TESTING PARAMETER INSTABILITY IN THE LINEAR MODEL WITH NON-STATIONARY REGRESSORS

SIMONE GROSE*

AND

BRETT INDER

Department of Econometrics and Business Statistics
Monash University
Clayton, Victoria 3168
AUSTRALIA

ABSTRACT

This paper extends the distributional theory for the problem of testing for structural change in the linear model when the timing of the change is unknown, and proposes a simple method of obtaining approximate critical values for the mean-Wald test. The results apply for a very wide range of regressor types, including integrated and trending regressors, and regressors that exhibit their own structural break. The proposed modification to the mean-Wald statistic thus provides a simple means of performing the test in a wide class of models.

JEL classification: C12, C22, C52

KEYWORDS: Structural change, Parameter instability, Asymptotic distribution, Mean Wald test, Generalised Brownian motion.

* Corresponding author: Dr. S. D. Grose, Department of Econometrics and Business Statistics, Monash University, Clayton, Australia 3168. Email: Simone.Grose@BusEco.monash.edu.au, Tel: +61-3-9905-2964, Fax: +61-3-9905-5474.

1. Introduction.

The development of the distributional theory for the problem of testing for structural change when the timing of the change is unknown has progressed enormously in the past few years, since Andrews (1993) developed the asymptotic theory for GMM-based variants of Quandt's (1960) “SupF” test. Andrews’ work covers estimation and testing for parameter instability in the context of both linear and nonlinear parametric models; however, his results rely heavily on an assumption that, in a linear regression context, requires both that $\mathrm{plim}T^{-1}\sum_1^T x_t x_t'$ be finite positive definite, and that, for any subsample, $\mathrm{plim}(sT)^{-1}\sum_1^{[sT]} x_t x_t' \equiv \mathrm{plim}T^{-1}\sum_1^T x_t x_t' \;\; \forall\, s \in (0,1)$ [1,2]. Such an assumption is clearly problematic in many if not most econometric contexts; ruling out, in particular, regressors that are non-stationary in the second moment. The asymptotic critical values for the SupF test tabulated in Andrews (1993) are thus not generally applicable.

There has, of course, been work on less limited regressor types, though usually in the context of the linear model only, and usually treating a specific “type” of regressor. For instance, Andrews and Ploberger (1994) introduced the average-exponential form of the LM test (the Exp-F test) in a general ML-based context, and used a “norming” matrix to permit deterministic trends, while Hansen (1992) considers the distributional theory of the Sup-F, Ave-F, and Nyblom's (1989) “L_c” test for parameter instability in the linear regression model (LRM) with I(1) regressors which may or may not include deterministic trends. However, Andrews and Ploberger’s tabulated critical values for the Exp-F_∞ and Ave-F (Exp-F_0) tests apply only for the case in which the regressors are strictly stationary, while Hansen’s effectively require that the stochastic process generating the regressors be known.

---

[1] Andrews (1993), Assumption 1(g).

[2] x_t is, as usual, a *k*-vector of regressors, and [.] denotes the integer part.

Hansen (1996) also derived the limiting distribution of, and tabulated critical values for, the Sup, Ave, and Exp-F tests and Nyblom's L_c test for models involving regressors that themselves exhibit a structural break. He found, as might be expected, that the null limiting distributions once again depend on the stochastic process generating the regressors (specifically, on the timing of the regressor break), making the tabulation of even asymptotic critical values impractical. Hansen accordingly suggests the use of bootstrap p-values for this case.

Andrews, Lee and Ploberger (1996) demonstrated the exact finite sample optimality of the Exp-F test under the assumption of iid normal errors and fixed regressors; and showed that their result also applies to certain cointegration models. However, while their Exp-F test is exactly similar, its null distribution naturally still depends on the regressors. They suggest critical values be obtained on a case-by-case basis by simulation. We also note the work of Bai and Perron (1998), in which the theory regarding break point estimation is extended to multiple breakpoints and more general regressors, including trending regressors. Their consideration of testing was, however, still restricted to the stationary-regressors case.

The primary aim of this paper is to extend the asymptotic theory for Wald-type tests of parameter instability without having to specify in advance the "type" of regressor being considered in any particular case. In particular, we derive the limiting distributions, under both the null and a suitable local alternative, of the mean and sup-Wald tests under assumptions encompassing stationary, trending and integrated regressors. The extension is important for both theoretical and practical purposes: as we have already observed, existing tabulated critical values for the sup and mean-Wald tests are appropriate only for a very restricted class of models, making the tests effectively inaccessible to the practitioner.

The secondary aim is to devise a suitable approximation or modification to the test statistics so as to avoid the need to tabulate critical values. Such an approximation might be considered almost essential in our case since, as might be expected, the limiting distributions derived in Section 3 depend explicitly on the asymptotic "behaviour" of the regressors. To this end, Section 4 derives the asymptotic mean and variance of the mean-Wald test, and uses them to construct a moment-adjusted statistic with an approximately chi-squared limiting distribution. The "closeness" of the chi-squared distribution to the actual limiting distribution of the modified mean-Wald test is then investigated by simulation, and found to be very good, particularly in the upper tail from which critical values are derived. The approximation thus provides a simple means of performing the mean-Wald test in a wide class of models with a variety of regressors, using critical values from the standard χ^2 tables.

Finally, in this paper we focus on testing for a one-off break, occurring at an unknown point in time, in the context of the linear model. We further restrict our attention to "pure" structural change, in which all the structural parameters are subject to change. The obvious extensions (testing for other forms of parameter instability, including multiple breaks; partial structural change, in which only some subset of the structural parameters are considered liable to suffer from instability; and parameter instability in nonlinear models) are relatively straightforward, and will be reserved for future work.

2. Model and test statistic.

Consider the linear regression model

$$y_t = \mathbf{x}_t'\beta_t + u_t, \tag{2.1}$$

with $$E(u_t | \mathscr{F}_{t-1}) = 0, \;\; E(u_t^2 | \mathscr{F}_{t-1}) = \sigma^2;$$

where $\mathbf{x}_t$ is a k-vector of linearly independent regressors, and $\mathscr{F}_{t-1}$ is the sigma-field generated by past u_t and present and past $\mathbf{x}_t$. Obviously, under the null hypothesis of no structural change ($\mathrm{H}_0\!: \beta_t = \beta$) model (2.1) reduces to the standard $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$, where $\mathbf{y}$ and $\mathbf{u}$ are T-vectors, and $\mathbf{X}$ is a $T \times k$ matrix of rank equal to k.

Now suppose a region of interest, Π, with closure in (0, 1); ie, $\Pi = [\pi_1, \pi_2] \subset (0,1)$; and assume a possible one-off change in the structural parameters at some unknown time $\tau = [\pi T]$, $\pi \in \Pi$. Let this be written as

$$\beta_t = \begin{cases} \beta + \delta, & t < \tau \\ \beta \quad , & t \geq \tau ; \end{cases}$$

so that the null hypothesis is now simply $\mathrm{H}_0 \colon \delta = 0$.

Partition the $T \times k$ matrix $\mathbf{X}$ for each of the $n = \tau_2 - \tau_1 + 1$ potential breakpoints in $\left[\tau_1 = [\pi_1 T], \tau_2 = [\pi_2 T]\right]$ as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_\tau \\ \overline{\mathbf{X}}_\tau \end{bmatrix} \begin{matrix} \scriptstyle \tau \times k \\ \scriptstyle (T-\tau)\times k \end{matrix} ;$$

and further define

$$\mathbf{Z}_\tau = \begin{bmatrix} \mathbf{X}_\tau \\ \mathbf{0} \end{bmatrix} \begin{matrix} \scriptstyle \tau \times k \\ \scriptstyle (T-\tau)\times k \end{matrix} .$$

Then, under $\mathrm{H}_1 \colon \delta \neq 0$, (2.1) can be written, for any given τ, as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}_\tau \delta + \mathbf{u} ; \tag{2.2}$$

and the corresponding Wald-type statistic for the test of H_0 *vs* H_1 is easily shown, for given τ and σ^2, to be

$$W_\tau = \sigma^{-2}\mathbf{y}'\overline{\mathbf{P}}_{\mathbf{X}}\mathbf{Z}_\tau(\mathbf{Z}_\tau'\overline{\mathbf{P}}_{\mathbf{X}}\mathbf{Z}_\tau)^{-1}\mathbf{Z}_\tau'\overline{\mathbf{P}}_{\mathbf{X}}\mathbf{y} \;\equiv\; \mathbf{y}'\mathbf{P}_\tau\mathbf{P}_\tau'\mathbf{y}\big/\sigma^2 , \tag{2.3}$$

where $k < \tau < T$, $\overline{\mathbf{P}}_{\mathbf{X}}$ is the usual orthogonal projector with respect to $\mathbf{X}$, and

$$\mathbf{P}_\tau = \overline{\mathbf{P}}_{\mathbf{X}}\mathbf{Z}_\tau(\mathbf{Z}_\tau'\overline{\mathbf{P}}_{\mathbf{X}}\mathbf{Z}_\tau)^{-1/2'} .$$

Clearly, $\mathbf{P}_\tau\mathbf{P}_\tau'$ is idempotent, with rank equal to k, and $\mathbf{P}_\tau'\mathbf{X} = \mathbf{0}$. Hence $W_\tau = \mathbf{u}'\mathbf{P}_\tau\mathbf{P}_\tau'\mathbf{u}\big/\sigma^2$ under H_0.

3. The asymptotic behaviour of W_τ.

The asymptotic properties of W_τ for *given* τ chosen independently of the sample are of course well known. Our interest here is in the limiting distributions, under both H_0

and H_1, of the various functions of W_τ that allow the testing of parameter instability without the need to resort to a restrictive point alternative. We will focus, in particular, on the SupF, which selects the largest W_τ observed on a prespecified range of feasible breakpoints, and the AveF, which takes a simple average of all W_τ, also over a prespecified range of feasible breakpoints.

3.1 Limiting distribution under H_0

In the following, $\Longrightarrow$ will used to denote weak convergence with respect to the uniform metric, $\xrightarrow{p}$ to denote convergence in probability, and $\xrightarrow{d}$ to denote convergence in distribution.

In addition, let $\mathbf{M}_\tau$ denote the "cumulative design" matrix $\sum_{t=1}^{\tau}\mathbf{x}_t\mathbf{x}_t' \equiv \mathbf{X}_\tau'\mathbf{X}_\tau$, $\mathbf{M}_0 = \mathbf{0}$, and observe that $\mathbf{X}'\mathbf{Z}_\tau \equiv \mathbf{Z}_\tau'\mathbf{Z}_\tau = \mathbf{M}_\tau$. Also let $m_t(\varphi) \triangleq \mathbf{x}_t'(y_t - \mathbf{x}_t'\varphi) = \mathbf{x}_t' u_t(\varphi)$ summarise the assumed data generating process (DGP) at time t for some parameter vector φ. Finally, let W_τ be also denoted $W_T(\pi)$, where $\pi = \tau/T \in (0, 1)$.

We now define the stochastic process we call Generalised Brownian Motion, and make the following assumption.

