999) and McConnell and Perez-Quiros (2000)). The break is significant at the 5% level and there is no support for an additional break, even at the 10% level. Estimates for the long-run growth rate, largest eigenvalue measure of persistence, and the conditional standard deviation are reported in Table 7.[16] Likelihood ratio tests of parameter restrictions suggest that the break corresponds to a change in the conditional standard deviation only, which is estimated to have dropped by more than 50%.[17]

The ILR confidence set is also reported in Figure 2. It covers a reasonably short interval of 1981Q4-1986Q1. Notably, as mentioned in the introduction, this interval is similar in length to the 67% confidence interval for the Great Moderation reported in Stock and Watson (2002) based on Bai's approach. For illustration, we compare our confidence set to the 95% confidence interval calculated by the QP approach using the same model and data. As discussed in the previous section, the QP confidence interval is based on the distribution of the break date estimator, as in Bai (1997), but is also applicable in the multivariate setting that we consider in the second part of our application. Aside from any concerns

[16]For easy comparison across models, we measure persistence by the (modulus of the) largest eigenvalue of the companion matrix for the stationary representations of an autoregressive model or a vector error correction model. For the AR(1) model, this is simply the autoregressive coefficient.

[17]Note that, for simplicity, we always consider the unrestricted model when constructing confidence sets, as this allows for a more straightforward comparison of results across models when certain parameter restrictions are rejected for only one model, but not for another.

Table 7: Autoregressive Model of U.S. Real GDP Growth: 1947Q3-2012Q1

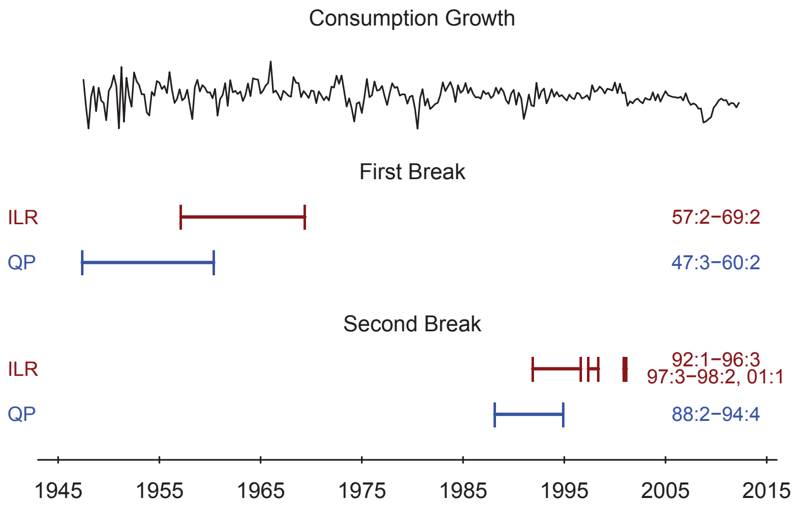
Regime	Break Date	LR Growth Rate	Largest Eig.	Cond. SD
1		3.460	0.330	4.487
2	1983Q2	2.712	0.538	2.195

we might have about its finite-sample coverage properties, the QP confidence interval is noticibly wider, running from 1969Q1-1984Q1, thus also including the possible "productivity growth slowdown" in the early 1970s (see, for example, Perron (1989) and Hansen (2001)). Therefore, the interval is much less informative about when the structural break occurred, including whether or not it was abrupt.

Figure 3 plots the consumption growth series. Although consumption is by far the largest expenditure component of U.S. real GDP, it is not as important for quarterly fluctuations in output given the volatility of other components, especially investment. Thus, it is not foregone conclusion that consumption growth will exhibit the same volatility reduction as output in the mid-1980s. Instead, it appears that there are breaks in consumption growth that do not manifest themselves in the overall behaviour of aggregate output. Indeed, applying Qu and Perron's (2007a) testing procedure to an AR(2) model estimated over the longest available sample period for conditional maximum likelihood of 1947Q4 to 2012Q1 with 15% trimming, we find evidence of two breaks (again the same as the number chosen by the modified BIC statistic) that are estimated to have occurred in 1958Q3 and 1993Q3, respectively. The breaks are significant at the 5% level and there is no support for additional breaks at the 10% level. Estimates for the long-run growth rate, largest eigenvalue measure of persistence, and the conditional standard deviation are reported in Table 8. Likelihood ratio tests of parameter restrictions for this model suggest that both of these are breaks in the conditional standard deviation of consumption growth, with the second break also corresponding to a decrease in the long-run growth rate and an increase in persistence.

The confidence sets for the two structural breaks in consumption growth are also reported in Figure 3. As with output growth, the ILR confidence sets are shorter than those based on the QP approach, although the differences in length are not as big as before. Notably, despite more similar lengths, the confidence sets still have different implications from each other about the timing of the breaks, with the ILR confidence sets shifted later in the

Figure 3: U.S. Consumption Growth and Confidence Sets for AR(2) Model



Notes: Sample period is 1947Q4 to 2012Q1. ILR refers to inverted likelihood ratio and QP refers to Qu and Perron (2007a).

sample period. However, in both cases, the confidence sets exclude the periods of a possible productivity growth slowdown in the early 1970s and the Great Moderation in the mid-1980s that correspond to the most widely-hypothesized breaks in U.S. economic activity. Given these apparently different breaks from output growth, it is an open question as to whether a multivariate model of output and consumption would lead to different or more precise inferences about structural breaks in these two series, as found, for example, by Bai, Lumsdaine, and Stock (1998). We turn to this question next.

