Econometric Theory, **21**, 2005, 962–990. Printed in the United States of America.
DOI: 10.1017/S0266466605050486

OPTIMAL TESTS FOR NESTED MODEL SELECTION WITH UNDERLYING PARAMETER INSTABILITY

BARBARA ROSSI
Duke University

This paper develops optimal tests for model selection between two nested models in the presence of underlying parameter instability. These are joint tests for both parameter instability and a null hypothesis on a subset of the parameters. They modify the existing tests for parameter instability to allow the parameter vector to be unknown. These test statistics are useful if one is interested in testing a null hypothesis on some parameters but is worried about the possibility that the parameters may be time varying. The paper provides the asymptotic distributions of this class of test statistics and their critical values for some interesting cases.

1. INTRODUCTION

This paper develops optimal tests for model selection between two nested models in the presence of underlying parameter instability in the data. The model selection procedure considered in this paper is hypothesis testing; in fact, when the competing models are nested, the problem of testing which model is best among the two is to test the significance of additional variables that are present only under the largest model. The tests proposed in this paper thus *jointly* test for *both* parameter instability *and* a null hypothesis on a subset of the parameters.

The main contribution of this paper is to address *simultaneously* the two problems of testing parameter instability and model selection among nested models. It is argued that tests for model selection fail to detect parameter instability and that tests for parameter instability are not designed to choose between nested models. If the goal is to jointly test parameter stability and select a model, then it is possible to identify a class of optimal tests. The optimal tests modify exist-

I thank M. Watson for the idea of this paper and for numerous discussions, suggestions, comments, and teaching. I am grateful to T. Clark, G. Chow, R. Gallant, F. Sowell, N. Swanson, E. Tamer, and A. Tarozzi and also to the co-editor and two referees and to seminar participants at the University of Virginia, ECARES Université Libre de Brussels, the 2001 Triangle Econometrics Conference, the 2002 NBER Summer Institute, the 2003 Summer Meetings of the Econometric Society, and the 2003 EC^2 Conference for comments. Financial support from IFS Summer Research, Princeton University, is gratefully acknowledged. All mistakes are mine. Address correspondence to Barbara Rossi, Department of Economics, Duke University, P.O. Box 90097, Durham, NC 27708, USA; e-mail: brossi@econ.duke.edu.

  0266-4666/05 $12.00

ing tests for parameter instability to allow them to reject the incorrect model. This is achieved by *imposing*, rather than *estimating*, the parameters of interest under the null, thus making the statistic *not* invariant to shifts in these parameters.

The tests presented in this paper are useful in situations in which one is interested not only in whether the explanatory variables proposed by some economic model are statistically significant in explaining the observed data, but also in whether this relationship is stable over time. For example, these tests would be useful if one is interested in testing whether inflation or exchange rates are random walks but is also worried about the possibility that parameters may be varying over time (see Clark and McCracken, 2005; Rossi, 2005).

The strand of research closest to this paper is that concerning tests for parameter instability, in particular the works by Chow (1960), Quandt (1960), Ploberger and Krämer (1990, 1992), Andrews (1993), Andrews and Ploberger (1994), Sowell (1996), Ghysels and Hall (1990), Ghysels, Guay, and Hall (1998), and Elliott and Müller (2003). However, these tests are designed to detect parameter instability *only*, whereas this paper is also concerned about testing hypotheses on the parameter vector and, hence, treats it as *unknown*.

An alternative way to deal with model selection issues in the presence of parameter instability is to do a two-stage procedure: first test whether there is parameter instability; then test which model, among the competing ones, is the best description of the data. In some special cases analyzed in this paper, that is, for the special weighting distributions over the local alternatives analyzed in Section 3, the test statistics in the two stages are asymptotically independent. In this case, it is easy to fix the size in each stage of the procedure so that the two-stage procedure will have an overall correct size asymptotically. However, this result is not true for general weighting distributions. In addition, two-stage tests have advantages and disadvantages. The advantage is that if we reject we know which part of the alternative we reject; the disadvantage is that the test will not have the optimal weighted average power for alternatives that are equally likely.

The paper is organized as follows. Section 2 derives the optimal tests for testing the joint hypothesis of parameter stability and model selection and provides their asymptotic distribution. Section 3 discusses special tests and reports their asymptotic critical values, and Section 4 compares their asymptotic local powers. Section 5 concludes. Proofs of the results are in Appendix A, whereas Appendix B contains the tables of asymptotic critical values.

2. MODEL SELECTION IN THE PRESENCE OF UNDERLYING PARAMETER INSTABILITY

2.1. Heuristics

To gain some intuition about the results in this paper, consider a simple example where the data generating process (DGP) is as follows and the time of the break is known:

$$y_t = \beta_t + \epsilon_t; \qquad \beta_t = \begin{cases} \beta_1 & \text{for } t = 1,2,\dots,\tau \\ \beta_2 & \text{for } t = \tau+1,\dots,T \end{cases}; \qquad \epsilon_t \sim \text{i.i.d. } N(0,\sigma_\epsilon^2). \tag{1}$$

If the researcher is interested in testing whether the parameter β_t is constant over time and equal to a specific value β_0, a possible test statistic would be

$$Chow_T^* = \frac{\sum_{t=1}^{T}(y_t-\beta_0)^2 - \left(\sum_{t=1}^{\tau}(y_t-\hat\beta_1)^2 + \sum_{t=\tau+1}^{T}(y_t-\hat\beta_2)^2\right)}{\frac{1}{T}\left(\sum_{t=1}^{\tau}(y_t-\hat\beta_1)^2 + \sum_{t=\tau+1}^{T}(y_t-\hat\beta_2)^2\right)}, \tag{2}$$

where $\hat\beta_1 = (1/\tau)\sum_{t=1}^{\tau} y_t$ and $\hat\beta_2 = [1/(T-\tau)]\sum_{t=\tau+1}^{T} y_t$ are the sample averages of y_t in the two subsamples. By adding and subtracting the full sample average $\hat\beta = (1/T)\sum_{t=1}^{T} y_t$ inside the square of the first addend on the numerator, (2) can be rewritten as

$$Chow_T^* = \frac{T(\hat\beta-\beta_0)^2}{\hat\sigma_\epsilon^2} + \frac{\sum_{t=1}^{T}(y_t-\hat\beta)^2 - \left(\sum_{t=1}^{\tau}(y_t-\hat\beta_1)^2 + \sum_{t=\tau+1}^{T}(y_t-\hat\beta_2)^2\right)}{\frac{1}{T}\left(\sum_{t=1}^{\tau}(y_t-\hat\beta_1)^2 + \sum_{t=\tau+1}^{T}(y_t-\hat\beta_2)^2\right)}, \tag{3}$$

where

$$\hat\sigma_\epsilon^2 \equiv \frac{1}{T}\left(\sum_{t=1}^{\tau}(y_t-\hat\beta_1)^2 + \sum_{t=\tau+1}^{T}(y_t-\hat\beta_2)^2\right) \underset{p}{\rightarrow} \sigma_\epsilon^2.$$

Thus, the test is decomposed in two components: the one on the left is a test on β, and the one on the right is the standard Chow test for structural break. Hence, the test achieves power in detecting deviations from β_0 by adding to the traditional test for structural break a component that is variant to constant shifts in the mean. The asymptotic distribution of the test can easily be found in this case because the two components are independent.[1] Let $B_1(.)$ denote a scalar Brownian motion and $BB_1(.)$ denote a scalar Brownian bridge and $\pi = [\tau/T]$. Note that the first component on the right-hand side in (3) is asymptotically the square of a standardized normal $(B_1(1)^2)$, whereas the distribution of the second component is known from Andrews (1993) to be $BB_1(\pi)^2/\pi(1-\pi)$. As a result, the asymptotic distribution of this modified Chow test will be

$$Chow_T^* \Rightarrow B_1(1)^2 + \frac{BB_1(\pi)^2}{\pi(1-\pi)}. \tag{4}$$

Thus, the first component, which makes the test powerful in detecting constant shifts in the mean, adds a chi-square component to the limiting distribution of a standard Chow test for parameter instability. This example provides an easy and intuitive explanation of the asymptotic distribution of the tests considered in this paper.

2.2. Framework

This section describes the class of models considered in this paper and the assumptions under which the results are valid. The parametric model applies to a stationary and ergodic time series process.

Assumption 1. For each T, the sequence $\{x_{t,T}\}_{t=1}^{T}$ consists of the first T elements of an r-dimensional stationary and ergodic process. The parameter space Θ is a compact subset of R^k. For notational simplicity, x_t will be used to denote $x_{t,T}$.

The class of local alternatives allows both for structural changes and for non-linear hypotheses on the parameters.

Assumption 2 (Local alternatives). The local alternatives (H_{AT}) are specified as

$$\theta_{t,T} = \theta^* + \frac{1}{\sqrt{T}}\, g\left(\gamma, \pi, \frac{t}{T}\right), \qquad (H_{AT}^{(1)})$$

$$a(\theta^*) = \frac{1}{\sqrt{T}}\, \bar{a}, \qquad (H_{AT}^{(2)})$$

where $g(\gamma,\pi,s)$, for $s \in [0,1]$, is a k-dimensional step function, $\gamma \in R^i$, $\pi \in (0,1)^j$ denotes the times of the structural changes as fractions of the sample size (j being the number of such breaks), $a(\theta^*) = 0$ is a possibly nonlinear restriction that identifies the true parameter value under the null hypothesis when there is no structural change, and $\bar{a}$ denotes its local alternative.

Hence, the parameter θ is unknown and possibly time-varying. The class of estimators considered here are extremum estimators that minimize the objective function $Q_T(\theta)$, which depends on both the data and the sample size. The focus will be on the restricted estimator $\tilde{\theta}$:

$$\tilde{\theta} = \arg\min_{\theta\in\Theta} \hat{Q}_T(\theta) \quad \text{s.t. } a(\theta) = 0, \quad \text{where } \hat{Q}_T(\theta) \equiv F_T(\theta)' W_T F_T(\theta), \qquad \textbf{(5)}$$

where $F_T(\theta) = (1/T)\sum_{t=1}^{T} f(x_t,\theta)$ is the sample analogue of $E(f(x_t,\theta))$, the moment condition that is equal to zero at the true parameter value, and $E(.)$ is the expected value function. The moment condition is such that $f\colon R^r \times R^k \to R^m$ and W_T is a (sequence of) positive semidefinite matrices. Note that our framework allows for both exactly and overidentified generalized method of moments (GMM).

The next assumptions are sufficient to ensure consistency and identification of the estimator (see Sowell, 1996, p. 1100; see also Andrews, 1993).[2] Furthermore, the class of estimators is restricted to efficient GMM estimators, and Assumption 6 provides a sufficient condition for efficiency.

Assumption 3 (Identification). $\lim_{T\to\infty} E[f(x,\theta)] = 0$ only if $\theta = \theta^*$.

Assumption 4 (Smoothness and boundedness). (i) $\theta^* \in interior(\Theta)$; (ii) $f(x,\theta)$ is continuously partially differentiable in a neighborhood Υ of θ^*, $\forall\theta \in \Theta$; (iii) the functions $f(x,\theta)$ and $\nabla_\theta f(x,\theta) \equiv (\partial/\partial\theta) f(x,\theta)$ are measurable functions of x $\forall\theta \in \Theta$ and $E[\|f(x,\theta^*)\|^2]$ is finite; (iv) $E[f(x_t,\theta^*)] = 0, E[f(x_t,\theta^*)'f(x_t,\theta^*)] < \infty$, and $\sup_{\theta\in\Theta}\|f(x_t,\theta)\| < \infty$ $\forall t = 1,\dots,T$ and $T = 1,2,\dots$; each element of $f(x_t,\theta_{t,T})$ is uniformly square integrable $\forall t = 1,\dots,T$ and $T = 1,2,\dots$; (v) $M = \lim_{T\to\infty} E[\nabla_\theta f(x,\theta^*)] \in R^{m\times k}$ has full column rank, where $\nabla_\theta f(x,\theta^*) = (\partial/\partial\theta) f(x,\theta)|_{\theta=\theta^*}$ and $M'W_T M$ is nonsingular; (vi) $\{x_t\}$ is strong mixing with strong mixing coefficients $\{\alpha(n)^{1-2/\beta}\} < \infty$ with $\beta > 2$, and the individual elements of $f(x_t,\theta_{t,T})$ have finite absolute moments $E[|f^{(i)}(x_t,\theta_{t,T})|^\beta]$ for $i = 1,\dots,m$.

Assumption 5 (Constraints). $a(\theta)$ is continuously partially differentiable in a neighborhood Υ of θ^*, $\forall\theta \in \Theta$; $A \equiv \nabla_\theta a(\theta^*) \in R^{r\times k}$ has rank $r \le k$.