DEFINITION. Let $G(\cdot)$ denote a zero-mean vector Gaussian process on [0, 1] with positive definite covariance kernel $E\left(G(r)G(s)'\right) = \mathscr{G}(s)$ for $r \geq s > 0$. Then $G(r) \sim \mathrm{N}\left(0, \mathscr{G}(r)\right)$, and $E\left(\left(G(r) - G(s)\right)G(s)'\right) = 0$; implying $G(r) - G(s) \sim \mathrm{N}\left(0, \mathscr{G}(r) - \mathscr{G}(s)\right)$. $G(\cdot)$ is then said to be a generalised Brownian motion (GBM) on [0, 1]. If it is also the case that $G(0) = 0$ and $\mathscr{G}(1) = \mathbf{I}_k$ then $G(\cdot)$ will be referred to as a standard GBM.

ASSUMPTION 3.1. For $\tau = [\pi T]$ and $\pi \in (0, 1]$,

(a) $E m_t(\beta) = 0$.

(b) $E(u_t | \mathscr{F}_{t-1}) = 0$ and $E(u_t^2 | \mathscr{F}_{t-1}) = \sigma^2$, where $\mathscr{F}_{t-1}$ is the sigma-field generated by past u_t and present and past $\mathbf{x}_t$.

(c) $\mathbf{R}_\tau \triangleq \mathbf{M}_T^{-1/2}\mathbf{M}_\tau^{1/2} \implies \mathcal{R}(\pi) = O(1)$ nonsingular.

(d) A central limit theorem (CLT) applies such that
$$\mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'\mathbf{u}/\sigma \equiv \left(\sigma^2 \sum_{t=1}^{\tau} \mathbf{x}_t\mathbf{x}_t'\right)^{-1/2} \sum_{t=1}^{\tau} \mathbf{x}_t u_t \xrightarrow{d} \mathrm{N}(0, \mathbf{I}_k).$$

REMARKS. Assumption 3.1(a) imposes weak exogeneity under H_0: β constant; and in effect establishes the null DGP; while Assumption 3.1(b) states that the model disturbances form a martingale difference sequence (mds) with constant conditional, and hence unconditional, variance.

Assumption 3.1(c) places quite weak restrictions on the nature of the regressors. It stipulates that every sub-sample cumulative design be the same order (in T) as that based on the entire sample; with the product $\mathbf{M}_T^{-1/2}\mathbf{M}_\tau^{1/2}$ converging to a nonstochastic matrix function of π as T tends to infinity. The assumption is accordingly satisfied by $\mathbf{X}$ stationary, step-stationary[3], and trending[4]. In particular, for the $\mathbf{X}$-stationary case we have $\mathrm{plim} T^{-1}\mathbf{M}_T = \mathrm{plim}(sT)^{-1}\mathbf{M}_{[sT]}$ finite positive definite ($s \in (0, 1]$), and thus $\mathcal{R}(\pi) = \sqrt{\pi}\, \mathbf{I}_k$. For notational convenience we will usually write $\mathcal{R}(\pi)$ as $\mathcal{R}_\pi$.

Assumption 3.1(c) does not, at present, permit integrated regressors, as in this case the appropriately scaled moment matrix converges (weakly) to a stochastic matrix. An $O_p(1)$ version of A3.1(c) would allow I(1) regressors, but all results must then be

---

[3] That is, $\boldsymbol{x}_t$ is stationary before and after an arbitrary structural change somewhere in the sample as per Hansen (1996).

[4] An example of a regressor which would violate A3.1(c) is $x_t = 1$ for $t < \gamma T$, $x_t = t$ for $t \geq \gamma T$, $\gamma \in (0,1)$. In this case $\mathbf{M}_\tau$ is only $O(T)$ for $\tau < \gamma T$, while $\mathbf{M}_T$ is $O(T^3)$. Hence $\mathcal{R}_{\pi<\gamma} = O_p(T^{-1})$.

made conditional on a particular realisation of the limit matrix $\mathcal{R}$. We consider this matter in more detail at the end of this Section.

The conditions under which Assumption 3.1(d) is satisfied obviously depend on the nature of the regressors. It is, for instance, trivially satisfied by **X** fixed, and $u_t \sim iid$. For **X** stochastic the obvious approach appeals to the martingale CLT, requiring, firstly, that $(m_t = \mathbf{x}_t u_t, \mathcal{F}_t)$ be a square integrable mds satisfying (i) the Lindeberg condition plus $\sum_{t=1}^{\tau} m_t m_t' \Big/ \sum_{t=1}^{\tau} E(m_t m_t') \xrightarrow{p} 1$; plus the further condition (ii) that $\sum_{t=1}^{\tau} E(m_t m_t') \Big/ \sigma^2 \sum_{t=1}^{\tau} E(\mathbf{x}_t \mathbf{x}_t') \longrightarrow 1$[5]. **X** weakly exogenous ($E(\mathbf{x}_t u_t) = 0$) with suitably bounded moments plus Assumption 3.1(b) would then suffice. Note that, since $\mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'\mathbf{u}$ is (almost) the partial sample OLSE of β, we are, in effect, looking for situations in which the asymptotic normality of OLS can be justified.

Finally, both Assumptions 3.1(c) and (d) require the exclusion of the endpoint $\pi = 0$. $\mathbf{R}_0 \equiv \mathbf{0}$ for any T obviously violates the positive definite part of A3.1(c), and A3.1(d) clearly requires $\tau = [\pi T]$ to be some non-zero fraction of a (growing) sample.

We now establish the following Lemma and Theorem.

LEMMA 3.1. *Under H_0, and Assumptions* 3.1*(a) – (d),*

$$\left(\sigma^2 \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'\right)^{-1/2} \sum_{t=1}^{\tau} \mathbf{x}_t u_t \equiv \mathbf{M}_T^{-1/2}\mathbf{Z}_\tau'\mathbf{u}/\sigma \implies G(\pi)$$

where $\pi = T^{-1}\tau \in [0,1]$, and $G(\pi)$ is a k-variate standard GBM on [0, 1] with variance-covariance matrix $\mathcal{G}(\pi) = \mathcal{R}(\pi)\mathcal{R}'(\pi)$.

---

[5] Alternatively, make the various expectations in (i) and (ii) conditional on $\mathcal{F}_{t-1}$; in which case (ii) requires convergence in probability.

PROOF OF LEMMA 3.1.

Let $g_\tau = \mathbf{M}_T^{-1/2}\mathbf{Z}_\tau'\mathbf{u}/\sigma \equiv \mathbf{R}_\tau\mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'\mathbf{u}/\sigma$, $\tau = [\pi T]$, $\pi \in [0,1]$; and rewrite Assumption 3.1(d) as $\mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'\mathbf{u}/\sigma \Longrightarrow \mathcal{Z}_k^{(\pi)}$, where $\mathcal{Z}_k \sim \mathrm{N}(0,\mathbf{I}_k)$ and the superscript indicates the underlying sample fraction. Then

$$g_\tau \Longrightarrow \mathcal{R}(\pi)\mathcal{Z}_k^{(\pi)} \equiv H(\pi)$$

by Assumption 3.1(c), (d), and the Continuous Mapping Theorem (CMT). Furthermore, for $\tau_j = [rT] > \tau_i = [sT]$,

$$\begin{aligned}
g_{\tau_j} - g_{\tau_i} &= \mathbf{M}_T^{-1/2}\sum_{t=\tau_i+1}^{\tau_j}\mathbf{x}_t u_t \equiv \mathbf{M}_T^{-1/2}\left(\mathbf{M}_{\tau_j} - \mathbf{M}_{\tau_i}\right)^{1/2}\left(\mathbf{M}_{\tau_j} - \mathbf{M}_{\tau_i}\right)^{-1/2}\sum_{t=\tau_i+1}^{\tau_j}\mathbf{x}_t u_t \\
&\equiv \left(\mathbf{R}_{\tau_j}\mathbf{R}_{\tau_j}' - \mathbf{R}_{\tau_i}\mathbf{R}_{\tau_i}'\right)^{1/2}\left(\sum_{t=\tau_i+1}^{\tau_j}\mathbf{x}_t\mathbf{x}_t'\right)^{-1/2}\sum_{t=\tau_i+1}^{\tau_j}\mathbf{x}_t u_t \\
&\Longrightarrow \mathcal{R}_{(r,s)}\mathcal{Z}_k^{(r,s)}
\end{aligned}$$

where $\mathcal{R}_{(r,s)} \equiv \left(\mathcal{R}_r\mathcal{R}_r' - \mathcal{R}_s\mathcal{R}_s'\right)^{1/2}$, and $\mathcal{Z}_k^{(r,s)}$ is independent of $\mathcal{Z}_k^{(s)}$ by A3.1(b). Invoking the "nonstochastic" part of A3.1(c) then immediately yields

(i) $\mathcal{R}(\pi)\mathcal{Z}_k \equiv \mathrm{N}\left(0,\mathcal{R}_\pi\mathcal{R}_\pi'\right)$, implying $g_\tau \xrightarrow{d} \mathrm{N}\left(0,\mathcal{R}_\pi\mathcal{R}_\pi'\right)$; and

(ii) $E_{r>s}\left\{\left(H(r) - H(s)\right)H'(s)\right\} = E\left\{\mathcal{R}_{(r,s)}\mathcal{Z}_k^{(r,s)}\mathcal{Z}_k^{(s)\prime}\mathcal{R}_s'\right\} = 0$

implying $E_{r>s}\left\{H(r)H'(s)\right\} \equiv V\left\{H(s)\right\}$.

Finally, $\mathcal{R}(0) = \mathbf{0}$, $\mathcal{R}(1) = \mathbf{I}_k$ implies $H(0) = 0$ and $E\left(H(1)H'(1)\right) = \mathcal{R}_1\mathcal{R}_1' = \mathbf{I}_k$ also. The limit process $H(\cdot)$ is thus a standard GBM on [0, 1].

COROLLARY 3.1. *Under H_0, and Assumptions* 3.1*(a) – (d)*

$$\begin{aligned}\mathbf{M}_\tau^{-1/2}(\mathbf{Z}_\tau' - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{X}')\mathbf{u} &= \mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'\mathbf{u} - \mathbf{M}_\tau^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{X}'\mathbf{u}\\ &= \mathbf{R}_\tau^{-1}\mathbf{M}_T^{-1/2}\mathbf{Z}_\tau'\mathbf{u} - \mathbf{R}_\tau'\mathbf{M}_T^{-1/2}\mathbf{X}'\mathbf{u}\\ &\Longrightarrow \mathcal{R}_\pi^{-1}\big(G(\pi)-\mathcal{G}(\pi)G(1)\big).\end{aligned}$$

THEOREM 1. *Let $W_T(\cdot)$ be indexed on $\pi\in\Pi=[\pi_1,\pi_2]\subset(0,1)$. Then, under H_0 and Assumptions* 3.1*(a) – (d),*

$$W_T(\cdot)\Longrightarrow Q(\cdot),$$

$$\mathrm{SupF}\triangleq\sup_{\pi\in\Pi}W_T(\pi)\xrightarrow{d}\sup_{\pi\in\Pi}Q(\pi),$$

and $$\mathrm{AveF}=\overline{W}_T(\pi_1,\pi_2)\triangleq\frac{1}{n}\sum_{t=\tau_1}^{\tau_2}W_T(t/T)\xrightarrow{d}\frac{1}{\pi_2-\pi_1}\int_{\pi_1}^{\pi_2}Q(r)dr;$$

where $Q(\cdot)$ is defined on $\pi\in\Pi$ by

$$\begin{aligned}Q(\pi) &= \big(G(\pi)-\mathcal{G}(\pi)G(1)\big)'\mathcal{R}_\pi^{-1\prime}(\mathbf{I}_k-\mathcal{R}_\pi'\mathcal{R}_\pi)^{-1}\mathcal{R}_\pi^{-1}\big(G(\pi)-\mathcal{G}(\pi)G(1)\big)\\ &\equiv\big(G(\pi)-\mathcal{G}(\pi)G(1)\big)'\big(\mathcal{G}(\pi)-\mathcal{G}^2(\pi)\big)^{-1}\big(G(\pi)-\mathcal{G}(\pi)G(1)\big).\end{aligned}\qquad(3.1)$$