4.2 Multivariate Model

Following Cochrane (1994), we assume that real GDP and consumption of nondurables and services have balanced long-run growth due to a common stochastic trend, possibly reflecting common shocks to productivity as suggested by a stochastic neoclassical growth model (see Bai, Lumsdaine, and Stock (1998) for a full theoretical motivation of this assumption). The empirical justification for the balanced-growth assumption comes from the apparent cointegrating relationship between these particular measures of consumption and output. If we impose a balanced long-run relationship corresponding to a cointegrating vector of $(1\ -1)$

Table 8: Autoregressive Model of U.S. Consumption Growth: 1947Q4-2012Q1

Regime	Break Date	LR Growth Rate	Largest Eig.	Cond. SD
1		3.067	-0.044	3.092
2	1958Q3	3.264	0.511	1.841
3	1993Q3	1.976	0.901	0.951

for the natural logarithms of consumption and output from 1947Q1 to 2012Q1, we find that we can reject a unit root with a p-value of 0.008 for an ADF test for the consumption rate, $c_t - y_t$ with a constant in the test regression and BIC for lag selection. Thus, there is empirical support for the idea that output and consumption (appropriately measured) have a balanced long-run relationship.

Assuming log output and consumption have a finite-order vector autoregressive representation, cointegration with known cointegrating vector $(1 \ -1)$ implies that the growth rates of output and consumption can be captured by the following VECM(p) model:

$$\Delta y_t = \gamma_y + \sum_{j=1}^{p} \zeta_{yy,j} \Delta y_{t-j} + \sum_{j=1}^{p} \zeta_{yc,j} \Delta c_{t-j} + \pi_y (c_{t-1} - y_{t-1}) + e_{yt}, \tag{14}$$

$$\Delta c_t = \gamma_c + \sum_{j=1}^{p} \zeta_{cy,j} \Delta y_{t-j} + \sum_{j=1}^{p} \zeta_{cc,j} \Delta c_{t-j} + \pi_c (c_{t-1} - y_{t-1}) + e_{ct}, \tag{15}$$

where $e_t \sim \mathcal{N}(0, \Omega)$. This form of cointegration also directly implies that the long-run consumption rate is constant and consumption and output share the same long-run growth rate. We parameterize these two long-run rates as follows:

$$E[c_t - y_t] = \kappa,$$

$$E[\Delta y_t] = E[\Delta c_t] = \mu.$$

It is possible then to solve for these two long-run parameters given estimates of the VECM parameters in (14) and (15) as follows:

$$\begin{bmatrix} \kappa \\ \mu \end{bmatrix} = \begin{bmatrix} -\pi_y & 1 - \sum_{j=1}^{p} (\zeta_{yy,j} + \zeta_{yc,j}) \\ -\pi_c & 1 - \sum_{j=1}^{p} (\zeta_{cy,j} + \zeta_{cc,j}) \end{bmatrix}^{-1} \begin{bmatrix} \gamma_y \\ \gamma_c \end{bmatrix}. \tag{16}$$

Using this relationship in (16), we can uncover structural breaks in the long-run consumption rate and the long-run growth rate by testing for structural breaks in the conditional mean parameters of the VECM. Bai, Lumsdaine, and Stock (1998) emphasize that this is a test for break in the long-run growth rate, μ, under the assumption of no break in unconditional mean of the cointegrating relationship, κ. However, we leave it as an empirical issue whether a common break in the conditional mean parameters corresponds to a break in the long-run consumption rate, long-run growth, or both.[18]

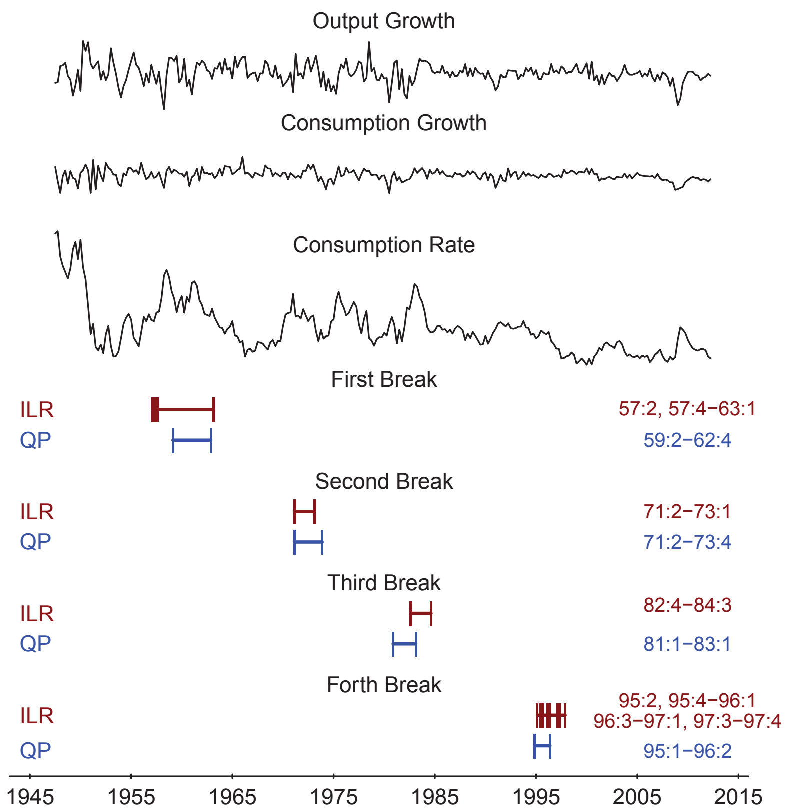
As with the univariate model for output growth, we find that the highest lag order selected by the modified BIC is $p = 1$. However, under the assumption of no breaks, the second lags of the growth rates are jointly significant at 5% level based on a likelihood ratio test (notably, the second lag of consumption growth in (15) has t-statistic of 2.1). Therefore, to avoid under-fitting, we consider $p = 2$.[19]