Assumption 6 (Efficiency in the class of GMM estimators). The asymptotic variance of the GMM estimator is efficient in the class of GMM estimators: $\{W_T^{-1}\}_{T=1}^{\infty} \underset{p}{\rightarrow} \Sigma \equiv \lim_{T\to\infty} E[TF_T(\theta^*)F_T(\theta^*)'] \in R^{m\times m}$.

When the alternative hypothesis of interest is *either* $H_{AT}^{(1)}$ *or* $H_{AT}^{(2)}$ then optimal tests are available. In the former case, an optimal test when the break date is known is the Chow (1960) test, and when the break date is unknown, a class of tests with optimal weighted average power is that of Andrews and Ploberger (1994); although Andrews' Sup-LR test (see Andrews, 1993) is not a member of that class, Andrews and Ploberger, 1993, show its asymptotic admissibility against alternatives that are sufficiently distant from the null hypothesis). In case the alternative is $H_{AT}^{(2)}$ *only*, the likelihood ratio test (and the asymptotically equivalent Wald and Lagrange multiplier [LM] tests) is asymptotically locally most powerful among all invariant tests, and, hence, it is optimal (see Engle, 1984).

However, when *both* hypotheses are of interest then considering separately tests for parameter instability and likelihood ratio tests is not sufficient anymore. This paper identifies a class of tests that are optimal, in the sense of having the highest asymptotic local power function for some specified alternatives. This class of tests is discussed in Section 2.3.

2.3. Optimal Tests

We are interested in constructing a LM test statistic for testing jointly alternatives $H_{AT}^{(1)}$ and $H_{AT}^{(2)}$. The test builds on partial sums of the form

$$F_{sT}(\tilde{\theta}) \equiv \frac{1}{T}\sum_{t=1}^{[sT]} f(x_t,\tilde{\theta}), \qquad s \in [0,1], \tag{6}$$

where the partial sums are evaluated at the restricted estimator vector, $\tilde{\theta}$. When Assumptions 1–6 are satisfied, the asymptotic distribution of the partial sums of sample moments under the null and the alternative hypotheses is stated in Results 1 and 2, which follow. For notational convenience, let $\overline{M} \equiv \Sigma^{-1/2}M \in R^{m\times k}$ and partition it as $\overline{M} = (\overline{M}_\beta, \overline{M}_\delta)$. Also, let $\Rightarrow$ denote weak convergence to the relevant stochastic process and $\underset{p}{\rightarrow}$ denote convergence in probability.

RESULT 1 (Distribution under the alternative hypothesis). *If Assumptions 1–6 are satisfied, then*

$$\sqrt{T}W_T^{1/2}F_{sT}(\tilde{\theta}) \Rightarrow Z(s) \equiv B_m(s) - s\overline{H}B_m(1) - s\overline{M}\overline{D}'\bar{a} - \overline{M}\int_0^s g(\gamma,\pi,r)\,dr + s\overline{M}D^{-1/2}HD^{1/2}\int_0^1 g(\gamma,\pi,r)\,dr, \tag{7}$$

where $B_m(\cdot)$ is an m-dimensional standard Brownian motion, $\overline{D}' \equiv D^{-1}A' \times (AD^{-1}A')^{-1}$, $D \equiv \overline{M}'\overline{M}$, $H \equiv I_k - D^{-1/2}A'(AD^{-1}A')^{-1}AD^{-1/2}$, I_k is a k-dimensional identity matrix, $\overline{H} \equiv \overline{M}D^{-1/2}HD^{-1/2}\overline{M}'$, and both H and $\overline{H}$ are idempotent with rank equal to $(k-r)$.

See Appendix A for proofs. Result 2 shows the asymptotic distribution of the standardized moment condition under the null hypothesis that there is no parameter instability in any of the coefficients and that a subset of parameters satisfies some restriction condition as follows.

Assumption 7 (Null hypothesis). Under the null hypothesis (H_0):

$$\theta_{t,T} = \theta^* \quad \text{for all } t, T, \tag{H_0}$$

where θ^* satisfies $a(\theta^*) = 0$.

RESULT 2. (Distribution under the null hypothesis). *If Assumptions 1 and 3–7 are satisfied then*

$$CW_T^{1/2}\sqrt{T}F_{sT}(\tilde{\theta}) \Rightarrow \begin{pmatrix} BB_{k-r}(s) \\ B_r(s) \\ B_{m-k}(s) \end{pmatrix} \tag{8}$$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

for an orthonormal matrix C such that $\overline{H} = C'\Lambda C$, $CC' = I_m$, and $\Lambda = \begin{pmatrix} I_{k-r} & 0 \\ 0 & 0 \end{pmatrix}$. Here $BB_{k-r}(s)$ is a $(k-r)$-dimensional Brownian bridge and $[B_r(s)', B_{m-k}(s)']'$ is an $(m-k+r)$-dimensional vector Brownian motion. The Brownian motions and the Brownian bridges are independent.

Note that, under the null hypothesis, the asymptotic distribution of the standardized partial sum of moment conditions is composed by both Brownian bridges and Brownian motions. The $(k-r)$-dimensional Brownian bridge component derives from the parameters that are not specified under the null. In fact, this component is a partial sum of mean zero moment conditions, where the zero mean is obtained by estimating the drift.[3]

The alternative hypothesis will add drift components to the moment conditions, as Result 1 shows. In particular, the drift components originate both from deviations from the parameter stability hypothesis and from deviations from the specified null hypothesis on the value of the parameters. For the local alternatives considered in this paper, the normalized partial sum of the sample moments evaluated under the null hypothesis converges to a stochastic process denoted by $Z(s)$. Under the local alternative, $Z(s)$ satisfies the following stochastic differential equation:

$$dCZ(s) = \begin{pmatrix} dBB_{k-r}(s) \\ dB_{m-k+r}(s) \end{pmatrix} + Cv(s)\,ds, \tag{9}$$

where $v(s) \equiv -\overline{M}\overline{D}'\bar{a} - \overline{M}g(\gamma,\pi,s) + \overline{M}D^{-1/2}HD^{1/2}(\int_0^1 g(\gamma,\pi,r)\,dr)$. Under the null hypothesis, the same expression holds with $v(s) = 0$. To get some insight, rearrange (7):[4]

$$\begin{aligned} Z(s) = {} & (I_m - \overline{H})\left(B_m(s) - \overline{M}\left(\int_0^s g(\gamma,\pi,v)\,dv\right) - s\overline{M}\overline{D}'\bar{a}\right) \\ & + \overline{H}\left(BB_m(s) - \overline{M}\int_0^s \left(g(\gamma,\pi,v) - \left(\int_0^1 g(\gamma,\pi,r)\,dr\right)\right)dv\right) \end{aligned} \tag{10}$$

so that

$$\begin{aligned} dCZ(s) &= (I-\Lambda)C[dB_m(s) - \overline{M}g(\gamma,\pi,s) - \overline{M}\overline{D}'\bar{a}] \\ &\quad + \Lambda C\left[dBB_m(s) - \overline{M}\left(g(\gamma,\pi,s) - \left(\int_0^1 g(\gamma,\pi,r)\,dr\right)\right)\right] \\ &= \begin{pmatrix} dBB_{k-r}(s) \\ dB_{m-k+r}(s) \end{pmatrix} - \begin{pmatrix} C^{(1)}\overline{M}\left(g(\gamma,\pi,s) - \left(\int_0^1 g(\gamma,\pi,r)\,dr\right)\right) \\ C^{(2)}\overline{M}(g(\gamma,\pi,s) + \overline{D}'\bar{a}) \end{pmatrix}, \end{aligned}$$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

where $C^{(1)}$ and $C^{(2)}$ are, respectively, the first $(k-r)$ and the last $(m-k+r)$ rows of C. Thus, the null hypothesis puts restrictions on *both* the Brownian motions and the Brownian bridge components. In fact, it is a joint hypothesis on parameter instability (affecting the Brownian bridge component) and on the parameters (affecting the Brownian motion component). This differs from the Sowell (1996) case (see the discussion following his eqn. (3), p. 1091), where the alternative *only* places restrictions over Brownian bridges. However, Sowell (1996) derived optimal tests in terms of the Radon–Nikodym derivative of the measure implied by the null hypothesis for *both* the Brownian motion and the Brownian bridge components (see the proof of his Thm. 3), so we can apply a similar argument. Thus, the test with the greatest weighted average power, according to some weighting functions $R(\eta,\pi)$ (on η for every π) and $J(\pi)$ (on π), rejects the joint null hypothesis of no structural break and $a(\theta^*)=0$ if

$$\iint \zeta(\eta,\pi)\, dR(\eta,\pi)\, dJ(\pi) \geq k_\alpha, \tag{11}$$

$$\text{where} \quad \zeta(\eta,\pi) = \exp\left\{\int_0^1 v(s)'\, dZ(s) - \frac{1}{2}\int_0^1 v(s)'v(s)\, ds\right\}, \tag{12}$$

$\eta \equiv [\bar{a},\gamma']' \in R^{2p\times 1}$, and k_α is defined so that the test has size α.

3. SPECIAL TESTS[5]

The leading case of the class of alternatives for structural break is that of alternatives that are linear in the parameters, that is, $g(\gamma,\pi,s) = \widetilde{G}(\pi,s)\gamma$. In the case of a single structural break, $\widetilde{G}(\pi,s) = 1(s\geq\pi)G$, where $1(s\geq\pi)$ is the indicator function, equal to one if $s\geq\pi$ and zero otherwise, and G is a $(k\times p)$ matrix identifying the p-dimensional vector of time-varying parameters, say, $G=[I_p \quad 0_{q\times p}]$. Let us define

$$A(\pi) = \begin{pmatrix} -\bar{D}\bar{M}' & 0 \\ -(1-\pi)G'A'\bar{D}\bar{M}' & G'\bar{M}' \end{pmatrix}\begin{pmatrix} Z(1) \\ Z(\pi)-\pi Z(1)\end{pmatrix}, \tag{13}$$

$$V(\pi) = \begin{pmatrix} \bar{D}D\bar{D}' & (1-\pi)\bar{D}D^{1/2}(I-H)D^{1/2}G \\ (1-\pi)G'D^{1/2}(I-H)D^{1/2}\bar{D}' & (1-\pi)G'D^{1/2}[I-(1-\pi)H]D^{1/2}G \end{pmatrix}. \tag{14}$$

The optimal test statistic described by (11) becomes $\iint \exp\{\eta'A(\pi) - \frac{1}{2}\eta'V(\pi)\eta\}\, dR(\eta,\pi)\, dJ(\pi)$. As in Sowell (1996), different choices of the weighting function $R(\eta,\pi)$ lead to different test statistics. The weighting function considered here is an $(r+p)$-dimensional multivariate normal distribution with zero mean and covariance $U(\pi)$. When the time of the break is not known and we are interested in the test statistic that gives equal weight to alternatives that are equally difficult to detect when π is known, so that $U(\pi)^{-1} = (1/c)V(\pi)$, then the test statistic in (11) becomes $\int_{\Pi}(\exp\{\frac{1}{2}(c/(1+c))\Phi^*(\pi)\})\,dJ(\pi)$ where Π is the support of $J(\pi)$ and $\Phi_T^*(\pi) = A(\pi)'V(\pi)^{-1}A(\pi)$. The latter is a Wald test for the fixed and known π scenario. The test statistic can be estimated as

$$TS_{c,T}^{AP*} = \int_{\Pi}\left(\exp\left\{\frac{1}{2}\,\frac{c}{1+c}\,\Phi_T^*(\pi)\right\}\right) dJ(\pi), \tag{15}$$

$$\Phi_T^*(\pi) = LM_1 + LM_2(\pi),$$

where

$$LM_1 \equiv \left(\frac{1}{T}\right)\left(W_T^{1/2}\sum_{t=1}^{T} f_t(\tilde{\theta})\right)'\hat{\Omega}_1\left(W_T^{1/2}\sum_{t=1}^{T} f_t(\tilde{\theta})\right)$$

$$\begin{aligned} LM_2(\pi) \equiv{}& \frac{1}{\pi(1-\pi)}\,\frac{1}{T}\left(\sum_{t=1}^{[T\pi]} f_t(\tilde{\theta}) - \pi\sum_{t=1}^{T} f_t(\tilde{\theta})\right)'\hat{\Sigma}^{-1/2\prime}\hat{\Omega}_2 \\ &\times \hat{\Sigma}^{-1/2}\left(\sum_{t=1}^{[T\pi]} f_t(\tilde{\theta}) - \pi\sum_{t=1}^{T} f_t(\tilde{\theta})\right), \end{aligned} \tag{16}$$

$$\hat{\Omega}_1 = \hat{C}_1'(\hat{C}_1\hat{C}_1')^{-1}\hat{C}_1, \qquad \hat{C}_1 \equiv (\hat{A}\hat{D}^{-1}\hat{A}')\hat{A}\hat{D}^{-1}\hat{\hat{M}}',$$

$$\hat{\Omega}_2 = \hat{C}_2'(\hat{C}_2\hat{C}_2')^{-1}\hat{C}_2, \qquad \hat{C}_2 \equiv G'\hat{\hat{M}}',$$

$$\hat{D} = \hat{\hat{M}}'\hat{\hat{M}}, \qquad \hat{A} = \frac{1}{T}\sum_{t=1}^{T}\nabla_\theta a(\tilde{\theta}), \qquad \hat{\hat{M}} = \hat{\Sigma}^{-1/2}\frac{1}{T}\sum_{t=1}^{T}\nabla_\theta f_t(\tilde{\theta}),$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\left(f_t(\tilde{\theta}) - \frac{1}{T}\sum_{t=1}^{T} f_t(\tilde{\theta})\right)\left(f_t(\tilde{\theta}) - \frac{1}{T}\sum_{t=1}^{T} f_t(\tilde{\theta})\right)'$$

if $f_t(.)$ are independent and identically distributed (i.i.d.); otherwise $\hat{\Sigma}$ is estimated with a Newey–West heteroskedasticity and autocorrelation consistent (HAC) estimator. The limiting distribution of this test statistic under the null hypothesis is described in the following proposition.