PROOF OF THEOREM 1. The first part of the Theorem follows instantly from the CMT, Corollary 3.1, and Assumption 3.1(c), on using $\overline{\mathbf{P}}_\mathbf{X}\mathbf{Z}_\tau=\mathbf{Z}_\tau-\mathbf{X}\mathbf{M}_T^{-1}\mathbf{M}_\tau$ and $\mathbf{M}_T^{-1/2}\mathbf{M}_\tau^{1/2}=\mathbf{R}_\tau$ to rewrite the numerator of $W_T(\pi)$ under H_0 as

$$\begin{aligned}\mathbf{u}'\mathbf{P}_\tau\mathbf{P}_\tau'\mathbf{u} &= \mathbf{u}'(\mathbf{Z}_\tau-\mathbf{X}\mathbf{M}_T^{-1}\mathbf{M}_\tau)(\mathbf{M}_\tau-\mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_\tau)^{-1}(\mathbf{Z}_\tau'-\mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{X}')\mathbf{u} &(3.2)\\ &= \mathbf{u}'(\mathbf{Z}_\tau\mathbf{M}_\tau^{-1/2\prime}-\mathbf{X}\mathbf{M}_T^{-1}\mathbf{M}_\tau^{1/2})(\mathbf{I}_k-\mathbf{M}_\tau^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{M}_\tau^{1/2})^{-1}(\mathbf{M}_\tau^{-1/2}\mathbf{Z}_\tau'-\mathbf{M}_\tau^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{X}')\mathbf{u}\\ &\equiv \mathbf{u}'(\mathbf{Z}_\tau\mathbf{M}_T^{-1/2\prime}\mathbf{R}_\tau^{-1\prime}-\mathbf{X}\mathbf{M}_T^{-1/2\prime}\mathbf{R}_\tau)(\mathbf{I}_k-\mathbf{R}_\tau'\mathbf{R}_\tau)^{-1}(\mathbf{R}_\tau^{-1}\mathbf{M}_T^{-1/2}\mathbf{Z}_\tau'-\mathbf{R}_\tau'\mathbf{M}_T^{-1/2}\mathbf{X}')\mathbf{u}.\end{aligned}$$

The limiting distributions of the Sup and AveF statistics then follow via the CMT.

REMARKS: 1. The Theorem establishes, subject to the stated assumptions, the limiting distributions of test statistics based on continuous functions of $W_T(\pi),\ \pi \in \Pi$; including, in particular, the SupF and AveF statistics; and so provides a basis for the simulation of critical values appropriate to the design being considered. These limiting distributions are functions only of the "trimming" parameters π_1 and π_2, selected by the practitioner; and of $\mathscr{G}(\pi) = \text{plim}\mathbf{M}_T^{-1/2}\mathbf{M}_\tau\mathbf{M}_T^{-1/2\prime}$. Thus $\mathscr{G}(\pi)$ can be consistently estimated for given π from the sample moments of $\mathbf{X}$, with the estimates then being used to simulate drawings from $G(\cdot)$. It is therefore possible in theory to obtain drawings from the asymptotic distributions of, and hence critical values for, test statistics based on $\{W_T(\pi); \pi \in \Pi\}$. The explicit exclusion of the endpoints 0 and 1 from the region Π is required primarily for convergence of the SupF statistic (see Andrews (1993)); but imposes no great restriction on, for instance, AveF, since neither W_τ nor $Q(\pi)$ are defined for $\pi = 0$ or $\pi = 1$.

2. Although we do not specifically consider the ExpF statistic of Andrews and Ploberger (1994), of which the AveF is a limiting case, its limiting distribution under Assumption 3.1 would similarly follow from the first part of Theorem 1 via the continuous mapping theorem.

3. The limit process $Q(\cdot)$ is a quadratic form in the "tied-down" vector process $G(\pi) - \mathscr{G}(\pi)G(1)$ which we will call the Generalized Brownian bridge (GBB). Alternatively, $Q(\cdot)$ can be written as $Q(\cdot) = q(\cdot)'q(\cdot)$, with $q(\cdot)$ defined on $\pi \in \Pi$ by

$$q(\pi) = (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}\big(G(\pi) - \mathscr{G}(\pi)G(1)\big).$$

$q(\cdot)$ is, of course, the limit process of $\mathbf{P}_\tau'\mathbf{u}/\sigma$, and can easily be shown (see the Corollary to LEMMA 4.1. below) to have a k-variate standard normal distribution for given π. Thus, as expected, $Q(\pi)$ is, for given $\pi \in (0,1)$, chi-square distributed with k degrees of freedom.

4. If the regressors are stationary in the sense of Andrews (1993, Assumption 1(g)), then $\mathscr{G}(\pi) = \pi\mathbf{I}_k$, $G(\cdot)$ reduces to a standard Brownian motion on [0, 1], and the limit

process of Theorem 1 above instantly simplifies to that described in Andrews (1993, Theorem 3). Theorem 1 is thus a generalization of Andrews (1993, Theorem 3) applied to pure structural change in the linear model.

5. Finally, consider replacement of Assumption 3.1(c) by Assumption 3.1(c′):

$$\text{(c}'\text{)} \qquad \mathbf{R}_\tau \triangleq \mathbf{M}_T^{-1/2}\mathbf{M}_\tau^{1/2} \implies \mathcal{R}(\pi) = O_p(1) \text{ nonsingular;}$$

so as to allow integrated regressors. Let "$\cdot\big|_{\mathcal{R}}$" denote the conditional distribution given a realization of the limit process $\mathcal{R}(\cdot)$; and observe that, while the preliminaries in the proof of Lemma 3.1 are unaltered under A3.1(a, b, c′ and d); ie, $g_{[sT]} \Longrightarrow \mathcal{R}(s)\mathcal{Z}_k^{(s)}$ etc; points (i) and (ii) no longer follow.

However, the asymptotic normality assumption A3.1(d) now requires that the regressors and disturbances be at least asymptotically uncorrelated[6]; a condition usually satisfied by assuming strict exogeneity ($E(\mathbf{x}_t u_s) = 0 \;\; \forall\, t, s$). It then follows that $\mathcal{R}(s)$ and $\mathcal{Z}_k^{(s)}$ are independent, immediately implying $\mathcal{R}(s)\mathcal{Z}_k^{(s)}\big|_{\mathcal{R}} \equiv \mathrm{N}\left(0, \mathcal{R}_s\mathcal{R}_s'\right)$ and $g_{[\pi T]}\big|_{\mathbf{X}} \xrightarrow{d} \mathrm{N}\left(0, \mathcal{R}_\pi\mathcal{R}_\pi'\right)$. The same argument applied to point (ii) of the proof of Lemma 3.1 then yields a conditional version of the Lemma and the subsequent Theorem.

In summary, if the explanatory variables include I(1) regressors then the results of Theorem 1 must be treated as conditional on a particular "realization" of the limit process $\mathcal{R}(\cdot)$. This has no effect on the marginal distribution of $Q(\pi)$ for given π, as this does not involve $\mathcal{R}(\cdot)$; however, the joint distribution of, say, $Q(\pi_i)$ and $Q(\pi_j)$, *does* involve $\mathcal{R}(\cdot)$; and thus so do the limiting distributions of the Sup and Ave-F statistics. Simulated critical values would therefore also be "conditional"; and although this does not necessarily make them unusable, it does mean that there would be no point in tabulating those obtained. Of course, it might be argued that this is pretty much the case anyway, even under A3.1(c): the limiting distributions still

[6] See Park and Phillips (1988, §5.1).

depend on $\mathcal{R}(\cdot)$; and hence, in general, on the "type" and number of regressors. We return to the problem of obtaining critical values in Section 4.

3.2 Local Power

Before turning to the practicalities of implementing tests based on $\{ W_T(\pi); \pi \in \Pi \}$, with their non-standard regressor-dependent limiting null distributions, consider the limiting behaviour of $W_T(\cdot)$ under the sequence of local alternatives described by the following assumption.

ASSUMPTION 3.2. As for Assumption 3.1, except A3.1(a) is replaced by

(a) $E\{m_t(\beta + \sigma \mathbf{M}_T^{-1/2\prime}\eta(t/T))\} = 0$, where $\eta(\cdot)$ is any step function or uniform limit of step functions defined on [0, 1].

REMARKS. Our alternative takes the form of a break or series of breaks in β of magnitude $\delta_{t,T} = \sigma \mathbf{M}_T^{-1/2\prime}\eta(t/T)$; that is, the rate at which the break magnitude diminishes with T is governed explicitly by the design. Obviously, for certain designs the factor $\mathbf{M}_T^{-1/2}$ could be replaced an "equivalent" function of T ($T^{-1/2}$ for stationary $\mathbf{X}$, $T^{-3/2}$ for a linear trend, *et cetera*), after the manner of Andrews and Ploberger's[7] normalization matrix.

Also notice that $\delta_{t,T} = \sigma \mathbf{M}_T^{-1/2\prime}\eta(t/T)$ includes everything between the "varying coefficient" alternative $y_t = \mathbf{x}_t\beta_t + u_t$, and a one-off step-break at, say, $\pi = \pi^*$. The latter would correspond to $\eta(\pi) = \xi\mathbf{1}(\pi \le \pi^*)$ where $\mathbf{1}(\cdot)$ is the indicator function and ξ is a non-zero constant; in which case the DGP under $\mathrm{H}_1\!: \delta \ne 0$ can be written in the style of (2.2) with $\mathbf{Z}_\tau$ replaced by the "correctly" partitioned $\mathbf{Z}_{\tau^*}$, $\tau^* = [\pi^* T]$.

---

[7] Andrews and Ploberger (1994).

Finally, we consider local power under the non-stochastic version of A3.1/3.2(c) ie; $\mathbf{R}_{[\pi T]} = \mathbf{M}_T^{-1/2}\mathbf{M}_{[\pi T]}^{1/2} \Longrightarrow \mathcal{R}(\pi)$ finite nonstochastic; $\pi \in [0,1]$; $\{\mathcal{R}(\pi): \pi \in (0,1]\}$ nonsingular. Hence $\mathbf{R}_{[\pi T]}\mathbf{R}'_{[\pi T]} = \mathbf{G}_{[\pi T]} \Longrightarrow \mathcal{G}(\pi)$, where $\mathcal{G}(\cdot) = \mathcal{R}(\cdot)\mathcal{R}'(\cdot)$ is also finite nonstochastic on [0, 1] and $\{\mathcal{G}(\pi): \pi \in (0,1]\}$ is positive definite. $\mathcal{G}(\pi)$ will similarly be denoted by the alternate notation $\mathcal{G}_\pi$ as convenient.