Applying Qu and Perron's (2007a) testing procedure for structural breaks to the VECM(2) model estimated over the longest available sample period for conditional maximum likelihood of 1947Q4 to 2012Q1 with 15% trimming, we find evidence of three breaks estimated in 1958Q1, 1982Q4, and 1996Q1 at the 5% level. The estimated timing of these breaks corresponds closely to the timing for the breaks in the univariate models of output growth and consumption growth. Thus, the first and third break likely correspond to a change in the behaviour of consumption, while the second break corresponds to the Great Moderation. However, in contrast to the univariate results, we now find evidence of four breaks estimated in 1961Q3, 1972Q4, 1982Q3, and 1996Q1 at the 10% level. The first, third, and fourth breaks again correspond closely to the breaks found in the univariate models. But the second break estimated in 1972Q4 appears to conform, at least in its timing, to the widely-hypothesized productivity growth slowdown that should affect both output and consumption and may be

[18]In a related empirical study, Cogley (2005) considers a time-varying parameter version of Cochrane's (1994) VECM model of output and consumption to investigate changes in the long-run growth rate and long-run consumption rate for the U.S. economy. He finds a gradual decline in the long-run growth rate from the mid-1960s to the early 1990s, followed by a gradual increase in long-run growth in the 1990s. He also finds that the consumption rate is very stable over the postwar period, although it gradually declines in the 1990s. However, Bayesian estimation of the time-varying parameter model imposes the strong prior that structural change is gradual, precluding the possibility of large, abrupt changes that are considered and found in our analysis.

[19]Because $p = 2$ was the lag order selected for the univariate model of consumption growth, allowing for a second lag also has the virtue of nesting the univariate models in our multivariate analysis.

Figure 4: U.S. Real GDP Growth, Consumption Growth, Log Consumption Rate, and Confidence Sets for VECM(2) Model



Notes: Sample period is 1947Q4 to 2012Q1. ILR refers to inverted likelihood ratio and QP refers to Qu and Perron (2007a).

better identified by the consideration of a multivariate model that imposes the same long-run growth rate for both series.

Figure 4 plots the output growth, consumption growth, and the consumption rate series over the postwar period. Visually, it is difficult to detect whether the estimated break in 1972Q4 corresponds to a break in the long-run growth rate or the long-run consumption rate. However, it is easier to see that the estimated break in 1996Q1 corresponds to a reduction in the long-run consumption rate in addition to a change in the behaviour of consumption growth detected in the univariate analysis. Indeed, the reasonable clarity of this change could explain the slight change in timing of the estimated break date from 1993Q3 for a change in consumption behaviour in the univariate analysis.

Table 9: Vector Error Correction Model of U.S. Real GDP and Consumption Growth: 1947Q4-2012Q1

Regime	Break Date	LR Growth Rate	LR Con. Rate	Largest Eig.	Cond. SDs	
1		3.188	-181.218	0.776	4.566	2.719
2	1961Q3	4.225	-188.367	0.783	3.068	1.571
3	1972Q4	2.826	-184.370	0.803	3.713	1.666
4	1982Q4	2.875	-188.825	0.807	1.496	1.273
5	1996Q1	1.585	-196.138	0.746	1.842	0.948

Table 9 reports the estimates of the long-run growth rate, long-run consumption rate, largest eigenvalue measure of persistence, and conditional standard deviations of output growth and consumption growth for the VECM(2) model with four structural breaks. Consistent with the univariate findings, the first break in the early 1960s clearly corresponds to a reduction in consumption growth volatility. The second break in the early 1970s corresponds to a reduction in the long-run growth rate of 1.4 annualized percentage points, in line with a productivity growth slowdown, more than a change in the long-run consumption rate or a change in volatility or persistence.[20] The third break in the mid-1980s clearly corresponds to a reduction in output growth volatility, consistent with the Great Moderation. The fourth break in the mid-1990s corresponds to an additional reduction in the long-run growth rate of 1.3 annualized percentage points, as well as to the reduction in the long-run consumption rate evident in Figure 4.[21] Interestingly, the largest eigenvalue measure of persistence remains remarkably stable over the full sample period. Likelihood ratio tests of parameter restrictions generally support our interpretation of the breaks, although it can be harder to relate how rejections of restrictions on intercept, slope, and/or the conditional variance/covariance parameters map into some of the parameters of interest. Thus, we report confidence sets for the unrestricted model that allows all parameters to change with each break.

The most striking result for the multivariate model is how precise the confidence sets

[20]It should be noted that our results in terms of the timing of break dates are almost identical if we consider *per capita* measures of output and consumption using U.S. civilian population data. Thus, breaks in long-run growth rates appear to reflect shifts in productivity growth rather than discrete changes in the population growth rate.