PROPOSITION 1. *Let Assumptions 1–6 hold. The test statistic for testing* $a(\theta^*) = 0$ *against* $H_{AT}^{(1)}$ *and* $H_{AT}^{(2)}$ *with the greatest average power according to the weighting function* $R(\eta,\pi) \sim N(0, cV(\pi)^{-1})$*, for* $V(\pi)$ *defined in (14), is the test statistic defined in (15). Its asymptotic distribution under the null hypothesis is*

$$TS_{c,T}^{AP*} \Rightarrow \int_{\Pi}\left(\exp\left\{\frac{1}{2}\,\frac{c}{1+c}\,\Phi^*(\pi)\right\}\right) dJ(\pi), \tag{17}$$

$$\Phi^*(\pi) \equiv \left(\frac{BB_p(\pi)'BB_p(\pi)}{\pi(1-\pi)} + B_r(1)'B_r(1)\right). \tag{18}$$

Proposition 1 shows that $TS_{c,T}^{AP*}$ is a weighted average of LM tests. As noted previously, the difference between the asymptotic distribution of the tests defined in this paper and that of the test for structural break only is that the latter does not have the $B_r(1)'B_r(1)$ component. This component arises from testing restrictions on θ over the whole sample. In fact, it corresponds to a centered chi-square with r degrees of freedom, the usual limiting distribution of the Wald test statistic for testing hypotheses on a parameter vector. Appendix A shows that both the tests for structural break and the classical tests obtain as special cases of (24).

Although this paper is mainly concerned about testing a null hypothesis on the parameters in the presence of possible parameter instability, instabilities may also affect other aspects of the model, namely, the overidentifying restrictions (OIRs). Hall and Sen (1999) and Sowell (1996) show that the population moment conditions can be decomposed into two orthogonal components: identifying restrictions—the part used in estimation—and overidentifying restrictions—the part unused in estimation. Hall and Sen (1999) propose tests for the structural stability of the OIRs. Their approach turns out to be useful to discriminate between situations in which the instability is confined to the parameters alone and those in which the instability permeates other aspects of the model. In what follows, we examine the relationship between the tests proposed by Hall and Sen (1999) and those proposed in the present paper, and we show that the Hall and Sen (1999) results do carry over to the tests proposed in this paper.

The tests proposed by Hall and Sen (1999) are as follows:

$$\sup O_T \equiv \sup_{\pi} O_T(\pi), \qquad av\, O_T \equiv \int_{\Pi} O_T(\pi)\, dJ(\pi),$$
$$\exp O_T \equiv \log\left[\int_{\Pi} \exp\left(\frac{1}{2}\, O_T(\pi)\right) dJ(\pi)\right],$$

where $O_T(\pi) \equiv T\bar{F}_T(\hat{\beta}_1, \hat{\beta}_2, \hat{\delta}, \pi)'\hat{\Gamma}\bar{F}_T(\hat{\beta}_1, \hat{\beta}_2, \hat{\delta}, \pi)$. Hall and Sen (1999) find that their test statistics are asymptotically independent of tests for parameter instability. They also show that their tests have no local power against parameter variation and tests for parameter variation have no local power against instability in the OIRs. This latter result follows from the fact that the components of $O_T(\pi)$ are orthogonal to the components of $LM_2(\pi)$. In fact, they show that

$$\bar{F}_T(\hat{\beta}_1, \hat{\beta}_2, \hat{\delta}, \pi) = [I_2 \otimes (I - \Sigma^{1/2}\overline{M}(\overline{M}'\overline{M})^{-1}\overline{M}'\Sigma^{-1/2})]\bar{F}_T(\beta_0, \beta_0, \delta_0, \pi) + o_p(1)$$

(see Hall and Sen, 1999, eqn. (A.4); Andrews, 1991, p. 848, for the more general case of subsets of parameters), where $\otimes$ denotes the Kronecker product. Thus, the rescaled moment conditions in $O_T(\pi)$ become

$$(I_2 \otimes \Sigma^{-1/2})\bar{F}_T(\hat{\beta}_1, \hat{\beta}_2, \hat{\delta}, \pi) = [I_2 \otimes (I - \overline{M}(\overline{M}'\overline{M})^{-1}\overline{M}')] \times (I_2 \otimes \Sigma^{-1/2})\bar{F}_T(\beta_0, \beta_0, \delta_0, \pi) + o_p(1),$$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

and it is clear that they are orthogonal to $LM_2(\pi)$, which instead builds on $\overline{M}$ (see equation (16)), as $\overline{M}'(I - \overline{M}(\overline{M}'\overline{M})^{-1}\overline{M}') = 0$.

Note that all of the preceding results still hold in our framework. Also, note that the components in LM_1 are orthogonal to the components of $O_T(\pi)$ too, as LM_1, like $LM_2(\pi)$, builds on $\overline{M}$. Thus, the results in Hall and Sen (1999) do carry over to the test statistics QLR_T^*, $Mean\text{-}Wald_T^*$, and $Exp\text{-}Wald_T^*$. More details are provided in Section 4.

From now until the end of this section, we specialize the preceding findings to situations in which the researcher is interested in testing hypotheses on a subset of the parameters. This is discussed in the following corollary.

COROLLARY 1 (Null hypotheses on subsets of parameters). *Let the parameter vector $\theta \in R^k$ be partitioned as $\theta = [\beta', \delta']'$, where $\beta \in R^p$ and $\delta \in R^q$. Let Assumptions 1 and 3–6 hold. Let Assumption 2 be replaced by Assumption 2′: $\beta_{t,T} = \beta^* + (1/\sqrt{T}) g_\beta(\gamma, \pi, t/T)$ and $\beta^* = \beta_0 + (1/\sqrt{T})\beta_A$. It follows that*

$$\sqrt{T}\, W_T^{1/2} F_{sT}(\beta_0, \tilde{\delta}) \Rightarrow B_m(s) - s\bar{P}_\delta B_m(1) - s(I_m - \bar{P}_\delta)\overline{M}_\beta \beta_A$$
$$- \overline{M}_\beta \int_0^s g_\beta(\gamma, \pi, r)\, dr + s\bar{P}_\delta \overline{M}_\beta \int_0^1 g_\beta(\gamma, \pi, r)\, dr, \quad \textbf{(19)}$$

where $\bar{P}_\delta \equiv \overline{M}_\delta(\overline{M}_\delta' \overline{M}_\delta)^{-1}\overline{M}_\delta' \in R^{m\times m}$. Also, $v(s)$ in (9) becomes $v(s) \equiv -\overline{M}_\beta g_\beta(\gamma, \pi, s) + \bar{P}_\delta \overline{M}_\beta (\int_0^1 g_\beta(\gamma, \pi, v)\, dv) - (I_m - \bar{P}_\delta)\overline{M}_\beta \beta_A$.

We will finally consider special cases of $TS_{c,T}^{AP*}$ that have been considered in the literature for tests for structural break only. Each of these special cases has greatest weighted average power against particular forms of parameter instability. We will analyze the form that the optimal test proposed in this paper assumes for these particular forms of parameter instability.

Andrews and Ploberger Test. Let $\beta_t = \beta_1(\pi)$ for $t = 1, 2, \ldots, [T\pi]$ and $\beta_t = \beta_2(\pi)$ for $t = [T\pi] + 1, \ldots, T$, where $[\cdot]$ denotes the greatest integer function. Also, to simplify notation, let $f_t(\beta_k) \equiv f_t(x_t, \beta_k, \delta)$, $f_t(\hat{\beta}_k) \equiv f_t(x_t, \hat{\beta}_k, \hat{\delta})$, $f_t(\beta_0) \equiv f_t(x_t, \beta_0, \tilde{\delta})$, $k = 1, 2$. Let $\theta(\pi) \equiv (\beta_1, \beta_2, \delta)$, $\hat{\theta}(\pi) \equiv (\hat{\beta}_1, \hat{\beta}_2, \hat{\delta})$ be the unrestricted GMM estimator under the hypothesis that there is a break at the fraction $[T\pi]$ of the sample and let $\tilde{\theta}(\pi)$ be the constrained estimator. Thus, the Wald test for a fixed and known π can be estimated as either[6]

$$\text{Wald: } \Phi_T^*(\pi) = T(R\hat{\theta}(\pi) - r)'(RV(\hat{\theta}(\pi))R')^{-1}(R\hat{\theta}(\pi) - r), \quad \textbf{(20)}$$

$$\text{Distance metric form: } \Phi_T^*(\pi) = \bar{Q}(\tilde{\theta}(\pi)) - \bar{Q}(\hat{\theta}(\pi)), \quad \text{or} \quad \textbf{(21)}$$

$$\text{Lagrange multiplier: } \Phi_T^*(\pi) = LM_1 + LM_2(\pi), \quad \textbf{(22)}$$

where notation is in Table 1. The table assumes that $f_t(.)$ consists of mean zero uncorrelated random variables. When $f_t(.)$ consists of mean zero but serially correlated random variables then consistent estimation of $\hat{\Sigma}_1$, $\hat{\Sigma}_2$, $\hat{\Sigma}$, and $\hat{\Gamma}$ requires a HAC estimator (e.g., see Newey and West, 1987). Note that (22) is particularly easy to calculate. It is simply the sum of the two LM tests to test $H_{AT}^{(1)}$ and $H_{AT}^{(2)}$ separately. Then, Proposition 2 follows.

PROPOSITION 2. *Let Assumptions 1, 2′, and 3–6 hold. The test statistic for testing* $\beta = \beta^*$ *against* $\beta_{t,T} = \beta^* + (1/\sqrt{T})\beta_A + (1/\sqrt{T})\gamma 1(s \geq \pi)$ *with the greatest average power according to the weighting function* $R(\eta,\pi) \sim N(0, cV(\pi)^{-1})$*, for* $V(\pi)$ *defined in (14) with either (20), (21), or (22). Its asymptotic distribution under the null hypothesis is*

$$TS_c^{AP*} \Rightarrow \int_\Pi \left(\exp\left\{ \frac{1}{2} \frac{c}{1+c} \Phi^*(\pi) \right\} \right) dJ(\pi), \tag{23}$$

$$\Phi^*(\pi) \equiv \left(\frac{BB_p(\pi)' BB_p(\pi)}{\pi(1-\pi)} + B_p(1)' B_p(1) \right). \tag{24}$$

As special cases, we have

$$(a)\ \ c \to \infty\text{: } \textit{Exp-Wald}_T^* \equiv \operatorname*{plim}_{c\to\infty} TS_{c,T}^{AP*} \Rightarrow \int_\Pi \left(\exp\left\{ \frac{1}{2} \Phi^*(\pi) \right\} \right) dJ(\pi), \tag{25}$$

$$(b)\ \ c \to 0\text{: } \textit{Mean-Wald}_T^* \equiv \operatorname*{plim}_{c\to 0} 2 \left(\frac{TS_{c,T}^{AP*} - 1}{c} \right) \Rightarrow \int_\Pi \Phi^*(\pi)\, dJ(\pi). \tag{26}$$

The special cases that correspond to extreme values of the parameter c are similar to those in Andrews and Ploberger (see also Andrews, Lee, and Ploberger, 1996). When $c \to \infty$ ($c \to 0$), more weight is assigned to alternatives about parameter instability further from (closer to) the null hypothesis.