THEOREM 2. *Suppose Assumptions 3.2(a) – (d) hold. Also suppose that the limit matrix* $\{\mathcal{G}(\pi): \pi \in [0,1]\}$ *is a continuous and differentiable matrix function of* π. *Then*

$$W_T(\cdot) \Longrightarrow Q^*(\cdot) = q^*(\cdot)'q^*(\cdot)\,;$$

where $q^*(\cdot)$ is defined on $\pi \in \Pi$ by

$$q^*(\pi) = (\mathbf{I}_k - \mathcal{R}'_\pi\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}\big(G(\pi) - \mathcal{G}(\pi)G(1) + \mathcal{H}(\pi)\big)$$

and
$$\mathcal{H}(\pi) = (\mathbf{I}_k - \mathcal{G}_\pi)\int_0^\pi \big(d\,\mathcal{G}(s)\big)\eta(s) - \mathcal{G}_\pi\int_\pi^1 \big(d\,\mathcal{G}(s)\big)\eta(s)\,. \tag{3.3}$$

PROOF OF THEOREM 2. In this case we have $\delta_{t,T} = \sigma\mathbf{M}_T^{-1/2\,\prime}\eta(t/T)$, implying DGP $y_t = \mathbf{x}_t\beta_t + u_t$ with $\beta_t = \beta + \delta_{t,T}$. This can be rewritten, by letting $\mathbf{Z}_t$ denote the $T \times k$ matrix comprised of the first t rows of $\mathbf{X}$, zeros thereafter, and dropping, for notational convenience, the second subscript on δ, as

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\beta + \mathbf{x}_2\delta_2 + \cdots + \mathbf{x}_{T-1}\delta_{T-1} + \mathbf{x}_T\delta_T + \mathbf{u}\\ &\equiv \mathbf{X}\beta + (\mathbf{Z}_2 - \mathbf{Z}_1)\delta_2 + \cdots + (\mathbf{Z}_T - \mathbf{Z}_{T-1})\delta_T + \mathbf{u}\,.\end{aligned}$$

Defining $\Delta\mathbf{Z}_t = \mathbf{Z}_t - \mathbf{Z}_{t-1}$ then yields

$$\mathbf{P}'_\tau\mathbf{y}/\sigma \;=\; \sum_{t=2}^{T}\mathbf{P}'_\tau\,\Delta\mathbf{Z}_t\,\delta_t/\sigma \;+ \mathbf{P}'_\tau\mathbf{u}/\sigma\,;$$

where the limit process of $\mathbf{P}'_\tau\mathbf{u}/\sigma$, $\tau = [\pi T]$, follows instantly from the first part of Theorem 1 as

$$\mathbf{P}_\tau'\mathbf{u}/\sigma \Longrightarrow q(\pi) = (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}\big(G(\pi) - \mathscr{G}(\pi)G(1)\big)\,. \tag{3.4}$$

Now, let $\mathbf{M}_t = \mathbf{X}'\mathbf{Z}_t \equiv \mathbf{Z}_t'\mathbf{Z}_t$, and note that $\mathbf{Z}_\tau'\mathbf{Z}_t = \mathbf{M}_\tau$ for $\tau \le t$; while for $\tau \ge t$ $\mathbf{Z}_\tau'\mathbf{Z}_t = \mathbf{M}_t$. Then, for $0 < \tau < T$,

$$\begin{aligned}\mathbf{P}_\tau'\mathbf{Z}_t\delta/\sigma &= (\mathbf{M}_\tau - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_\tau)^{-1/2}(\mathbf{Z}_\tau'\mathbf{Z}_t - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_t)\,\delta/\sigma \\ &= \begin{cases}(\mathbf{M}_\tau - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_\tau)^{-1/2}(\mathbf{M}_\tau - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_t)\,\delta/\sigma & ,\ \tau \le t \\ (\mathbf{M}_\tau - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_\tau)^{-1/2}(\mathbf{M}_t - \mathbf{M}_\tau\mathbf{M}_T^{-1}\mathbf{M}_t)\,\delta/\sigma & ,\ \tau \ge t\,.\end{cases}\end{aligned}$$

A certain amount of rearrangement, and applying Assumption 3.2(a), then yields, for $\tau \le t$,

$$\begin{aligned}\mathbf{P}_\tau'\mathbf{Z}_t\delta_t/\sigma &= (\mathbf{I}_k - \mathbf{M}_\tau^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{M}_\tau^{1/2})^{-1/2}\mathbf{M}_\tau^{1/2\prime}\mathbf{M}_t^{-1/2\prime} \\ &\qquad\times(\mathbf{I}_k - \mathbf{M}_t^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{M}_t^{1/2})\mathbf{M}_t^{1/2\prime}\mathbf{M}_T^{-1/2\prime}\eta(t/T) \\ &\equiv (\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{-1/2}\mathbf{R}_\tau'\mathbf{R}_t^{-1\prime}(\mathbf{I}_k - \mathbf{R}_t'\mathbf{R}_t)\mathbf{R}_t'\eta(t/T)\,;\end{aligned}$$

while for $\tau \ge t$,

$$\mathbf{P}_\tau'\mathbf{Z}_t\delta_t/\sigma = (\mathbf{I}_k - \mathbf{M}_\tau^{1/2\prime}\mathbf{M}_T^{-1}\mathbf{M}_\tau^{1/2})^{1/2\prime}\mathbf{M}_\tau^{-1/2}\mathbf{M}_t^{1/2}\ \mathbf{M}_t^{1/2\prime}\mathbf{M}_T^{-1/2\prime}\eta(t/T)\,.$$

Hence, for $0 < \tau < T$,

$$\mathbf{P}_\tau'\mathbf{Z}_t\delta_t/\sigma = \begin{cases}(\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{-1/2}\mathbf{R}_\tau'(\mathbf{I}_k - \mathbf{R}_t\mathbf{R}_t')\eta(t/T)\,, & \tau < t \\ (\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{1/2\prime}\mathbf{R}_\tau^{-1}\mathbf{R}_t\mathbf{R}_t'\eta(t/T)\,, & \tau \ge t\end{cases} \tag{3.5}$$

implying, for $2 \le t \le T$,

$$\mathbf{P}_\tau'\Delta\mathbf{Z}_t\delta_t/\sigma = \begin{cases}-(\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{-1/2}\mathbf{R}_\tau'\Delta\mathbf{G}_t\eta(t/T)\,, & t > \tau \\ (\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{1/2\prime}\mathbf{R}_\tau^{-1}\Delta\mathbf{G}_t\eta(t/T)\,, & t \le \tau\end{cases}$$

where $\mathbf{R}_t\mathbf{R}_t' = \mathbf{G}_t$ and $\mathbf{R}_t\mathbf{R}_t' - \mathbf{R}_{t-1}\mathbf{R}_{t-1}' = \mathbf{M}_T^{-1/2}\mathbf{x}_t\mathbf{x}_t'\mathbf{M}_T^{-1/2\prime} = \Delta\mathbf{G}_t$. Thus

$$\sum_{t=2}^{T}\mathbf{P}_\tau'\Delta\mathbf{Z}_t\,\delta_t/\sigma = (\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{1/2'}\mathbf{R}_\tau^{-1}\sum_{t=2}^{\tau}\Delta\mathbf{G}_t\eta(t/T)$$
$$- (\mathbf{I}_k - \mathbf{R}_\tau'\mathbf{R}_\tau)^{-1/2}\mathbf{R}_\tau'\sum_{t=\tau+1}^{T}\Delta\mathbf{G}_t\eta(t/T)\,. \tag{3.6}$$

Now, let $\pi = \tau/T$ and $s = t/T$; and recall that $\mathbf{R}_{[sT]}\mathbf{R}_{[sT]}' = \mathbf{G}_{[sT]} \Longrightarrow \mathcal{G}(s)$ where $\mathcal{G}(\cdot) = \mathcal{R}(\cdot)\mathcal{R}'(\cdot)$ is continuous differentiable nonstochastic by assumption. Hence, for arbitrary h $(0 < h \le s)$,

$$\frac{\mathbf{G}_{[sT]} - \mathbf{G}_{[(s-h)T]}}{h} \Longrightarrow \frac{\mathcal{G}(s) - \mathcal{G}(s-h)}{h}\,.$$

Setting $h = 1/T$ on the LHS of this expression, and observing that $\Delta\mathbf{G}_t = \Delta\mathbf{G}_{[sT]} = \mathbf{G}_{[sT]} - \mathbf{G}_{[(s-\frac{1}{T})T]}$, then yields

$$T\Delta\mathbf{G}_{[sT]} \Longrightarrow \lim_{h\to 0}\frac{\mathcal{G}(s) - \mathcal{G}(s-h)}{h} \equiv \frac{d\mathcal{G}(s)}{ds}\,,$$

by definition. Hence, for $0 \le r_1 < r_2 \le 1$,

$$\sum_{t=[r_1T]+1}^{[r_2T]}\Delta\mathbf{G}_t\eta(t/T) = \frac{1}{T}\sum_{t=[r_1T]+1}^{[r_2T]}T\Delta\mathbf{G}_t\eta(t/T) \Longrightarrow \int_{r_1}^{r_2}\frac{d\mathcal{G}(s)}{ds}\eta(s)ds \equiv \int_{r_1}^{r_2}\big(d\mathcal{G}(s)\big)\eta(s)\,.$$

Applying this result, and Assumption 3.2(b), to (3.6) then yields

$$\sum_{t=2}^{T}\mathbf{P}_\tau'\Delta\mathbf{Z}_t\,\delta_t/\sigma \Longrightarrow (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{1/2'}\mathcal{R}_\pi^{-1}\int_0^\pi\big(d\mathcal{G}(s)\big)\eta(s)$$
$$- (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi'\int_\pi^1\big(d\mathcal{G}(s)\big)\eta(s)\,,$$

which in combination with (3.4) establishes

$$\mathbf{P}_\tau'\mathbf{y}/\sigma \Longrightarrow q^*(\pi) = (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}\big(G(\pi) - \mathcal{G}(\pi)G(1)\big)$$
$$+ (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{1/2'}\mathcal{R}_\pi^{-1}\int_0^\pi\big(d\mathcal{G}(s)\big)\eta(s)$$
$$- (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi'\int_\pi^1\big(d\mathcal{G}(s)\big)\eta(s)\,.$$

Rearranging the second and third terms and applying the CMT then establishes the Theorem.

COROLLARY. *Suppose Assumptions 3.2(a) – (d) hold, except now* $\eta(\pi) = \xi\mathbf{1}(\pi \le \pi^*)$.

Then the limiting behaviour of $W_T(\cdot)$ *is as described by Theorem* 2 *above, with* $\mathcal{H}(\cdot)$ *now given by*

$$\mathcal{H}(\pi) = \begin{cases} \mathcal{G}_\pi(\mathbf{I}_k - \mathcal{G}_{\pi^*})\xi \ , & \pi \le \pi^* \\ (\mathbf{I}_k - \mathcal{G}_\pi)\mathcal{G}_{\pi^*}\xi \ , & \pi > \pi^* \ . \end{cases} \tag{3.7}$$

PROOF. The corollary is readily obtained on substitution of $\eta(\pi) = \xi\mathbf{1}(\pi \le \pi^*)$ into (3.3). Alternatively, since the DGP is now just $\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}^*\delta + \mathbf{u}$, where $\mathbf{Z}^*$ is the $T \times k$ matrix comprised of the first $\tau^* = [\pi^* T]$ rows of **X**, zeros thereafter, it follows that $\mathbf{P}_\tau'\mathbf{y}/\sigma = \mathbf{P}_\tau'\mathbf{Z}^*\delta/\sigma + \mathbf{P}_\tau'\mathbf{u}/\sigma$, where the limit process of $\mathbf{P}_\tau'\mathbf{Z}^*\delta/\sigma$ proceeds directly from (3.5) on replacing s by π* and applying $\mathbf{R}_{[\pi T]} \Longrightarrow \mathcal{R}(\pi)$. Rearranging the resulting expression as

$$\mathbf{P}_\tau'\mathbf{Z}^*\delta/\sigma \Longrightarrow \begin{cases} (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}\mathcal{G}_\pi(\mathbf{I}_k - \mathcal{G}_{\pi^*})\xi \ , & \pi \le \pi^* \\ (\mathbf{I}_k - \mathcal{R}_\pi'\mathcal{R}_\pi)^{-1/2}\mathcal{R}_\pi^{-1}(\mathbf{I}_k - \mathcal{G}_\pi)\mathcal{G}_{\pi^*}\xi \ , & \pi > \pi^* \end{cases}$$

then establishes the Corollary.