[21]Given the assumption of balanced growth for output and consumption, the magnitude of the estimated reduction in the long-run growth rate in Table 9 is influenced by changes in the average growth rates for both output and consumption. However, it should be noted that, on its own, the average growth rate for output declined by 1.1 annualized percentage points between regimes 4 and 5, quite consistent with a sizeable growth slowdown of 1.3 annualized percentage points reported in the table.

are in Figure 4. This finding is consistent with our Monte Carlo results for the system of equations and with the analysis in Bai, Lumsdaine, and Stock (1998) and Qu and Perron (2007a) on the usefulness of multivariate inference about break dates. The ILR confidence sets are much shorter than in the univariate analysis, although it should be noted that they cover or nearly cover the entire admissible sets for the first two breaks given estimated break dates and 15% trimming. Perhaps surprisingly, the QP confidence intervals, which are not affected by trimming, are even shorter than the ILR confidence sets in three of the four cases. Of course, again, the Monte Carlo analysis should leave us with serious concerns about the finite-sample coverage properties of the QP confidence intervals. But the general implication of the results is that the structural changes were abrupt. Notably, too, the increased precision in confidence sets is not just for breaks in parameters that are common to both output and consumption in the VECM model, such as the long-run growth rate. The confidence sets for the Great Moderation, which appears to be a much more important phenomenon for output growth than for consumption growth, also become a lot more precise with the length of the ILR confidence set for this break shrinking from 18 quarters to 8 quarters.[22] Thus, the improvement in inferences arises from both the model structure and from the additional multivariate information.

5 Conclusion

We have proposed a likelihood-ratio-based approach to constructing confidence sets for the timing of structural breaks. In particular, the confidence set includes all hypothesized break dates that cannot be rejected based on a likelihood ratio test. The asymptotic validity for this approach is established for a broad setting of a system of multivariate linear regression equations under the assumption of a slowly-shrinking magnitude of a break, with the asymptotic expected length of the 95% confidence sets being about half that of standard methods employed in the literature. Monte Carlo analysis supports the finite-sample performance of the proposed approach in a number of realistic experiments, including given small breaks.

[22] This result is not due to trimming reducing the admissible set of break dates given four breaks. In particular, we re-calculated the ILR confidence sets using 10% trimming instead of 15% trimming and found that the confidence sets for the third and fourth breaks were unchanged.

An application to U.S. real GDP and consumption demonstrates the empirical relevance of the performance gains of the proposed approach. Specifically, the analysis provides much more precise inferences about the timing of the "productivity growth slowdown" in the early 1970s and the "Great Moderation" in the mid-1980s than previously found in the literature. It also suggests the presence of an additional large, abrupt decline in the long-run growth rate of the U.S. economy in the mid-1990s, at least when taking cointegration between output and consumption into account.

A Appendix

Proof of Proposition 1. Following Qu and Perron (2007a,b), we consider the jth break date τ_j without loss of generality. The log-profile likelihood ratio subject to the restrictions $g(\beta, \Sigma) = 0$ under the null hypothesis $H_0 : \tau_j = \tau_j^0$ and the alternative hypothesis $H_1 : \tau_j \neq \tau_j^0$ is given by

$$\begin{aligned}
LR_j(\tau_j^0) &= -2\left[l_j^r\left(\tau_j^0, \hat{\beta}(\tau_j^0), \hat{\Sigma}(\tau_j^0)\right) - l_j^r\left(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma}\right)\right] \\
&= \underbrace{-2\left\{l_j^r\left(\tau_j^0, \hat{\beta}(\tau_j^0), \hat{\Sigma}(\tau_j^0)\right) - l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right)\right\}}_{-\max_{\beta_j,\Sigma_j} lr_j^r(\tau_j^0, \beta_j, \Sigma_j)} + \underbrace{2\left\{l_j^r\left(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma}\right) - l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right)\right\}}_{\max_{\tau_j,\beta_j,\Sigma_j} lr_j^r(\tau_j, \beta_j, \Sigma_j)} \\
&= \max_{\tau_j} \quad lr_j(\tau_j, \beta_j^0, \Sigma_j^0) + o_p(1), \qquad\qquad (A.1)
\end{aligned}$$

where the maximization is taken over C_M. The second line in (A.1) results from adding and subtracting the log-likelihood at the true values $l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right)$ to the first line.[23] The equality of the second and the third lines in (A.1) follows from Theorem 1 in Qu and Perron (2007a).

We focus on the term $lr_j(\tau_j, \beta_j^0, \Sigma_j^0) = -2\left[l_j\left(\tau_j^0, \beta_j^0, \Sigma_j^0\right) - l_j\left(\tau_j, \beta_j^0, \Sigma_j^0\right)\right]$ in the third line of (A.1) in order to find the asymptotic distribution of $LR_j(\tau_j^0)$. Letting $lr_j(\tau_j, \beta_j^0, \Sigma_j^0) = lr_j(\tau_j - \tau_j^0)$ and $r = \tau_j - \tau_j^0$,

$$\begin{aligned}
lr_j(r) &= 0 \text{ for } r = 0 \\
lr_j(r) &= 2\Bigg(-\frac{r}{2}(log|\Sigma_j^0| - log|\Sigma_{j+1}^0|) \\
&\quad -\frac{1}{2}\sum_{t=\tau_j^0+r}^{\tau_j^0} (y_t - x_t'\beta_{j+1}^0)(\Sigma_{j+1}^0)^{-1}(y_t - x_t'\beta_{j+1}^0) - (y_t - x_t'\beta_j^0)(\Sigma_j^0)^{-1}(y_t - x_t'\beta_j^0)\Bigg) \text{ for } r < 0, \\
lr_j(r) &= 2\Bigg(-\frac{r}{2}(log|\Sigma_j^0| - log|\Sigma_{j+1}^0|) \\
&\quad -\frac{1}{2}\sum_{t=\tau_j^0+1}^{\tau_j^0+r} (y_t - x_t'\beta_j^0)(\Sigma_j^0)^{-1}(y_t - x_t'\beta_j^0) - (y_t - x_t'\beta_{j+1}^0)(\Sigma_{j+1}^0)^{-1}(y_t - x_t'\beta_{j+1}^0)\Bigg) \text{ for } r > 0.
\end{aligned}$$