Andrews (1993) Sup-LR *Test.* A test statistic commonly considered in the literature of structural breaks is the Quandt likelihood ratio (QLR) test statistic (or Sup-LR test), which is the supremum (over all possible break dates) of the Chow statistic designed for these alternatives for a fixed break date. Andrews (1993) derived its asymptotic distribution. The modified QLR test statistic for the alternatives specified in this paper can be obtained by letting $c/(1+c) \to \infty$ in (17), which gives

$$QLR_T^* = \sup_\pi \Phi_T^*(\pi). \tag{27}$$

The limiting distribution of (27) under the null hypothesis is given in the following proposition.

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

Table 1. Notation related to equations (20), (21), and (22)

$$\bar{Q}(\beta_1,\beta_2,\delta) \equiv \bar{F}_T(\beta_1,\beta_2,\delta,\pi)'\hat{\Gamma}\bar{F}_T(\beta_1,\beta_2,\delta,\pi)$$

Wald test

$$\hat{\theta}(\pi) = \arg\min_{\beta_1,\beta_2,\delta} \bar{Q}(\beta_1,\beta_2,\delta)$$

$$\bar{F}_T(\beta_1,\beta_2,\delta,\pi) = \left(\frac{1}{T}\sum_{t=1}^{[T\pi]} f_t(\beta_1)', \frac{1}{T}\sum_{t=[T\pi]+1}^{T} f_t(\beta_2)'\right)' \in R^{2m\times 1}$$

$$R \equiv \begin{pmatrix} I_p & -I_p & 0_{p\times q} \\ \pi I_p & (1-\pi)I_p & 0_{p\times q}\end{pmatrix}, \qquad r \equiv \begin{pmatrix} 0_{p\times 1} \\ \beta_0 \end{pmatrix}$$

$$V(\hat{\theta}(\pi)) = [M(\pi)'\hat{\Gamma}M(\pi)]^{-1}$$

$$M(\pi) \equiv \begin{pmatrix} \pi M_\beta & 0 & \pi M_\delta \\ 0 & (1-\pi)M_\beta & (1-\pi)M_\delta \end{pmatrix} \in R^{2m\times 1}$$

$$\hat{\Gamma} \equiv \begin{pmatrix} \pi\hat{\Sigma}_1 & 0 \\ 0 & (1-\pi)\hat{\Sigma}_2 \end{pmatrix}^{-1} \underset{p}{\rightarrow} \Gamma = \begin{pmatrix} \pi\Sigma & 0 \\ 0 & (1-\pi)\Sigma \end{pmatrix}^{-1} \in R^{2m\times 2m}$$

$$\hat{\Sigma}_1 = \frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]}\left(f_t(\hat{\beta}_1) - \frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]} f_t(\hat{\beta}_1)\right)\left(f_t(\hat{\beta}_1) - \frac{1}{[T\pi]}\sum_{t=1}^{[T\pi]} f_t(\hat{\beta}_1)\right)'$$

$$\hat{\Sigma}_2 = \frac{1}{T-[T\pi]}\sum_{t=[T\pi]+1}^{T}\left(f_t(\hat{\beta}_2) - \frac{\sum_{t=[T\pi]+1}^{T} f_t(\hat{\beta}_2)}{T-[T\pi]}\right)\left(f_t(\hat{\beta}_2) - \frac{\sum_{t=[T\pi]+1}^{T} f_t(\hat{\beta}_2)}{T-[T\pi]}\right)'$$

Lagrange multiplier test

$$LM_1 \equiv 1/T\left(W_T^{1/2}\sum_{t=1}^{T} f_t(\beta_0)\right)'\hat{\Omega}_1\left(W_T^{1/2}\sum_{t=1}^{T} f_t(\beta_0)\right)$$

$$LM_2(\pi) \equiv \frac{1}{\pi(1-\pi)}\frac{1}{T}\left(\sum_{t=1}^{[T\pi]} f_t(\beta_0) - \pi\sum_{t=1}^{T} f_t(\beta_0)\right)'$$

$$\times\,\hat{\Sigma}^{-1/2\prime}\hat{\Omega}_2\hat{\Sigma}^{-1/2}\left(\sum_{t=1}^{[T\pi]} f_t(\beta_0) - \pi\sum_{t=1}^{T} f_t(\beta_0)\right)$$

$$\hat{\Omega}_1 = C_1'(C_1C_1')^{-1}C_1, \qquad C_1 = \overline{M}_\beta'(I-\bar{P}_\delta)$$

$$\hat{\Omega}_2 = C_2'(C_2C_2')^{-1}C_2, \qquad C_2 = \overline{M}_\beta'$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\left(f_t(\beta_0) - \frac{1}{T}\sum_{t=1}^{T} f_t(\beta_0)\right)\left(f_t(\beta_0) - \frac{1}{T}\sum_{t=1}^{T} f_t(\beta_0)\right)'$$

Distance metric test

$$\tilde{\theta}(\pi) = \arg\min_{\beta_1,\beta_2,\delta} \bar{Q}(\beta_1,\beta_2,\delta) \quad \text{s.t. } R\theta(\pi) = r$$

Note: Table 1 assumes that $f_t(.)$ consists of mean zero uncorrelated random variables. When $f_t(.)$ consists of mean zero but serially correlated random variables, then consistent estimation of $\hat{\Sigma}_1$, $\hat{\Sigma}_2$, $\hat{\Sigma}$, and $\hat{\Gamma}$ requires a HAC estimator (e.g. Newey and West, 1987).

PROPOSITION 3. *Let Assumptions 1, 2′, and 3–6 hold. The test statistic for testing* $\beta = \beta^*$ *against* $\beta_{t,T} = \beta^* + (1/\sqrt{T})\beta_A + (1/\sqrt{T})\gamma 1(s \geq \pi)$ *with the greatest average power according to the weighting function* $R(\eta,\pi) \sim N(0, cV(\pi)^{-1})$, *for* $V(\pi)$ *defined in (14) and c such that* $c/(1+c) \to \infty$, *is (27), whose asymptotic distribution under the null hypothesis is*

$$QLR_T^* \Rightarrow \sup_{\pi} \Phi^*(\pi). \tag{28}$$

Nyblom (1989) Test. Another test for parameter instability is that considered by Nyblom (1989) and Nyblom and Mäkeläinen (1983). These authors derive the locally most powerful invariant (to translations and scale transformations) test for constancy of the parameter process against the alternative that the parameters follow a random walk process:[7]

$$\beta_t = \beta_{t-1} + e_t, \qquad e_t \sim N\left(0, \frac{1}{T^2}\sigma_e^2 \bar{\Gamma}\right). \tag{29}$$

The modified Nyblom test statistic for testing whether β_t is equal to β_0 is

$$Nyblom_T^* = \int_0^1 (T \cdot \nabla_\beta Q_{1,[\pi T]}(\beta_0, \tilde{\delta})' \hat{\Omega}_N^{-1} \nabla_\beta Q_{1,[\pi T]}(\beta_0, \tilde{\delta})) J(\pi)\, d\pi, \tag{30}$$

where $\hat{\Omega}_N = \overline{M}_\beta'(I_m - \bar{P}_\delta')\overline{M}_\beta \in R^{p \times p}$ and the gradient of the objective function is defined as

$$\nabla_\beta Q_{1,[\pi T]}(\beta_0, \tilde{\delta}) \equiv \frac{1}{T} \sum_{t=1}^{[T\pi]} \nabla_\beta F_T(\beta_0, \tilde{\delta})' \Sigma^{-1} f_t(x_t, \beta_0, \tilde{\delta}).$$

Note that (30) is a generalization of the locally best invariant test statistic proposed by Nabeya and Tanaka (1988) for the case in which β is known and equal to β_0. The test proposed in this paper is more general than that of Nabeya and Tanaka, as estimation is not restricted to the ordinary least squares case and β can be a vector. Appendix A shows that the asymptotic distribution of the modified Nyblom statistic under the null hypothesis is as follows.

PROPOSITION 4. *Let Assumptions 1, 2′, and 3–6 hold. The test statistic for testing* $\beta = \beta_0$ *against* $\beta_{t,T} = \beta_0 + (1/\sqrt{T})\beta_A + \beta_{t-1,T} + e_t$, $e_t \sim N(0, (1/T^2)\sigma_e^2 \bar{\Gamma})$ *with the greatest average power according to the weighting function* $R(\eta,\pi) \sim N(0, cV(\pi)^{-1})$, *for* $V(\pi)$ *defined in (14), is (30). Its asymptotic distribution under the null hypothesis is*

$$Nyblom_T^* \Rightarrow \int_0^1 B_p(\pi)' B_p(\pi) J(\pi)\, d\pi. \tag{31}$$

Tables B1–B4 in Appendix B report critical values for the optimal tests for $J(\pi)$ uniformly distributed on $[0.15, 0.85]$.[8] The significance levels considered in the tables are 10%, 5%, 2.5%, and 1%. The critical values are obtained by simulating the asymptotic distributions described in this section. The number of Monte Carlo replications is 5,000. Notice that all the values in Tables B1–B3 are higher than those for the corresponding tests for structural break only, the reason being that the optimal tests add the nonnegative component $B_p(1)'B_p(1)$ (see equation (24)).

4. ASYMPTOTIC LOCAL POWER ANALYSIS

The local power properties of the optimal tests derived previously can be compared with those of tests for parameter instability only and those of tests for $a(\theta^*) = 0$ only. The comparison can be made both theoretically and by Monte Carlo simulations.

Let us first consider the theoretical local power properties of the various tests. To facilitate a comparison with the tests existing in the literature, we focus on the tests discussed in the second part of Section 3, and, for brevity, we analyze only (25).[9] Let $\tilde{\theta} = (\beta_0, \tilde{\delta})$. From (22) and (25), and using the notation in Table 1, we have that

$$\log(\text{Exp-Wald}_T^*) = \frac{1}{2} LM_1 + \log \int_\Pi \left(\exp\left\{ \frac{1}{2} LM_2(\pi) \right\} \right) dJ(\pi).$$

Appendix A shows that

(a) $\Omega_1^{1/2}(1/\sqrt{T}) W_T^{1/2} \sum_{t=1}^{T} f_t(\tilde{\theta}) \Rightarrow Z_p^{(1)}$;
$Z_p^{(1)} \equiv B_p(1) - \Omega_1^{1/2}(I - \bar{P}_\delta)\overline{M}_\beta \beta_A - \Omega_1^{1/2}(I - \bar{P}_\delta)\overline{M}_\beta \int_0^1 g_\beta(s)\, ds.$

(b) $\Omega_2^{1/2}((1/\sqrt{T}) W_T^{1/2}[\sum_{t=1}^{[T\pi]} f_t(\tilde{\theta}) - \pi \sum_{t=1}^{T} f_t(\tilde{\theta})]) \Rightarrow Z_p^{(2)}(\pi)$;
$Z_p^{(2)}(\pi) \equiv BB_p(\pi) - (1-\pi)\Omega_2^{1/2}\overline{M}_\beta \int_0^\pi g_\beta(\gamma, \pi, r)\, dr +$
$\pi \Omega_2^{1/2} \overline{M}_\beta \int_\pi^1 g_\beta(\gamma, \pi, r)\, dr.$[10]

(c) Under the null hypothesis, LM_1 and $LM_2(\pi)$ are asymptotically independent; this follows from the fact that $B_p(1)$ and $BB_p(\pi)$ are independent. Thus, if one performs two tests, LM_1 and $\int_\Pi (\exp\{\frac{1}{2} LM_2(\pi)\})\, dJ(\pi)$, each at size $1 - \sqrt{1-\alpha}$, then the joint test will have size α. However, this two-stage test, by construction, will not have the highest weighted average power according to the weight function in Proposition 1.

(d) $LM_1 \Rightarrow Z_p^{(1)\prime} Z_p^{(1)}$, and thus it may have no power to detect structural breaks (the alternative described in $H_{AT}^{(1)}$). In fact, $Z_p^{(1)}$ does not depend on $\int_0^\pi g_\beta(s)\, ds$. Thus, for example, the power versus alternatives where the break function is such that $\int_0^1 g_\beta(s)\, ds = 0$ will be equal to the size.

(e) $LM_2 \Rightarrow Z_p^{(2)}(\pi)' Z_p^{(2)}(\pi)$, and thus it has no power to detect constant shifts in the parameters (the alternative described in $H_{AT}^{(2)}$). In fact, $Z_p^{(2)}$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

does not depend on β_A. Thus, for example, the power versus alternatives in which there is no break but $\beta_A \neq 0$ will be equal to the size.

(f) The test for parameter instability only can be obtained by substituting $a(\theta) = 0$, $A = 0$, $\bar{H} = I$ in the proof of Result 1, so that asymptotically it behaves like $Z_p^{(2)}(\pi)$, which is the same as in Andrews (1993), and conclusions similar to those in (d) hold. In fact, upon inspection, it is clear that LM_1 is the standard LM test for testing $\beta = \beta_0$, whereas $LM_2(\pi)$ has the same asymptotic distribution as the LM test for parameter instability.