REMARKS: 1. Obviously, when H_0 is true we have $\eta(\cdot) = 0$, and thus $Q^*(\cdot) \equiv Q(\cdot)$. Otherwise $\{ Q^*(\pi),\ \pi \in \Pi \subset (0,1) \}$ is well-defined, and distinct from the null distribution provided π is bounded away from zero or one, since $\mathcal{H}(\pi) = \mathbf{0}$ for π equal to zero or one, but is otherwise non-zero in general. The presence of the "noncentrality" factor $\mathcal{H}(\pi)$ in the limiting process $Q^*(\pi)$ thus establishes that tests based on $\{ W_T(\pi),\ \pi \in \Pi \}$, such as $\sup_{\pi\in\Pi} W_T(\pi)$ and $\overline{W_T}(\pi_1, \pi_2)$, have non-trivial power.

The exception, of course, is that of a single break occurring at the very beginning or end of the sample. This can be seen most easily from the Corollary: setting $\pi^* = 0$ or 1 in (3.7) obviously makes $\mathcal{H}(\pi)$ identically zero for all π.

2. Once again, if the regressors are stationary such that $\mathcal{G}(\pi) = \pi\mathbf{I}_k$ then the limiting process described here reduces to a variant of that described in Andrews (1993). In particular, Theorem 2 is a generalization of Andrews (1993, equation 5.4) applied to pure structural change in the linear model.

4. The Chi-square approximation

Reconsider the null limiting distributions of Theorem 1. In general these are non-standard, and depend not only on the dimension of the parameter vector β, and the upper and lower bounds π_1 and π_2 , but on the asymptotic "behaviour" of the regressors as well, through the limit matrix $\mathcal{R}_\pi$. The distributions can be simulated, for a given regressor "type", but tabulation of critical values is relatively pointless. In consequence, the chosen test must be performed with critical values obtained by simulation on a case-by-case basis; or via some form of bootstrap (see, for instance, Hansen, 1996); or by approximation.

It is, of course, well known that, subject to the usual regularity conditions, the statistic W_τ with breakpoint τ fixed independently of the sample converges in distribution under H_0 to a chi-square with k degrees of freedom. The same result, relying only on Assumptions 3.1, can be seen to proceed directly from the first part of Theorem 1. While this does not particularly assist with the problem of obtaining critical values for tests, such as the Sup or AveF, that are constructed as functions of W_τ over a range of possible breakpoints τ, it does suggest that a chi-square approximation may be worth investigating.

Accordingly, consider the AveF statistic, $\overline{W}_T$, and its null limiting distribution as established in Theorem 1. Denote the latter by $S \equiv S(\pi_1, \pi_2)$.

The "chi-square-adjusted" version of the AveF statistic proceeds by the simple expedient of matching the first and second moments of its limiting distribution under H_0 to those of the $\chi^2_{(k)}$. That is, we construct

$$\overline{W}_T^{\,c} = k + \sqrt{2k} Z_{\overline{W}}, \text{ where } Z_{\overline{W}} = \frac{\overline{W}_T - E(S)}{\sqrt{V(S)}};$$

implying
$$\overline{W}_T^{\,c} = k + \sqrt{\frac{2k}{V(S)}} (\overline{W}_T - E(S)). \tag{4.1}$$

Then, under H_0 and the assumptions of Section 3.1,

$$\overline{W}_T^{\,c}(\pi_1, \pi_2) \xrightarrow{d} S_c(\pi_1, \pi_2) \triangleq k + \sqrt{\frac{2k}{V(S)}} \big(S(\pi_1, \pi_2) - E(S)\big);$$

where, by construction, $E(S_c) = k$, and $V(S_c) = 2k$. The proposed test procedure is then to reject H_0 for observed values of $\overline{W}_T^{\,c}$ in excess of the chosen percentile of the $\chi^2_{(k)}$ distribution. This will be called the "AveF$_c$" test.

Naturally, the AveF$_c$ test will have actual asymptotic size close to nominal provided the S_c and $\chi^2_{(k)}$ distributions are sufficiently congruent. It thus remains to obtain the mean and variance of $S(\pi_1, \pi_2)$; and to evaluate the extent to which the $S_c(\pi_1, \pi_2)$ distribution can be said to "match" the $\chi^2_{(k)}$.

4.1 Mean and variance of the AveF limiting distribution under H_0

Let $\mathbb{G}(\pi)$ denote the Generalized Brownian bridge $G(\pi) - \mathscr{G}(\pi)G(1)$, and recall that $Q(\cdot) = q(\cdot)' q(\cdot)$ where $q(\pi) = (\mathbf{I}_k - \mathcal{R}_\pi' \mathcal{R}_\pi)^{-1/2} \mathcal{R}_\pi^{-1} \mathbb{G}(\pi) \equiv (\mathscr{G}_\pi - \mathscr{G}_\pi^{\,2})^{-1/2} \mathbb{G}(\pi)$, $\pi \in \Pi$. Also observe that, for

$$S \equiv S(\pi_1, \pi_2) \triangleq \frac{1}{\pi_2 - \pi_1} \int_{\pi_1}^{\pi_2} Q(r)dr\,; \tag{4.2}$$

$$E(S) = \frac{1}{\pi_2 - \pi_1} \int_{\pi_1}^{\pi_2} E\big(Q(r)\big)dr\,, \tag{4.3}$$

and $$E(S^2) = \frac{1}{(\pi_2 - \pi_1)^2} E \int_{\pi_1}^{\pi_2} Q(r)dr \int_{\pi_1}^{\pi_2} Q(r)dr \equiv \frac{1}{(\pi_2 - \pi_1)^2} \int_{\pi_1}^{\pi_2}\int_{\pi_1}^{\pi_2} E\big(Q(r)Q(s)\big)dsdr\,. \tag{4.4}$$

Rewrite $\mathbb{G}(\pi)$ in the usual manner as $\mathbb{G}(\pi) = (\mathbf{I}_k - \mathscr{G}_\pi)\, G(\pi) - \mathscr{G}_\pi D(\pi)$, where $D(\pi) \triangleq G(1) - G(\pi)$ is by definition independent of $G(\pi)$. Then

$$\begin{aligned}
\mathbb{G}(r)\mathbb{G}(s)' &= \big[(\mathbf{I}_k - \mathscr{G}_r)G(r) - \mathscr{G}_r D(r)\big]\big[(\mathbf{I}_k - \mathscr{G}_s)G(s) - \mathscr{G}_s D(s)\big]' \\
&= \big[(\mathbf{I}_k - \mathscr{G}_r)\big(D(r,s) + G(s)\big) - \mathscr{G}_r D(r)\big] \\
&\quad \times \big[G'(s)(\mathbf{I}_k - \mathscr{G}_s) - \big(D(r) + D(r,s)\big)\mathscr{G}_s\big]
\end{aligned}$$

where, for $r \geq s$, $G(s)$, $D(r)$, and $D(r,s) \triangleq G(r) - G(s)$ are mutually independent, with variance-covariance matrices $\mathscr{G}(s)$, $\mathbf{I}_k - \mathscr{G}(r)$, and $\mathscr{G}(r) - \mathscr{G}(s)$, respectively. We therefore have the following Lemma and Corollary.

LEMMA 4.1. *For* $r, s \in [0,1]$, *and* $r \geq s$,

$$E\big(\mathbb{G}(r)\mathbb{G}(s)'\big) = \big(\mathbf{I}_k - \mathscr{G}_r\big)\mathscr{G}_s\,.$$

COROLLARY 4.1. *For* $r \in [0,1]$, *and* $\pi \in \Pi$,

$$\mathbb{G}(r) \sim \mathrm{N}\big(\mathbf{0}, \mathscr{G}_r - \mathscr{G}_r^2\big) \;\Rightarrow\; q(\pi) \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_k)\,.$$

The marginal distribution, and expectation, of $Q(\cdot) = q(\cdot)'q(\cdot)$ is now obvious: for fixed $r \in \Pi \subset (0,1)$ $Q(r) \sim \chi^2_{(k)}$, and $E\big(Q(r)\big) = k$.

The covariance kernel $E\big(Q(r)Q(s)\big)$ could similarly be derived by writing $Q(r)Q(s)$ $= \mathbb{G}(r)'(\mathscr{G}_r - \mathscr{G}_r^2)^{-1}\mathbb{G}(r)\mathbb{G}(s)'(\mathscr{G}_s - \mathscr{G}_s^2)^{-1}\mathbb{G}(s)$ in terms of $G(s)$, $D(r)$, and $D(r, s)$, and collecting all terms with non-zero expectation, after the manner of LEMMA 4.1 above. However, a much less tedious approach uses the following Lemma to rewrite $Q(r)$ in terms of a GBM on [0, 1).

LEMMA 4.2. The vector process $\mathcal{K}(\cdot)$ defined on [0, 1) by

$$\mathcal{K}(\pi) \triangleq (\mathbf{I}_k - \mathcal{G}_\pi)^{-1}\mathbb{G}(\pi) \sim \mathrm{N}\left(\mathbf{0}, \mathcal{G}_\pi(\mathbf{I}_k - \mathcal{G}_\pi)^{-1}\right)$$

has covariance kernel $E\left(\mathcal{K}(r)\mathcal{K}'(s)\right) = \mathcal{G}_s(\mathbf{I}_k - \mathcal{G}_s)^{-1} \equiv V\left(\mathcal{K}(s)\right)$ for $r \geq s$. $\mathcal{K}(\cdot)$ is thus a GBM on [0, 1).

PROOF OF LEMMA 4.2. That the marginal distribution of $\mathcal{K}(\pi)$ is as stated is evident on simplifying $(\mathbf{I}_k - \mathcal{G}_\pi)^{-1}(\mathcal{G}_\pi - \mathcal{G}_\pi^2)(\mathbf{I}_k - \mathcal{G}_\pi)^{-1}$. The covariance kernel is just

$$\begin{aligned} E\left(\mathcal{K}(r)\mathcal{K}'(s)\right) &= (\mathbf{I}_k - \mathcal{G}_r)^{-1} E\left(\mathbb{G}(r)\mathbb{G}(s)'\right)(\mathbf{I}_k - \mathcal{G}_s)^{-1} \\ &\equiv (\mathbf{I}_k - \mathcal{G}_r)^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{G}_s(\mathbf{I}_k - \mathcal{G}_s)^{-1} \end{aligned}$$

for $r \geq s$ by LEMMA 4.1. Likewise, for $r \leq s$ we find that $E\left(\mathbb{G}(r)\mathbb{G}(s)'\right) = \mathcal{G}_r\left(\mathbf{I}_k - \mathcal{G}_s\right)$, implying $E\left(\mathcal{K}(r)\mathcal{K}'(s)\right) = (\mathbf{I}_k - \mathcal{G}_r)^{-1}\mathcal{G}_r \equiv V\left(\mathcal{K}(r)\right)$.