[23]Note that $l_j^r(\hat{\tau}_j, \hat{\beta}, \hat{\Sigma}) = l_j^r(\hat{\tau}_j)$ in (4).

Then, letting $s = v_T^2(\tau_j - \tau_j^0)$, with v_T defined in Assumption 7, the proof of Theorem 3 in Qu and Perron (2007b) shows that, for $s \leq 0$,

$$lr_j([\frac{s}{v_T^2}]) \Rightarrow 2\left(-\frac{|s|}{2}\Xi_{1,j} + \Lambda_{1,j}W_{1,j}(s)\right), \tag{A.2}$$

and, for $s > 0$,

$$lr_j([\frac{s}{v_T^2}]) \Rightarrow 2\left(-\frac{|s|}{2}\Xi_{2,j} + \Lambda_{2,j}W_{2,j}(s)\right), \tag{A.3}$$

where

$$\Lambda_{1,j} = \left(\frac{1}{4}vec(A_{1,j})'\Omega_{1,j}^0 vec(A_{1,j}) + \delta_j'\Pi_{1,j}\delta_j\right)^{1/2}, \tag{A.4}$$

$$\Lambda_{2,j} = \left(\frac{1}{4}vec(A_{2,j})'\Omega_{2,j}^0 vec(A_{2,j}) + \delta_j'\Pi_{2,j}\delta_j\right)^{1/2}, \tag{A.5}$$

$$\Xi_{1,j} = \left(\frac{1}{2}tr(A_{1,j}^2) + \delta_j'Q_{1,j}\delta_j\right), \tag{A.6}$$

$$\Xi_{2,j} = \left(\frac{1}{2}tr(A_{2,j}^2) + \delta_j'Q_{2,j}\delta_j\right). \tag{A.7}$$

Note that $W_{1,j}(0) = W_{2,j}(0) = 0$ because $W_{1,j}(s)$ and $W_{2,j}(s)$ are independent and starting at $s = 0$.

Qu and Perron (2007a) derive a Bai-type distribution of $\hat{\tau} - \tau_0$ by taking the *arg max* of (A.2) and (A.3) over C_M and using the continuous mapping theorem. Here, instead, we are deriving the distribution of the likelihood ratio by taking the *max* of (A.2) and (A.3) over C_M. Thus, under the null hypothesis $H_0 : \tau_j = \tau_j^0$ we have

$$LR_j(\tau_j^0) \Rightarrow \max_s \begin{cases} 2\left(-\frac{|s|}{2}\Xi_{1,j} + \Lambda_{1,j}W_j(s)\right) & \text{for} \quad s \leq 0 \\ 2\left(-\frac{|s|}{2}\Xi_{2,j} + \Lambda_{2,j}W_j(s)\right) & \text{for} \quad s > 0, \end{cases}$$

where we can simplify this expression to relate it to a known distribution from Bhattacharya and Brockwell (1976). Let $LR_j(\tau_j^0) = \xi = \max[\xi_1, \xi_2]$ where $\xi_1 = \max_{s \leq 0} 2\left(-\frac{|s|}{2}\Xi_{1,j} + \Lambda_{1,j}W_j(s)\right)$ and $\xi_2 = \max_{v>0} 2\left(-\frac{|s|}{2}\Xi_{2,j} + \Lambda_{2,j}W_j(s)\right)$. By a change in variables $s = (\Lambda_{1,j}^2/\Xi_{1,j}^2)v$ and the distributional equality with $W(a^2x) \equiv aW(x)$, for $s \leq 0$,

$$\xi_1 = \sup_{s\leq 0} 2\left(-\frac{|s|}{2}\Xi_{1,j} + \Lambda_{1,j}W_j(s)\right) = \max_{v\leq 0}\frac{\Lambda_{1,j}^2}{\Xi_{1,j}}2\left(-\frac{|v|}{2} + W_j(v)\right) = 2\omega_{1,j} \times \overline{\xi}_1, \qquad \text{(A.8)}$$

where $\overline{\xi}_1 = \max_{v\leq 0}\left(-\frac{|v|}{2} + W_j(v)\right)$ and

$$\frac{\Lambda_{1,j}^2}{\Xi_{1,j}} = \frac{\Lambda_{1,j}^2 v_T^2}{\Xi_{1,j}v_T^2} = \frac{\Gamma_{1,j}^2}{\Psi_{1,j}} \equiv \omega_{1,j}.$$