To verify these insights, we perform some Monte Carlo simulations. A variety of DGPs is considered, paying particular attention to situations where the standard tests fail to detect the alternative hypothesis. For simplicity, only a univariate model is considered:

$$y_t = \beta_{t,T} + \epsilon_t \qquad \epsilon_t \sim N(0,1), \qquad T = 100, \qquad \beta_0 = 0 \tag{32}$$

The likelihood ratio LR_1 tests whether the parameter equals β_0, whereas parameter instability tests check whether $\beta_{t,T}$ is constant; optimal tests jointly test the two hypotheses. The parameter instability tests (TVP) considered here are the Andrews and Ploberger exponential Wald tests (*Exp-Wald*$_T$), the Nyblom test (*Nyblom*$_T$), and the Quandt likelihood ratio (*QLR*$_T$). The optimal tests are *Exp-Wald*$_T^*$, *QLR*$_T^*$, and *Nyblom*$_T^*$ defined in Section 3. The nominal size is 5%.

We consider the following DGPs.[11]

Design 1. $\beta_{t,T} = \beta_0 + \beta_A(1/\sqrt{T})\ \forall t$

Figure 1a shows the asymptotic local power of the tests as a function of β_A. It shows that when the parameter is not time-varying, the likelihood ratio LR_1 is the most powerful test, according to the Neyman and Pearson lemma. The test designed to detect structural break, *Exp-Wald*$_T$, has a flat power function around the size of the test, whereas the *Exp-Wald*$_T^*$ test is almost as powerful as the LR_1 test.

Design 2. $\beta_{t,T} = \beta_0 + (1/\sqrt{T})\beta_A 1(t > [T/2])$

Design 2 involves a single break in the data. This particular alternative is both a deviation of the parameter vector from the null hypothesis and a structural break, so all the tests (the most powerful likelihood ratio test, LR^*,[12] the TVP, and the optimal tests) should detect it. This is in fact what Figure 1b shows.

Design 3. $\beta_{t,T} = \begin{cases} \beta_0 + \beta_A(1/\sqrt{T}) & \text{for } t = 1 \text{ up to } t = [T/2] \\ \beta_0 - \beta_A(1/\sqrt{T}) & \text{for } t = [T/2]+1 \text{ up to } T \end{cases}$

Figure 1c shows that, in this design, the shift in the parameter vector is not detected by a simple likelihood ratio (LR_1) because the statistic on which it is

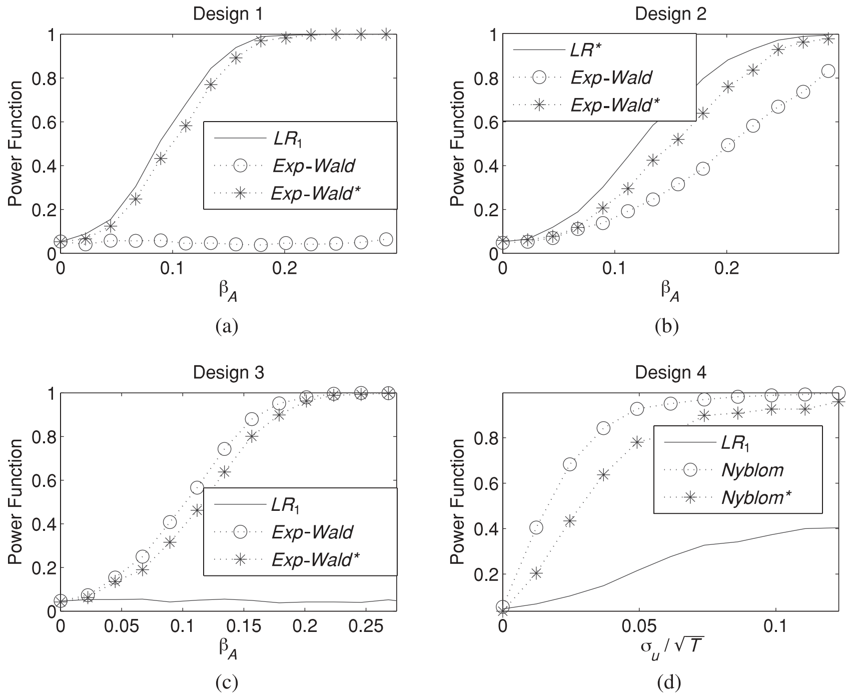

FIGURE 1. Asymptotic power functions of 5% tests for Designs 1–4.

based (the average of the observations) is invariant to it; in fact, notwithstanding the structural break, the average over the whole sample is asymptotically equal to β_0. Although the TVP test is the most powerful, the optimal test is powerful too.

Design 4. $\beta_t = \beta_0 + \beta_{t-1} + u_t$, where $u_t \sim N(0, \sigma_u^2/T^2)$ is independent from ϵ_t and $\sigma_u^2 \geq 0$

The asymptotic local power functions for this design are depicted in Figure 1d as functions of the parameter $\sigma_u^2 \geq 0$. When $\sigma_u^2 = 0$ then β_t is constant, whereas when $\sigma_u^2 \neq 0$ then β_t is a random walk with no drift. The test designed for this hypothesis is the Nyblom test; the LR_1 test is also powerful. The reason is that LR_1 is detecting deviations from the null hypothesis by comparing the sample average with the null hypothesis and the sample average is not a consistent estimate of the true parameter value. Note that the optimal Nyblom test is powerful too.

The results of the simulations suggest the following conclusions. First, *the tests that maintain some power across all the designs considered here are the optimal tests*. For all the other tests there is at least one design (a particular

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

direction away from the null hypothesis) in which the power is flat around the size of the test. Hence, they are not "robust" across designs, whereas the optimal tests are. Second, let us consider a two-stage testing procedure, where the first stage tests whether there is a structural break (by using either QLR_T or *Exp-Wald*$_T$) and the second stage, conditionally on the first stage, tests hypotheses on the parameters (by using the LR_1 test). Let the tests be labeled "*Seq.QLR*" and "*Seq.Exp-Wald*," respectively. In the special cases considered in Section 3 (obtained with particular weighting matrices), the two stages of the test are asymptotically independent. By choosing a size equal to $1 - \sqrt{1 - 0.95}$, the joint significance level will be the desired nominal level, 0.95. Figure 2 shows that there is no clear ranking between the sequential tests and the optimal tests. The power ranking will depend on the direction of the alternative hypothesis. However, by construction, the optimal tests will have the greatest average local power. Two-stage independent tests have advantages and disadvantages. The advantages are that if we reject we know which part of the alternative we reject and that the first-stage test could be used if the researcher is unsure about which elements of the parameter vector are subject to instability. The disadvantage is that they will not have the optimal weighted average power for alternatives that

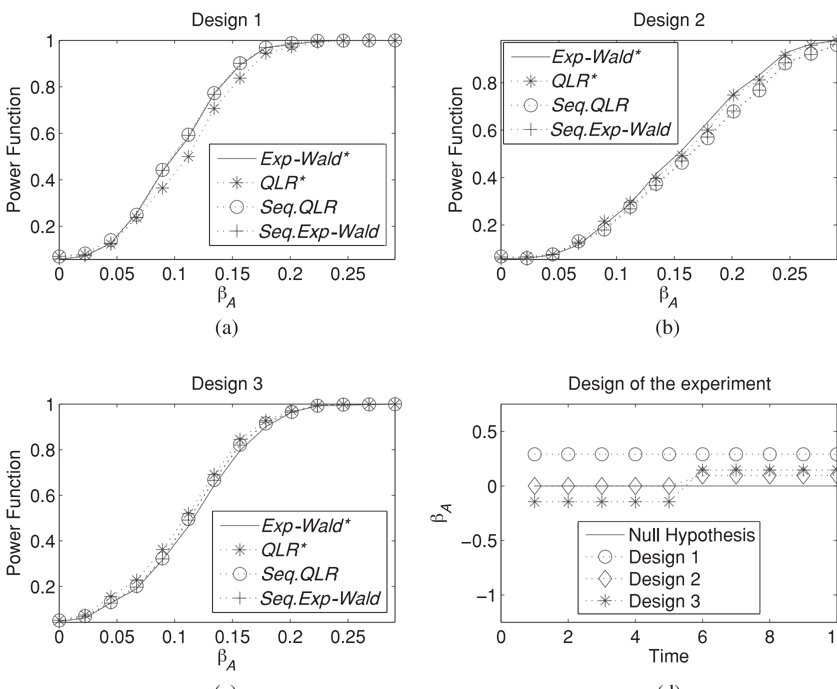

FIGURE 2. Comparison of asymptotic power functions of 5% selected optimal tests with the naive sequential test (across different designs).

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

are equally likely; in other words, if we want tests that are invariant to nonsingular linear transformations of the hypothesis, we cannot construct the test as formed by two independent components, as two-stage tests are not invariant to these transformations.

Finally, to investigate the properties of the Hall and Sen (1999) test for OIRs, we consider the following experiment.

Design 5. This design introduces instability in the OIRs. We introduce one OIR by using the following set of moment conditions: $f_t = (\epsilon_t, z_t^* \epsilon_t)'$ where z_t^* is an instrument such that $z_t^* = z_t + \epsilon_t(\beta_A \cdot 1(t \leq \frac{1}{2}T) - \beta_A \cdot 1(t > \frac{1}{2}T))$ and $z_t \sim N(0,1)$ is independent of ϵ_t. Note that when $\beta_A = 0$ then the OIR is valid and stable; on the other hand, it becomes unstable when $\beta_A \neq 0$. Figure 3 compares the power functions of the $Exp\text{-}Wald_T$, the $Exp\text{-}Wald_T^*$, and the $Exp\text{-}O_T$ tests for this design (see Figure 3d). To explore the properties of the $Exp\text{-}O_T$ test in the presence of parameter instability or constant shifts in the parameters, the figure also compares these tests in Designs 1–3 (where all the tests in Figure 3 now build on the two-dimensional moment condition). Figure 3 clearly shows that the $Exp\text{-}Wald_T$, $Exp\text{-}Wald_T^*$, and $Exp\text{-}O_T$ tests have power only against devia-

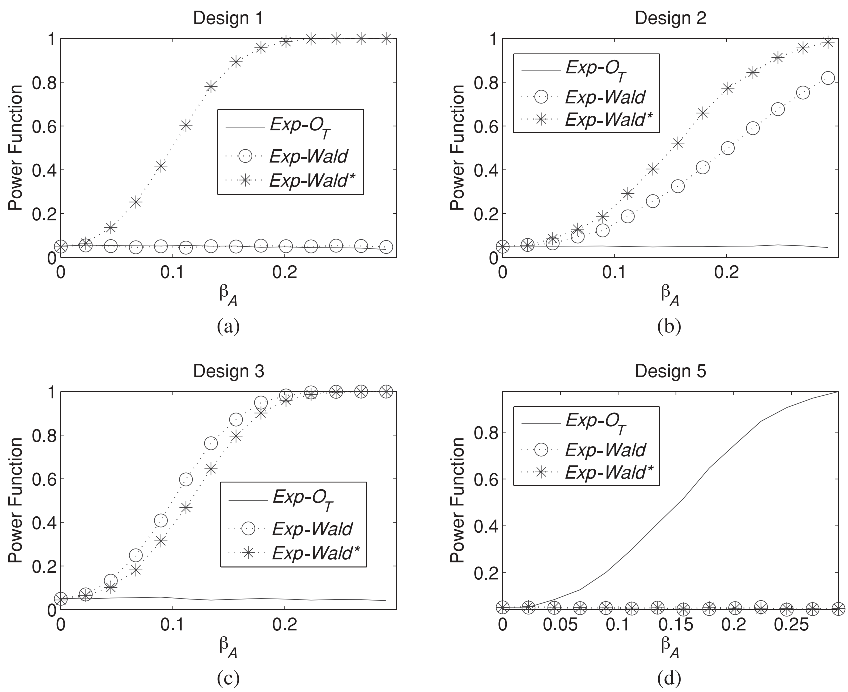

FIGURE 3. Comparison of asymptotic power functions of 5% selected optimal, TVP, and OIRs tests (across different designs).

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

tions from their specific null hypotheses. In particular, the $Exp\text{-}Wald_T$ and $Exp\text{-}Wald_T^*$ do not have power against instabilities in the OIRs, and the $Exp\text{-}O_T$ test does not have power against parameter instability or against the joint hypotheses considered in this paper, $(H_{AT}^{(1)})$ and $(H_{AT}^{(2)})$.