We are now ready to prove the final Lemma and Theorem.

LEMMA 4.3. *For* $Q(\cdot) = q(\cdot)'q(\cdot)$ *as given in Theorem* 1, $r, s \in (0,1)$, *and* $r \geq s$,

$$E\left(Q(r)Q(s)\right) = k^2 + 2\,\mathrm{tr}\left(\mathcal{G}_r^{-1}\left(\mathbf{I}_k - \mathcal{G}_r\right)\mathcal{G}_s\left(\mathbf{I}_k - \mathcal{G}_s\right)^{-1}\right).$$

PROOF OF LEMMA 4.3. First note that, in terms of $\mathcal{K}(\cdot)$,

$$\begin{aligned} Q(r) &= \mathbb{G}(r)'(\mathcal{G}_r - \mathcal{G}_r^2)^{-1}\mathbb{G}(r) \\ &= \mathcal{K}(r)'(\mathbf{I}_k - \mathcal{G}_r)(\mathcal{G}_r - \mathcal{G}_r^2)^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{K}(r) \\ &= \mathcal{K}(r)'\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{K}(r)\,; \end{aligned}$$

implying

$$
\begin{aligned}
Q(r)Q(s) &= \mathcal{K}'(r)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{K}(r)\mathcal{K}'(s)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\mathcal{K}(s) \\
&= \left(\mathcal{D}'(r,s) + \mathcal{K}'(s)\right)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\left(\mathcal{D}(r,s) + \mathcal{K}(s)\right) \\
&\qquad \times \mathcal{K}'(s)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\mathcal{K}(s)
\end{aligned}
$$

where, for $r > s$, $\mathcal{D}(r,s) = \mathcal{K}(r) - \mathcal{K}(s)$ is independent of $\mathcal{K}(s)$. Thus

$$
\begin{aligned}
E\left(Q(r)Q(s)\right) &= E\left(\mathcal{D}'(r,s)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{D}(r,s)\mathcal{K}'(s)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\mathcal{K}(s)\right) \\
&\quad + E\left(\mathcal{K}'(s)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\mathcal{K}(s)\mathcal{K}'(s)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\mathcal{K}(s)\right).
\end{aligned}
$$

Applying the usual results[8] regarding expectations of products of quadratic forms in independent normal variables then yields

$$
\begin{aligned}
E\left(Q(r)Q(s)\right) &= \mathrm{tr}\left(V\left(\mathcal{D}(r,s)\right)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\right)\mathrm{tr}\left(V\left(\mathcal{K}(s)\right)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\right) \\
&\quad + \mathrm{tr}\left(V\left(\mathcal{K}(s)\right)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)\right)\mathrm{tr}\left(V\left(\mathcal{K}(s)\right)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\right) \\
&\quad + 2\,\mathrm{tr}\left(V\left(\mathcal{K}(s)\right)\mathcal{G}_r^{-1}(\mathbf{I}_k - \mathcal{G}_r)V\left(\mathcal{K}(s)\right)\mathcal{G}_s^{-1}(\mathbf{I}_k - \mathcal{G}_s)\right),
\end{aligned}
$$

where $V\left(\mathcal{K}(s)\right) = \mathcal{G}_s(\mathbf{I}_k - \mathcal{G}_s)^{-1}$ and $V\left(\mathcal{D}(r,s)\right) = \mathcal{G}_r(\mathbf{I}_k - \mathcal{G}_r)^{-1} - \mathcal{G}_s(\mathbf{I}_k - \mathcal{G}_s)^{-1}$. Substituting these into the preceding expression then establishes the Lemma.

THEOREM 3. *Let $\mathcal{V}(r,s)$ denote the scalar quantity*

$$
\mathrm{tr}\left(\mathcal{G}_r^{-1}\left(\mathbf{I}_k - \mathcal{G}_r\right)\mathcal{G}_s\left(\mathbf{I}_k - \mathcal{G}_s\right)^{-1}\right) \equiv \mathrm{tr}\left(\left(\mathcal{G}_r^{-1} - \mathbf{I}_k\right)\left(\mathcal{G}_s^{-1} - \mathbf{I}_k\right)^{-1}\right).
$$

Then, for $0 < \pi_1 < \pi_2 < 1$ and $S \equiv S(\pi_1, \pi_2)$ as per (4.2)*,*

$$
E(S) = k\,,
$$

and

$$
V(S) = \frac{2}{(\pi_2 - \pi_1)^2}\left\{\int_{\pi_1}^{\pi_2}\left(\int_{\pi_1}^{r}\mathcal{V}(r,s)ds + \int_{r}^{\pi_2}\mathcal{V}(s,r)ds\right)dr\right\}. \tag{4.5}
$$

---

[8] For independent n-vectors $x \sim \mathrm{N}(0,V_x)$ and $y \sim \mathrm{N}(0,V_y)$, and symmetric $n \times n$ A and B, $E(x'Ax\,y'By) = \mathrm{tr}(V_xA)\mathrm{tr}(V_yB)$ and $E(x'Ax\,x'Bx) = \mathrm{tr}(V_xA)\mathrm{tr}(V_xB) + 2\mathrm{tr}(V_xAV_xB)$.

PROOF OF THEOREM 3. $E(S)=k$ obtains instantly on substituting $E\big(Q(r)\big)=k$ into (4.3). With regard to the variance, rewrite (4.4) with $r \geq s$ and $r < s$ treated separately, as

$$E(S^2)=\frac{1}{(\pi_2-\pi_1)^2}\left\{\int_{\pi_1}^{\pi_2}\int_{\pi_1}^{r}E_{r\geq s}\big(Q(r)Q(s)\big)dsdr+\int_{\pi_1}^{\pi_2}\int_{r}^{\pi_2}E_{r<s}\big(Q(r)Q(s)\big)dsdr\right\};$$

where $E_{r\geq s}\big(Q(r)Q(s)\big)=k^2+2\mathscr{V}(r,s)$ and $E_{r<s}\big(Q(r)Q(s)\big)=k^2+2\mathscr{V}(s,r)$ from Lemma 4.3. Hence

$$\begin{aligned}E(S^2)&=\frac{k^2}{(\pi_2-\pi_1)^2}\int_{\pi_1}^{\pi_2}\int_{\pi_1}^{\pi_2}dsdr+\frac{2}{(\pi_2-\pi_1)^2}\left\{\int_{\pi_1}^{\pi_2}\int_{\pi_1}^{r}\mathscr{V}(r,s)dsdr+\int_{\pi_1}^{\pi_2}\int_{r}^{\pi_2}\mathscr{V}(s,r)dsdr\right\}\\&=k^2+\frac{2}{(\pi_2-\pi_1)^2}\left\{\int_{\pi_1}^{\pi_2}\left(\int_{\pi_1}^{r}\mathscr{V}(r,s)ds+\int_{r}^{\pi_2}\mathscr{V}(s,r)ds\right)dr\right\};\end{aligned}$$

which on substitution into $V(S)=E(S^2)-k^2$ establishes the Theorem.

REMARK. We see that, while the mean of $S(\pi_1, \pi_2)$ involves only the dimension of the regressor vector, its variance depends also on the bounds π_1 and π_2, and on the limiting moment matrix $\mathscr{G}(\cdot)$. However, as already noted above, $\mathscr{G}(\pi)$ can be consistently estimated under Assumption 3.1(c) via the sample quantity $\mathbf{M}_T^{-1/2}\mathbf{M}_\tau\mathbf{M}_T^{-1/2\prime}$. Conditioning on the observed regressors accomplishes the same result under A3.1(c′). $V(S)$, and hence $\overline{W}_T^c$, can therefore be computed on the basis of sample information alone. We return to this matter in more detail in Section 4.3.

On the other hand, if the stochastic process generating the regressors is known then it may be worth extending the solution of $V(S)$ – particularly for certain simple cases in which a closed form expression for $\mathscr{G}(\cdot)$ is available. Accordingly, we close this section by restating Lemma 4.3 and Theorem 3 for the special case in which the regressors are stationary in the sense required by Andrews (1993). The proof of Corollary 4.4 can be found in the Appendix.

COROLLARY 4.3. For $r \geq s$ and stationary regressors ∍ $\mathscr{G}(\pi) = \pi \mathbf{I}_k$,

$$E\big(Q(r)Q(s)\big) = k^2 + 2k\,\frac{s(1-r)}{r(1-s)}\,. \tag{4.6}$$

COROLLARY 4.4. For stationary regressors ∍ $\mathscr{G}(\pi) = \pi \mathbf{I}_k$,

$$\begin{aligned} V(S) = 2k &- \frac{8k}{(\pi_2 - \pi_1)} \\ &+ \frac{2k}{(\pi_2-\pi_1)^2}\left\{\big(2(1-\pi_2)+\ln\pi_2\big)\ln\frac{1-\pi_1}{1-\pi_2} + \big(2\pi_1 + \ln(1-\pi_1)\big)\ln\frac{\pi_2}{\pi_1}\right\} \\ &+ \frac{2k}{(\pi_2-\pi_1)^2}\left\{\,\mathrm{Li}_2(\pi_2) - \mathrm{Li}_2(1-\pi_2) - \mathrm{Li}_2(\pi_1) + \mathrm{Li}_2(1-\pi_1)\right\} \end{aligned}$$

where $\mathrm{Li}_2(\cdot)$ denotes the dilogarithm $\mathrm{Li}_2(z) = \int_z^0 \frac{\ln(1-u)}{u}\,du = \sum_{k=1}^{\infty}\frac{z^k}{k^2}$.

REMARK. Under symmetric trimming ($\pi_2 = 1 - \pi_1$) $V(S)$ for stationary regressors simplifies to

$$V(S) = 2k - \frac{8k}{(\pi_2-\pi_1)} + \frac{4k}{(\pi_2-\pi_1)^2}\left\{\big(2\pi_1 + \ln\pi_2\big)\ln\frac{\pi_2}{\pi_1} + \mathrm{Li}_2(\pi_2) - \mathrm{Li}_2(\pi_1)\right\}.$$

Thus even for this simplest case we do not quite obtain a closed form solution. We conclude that for more complex regressor models it may well be more straightforward to compute (4.5) numerically, even assuming known limiting matrix $\mathscr{G}(\cdot)$.

4.2 Comparison: the S_c limiting distribution *vs* the chi-square

The question of whether, under H_0, the $\chi^2_{(k)}$ distribution is a close enough match for the limiting distribution of $\overline{W}_T^c$ to provide "reasonably accurate" approximate asymptotic critical values is examined by simulating the respective distributions for a variety of regressor types. The six designs included were:

(i) a constant plus a variable distributed iid N(0,1);

(ii) a constant plus a deterministic linear trend;

(iii) design (ii) plus a variable distributed iid N(0,1);

(iv) a constant plus a variable distributed iid N(0,1) for the first half of the sample, N(1,1) for the second half;

(v) a constant plus a variable distributed iid N(0,1) for the first half of the sample, N(0,2) for the second half; and

(vi) a constant, plus a random walk variable generated as the cumulative sum of a vector of iid N(0,1) variates.

These designs will be referred to as the "stationary", "linear trend", "trend + N(0,1)", "step-stationary", "step-heteroscedastic", and "random walk" designs, respectively. In addition, it will be convenient to refer to the limiting process $S(\pi_1, \pi_2)$ as the "AveQ" distribution; its chi-square adjusted variant $S_c(\pi_1, \pi_2)$ is then "AveQ$_c$".