Similarly, for $s > 0$ with $s = (\Lambda_{2,j}^2/\Xi_{2,j}^2)v$,

$$\xi_2 = \max_{s>0} 2\left(-\frac{|s|}{2}\Xi_{2,j} + \Lambda_{2,j}W_j(s)\right) = \max_{v>0}\frac{\Lambda_{2,j}^2}{\Xi_{2,j}}2\left(-\frac{|v|}{2} + W_j(v)\right) = 2\omega_{2,j} \times \overline{\xi}_2, \qquad \text{(A.9)}$$

where $\overline{\xi}_2 = \max_{v<0}\left(-\frac{|v|}{2} + W_j(v)\right)$ and

$$\frac{\Lambda_{2,j}^2}{\Xi_{2,j}} = \frac{\Lambda_{2,j}^2 v_T^2}{\Xi_{2,j}v_T^2} = \frac{\Gamma_{2,j}^2}{\Psi_{2,j}} \equiv \omega_{2,j}.$$

Thus, we have the simplified expression for the distribution of the likelihood ratio under the null hypothesis:

$$LR_j(\tau_j^0) \Rightarrow \max_s \begin{cases} 2\omega_{1,j}\left(-\frac{|v|}{2} + W_j(v)\right) & \text{for} \quad v \leq 0 \\ 2\omega_{2,j}\left(-\frac{|v|}{2} + W_j(v)\right) & \text{for} \quad v > 0. \end{cases}$$

Bhattacharya and Brockwell (1976) show that $\overline{\xi}_1$ and $\overline{\xi}_2$ in (A.8) and (A.9) are iid exponential random variables with respective distribution functions $P(\overline{\xi}_1 \leq x) = 1 - \exp(-x)$ for $x \leq 0$ and $P(\overline{\xi}_2 \leq x) = 1 - \exp(-x)$ for $x > 0$. Thus,

$$\begin{aligned} P(\xi \leq x) &= P(\max[2\omega_{1,j}\overline{\xi}_1, 2\omega_{2,j}\overline{\xi}_2] \leq x) \\ &= P(2\omega_{1,j}\overline{\xi}_1 \leq x)P(2\omega_{2,j}\overline{\xi}_2 \leq x) \\ &= P\left(\overline{\xi}_1 \leq \frac{x}{2\omega_{1,j}}\right)P\left(\overline{\xi}_2 \leq \frac{x}{2\omega_{2,j}}\right) \\ &= \left(1 - \exp(-\frac{x}{2\omega_{1,j}})\right)\left(1 - \exp(-\frac{x}{2\omega_{2,j}})\right). \end{aligned}$$

Then, using the distribution of the profile likelihood ratio for the break date τ_j, we can construct a $1-\alpha$ confidence set $C_{j,1-\alpha} = \{\tau_j | LR_j(\tau_j) \leq \kappa_{\alpha,j}\}$ by inverting the α-level likelihood ratio test. The probability of coverage $C_{j,1-\alpha}$ for any τ_j^0 is given by $P_{\tau_j^0}\left(\tau_j^0 \in C_{j,1-\alpha}\right)$, where we can easily calculate a critical value $\kappa_{\alpha,j}$ such that

$$P_{\tau_j^0}\left(\tau_j^0 \in C_{j,1-\alpha}\right) = \left(1 - \exp(-\kappa_{\alpha,j}/2\omega_{1,j})\right)\left(1 - \exp(-\kappa_{\alpha,j}/2\omega_{2,j})\right) = 1 - \alpha. \tag{A.10}$$

Note that $\kappa_{\alpha,j}$ will be unique because for all $\kappa > 0$ the CDF is a strictly increasing function $\frac{d(1-\exp(-\kappa/2\omega_{1,j}))(1-\exp(-\kappa/2\omega_{2,j}))}{d\kappa} > 0$. ■

Lemma 1 *Under the null hypothesis $H_0 : \tau = \tau_0$, if $lr(\hat{\tau} - \tau_0) \Rightarrow \overline{\xi} = \max_v \left(-\frac{1}{2}|v| + W(v)\right)$ for $v \in (-\infty, \infty)$, then $E_{\tau_0}\left[\lambda\{\tau | lr(\hat{\tau} - \tau) \leq x\}\right] = 4(1 - exp(-x))\left\{x - \frac{1}{2}(1 - exp(-x))\right\}$, where λ denotes a Lebesque measure.*

Proof of Lemma 1. As shown in Bhattacharya and Brockwell (1976), the CDF of $\overline{\xi} = \max_v \left(-\frac{1}{2}|v| + W(v)\right)$ is given by $P(\overline{\xi} \leq x) = (1 - exp(-x))^2$. Then letting $C_{1-\alpha} = \{\tau | lr(\hat{\tau} - \tau) \leq \kappa_\alpha\}$, Siegmund (1986) shows that the expected length for a $1-\alpha$ confidence set $C_{1-\alpha}$ is given by

$$\begin{aligned} E_{\tau^0}[\lambda\{C_{1-\alpha}\}] &= E_{\tau^0}[\lambda\{\tau | \tau \in C_{1-\alpha}\}] \\ &= \int_{-\infty}^{\infty} P_{\tau^0}\left(\tau \in C_{1-\alpha}\right) d\tau \\ &= 4(1-\alpha)^{1/2}\left\{-\log[1 - (1-\alpha)^{1/2}] - \frac{1}{2}(1-\alpha)^{1/2}\right\}. \end{aligned} \tag{A.11}$$

See Siegmund (1986) for more details.