5. SUMMARY COMMENTS

This paper shows that there exists a class of locally most powerful tests for testing the joint hypothesis of model selection between two nested models and parameter stability. This paper introduces this class of tests, states the assumptions under which they are valid, and works out their asymptotic distributions. It also derives some special cases that apply for specific forms of parameter instability. These tests are easy to calculate, and this paper reports their (asymptotic) critical values. Joint tests such as the ones developed in this paper could also be adapted to the case of multiple breaks, along the lines of Bai and Perron (1998) and Elliott and Müller (2003). We leave this issue for future research.

NOTES

1. In fact, the standard Chow test can be rewritten as a Wald test: $(\tau(T-\tau)/T)((\hat{\beta}_1-\hat{\beta}_2)^2/\sigma_\epsilon^2)+op(1)$ and $\hat{\beta}$ and $(\hat{\beta}_1-\hat{\beta}_2)$ are independent. To see why, note that $\text{cov}(\hat{\beta}_1-\hat{\beta}_2,\hat{\beta}) = \text{cov}(\hat{\beta}_1-\hat{\beta}_2,\pi\hat{\beta}_1+(1-\pi)\hat{\beta}_2) = \pi\text{var}(\hat{\beta}_1)-(1-\pi)\text{var}(\hat{\beta}_2) = \pi(1/\pi)\sigma_\epsilon^2-(1-\pi)(1/(1-\pi))\sigma_\epsilon^2 = 0$.

2. The assumptions used in this paper are stronger than necessary, and the results are expected to hold if the assumptions are relaxed as in Andrews (1993). I thank G. Elliott and U. Muller for pointing this out.

3. For example, when the process is univariate and such that $y_t=\beta_0+\epsilon_t$, $\epsilon_t\sim$ i.i.d. $N(0,1)$, $\beta_0=0$ (as in the introductory example at the beginning of the paper) then the partial sum of moments is

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}(y_t-\hat{\beta}) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}y_t-\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}\frac{1}{T}\sum_{t=1}^{T}y_t = \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}y_t-s\frac{1}{\sqrt{T}}\sum_{t=1}^{T}y_t,$$

and the origin of the Brownian bridge is evident. If there were no restrictions under the null hypothesis, then the asymptotic distribution of $CW_T^{1/2}\sqrt{T}F_{sT}(\beta_0,\tilde{\delta})$ would be $(BB_q(s)'\,B_{m-q}(s)')'$, which is the Sowell (1996) result. When there are restrictions on a subset of p parameters under the null hypothesis, these will show up as p-Brownian motions, in addition to the previous components. These are Brownian motions because they are the limiting distribution of a partial sum of mean zero moment conditions, where the zero mean is obtained by *imposing*, rather than *estimating*, the drift. In the previous example, in this case the partial sum of moments is

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}(y_t-\beta_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}y_t,$$

and the origin of the Brownian motion is clear. The $B_{m-q}(s)$ component corresponds to the overidentified moment restrictions.

4. The result follows because $\overline{M}D^{-1/2}HD^{1/2}=\overline{M}D^{-1/2}HD^{-1/2}\overline{M}'\overline{M}=\overline{H}\overline{M}$, $\overline{H}\overline{M}\overline{D}'=0$ and $(I-\overline{H})\overline{M}\overline{D}'=\overline{M}\overline{D}'$, which can be verified by direct calculations.

5. I consider only two-sided alternatives here; one may generalize the argument to one-sided alternatives.

6. For completeness, let us mention that the LM statistic can also be obtained as $T\cdot\nabla_\beta\bar{Q}(\beta_0,\beta_0,\tilde{\delta})'\hat{\Omega}^{-1}\nabla_\beta\bar{Q}(\beta_0,\beta_0,\tilde{\delta})$ where $\hat{\Omega}$ is a consistent estimator of $E(T\cdot\nabla_\beta\bar{Q}(\tilde{\theta})\nabla_\beta\bar{Q}(\tilde{\theta})') \in R^{2p\times 2p}$. However, the LM formula provided in the main text is easier to calculate.

7. The notation is the same as in Nyblom and Mäkeläinen (1983); $\bar{\bar{\Gamma}}$ is a known matrix and σ_e^2 is a scalar. See also King (1980), King and Hillier (1985), and Stock and Watson (1998).

8. Trimming values are required. See Andrews (1993).

9. A similar analysis applies to the optimal mean Wald, QLR, and Nyblom tests.

10. See also Appendix A for more details. Note that $-\overline{M}_\beta\int_0^\pi g_\beta(\gamma,\pi,r)\,dr + \pi\overline{M}_\beta\int_0^1 g_\beta(\gamma,\pi,r)\,dr = -(1-\pi)\overline{M}_\beta\int_0^\pi g_\beta(\gamma,\pi,r)\,dr + \pi\overline{M}_\beta\int_\pi^1 g_\beta(\gamma,\pi,r)\,dr.$

11. Note that different Monte Carlo experiments could be designed in which all models are possibly (dynamically) misspecified, as in Corradi and Swanson (2001). This setup would not be the one for which the optimal tests proposed in this paper are designed, so it is not investigated.

12. LR^* is the likelihood ratio test for testing $\beta_2 = \beta_0$ conditional on knowing that $\beta_1 = \beta_0$ (see the example at the beginning of Section 2).

REFERENCES

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Andrews, D.W.K., I. Lee, & W. Ploberger (1996) Optimal change-point tests for normal linear regression. *Journal of Econometrics* 70, 9–38.

Andrews, D.W.K. & W. Ploberger (1993) Admissibility of the likelihood ratio test when a nuisance parameter is present only under the alternative. *Annals of Statistics* 23, 1609–1629..

Andrews, D.W.K. & W. Ploberger (1994) Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62, 1383–1414.

Bai, J. & P. Perron (1998) Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.

Chow, G.C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28, 591–605.

Clark, T.E. & M.W. McCracken (2005) The power of tests of predictive ability in the presence of structural breaks. *Journal of Econometrics* 124, 1–31.

Corradi, V. & N. Swanson (2005) Bootstrap conditional distribution tests in the presence of dynamic misspecification. *Journal of Econometrics*, forthcoming.

Elliott, G. & U.K. Müller (2003) Optimally testing general breaking processes in linear time series models. University of California at San Diego Discussion Paper 2003-07.

Engle, R. (1984) Wald, likelihood ratio, and Lagrange multiplier tests in econometrics. In Z. Griliches & M. Intriligator (eds.), *Handbook of Econometrics*, vol. 2, pp. 775–826. North-Holland.

Ghysels, E., A. Guay, & A. Hall (1998) Predictive tests for structural change with unknown breakpoint. *Journal of Econometrics* 82, 209–233.

Ghysels, E. & A.R. Hall (1990) A test for structural stability of Euler conditions parameters estimated via the generalized method of moments estimator. *International Economic Review* 31, 355–364.

Hall, A.R. & A. Sen (1999) Structural stability testing in models estimated by generalized method of moments. *Journal of Business & Economic Statistics* 17, 335–348.

King, M. (1980) Robust tests for spherical symmetry and their application to least square regression. *Annals of Statistics* 80, 1265–1271.

King, M. & G. Hillier (1985) Locally best invariant tests of the error covariance matrix of the linear regression model. *Journal of the Royal Statistical Society, series B* 47, 98–102.

Nabeya, S. & K. Tanaka (1988) Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. *Annals of Statistics* 16, 218–235.

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

Newey, W. & D. McFadden (1994) Large sample estimation and hypothesis testing. In R. Engle & D. McFadden (eds.), *Handbook of Econometrics*, vol. 4, pp. 2111–2245. North-Holland.
Newey, W. & K. West (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
Nyblom, J. (1989) Testing for the constancy of parameters over time. *Journal of the American Statistical Association* 84, 223–230.
Nyblom, J. & T. Mäkeläinen (1983) Comparisons of tests for the presence of random walk coefficients in a simple linear model. *Journal of the American Statistical Association* 78, 856–864.
Ploberger, W. & W. Krämer (1990) The local power of the cusum and cusum of squares tests. *Econometric Theory* 6, 335–347.
Ploberger, W. & W. Krämer (1992) The CUSUM test with OLS residuals. *Econometrica* 60, 271–286.
Quandt, R.E. (1960) Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55, 324–330.
Rossi, B. (2005) Are Exchange Rates Really Random Walks? Some Evidence Robust to Parameter Instability. *Macroeconomic Dynamics*, forthcoming.
Sowell, F. (1996) Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* 64, 1085–1107.
Stock, J. & M. Watson (1998) Median unbiased estimation of coefficient variance in a time-varying parameter model. *Journal of the American Statistical Association* 93, 349–358.

APPENDIX A: Proofs

Proof of Result 1. To simplify notation, let $f_t(x_t, \tilde{\theta})$ be denoted as $f_t(\tilde{\theta})$ and $\theta_{t,T}$ be denoted by θ_t. The restricted estimator $\tilde{\theta}$ satisfies the following first-order conditions for minimizing the Lagrangian $Q(\theta) + a(\theta)'\lambda$, where λ is the $(r \times 1)$ vector of LMs:

$$0 = \nabla_\theta Q(\tilde{\theta}) + \nabla_\theta a(\tilde{\theta})'\tilde{\lambda}, \quad \textbf{(A.1)}$$

$$0 = a(\tilde{\theta}).$$

Take a mean value expansion of $f_t(\tilde{\theta})$ around θ^*:

$$f_t(\tilde{\theta}) = f_t(\theta^*) + \nabla_\theta f_t(\bar{\theta}) \cdot (\tilde{\theta} - \theta^*), \quad \textbf{(A.2)}$$

where $\bar{\theta}$ is a intermediate point (in euclidean distance) between $\tilde{\theta}$ and θ^*, and by consistency of $\tilde{\theta}$, $\bar{\theta} \underset{p}{\to} \theta^*$. Summing (A.2) from $t = 1$ to $[sT]$ gives $F_{sT}(\tilde{\theta}) = F_{sT}(\theta^*) + \nabla_\theta F_{sT}(\bar{\theta}) \cdot (\tilde{\theta} - \theta^*)$, which, evaluated at $s = 1$ and premultiplied by $\nabla_\theta F_T(\tilde{\theta})' W_T$, gives

$$\nabla_\theta Q(\tilde{\theta}) = \nabla_\theta F_T(\tilde{\theta})' W_T F_T(\theta^*) + \nabla_\theta F_T(\tilde{\theta})' W_T \nabla_\theta F_T(\bar{\theta}) \cdot (\tilde{\theta} - \theta^*). \quad \textbf{(A.3)}$$

Another mean value expansion of $a(\tilde{\theta})$ around θ^* gives

$$a(\tilde{\theta}) = a(\theta^*) + A(\bar{\theta})(\tilde{\theta} - \theta^*). \quad \textbf{(A.4)}$$

Thus, combining (A.1), (A.3), and (A.4) and $A(\bar{\theta}) \underset{p}{\to} A$:

$$\begin{pmatrix} -\nabla_\theta F_T(\tilde{\theta})' W_T F_T(\theta^*) \sqrt{T} \\ -a(\theta^*) \sqrt{T} \end{pmatrix} = \begin{pmatrix} \nabla_\theta F_T(\tilde{\theta})' W_T \nabla_\theta F_T(\bar{\theta}) & A' \\ A & 0 \end{pmatrix} \begin{pmatrix} (\tilde{\theta} - \theta^*) \sqrt{T} \\ \lambda \sqrt{T} \end{pmatrix} + o_p.$$

(A.5)

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

Define $D \equiv \bar{M}'\bar{M}$, $A \equiv A(\theta^*)$, $P \equiv D^{-1/2}A'(AD^{-1}A')^{-1}AD^{-1/2}$, and $H \equiv I - P$. Solving (A.5) for $(\tilde{\theta} - \theta^*)$ gives

$$\sqrt{T}(\tilde{\theta} - \theta^*) = -D^{-1/2}HD^{-1/2}\nabla_\theta F_T(\tilde{\theta})' W_T F_T(\theta^*)\sqrt{T} - D^{-1}A'(AD^{-1}A')^{-1}a(\theta^*)\sqrt{T} + o_p. \quad \textbf{(A.6)}$$

By substituting (A.6) in (A.2), summing from $t = 1$ to $[sT]$, and premultiplying by $\sqrt{T}W_T^{1/2}$, we have

$$\sqrt{T}W_T^{1/2}F_{sT}(\tilde{\theta}) = \sqrt{T}W_T^{1/2}F_{sT}(\theta^*) - \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1/2}HD^{-1/2}\nabla_\theta F_T(\tilde{\theta})'W_T F_T(\theta^*) - \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1}A'(AD^{-1}A')^{-1}a(\theta^*) + o_p. \quad \textbf{(A.7)}$$