AveQ and AveQ$_c$ were simulated over 10000 repetitions for each of the six designs. Each AveQ "draw" was obtained by simulating the limit process $Q(\pi)$ (equation (3.1)) over the grid $\Pi(N)$ defined by $[\pi_1, 1-\pi_1] \cap \{\pi = j/N : j = 0,1,\ldots,N\}$ with $\pi_1 = 0.1$. We set $N = 5000$ for designs (i) – (v), since for these cases $\mathscr{G}(\cdot)$ and hence $V(S)$ are nonstochastic and need only be computed once. For design (vi) we set $N = 2500$, since in this case Theorem 1 holds only conditionally, making it necessary to recompute the process $\mathscr{G}(\cdot)$, and in particular the limiting variance $V(S)$, every repetition. In either case $V(S)$ was computed over the $n = 4001$ (2001) possible breakpoints in $\Pi(N)$ by a double sum over $\mathscr{V}(i/T, j/T)$ along the lines of (4.9) below. The corresponding AveQ$_c$ draw follows via transformation (4.1). The simulated AveQ$_c$ distribution functions for each design are compared with the corresponding chi-square cdf (cumulative distribution function) in Figure 1.

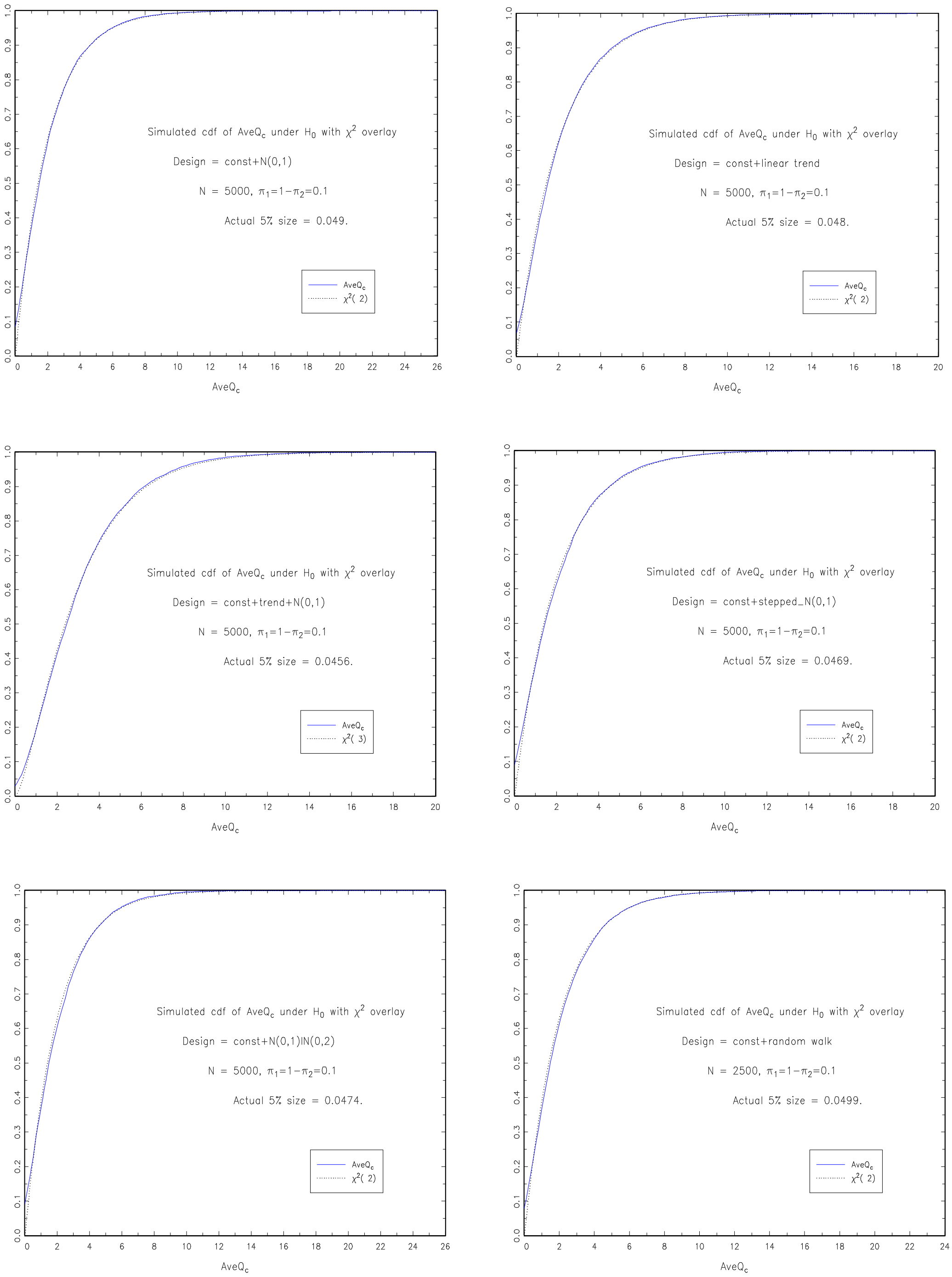

Figure 1. Simulated cdf of the AveQ$_c$ statistic for designs (i) iid N(0,1); (ii) linear trend; (iii) trend plus an iid N(0,1); (iv) stepped iid normal; (v) step-heteroscedastic; and (vi) random walk.

We see that, in each case, the greatest difference in the cdfs is in the left-hand tail, due to the non-zero probability of obtaining $S_c < 0$ after applying transformation (4.1) to S. However, in the right-hand tail the correspondence is remarkably close. We also note that the probability of obtaining $S_c < 0$ diminishes as the number of regressors increases; implying that the worst case for this kind of "pile-up" at zero would correspond to a single regressor ($k = 1$) design. Otherwise the chi-square distribution with k degrees of freedom appears to be a fairly good approximation for the actual limiting distribution of $\overline{W}_T^c$.

The "actual 5% sizes" reported are the observed probability that S_c exceeds the 95th percentile of the $\chi^2_{(k)}$ distribution. The reported probabilities thus estimate the true asymptotic size of the AveF_c test conducted with chi-square critical values for each of the designs. We note that, while all of these simulated sizes are below the nominal 5% significance level, none are significantly different[9] from 5%. The AveF_c test thus seems likely to have slightly lower size "asymptotically" (and hence, we expect, slightly lower power) than would properly be the case. The difference, however, would appear to be fairly marginal. We conclude that the chi-square approximation is a reasonable solution to the problem of obtaining critical values for the AveF test.

4.3 Practical issues: implementing the test

The greatest practical obstacle to use of the chi-square approximation is the need to compute the limiting variance $V(S)$. In theory, if the process generating the regressors is known then $V(S)$ can be derived as a known function of π_1 and π_2 – though the derivation would have to repeated for each case. On the other hand, even the simplest (stationary regressor) case does not allow a closed form solution for $V(S)$. Analogous results for other regressor types are not likely to be any less complex; and, of course, would apply only to the regressor "model" under consideration. Furthermore, the use of such a result in practice necessarily assumes that the chosen regressor model is

[9] Assuming the usual standard error of estimation of a proportion; in this case equal to 0.0022.

correct, at least to the extent of yielding the appropriate limiting moment matrix. Use of a sample-based alternative is therefore arguably more robust, as well as being more general, since it makes no assumptions beyond those of Section 3.1; and, in particular, does not require us to possess specific information regarding the process, stochastic or otherwise, generating the regressors.

Accordingly, consider the estimation of $V(S)$ via the sample quantities $\tilde{\mathscr{G}}(s)$ and $\tilde{\mathscr{G}}(r)$, where $\tilde{\mathscr{G}}(r) \triangleq \mathbf{M}_T^{-1/2}\mathbf{M}_{[rT]}\mathbf{M}_T^{-1/2\prime} \xrightarrow{p} \mathscr{G}(r)$ under A3.1(c). Then

$$\tilde{\mathscr{V}}(r,s) = \mathrm{tr}\left(\left(\tilde{\mathscr{G}}_s^{-1} - \mathbf{I}_k\right)^{-1}\left(\tilde{\mathscr{G}}_r^{-1} - \mathbf{I}_k\right)\right) \xrightarrow{p} \mathscr{V}(r,s);$$

implying

$$\tilde{V}(S) = \frac{2}{(\pi_2 - \pi_1)^2}\left\{\int_{\pi_1}^{\pi_2}\left(\int_{\pi_1}^{r}\tilde{\mathscr{V}}(r,s)ds + \int_{r}^{\pi_2}\tilde{\mathscr{V}}(s,r)ds\right)dr\right\} \tag{4.7}$$

is a consistent estimator of $V(S)$. The limiting distribution of $\overline{W}_T^c$ is accordingly unaltered by the use of $\tilde{V}(S)$ rather than $V(S)$. That is,

$$\overline{W}_T^c(\pi_1,\pi_2) = k + \sqrt{\frac{2k}{\tilde{V}(S)}}(\overline{W}_T - k) \xrightarrow{d} S_c(\pi_1,\pi_2) \tag{4.8}$$

under Assumption 3.1; with $\overline{W}_T^c$ now denoting our “estimated” statistic, and $E(S)$ replaced by k.

Replacing A3.1(c) by A3.1(c′) implies $\tilde{V}(S)$ converges only weakly to the now stochastic quantity $V(S)$; however this is sufficient for (4.8) to hold conditionally, as discussed in the Remarks concluding Section 3.1.

With regard to the computation of $\tilde{V}(S)$, first note that (4.7) can also be written, because of the “stepped” nature of $\tilde{\mathscr{V}}(r,s)$, as

$$\tilde{V}(S) = \frac{2}{n^2}\left\{\sum_{i=\tau_1}^{\tau_2}\sum_{j=\tau_1}^{i}\tilde{\mathcal{V}}(i/T, j/T) + \sum_{i=\tau_1}^{\tau_2}\sum_{j=i+1}^{\tau_2}\tilde{\mathcal{V}}(j/T, i/T)\right\}; \tag{4.9}$$

where $\tau_1 = [\pi_1 T]$, $\tau_2 = [\pi_2 T]$, and $n = \tau_2 - \tau_1 + 1$ is the number of breakpoints possible in the trimmed sample[10]. Rewriting $\tilde{\mathcal{V}}(r,s)$ as $\mathrm{tr}\left(\mathbf{D}_{[rT]}\mathbf{D}_{[sT]}^{-1}\right)$ where $\mathbf{D}_{[rT]} \triangleq \mathbf{M}_{[rT]}^{-1} - \mathbf{M}_T^{-1}$, $r \in [\pi_1, \pi_2]$, then yields

$$\begin{aligned}
\tilde{V}(S) &= \frac{2}{n^2}\left\{\sum_{i=\tau_1}^{\tau_2}\sum_{j=\tau_1}^{i}\mathrm{tr}\left(\mathbf{D}_i\mathbf{D}_j^{-1}\right) + \sum_{i=\tau_1}^{\tau_2}\sum_{j=i+1}^{\tau_2}\mathrm{tr}\left(\mathbf{D}_i^{-1}\mathbf{D}_j\right)\right\} \\
&= \frac{2}{n^2}\left\{nk + \sum_{i=\tau_1}^{\tau_2}\sum_{j=\tau_1}^{i-1}\mathrm{tr}\left(\mathbf{D}_i\mathbf{D}_j^{-1}\right) + \sum_{j=\tau_1}^{\tau_2}\sum_{i=j+1}^{\tau_2}\mathrm{tr}\left(\mathbf{D}_i\mathbf{D}_j^{-1}\right)\right\} \\
&= \frac{2k}{n} + \frac{4}{n^2}\sum_{i=\tau_1+1}^{\tau_2}\sum_{j=\tau_1}^{i-1}\mathrm{tr}\left(\mathbf{D}_i\mathbf{D}_j^{-1}\right).
\end{aligned} \tag{4.10}$$