Because we can find a critical value κ_α such that

$$P\left(\overline{\xi} \leq \kappa_\alpha\right) = (1 - exp(-\kappa_\alpha))^2 = 1 - \alpha,$$

it implies that

$$\kappa_\alpha = -\log[1 - (1-\alpha)^{1/2}]. \tag{A.12}$$

Then, by substituting (A.12) into (A.11), we can express the expected length for a $1-\alpha$ confidence set as a function of the critical value κ_α rather than the level $1-\alpha$ as follows:

$$E_{\tau^0}[\lambda\{C_{1-\alpha}\}] = 4(1-exp(-\kappa_\alpha))\left\{\kappa_\alpha - \frac{1}{2}(1-exp(-\kappa_\alpha))\right\}. \tag{A.13}$$

■

Proof of Proposition 2. For the general case, as in our setup under Assumptions 1-8, first consider the period before the true jth break date, $\tau_j - \tau_j^0 \le 0$ (i.e. $v \le 0$). Given a critical value $\kappa_{\alpha,j}$, the expected length of a $1-\alpha$ confidence set in the segment $\tau_j - \tau_j^0 \le 0$ can be shown to be the following:

$$\begin{aligned} & E_{\tau_j^0}\left[\lambda\{\tau_j | LR_j(\tau_j) \le \kappa_{\alpha,j},\ \ \hat{\tau}_j - \tau_j \le 0\}\right] \\ = \; & E_{\tau_j^0}\left[\lambda\{\tau_j \mid \frac{LR_j(\tau_j)}{2\omega_{1,j}} \le (\kappa_{\alpha,j}/2\omega_{1,j}),\ \ \hat{\tau}_j - \tau_j \le 0\}\right] \\ = \; & \underbrace{\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)}_{(i)} \underbrace{2(1-exp(-\kappa_{\alpha,j}/2\omega_{1,j}))\left\{\kappa_{\alpha,j}/2\omega_{1,j} - \frac{1}{2}(1-exp(-\kappa_{\alpha,j}/2\omega_{1,j}))\right\}}_{(ii)} \end{aligned} \tag{A.14}$$

The expression (i) in the third line of (A.14) is used for re-scaling because the expected length of the confidence set is measured on $v \in (-\infty, 0]$ and

$$\begin{aligned} \tau_j - \tau_j^0 &= r = s/v_T^2 = (\Lambda_{1,j}^2/\Xi_{1,j}^2)v/v_T^2 = (\Lambda_{1,j}^2 v_T^2/\Xi_{1,j}^2 v_T^4)v \\ &= \left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right) v. \end{aligned} \tag{A.15}$$

Note that from Proposition 1 the second line in (A.14) implies that

$$\frac{LR_j(\tau_j)}{2\omega_{1,j}} \Rightarrow \overline{\xi} = \max_v \ \left(-\tfrac{1}{2}|v| + W_j(v)\right) \quad \text{for} \quad v \le 0. \tag{A.16}$$

Thus, the expression (ii) in the third line of (A.14) is calculated for $P(\overline{\xi} \le \frac{\kappa_{\alpha,j}}{2\omega_{1,j}})$ by substituting the critical value $\kappa_{\alpha,j}/2\omega_{1,j}$ into half of the expected length in Lemma 1 given that we are considering $v \le 0$. The expected length for $v > 0$ is calculated in a similar fashion such that the expected length for the entire $1-\alpha$ likelihood-ratio-based confidence set is given by

$$\begin{aligned} & 2\left(\Gamma_{1,j}^2/\Psi_{1,j}^2\right)(1-exp(-\kappa_{\alpha,j}/2\omega_{1,j}))\left\{\kappa_{\alpha,j}/2\omega_{1,j}-\frac{1}{2}(1-exp(-\kappa_{\alpha,j}/2\omega_{1,j}))\right\} \\ + & \; 2\left(\Gamma_{2,j}^2/\Psi_{2,j}^2\right)(1-exp(-\kappa_{\alpha,j}/2\omega_{2,j}))\left\{\kappa_{\alpha,j}/2\omega_{2,j}-\frac{1}{2}(1-exp(-\kappa_{\alpha,j}/2\omega_{2,j}))\right\}. \end{aligned}$$

Note that as either $\omega_{1,j}$ or $\omega_{2,j}$ gets larger (i.e., the magnitude of a structural break is larger), the expected length becomes shorter because there is more precise information about the timing of the structural break. ■

Proof of Corollary 1. If there is no break in variance, $\Sigma_j = \Sigma$ for all j and $B_{1,j} = B_{2,j} = 0$. In addition, if the errors form a martingale difference sequence, $\Pi_{1,j} = Q_{1,j}$ and $\Pi_{2,j} = Q_{2,j}$. From these simplifications, $\omega_{1,j} = \omega_{2,j} = 1$, $\left(\frac{\Gamma_{1,j}}{\Psi_{1,j}}\right)^2 = \frac{1}{\Delta\beta_j' Q_1 \Delta\beta_j}$, and $\left(\frac{\Gamma_{2,j}}{\Psi_{2,j}}\right)^2 = \frac{1}{\Delta\beta_j' Q_2 \Delta\beta_j}$. Then, by substituting these values into the critical value and the expected length in Proposition 1, we can find the results in Corollary 1. The results in Remarks 1 and 2 follow in the same way. ■

Proof of Corollary 2. If there is no break in conditional mean, $\Delta\beta_j = 0$ and, in addition, if the standardized errors, η_t, are identically Normally distributed, $\eta_t\eta_t'$ has a Wishart distribution with $var(vec(\eta_t\eta_t')) = I_{n^2} + K_n$, where K_n is the commutation matrix. Then, $\Omega_{1,j} = \Omega_{2,j} = \Omega = I_{n^2} + K_n$. Furthermore, because K_n is an idempotent matrix,

$$\begin{aligned} & vec(B_{1,j})'\Omega^0 vec(B_{1,j})/4 \\ = & \; vec(B_{1,j})'(I_{n^2} + K_n)vec(B_{1,j})/4 \\ = & \; vec(B_{1,j})'vec(B_{1,j})/2. \end{aligned}$$