Next, a mean value expansion of $F_{sT}(\theta^*)$ around θ_t implies

$$F_{sT}(\theta^*) = F_{sT}(\theta_t) + \frac{1}{T}\sum_{t=1}^{[sT]} \nabla_\theta f_t(\bar{\theta}_t)(\theta^* - \theta_t), \quad \textbf{(A.8)}$$

where $\bar{\theta}_t$ is an intermediate point between θ_t and θ^*. Substituting (A.8) in (A.7), we have

$$\begin{aligned}\sqrt{T}W_T^{1/2}F_{sT}(\tilde{\theta}) &= \sqrt{T}W_T^{1/2}F_{sT}(\theta_t) + \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta}_t)(\theta^* - \theta_t) \\ &- \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1/2}HD^{-1/2}\nabla_\theta F_T(\tilde{\theta})'W_T F_T(\theta_t) \\ &- \frac{1}{T}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1/2}HD^{-1/2}\nabla_\theta F_T(\tilde{\theta})'W_T \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\nabla_\theta f_t(\bar{\theta}_t)(\theta^* - \theta_t) \\ &- \frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1}A'(AD^{-1}A')^{-1}a(\theta^*) + o_p. \quad \textbf{(A.9)}\end{aligned}$$

Letting $T \to \infty$, we have

$$\sqrt{T}W_T^{1/2}F_{sT}(\theta_t) \Rightarrow B_m(s),$$

$$\frac{1}{T}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta}_t)\sqrt{T}(\theta^* - \theta_t) \underset{p}{\to} -\bar{M}\int_0^s g(\gamma, \pi, r)\,dr, \quad \text{(included } s = 1\text{)},$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[sT]} W_T^{1/2}\nabla_\theta f_t(\bar{\theta})D^{-1/2}HD^{-1/2}\nabla_\theta F_T(\tilde{\theta})'W_T F_T(\theta_t) \Rightarrow s\bar{M}D^{-1/2}HD^{-1/2}\bar{M}'B_m(1),$$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} W_T^{1/2} \nabla_\theta f_t(\bar{\theta}) D^{-1/2} M D^{-1/2} \nabla_\theta F_T(\tilde{\theta})' W_T \frac{1}{T} \sum_{t=1}^{T} \nabla_\theta f_t(\bar{\theta}_t)(\theta^* - \theta_t)$$
$$\underset{p}{\to} -s\overline{M} D^{-1/2} H D^{-1/2} \overline{M}' \overline{M} \int_0^1 g(\gamma, \pi, r)\, dr,$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[sT]} W_T^{1/2} \nabla_\theta f_t(\bar{\theta}) D^{-1} A' (A D^{-1} A')^{-1} a(\theta^*) \underset{p}{\to} s\overline{M} D^{-1} A' (A D^{-1} A')^{-1} \bar{a}.$$

By substituting the preceding expressions in (A.9), we have

$$\sqrt{T} W_T^{1/2} F_{sT}(\tilde{\theta}) \Rightarrow B_m(s) - s\overline{H} B_m(1) - s\overline{M} D^{-1} A' (A D^{-1} A')^{-1} \bar{a}$$
$$- \overline{M} \int_0^s g(\gamma, \pi, r)\, dr + s\overline{M} D^{-1/2} H D^{1/2} \int_0^1 g(\gamma, \pi, r)\, dr, \qquad \textbf{(A.10)}$$

where $\overline{H} \equiv \overline{M} D^{-1/2} H D^{-1/2} \overline{M}'$, which proves Result 1. ■

Proof of Result 2. To prove Result 2, note that under the null hypothesis $\bar{a} = 0$ and $g(.) = 0$ so that only the first two components on the right-hand side of (A.10) are relevant. Note also that $\overline{H}$ is a projection matrix with rank $(k - r)$ so that $\overline{H} = C' \Lambda C$, where $\Lambda = \begin{pmatrix} I_{k-r} & 0 \\ 0 & 0 \end{pmatrix}$ and C is an orthonormal matrix such that $CC' = I_m$. Thus $C[B_m(s) - s\overline{H} B_m(1)] = CB_m(s) - s\Lambda C B_m(1)$, which has the same distribution as $B_m(s) - s\Lambda B_m(1) = (BB_{k-r}(s)', B_{m-(k-r)}(s)')'$, because C is orthonormal. Hence, Result 2 follows. ■

Proof of Corollary 1. Let $D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$ and $D^{-1} = \begin{pmatrix} D_{11}^{-} & D_{12}^{-} \\ D_{21}^{-} & D_{22}^{-} \end{pmatrix}$. Also, let the restrictions be linear restrictions on subsets of the parameters, so that $A = [I_{p \times p} \vdots 0_{p \times q}]$. Let $\overline{M}_\theta = [\overline{M}_\beta \vdots \overline{M}_\delta]$, $\bar{P}_\delta \equiv \overline{M}_\delta (\overline{M}_\delta' \overline{M}_\delta) \overline{M}_\delta'$. Corollary 1 follows from (A.10) by using the following results (a)–(e). (Results (a)–(d) follow from direct calculation. Details are provided in an Appendix available upon request.)

(a) $\overline{H} = \bar{P}_\delta$
(b) $\overline{M} D^{-1/2} H D^{1/2} = \bar{P}_\delta \overline{M}$
(c) $\overline{M} D^{-1} A' (A D^{-1} A')^{-1} = (I - \bar{P}_\delta) \overline{M}_\beta$
(d) $\tilde{\beta} = \beta_0$
(e) $g(.) = [g_\beta(.)\ 0_{q \times p}]$ ■

Proof of (13) and (14). Let Assumption 2 hold and let the class of alternatives be linear in the parameters: $g(\gamma, \pi, s) = \widetilde{G}(\pi, s)\gamma$. Thus $v(s)$, defined following (9), becomes

$$v(s) = \left(-\overline{M}\overline{D}' \quad -\overline{M}\widetilde{G}(\pi, r) + \overline{M} D^{-1/2} H D^{1/2} \left(\int_0^1 \widetilde{G}(\pi, r)'\, dr \right) \right) \begin{pmatrix} \bar{a} \\ \gamma \end{pmatrix}.$$

Let $\eta \equiv (\bar{a}, \gamma')'$ and define $a(s)$ to be such that $v(s)' = \eta' a(s)$, that is:

$$a(s) = \begin{pmatrix} -\overline{D}\overline{M}' \\ -\widetilde{G}(\pi, r)' \overline{M}' + \left(\int_0^1 \widetilde{G}(\pi, r)'\, dr \right)' D^{1/2} H D^{-1/2} \overline{M}' \end{pmatrix}.$$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

The term $A(\pi)$ is defined as $\int_0^1 a(s)\,dZ(s)$, and $V(\pi)$ is defined as $\int_0^1 a(s)a(s)'\,ds$. When there is only one break, and $\tilde{G}(\pi,s) = 1(s \geq \pi)G$, then direct calculations show that

$$A(\pi) = \begin{pmatrix} -\bar{D}\overline{M}'Z(1) \\ -G'\overline{M}'[Z(1) - Z(\pi)] + (1-\pi)G'D^{1/2}HD^{-1/2}\overline{M}'Z(1) \end{pmatrix}$$

$$= \begin{pmatrix} -\bar{D}\overline{M}' & 0 \\ -(1-\pi)G'A'\bar{D}\overline{M}' & G'\overline{M}' \end{pmatrix} \begin{pmatrix} Z(1) \\ Z(\pi) - \pi Z(1) \end{pmatrix}, \tag{A.11}$$

$$V(\pi) = \begin{pmatrix} \bar{D}D\bar{D}' & (1-\pi)\bar{D}D^{1/2}(I-H)D^{1/2}G \\ (1-\pi)G'D^{1/2}(I-H)D^{1/2}\bar{D}' & (1-\pi)G'D^{1/2}[I-(1-\pi)H]D^{1/2}G \end{pmatrix}. \tag{A.12}$$

■

Proof of Propositions 1–3. When the weighting function is an $(r+p)$-dimensional multivariate normal distribution with zero mean and covariance $U(\pi)$ then in this case, and for two-sided alternatives, the optimal tests in (11) simplify to (by completing the square and integrating out the parameter vector)

$$TS = \int \left(\frac{|U(\pi)^{-1}|^{1/2}}{|V(\pi) + U(\pi)^{-1}|^{1/2}} \exp\left\{ \frac{1}{2} A(\pi)'(V(\pi) + U(\pi)^{-1})^{-1} A(\pi) \right\} \right) dJ(\pi). \tag{A.13}$$

When $U(\pi)^{-1} = (1/c)V(\pi)$ then (up to a constant factor that does not matter)

$$TS = \int \left(\frac{|U(\pi)^{-1}|^{1/2}}{|V(\pi) + U(\pi)^{-1}|^{1/2}} \exp\left\{ \frac{1}{2} \Phi^*(\pi) \right\} \right) dJ(\pi),$$

$$\Phi^*(\pi) = A(\pi)'V(\pi)^{-1}A(\pi). \tag{A.14}$$

By using (A.12) and standard formulas for the inverse of a partitioned matrix,

$$V(\pi)^{-1} = \begin{pmatrix} \frac{1}{\pi} AD^{-1}A' & -\frac{1}{\pi} AD^{-1}G \\ -\frac{1}{\pi} G'D^{-1}A' & \frac{1}{\pi(1-\pi)} G'D^{-1}G \end{pmatrix}. \tag{A.15}$$

By combining (A.15) with (A.11) and (A.14), one finds that

$$\Phi^*(\pi) = \left(Z(1)' \vdots \frac{\{Z(\pi) - \pi Z(1)\}'}{\sqrt{\pi(1-\pi)}} \right) \begin{pmatrix} C_1'(C_1C_1')^{-1}C_1 & 0 \\ 0 & C_2'(C_2C_2')^{-1}C_2 \end{pmatrix}$$
$$\times \begin{pmatrix} Z(1) \\ \frac{Z(\pi) - \pi Z(1)}{\sqrt{\pi(1-\pi)}} \end{pmatrix}, \tag{A.16}$$

where $C_1 \equiv (AD^{-1}A')^{-1}AD^{-1}\overline{M}'$ has dimension $(r \times m)$ and $C_2 \equiv G'\overline{M}'$ has dimension $(p \times m)$. Notice that $C_1(I - \bar{H}) = C_1$ so that $C_1Z(1) = C_1B_m(1)$. Thus, $Z_r(1) \equiv$

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

$(C_1C_1')^{-1/2}C_1Z(1)$ is an r-vector of independent standard normals and $\{Z_p(\pi) - \pi Z_p(1)\} \equiv (C_2C_2')^{-1/2}C_2\{Z(\pi) - \pi Z(1)\}$ is a p-vector of independent Brownian bridges because $\{(C_1C_1')^{-1/2}C_1\}\{(C_1C_1')^{-1/2}C_1\}' = I_p$ (same for C_2). Hence:

$$A(\pi)'V(\pi)^{-1}A(\pi) = Z_r(1)'Z_r(1) + \frac{\{Z_p(\pi) - \pi Z_p(1)\}'\{Z_p(\pi) - \pi Z_p(1)\}}{\pi(1-\pi)}. \quad \textbf{(A.17)}$$

Thus, under the null hypothesis:

$$\Phi^*(\pi) = B_r(1)'B_r(1) + \frac{BB_p(\pi)'BB_p(\pi)}{\pi(1-\pi)}. \quad \textbf{(A.18)}$$

Proposition 1 thus follows from Result 1 and the continuous mapping theorem, and Propositions 2 and 3 follow directly from Proposition 1, Corollary 1, and the results in Andrews and Ploberger (1994). ■

Asymptotic Local Power. Under the alternative hypothesis, and using (10):

$$Z(1) = (I - \bar{H})\left\{B_m(1) - \overline{M}\bar{D}'\bar{a} - \overline{M}\int_0^1 g(s)\,ds\right\},$$

$$Z(\pi) - \pi Z(1) = BB_m(\pi) - (1-\pi)\overline{M}\int_0^\pi g(s)\,ds + \pi\overline{M}\int_\pi^1 g(s)\,ds,$$

and substituting these into (A.16):

$$\Phi^*(\pi) = Z_r^{(1)\prime}Z_r^{(1)} + Z_p^{(2)}(\pi)'Z_p^{(2)}(\pi), \quad \textbf{(A.19)}$$

where

$$Z_r^{(1)} \equiv B_r(1) - (C_1C_1')^{-1/2}C_1\overline{M}\left(\int_0^1 g(s)\,ds + \bar{D}'\bar{a}\right), \quad \textbf{(A.20)}$$

$$Z_p^{(2)}(\pi) \equiv \frac{\{B_p(\pi) - \pi B_p(1)\}}{\sqrt{\pi(1-\pi)}} - (C_2C_2')^{-1/2}C_2\overline{M}\left(\left(\frac{1-\pi}{\pi}\right)^{1/2}\int_0^\pi g(s)\,ds - \left(\frac{\pi}{1-\pi}\right)^{1/2}\int_\pi^1 g(s)\,ds\right). \quad \textbf{(A.21)}$$

Note that when $A = [I_p \quad 0_{p\times q}] = G$, $g(.) = [I_p \quad 0_{q\times p}]g_\beta(.)$, then $r = p$, $C_1 \equiv \overline{M}_\beta'(I - \bar{P}_\delta)$ (see (c) in the proof of Corollary 1) and $C_2 \equiv \overline{M}_\beta'$; note also that $\overline{M}_\beta'(I - \bar{P}_\delta)\overline{M}_\beta = \nabla_{\beta\beta}Q - \nabla_{\beta\delta}Q(\nabla_{\delta\delta}Q)^{-1}\nabla_{\delta\beta}Q$. In addition, note that when π is fixed and $\bar{a} = 0$, which is the case examined by Chow (1960) for *testing the existence of structural breaks only*, only $[I \vdots 0]A(\pi)$ and $[I \vdots 0]V(\pi)[I' \vdots 0']'$ are relevant so that the test becomes $\Phi(\pi) \Rightarrow BB_p(\pi)'BB_p(\pi)/\pi(1-\pi)$, which is the result of Andrews (1993) (see also Sowell, 1996). Notice also that when $\pi = 1$, which is *the case without structural break*, the result is the classical test statistic for tests on a subset of p parameters: $B_p(1)'B_p(1) \sim \chi^2_{(p)}$ because $BB_p(1) = 0$ and $B_p(1)$ is a p-dimensional multivariate standard normal distribution.