$\tilde{V}(S)$ is clearly not difficult to compute. Indeed, the double sum in (4.10) is generally most efficiently obtained by constructing the $n \times n$ matrix of traces $\tilde{\mathbf{D}}'\overline{\mathbf{D}}$, where $\tilde{\mathbf{D}} = \left[vec\mathbf{D}_{\tau_1} \quad vec\mathbf{D}_{\tau_1+1} \quad \cdots \quad vec\mathbf{D}_{\tau_2}\right]$ and $\overline{\mathbf{D}} = \left[vec\mathbf{D}_{\tau_1}^{-1} \quad vec\mathbf{D}_{\tau_1+1}^{-1} \quad \cdots \quad vec\mathbf{D}_{\tau_2}^{-1}\right]$; and summing over the $\frac{1}{2}n(n-1)$ elements below the principal diagonal. Though it appears cumbersome, this procedure takes less than a second for the sample sizes typically encountered in practical work; particularly since the component matrices ($\mathbf{M}_{[rT]}^{-1},\ r \in [\pi_1, \pi_2]$, and $\mathbf{M}_T^{-1}$) would normally have already been computed in the construction of the test statistic. This is particularly evident if we use (3.2) to rewrite (2.3) as

$$W_\tau = W_T(\tau/T) = \hat{\sigma}^{-2}\mathbf{y}'(\mathbf{Z}_\tau\mathbf{M}_\tau^{-1} - \mathbf{X}\mathbf{M}_T^{-1})\mathbf{D}_\tau^{-1}(\mathbf{M}_\tau^{-1}\mathbf{Z}_\tau' - \mathbf{M}_T^{-1}\mathbf{X}')\mathbf{y},$$

where $\hat{\sigma}^2 = \mathbf{y}'\overline{\mathbf{P}}_{\mathbf{X}}\mathbf{y}$ is a consistent estimator of σ^2 under H_0, and $\mathbf{D}_\tau^{-1} = (\mathbf{M}_\tau^{-1} - \mathbf{M}_T^{-1})^{-1}$. The AveF statistic $\overline{W_T}(\pi_1, \pi_2)$ is then computed as

$$\overline{W_T}(\pi_1, \pi_2) = \frac{1}{n}\sum_{\tau=\tau_1}^{\tau_2} W_\tau,$$

with the $k^2 \times n$ matrices $\tilde{\mathbf{D}}$ and $\overline{\mathbf{D}}$ being constructed as by-products.

However, for the most part it turns out to be more efficient to compute $\overline{W}_T(\pi_1, \pi_2)$ as the ratio of quadratic forms $\mathbf{y}'\overline{\mathbf{P}}_\mathbf{Z}\mathbf{y} / \mathbf{y}'\overline{\mathbf{P}}_\mathbf{X}\mathbf{y}$, where the $T \times T$ matrix

$$\overline{\mathbf{P}}_\mathbf{Z} = \frac{1}{n}\sum_{\tau=\tau_1}^{\tau_2}(\mathbf{Z}_\tau\mathbf{M}_\tau^{-1} - \mathbf{X}\mathbf{M}_T^{-1})\mathbf{D}_\tau^{-1}(\mathbf{M}_\tau^{-1}\mathbf{Z}_\tau' - \mathbf{M}_T^{-1}\mathbf{X}')$$

is constructed, along with $\tilde{\mathbf{D}}$ and $\overline{\mathbf{D}}$, from scratch. The steps involved in computing the AveF$_c$ test statistic and performing the test are then as follows.

1. Obtain $\hat{\sigma}^2$ and $\mathbf{M}_T$ from the OLS regression on $\mathbf{y}$ on $\mathbf{X}$, and select π_1 and π_2.

2. Compute, for each $\tau = [\pi_1 T], \ldots, [\pi_2 T]$, $\mathbf{Z}_\tau$, $\mathbf{M}_\tau$, $\mathbf{D}_\tau$, *et cetera*, and hence construct $\overline{\mathbf{P}}_\mathbf{Z}$, $\tilde{\mathbf{D}}$ and $\overline{\mathbf{D}}$.

3. Calculate the AveF statistic $\overline{W}_T(\pi_1, \pi_2) = \mathbf{y}'\overline{\mathbf{P}}_\mathbf{Z}\mathbf{y} / \hat{\sigma}^2$; and estimate its variance using (4.10).

4. Compare the AveF$_c$ statistic $\overline{W}_T^c(\pi_1, \pi_2) = k + \sqrt{\frac{2k}{\tilde{V}(S)}}\left(\overline{W}_T(\pi_1, \pi_2) - k\right)$ with critical values from the χ^2 distribution with k degrees of freedom.

5. Conclusion

This paper derives the general distributional theory for the Sup- and AveF tests for a structural break of unknown timing in the parameters of the linear regression model under assumptions that permit most common types of regressors. With the limiting distributions heavily dependent on a design-dependent limiting matrix, we then propose and evaluate a sample-based chi-square "adjustment" to the AveF test to allow the use of standard chi-square critical values. The chi-square distribution

---

[10] Naturally, $\tilde{V}(S)$ as given in (4.9) is identical to the expression that would be obtained for $V(\overline{W}_T)$ under the assumption of non-stochastic $\mathbf{X}$ and normally distributed disturbances.

function is found to be a reasonably accurate approximation to the true limiting distribution of the two-moment modified mean Wald statistic for all the designs considered. The problem of obtaining useable critical values for a diagnostic-style test of structural change might thus be described as (almost) more-or-less solved, at least for these and similar designs in the context of the linear model.

Of course, more work needs to be undertaken to see how well the $\chi^2_{(k)}$ distribution matches the *finite sample* distribution of $\overline{W}_T^c$. In other words, will the use of chi-square critical values in conjunction with the $\overline{W}_T^c$ statistic also result in *finite sample* sizes acceptably close to nominal? Furthermore, while the AveF_c test necessarily has the same asymptotic local power as the AveF, the effect of the chi-square adjustment on finite sample power should also be investigated.

Finally, we have restricted our attention to one-off change in the structural parameters in the linear regression model. The consequences of relaxing each or all of the assumptions involved would provide ample scope for further work.

Appendix

PROOF OF COROLLARY 4.4. Direct substitution of (4.6) into (4.5) yields

$$V(S) = \frac{2k}{(\pi_2 - \pi_1)^2} \left\{ \int_{\pi_1}^{\pi_2} \left(\frac{1-r}{r} \int_{\pi_1}^{r} \frac{s}{1-s} ds + \frac{r}{1-r} \int_{r}^{\pi_2} \frac{1-s}{s} ds \right) dr \right\}.$$

Using $\int \frac{1-s}{s} ds = \ln s - s$ and $\int \frac{s}{1-s} ds = -s - \ln(1-s)$ we obtain, for $0 < a < b < 1$,

$$\begin{aligned}\int_a^b \frac{r}{1-r} \int_r^b \frac{1-s}{s} ds\, dr &= \int_a^b \frac{r}{1-r} \{\ln b - b + r - \ln r\} dr \\ &= \frac{a^2}{2} - \frac{b^2}{2} + \left\{ a - b + \ln \frac{1-a}{1-b} \right\} (1 - b + \ln b) - \int_a^b \frac{r}{1-r} \ln r \, dr\end{aligned}$$

and

$$\int_a^b \frac{1-r}{r}\int_a^r \frac{s}{1-s}\,ds\,dr = \int_a^b \frac{1-r}{r}\left\{a+\ln(1-a)-r-\ln(1-r)\right\}dr$$
$$= \frac{(b-1)^2}{2}-\frac{(a-1)^2}{2}+\left\{a-b+\ln\frac{b}{a}\right\}\left(a+\ln(1-a)\right)-\int_a^b \frac{1-r}{r}\ln(1-r)\,dr\,.$$

The remaining integrals are solved in terms of the dilogarithm function as

$$\int_a^b \frac{r}{1-r}\ln r\,dr = \int_a^b \left\{\frac{\ln r}{1-r}-\ln r\right\}dr = \left[\mathrm{Li}_2(1-r)\right]_a^b + \left[r-r\ln r\right]_a^b,\ \text{ and}$$

$$\int_a^b \frac{1-r}{r}\ln(1-r)\,dr = \int_a^b\left\{\frac{\ln(1-r)}{r}-\ln(1-r)\right\}dr = \left[-\mathrm{Li}_2(r)\right]_a^b + \left[r+(1-r)\ln(1-r)\right]_a^b;$$

so that, after rearranging and collecting terms,

$$\int_a^b \frac{r}{1-r}\int_r^b \frac{1-s}{s}\,ds\,dr = 2(a-b)+\frac{1}{2}(b-a)^2+a\ln\frac{b}{a}+(1-b)\ln\frac{1-a}{1-b}$$
$$+\ln b\ln\frac{1-a}{1-b}\ -\mathrm{Li}_2(1-b)+\mathrm{Li}_2(1-a)$$

and
$$\int_a^b \frac{1-r}{r}\int_a^r \frac{s}{1-s}\,ds\,dr = 2(a-b)+\frac{1}{2}(b-a)^2+a\ln\frac{b}{a}+(1-b)\ln\frac{1-a}{1-b}$$
$$+\ln(1-a)\ln\frac{b}{a}\ -\mathrm{Li}_2(a)+\mathrm{Li}_2(b)\,.$$

Hence

$$\int_a^b \frac{r}{1-r}\int_r^b \frac{1-s}{s}\,ds\,dr + \int_a^b \frac{1-r}{r}\int_a^r \frac{s}{1-s}\,ds\,dr$$
$$=(b-a)^2-4(b-a)+2a\ln\frac{b}{a}+2(1-b)\ln\frac{1-a}{1-b}$$
$$+\ln b\ln\frac{1-a}{1-b}+\ln(1-a)\ln\frac{b}{a}+\mathrm{Li}_2(b)-\mathrm{Li}_2(1-b)-\mathrm{Li}_2(a)+\mathrm{Li}_2(1-a)\,.$$

Replacing a by π_1 and b by π_2 then yields the stated result.

References

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 and W. Ploberger, 1996, Optimal changepoint tests for normal linear regression, Journal of Econometrics 70, 9-38.

Andrews, D. W. K. and W. Ploberger, 1994, Optimal tests when a nuisance parameter is present only under the alternative, Econometrica 62, 1383-1414.

Bai, J. and P. Perron, 1998, Estimating and testing linear models with multiple structural changes, Econometrica 66, 47-78.

Hansen, B. E., 1992, Tests for parameter instability in regressions with I(1) processes, Journal of Business and Economic Statistics 10, 321-335.

Hansen, B. E., 1996, Testing for structural change in conditional models, unpublished manuscript, Dept. of Economics, Boston College.

Nyblom, J., 1989, Testing for the constancy of parameters over time, Journal of the American Statistical Association 84, 223-230.

Park, J. Y. and P. C. B. Phillips, 1988, Statistical inference in regressions with integrated processes: Part I, Econometric Theory 4, 468-497.

Quandt, R., 1960, Tests of the hypothesis that a linear regression system obeys two separate regimes, Journal of the American Statistical Association 55, 324-330.