Thus,

$$\begin{aligned}
\omega_{1,j} &= \frac{\Gamma_{1,j}^2}{\Psi_{1,j}} \\
&= \frac{\frac{1}{4}vec(B_{1,j})'\Omega_{1,j}^0 vec(B_{1,j})}{\frac{1}{2}tr(B_{1,j}^2)} \\
&= \frac{\frac{1}{2}vec(B_{1,j})'vec(B_{1,j})}{\frac{1}{2}tr(B_{1,j}^2)} \\
&= 1
\end{aligned}$$

because $vec(B_{1,j})'vec(B_{1,j}) = tr(B_{1,j}^2)$. Similarly, $\omega_{2,j} = 1$. Then, $\frac{\Gamma_{1,j}^2}{\Psi_{1,j}^2} = \frac{2}{tr(B_1^2)}$, and $\frac{\Gamma_{2,j}^2}{\Psi_{2,j}^2} = \frac{2}{tr(B_2^2)}$. ■

References

Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61(4), 821–56.

Andrews, D. W. K., and J. C. Monahan (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60(4), 953–66.

Bai, J. (1997): "Estimation Of A Change Point In Multiple Regression Models," *Review of Economics and Statistics*, 79(4), 551–563.

——— (2000): "Vector Autoregressive Models with Structural Changes in Regression Coefficients and in Variance-Covariance Matrices," *Annals of Economics and Finance*, 1, 301–336.

Bai, J., R. L. Lumsdaine, and J. H. Stock (1998): "Testing for and Dating Common Breaks in Multivariate Time Series," *Review of Economic Studies*, 65(3), 395–432.

Bai, J., and P. Perron (1998): "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66(1), 47–78.

——— (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18(1), 1–22.

——— (2006): "Multiple Structural Change Models: A Simulation Analysis," in *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, ed. by D. Corbea, S. Durlauf, and B. Hansen, pp. 212–237. Cambridge University Press, Cambridge.

Bhattacharya, P., and P. Brockwell (1976): "The minimum of an additive process with applications to signal estimation and storage theory," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 37, 51–75.

Cobb, G. W. (1978): "The Problem of the Nile: Conditional Solution to a Changepoint Problem," *Biometrika*, 65(2), 243–251.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *The Quarterly Journal of Economics*, 109(1), 241–65.

Cogley, T. (2005): "How fast can the new economy grow? A Bayesian analysis of the evolution of trend growth," *Journal of Macroeconomics*, 27(2), 179–207.

Dümbgen, L. (1991): "The Asymptotic Behavior Of Some Nonparametric Change-Point," *The Annals of Statistics*, 19(3), 1471–1495.

Elliott, G., and U. K. Müller (2007): "Confidence Sets for the Date of a Single Break in Linear Time Series Regressions," *Journal of Econometrics*, 141, 1196–1218.

Fisher, R. A. (1930): "Inverse Probability," *Proceedings of the Cambridge Philosophical Society*, 28, 257–261.

Hansen, B. E. (2000): "Sample Splitting and Threshold Estimation," *Econometrica*, 68(3), 575–603.

——— (2001): "The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity," *Journal of Economic Perspectives*, 15(4), 117–128.

Kim, C.-J., J. Morley, and J. Piger (2008): "Bayesian counterfactual analysis of the sources of the great moderation," *Journal of Applied Econometrics*, 23(2), 173–191.

Kim, C.-J., and C. R. Nelson (1999): "Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle," *Review of Economics and Statistics*, 81(4), 608–616.

Kurozumi, E., and P. Tuvaandorj (2011): "Model selection criteria in multivariate models with multiple structural changes," *Journal of Econometrics*, 164(2), 218–238.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed since the Early 1980's?," *American Economic Review*, 90(5), 1464–1476.

Morley, J. (2009): "The Great Moderation: What Caused It and Is It Over?," Macro Focus Vol. 4, No. 8, Macroeconomic Advisers.

PERRON, P. (1989): “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis,” *Econometrica*, 57(6), 1361–1401.

PICARD, D. (1985): “Testing and Estimating Change-Points in Time Series,” *Advanced Applied Probability*, 17, 841–867.

QU, Z., AND P. PERRON (2007a): “Estimating and Testing Structural Changes in Multivariate Regressions,” *Econometrica*, 75(2), 459–502.

——— (2007b): “Supplement to Eestimating and Testing Structural Changes in Multivariate Regressions,” *Econometrica*, Vol. 75, No. 2.

SIEGMUND, D. (1986): “Boundary Crossing Probabilities and Statistic Applications,” *Annals of Statistics*, 14(2), 361–404.

——— (1988): “Confidence Set in Change-point Problems,” *International Statistical Review*, 56(1), 31–48.

STOCK, J. H., AND M. W. WATSON (2002): “Has the Business Cycle Changed, and Why?,” *NBER Macroeconomics Annual*, 17, 159–218.

WHELAN, K. (2000): “A Guide to the Use of Chain Aggregated NIPA Data,” Finance and Economics Discussion Series 2000-35, Board of Governors of the Frederal Reserve System.