The proof that (20), (21), and (22) are asymptotically equivalent under both the null hypothesis and the local alternatives follows from applying results similar to those in Andrews (1993) and Newey and McFadden (1994). ■

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

Proof of Proposition 4. The (modified) Nyblom test statistic for testing both parameter instability and that the parameter vector is equal to some value β_0 was defined as

$$Nyblom_T^* = \int_0^1 (T \cdot \nabla_\beta Q_{1,[\pi T]}(\beta_0,\tilde{\delta})'\hat{\Omega}_N^{-1}\nabla_\beta Q_{1,[\pi T]}(\beta_0,\tilde{\delta}))J(\pi)\,d\pi, \tag{A.22}$$

where $\hat{\Omega}_N \in R^{2p\times 2p}$ is a consistent estimate of the asymptotic variance of $\nabla_\beta Q(\beta_0,\tilde{\delta})$ and the gradient function is defined as

$$\nabla_\beta Q_{1,[\pi T]}(\beta_0,\tilde{\delta}) \equiv \frac{1}{T}\sum_{t=1}^{[T\pi]} \nabla_\beta F_T(\beta_0,\tilde{\delta},\pi)'\Sigma^{-1}f_t(x_t,\beta_0,\tilde{\delta}). \tag{A.23}$$

Notice that $(1/T)\sum_{t=1}^{[T\pi]} f_t(x_t,\beta_0,\tilde{\delta})$ is the first component of $\bar{F}_T(\beta_0,\tilde{\delta},\pi)$, so that one would expect the asymptotics to be driven by $B(\pi)$. In fact, let $\tilde{\delta}$ be estimated on observations $1,2,\ldots,T$ and take a mean value expansion to obtain

$$\begin{aligned}
\sqrt{T}\nabla_\beta Q_{1,[\pi T]}(\beta_0,\tilde{\delta}) &= \nabla_\beta F_T(\theta_0)'\hat{\Sigma}^{-1}\frac{1}{T}\sum_{t=1}^{[T\pi]} f_t(\theta_0)\sqrt{T} \\
&\quad - \frac{1}{T}\sum_{t=1}^{[T\pi]} \nabla_\beta F_T(\theta_0)'\hat{\Sigma}^{-1}\nabla_\delta f_t(\theta_0) \\
&\quad \times (\nabla_\delta F_T(\theta_0)'\hat{\Sigma}^{-1}\nabla_\delta F_{1,T\pi}(\theta_0))^{-1}\nabla_\delta F_T(\theta_0)'\hat{\Sigma}^{-1}F_{1,T\pi}(\theta_0)\sqrt{T} \\
&= \overline{M}_\beta'(I_m - \bar{P}_\delta)\Sigma^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\pi]} f_t(\theta_0) + op(1),
\end{aligned}$$

where $F_{1,[T\pi]}(\theta_0) \equiv (1/T)\sum_{t=1}^{[T\pi]} f_t(x_t,\theta_0)$ has the following asymptotic distribution:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\pi]} f_t(\theta_0) \Rightarrow \Sigma^{1/2}B_m(\pi).$$

Hence, (A.23) is such that

$$\nabla_\beta Q_{1,[\pi T]}(\beta_0,\tilde{\delta}) \Rightarrow \overline{M}_\beta'(I - \bar{P}_\delta)B_m(\pi),$$

$$\begin{aligned}
Nyblom_T^* &= \int_0^1 T\cdot(\nabla_\beta Q_{1,[T\pi]}(\beta_0,\tilde{\delta})'\hat{\Omega}_N^{-1}\nabla_\beta Q_{1,[T\pi]}(\beta_0,\tilde{\delta}))J(\pi)\,d\pi \\
&\Rightarrow \int_0^1 B_m(\pi)'(I_m - \bar{P}_\delta)\overline{M}_\beta(\overline{M}_\beta'(I_m - \bar{P}_\delta)\overline{M}_\beta)^{-1}\overline{M}_\beta'(I_m - \bar{P}_\delta) \\
&\quad \times B_m(\pi)J(\pi)\,d\pi \\
&= \int_0^1 B_p(\pi)'B_p(\pi)J(\pi)\,d\pi.
\end{aligned} \tag{A.24}$$

Notice that, like the (modified) Andrews and Ploberger case for $c \to 0$, this statistic is a special case of (A.13); in fact, the $Nyblom_T^*$ and the modified $Mean\text{-}Wald_T^*$ statistics simply use two different weighting matrices. Notice that in the structural break case only, the test statistic is constructed on the basis of the first component of $\bar{F}_T(\hat{\theta},\pi)$ and the estimation of β transforms the Brownian motion in (A.24) into a Brownian bridge, thus originating the Nyblom test statistic: $\int_0^1 B_p(\pi)'B_p(\pi)J(\pi)\,d\pi$. ■

https://doi.org/10.1017/S0266466605050486 Published online by Cambridge University Press

APPENDIX B: Asymptotic Critical Values for the Optimal Test Statistics

Tables B1–B4 report critical values for the optimal tests and the QLR_T^* test considered in Section 3. The significance levels considered in the tables are 10%, 5%, 2.5%, and 1%. The critical values are obtained by simulating the asymptotic distributions described in Section 3. The number of Monte Carlo replications is 5,000. (The trimming values considered are only 15% and 85% of the available sample, and the grid of points is quite sparse; basically each observation is a point in the grid.)

TABLE B1. Asymptotic critical values of the Exp-Wald* statistic

p	0.10	0.05	0.025	0.01	p	0.10	0.05	0.025	0.01
1	2.449	3.134	3.817	4.6727	16	22.852	24.626	26.219	27.891
2	4.204	5.015	5.8842	6.8129	17	24.198	26.094	27.681	29.642
3	5.656	6.738	7.7042	8.9198	18	25.589	27.414	29.158	31.221
4	7.095	8.191	9.3191	10.421	19	26.603	28.352	30.111	31.930
5	8.744	9.824	10.9535	12.194	20	28.108	30.075	31.882	33.854
6	10.026	11.203	12.4487	14.039	25	34.138	36.177	37.908	40.342
7	11.42	12.630	13.8575	15.173	30	39.955	42.167	44.374	46.574
8	12.87	14.225	15.3435	16.751	35	45.947	48.183	50.325	53.100
9	14.138	15.537	16.9444	18.628	40	51.820	54.476	56.542	59.338
10	15.426	16.761	18.3168	19.786	50	63.098	65.688	68.191	71.299
11	16.758	18.467	19.5883	21.547	60	74.534	77.369	80.178	83.654
12	17.915	19.582	21.1368	22.686	70	86.038	89.291	92.120	94.955
13	19.288	20.945	22.7773	24.986	80	97.348	100.647	103.451	107.566
14	20.691	22.285	23.8681	25.579	90	108.533	111.558	114.873	118.682
15	21.626	23.385	24.799	26.675	100	119.749	123.685	126.987	130.755

TABLE B2. Asymptotic critical values of the Mean-Wald* statistic

p	0.10	0.05	0.025	0.01	p	0.10	0.05	0.025	0.01
1	4.263	5.364	6.675	8.151	16	41.340	44.565	47.377	50.509
2	7.292	8.743	10.301	12.190	17	43.575	46.872	49.775	54.276
3	10.014	11.920	13.569	15.653	18	46.090	49.403	52.342	56.298
4	12.422	14.362	16.138	18.346	19	48.138	51.509	54.585	58.149
5	15.539	17.523	19.338	21.712	20	50.814	54.100	58.022	62.096
6	17.753	19.877	22.094	24.777	25	61.834	66.202	69.063	73.344
7	20.105	22.389	24.434	27.024	30	72.719	76.963	80.804	86.127
8	22.858	25.397	27.349	30.508	35	84.301	88.682	92.603	97.196
9	25.369	27.844	30.348	32.944	40	95.549	99.918	104.361	109.622
10	27.318	30.039	32.374	35.668	50	116.933	121.688	126.194	131.524
11	30.130	32.994	35.295	38.451	60	138.852	144.073	148.816	155.288
12	32.078	34.880	37.380	41.187	70	160.519	166.723	171.577	178.353
13	34.609	37.691	40.609	44.264	80	182.513	188.975	193.457	200.146
14	37.205	40.418	43.179	46.328	90	204.115	210.106	215.101	222.797
15	38.986	42.184	44.786	47.969	100	225.565	232.309	239.094	246.113

TABLE B3. Asymptotic critical values of the QLR* statistic

p	0.10	0.05	0.025	0.01	p	0.10	0.05	0.025	0.01
1	8.138	9.826	11.332	13.481	16	51.435	55.160	58.421	61.567
2	12.196	14.225	16.088	17.942	17	54.100	58.121	61.517	65.132
3	15.562	17.640	19.922	22.482	18	56.910	60.836	64.469	68.850
4	18.611	21.055	22.989	25.997	19	59.116	62.716	66.236	70.014
5	22.157	24.550	26.590	29.481	20	62.153	66.023	69.749	73.997
6	24.817	27.377	29.781	33.217	25	74.408	78.648	82.569	87.093
7	27.754	30.414	32.878	35.948	30	86.205	90.715	94.942	100.049
8	30.723	33.717	36.260	38.962	35	98.043	102.820	107.078	112.566
9	33.553	36.552	39.228	42.905	40	110.049	115.263	119.674	125.411
10	36.173	39.020	41.891	45.266	50	132.900	138.324	143.162	150.215
11	38.889	42.332	44.843	48.936	60	155.959	161.557	167.068	175.141
12	41.334	44.820	47.747	51.389	70	179.054	185.749	191.606	196.879
13	44.017	47.589	51.386	56.251	80	201.808	208.198	214.228	222.182
14	46.877	50.193	53.394	56.935	90	224.120	230.100	236.930	245.060
15	49.059	52.347	55.562	59.082	100	246.614	254.605	261.541	268.573

TABLE B4. Asymptotic critical values of the Nyblom* statistic

p	0.10	0.05	0.025	0.01	p	0.10	0.05	0.025	0.01
1	1.103	1.404	1.803	2.251	16	10.532	11.418	12.234	13.187
2	1.876	2.292	2.676	3.249	17	11.107	12.011	12.840	13.975
3	2.589	3.143	3.590	4.173	18	11.760	12.650	13.492	14.596
4	3.178	3.750	4.246	4.862	19	12.151	13.151	13.986	14.992
5	3.972	4.550	5.099	5.741	20	12.911	13.807	14.938	16.064
6	4.545	5.157	5.773	6.564	25	15.732	16.819	17.787	18.856
7	5.155	5.752	6.349	7.079	30	18.476	19.565	20.728	22.147
8	5.869	6.555	7.089	8.025	35	21.309	22.483	23.648	24.846
9	6.487	7.183	7.884	8.552	40	24.175	25.391	26.621	28.003
10	6.959	7.707	8.453	9.209	50	29.577	30.859	32.004	33.626
11	7.684	8.513	9.170	9.963	60	35.149	36.551	37.934	39.475
12	8.182	8.970	9.676	10.744	70	40.545	42.203	43.624	45.210
13	8.869	9.659	10.510	11.373	80	46.055	47.764	49.146	50.823
14	9.504	10.334	11.173	12.047	90	51.525	53.249	54.652	56.297
15	9.964	10.816	11.542	12.470	100	56.916	58.897	60.602	62.614