Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Econometrics 115 (2003) 91 – 123
JOURNAL OF Econometrics
www.elsevier.com/locate/econbase

Structural change tests for simulated method of moments

Eric Ghysels[a], Alain Guay[b,*]

[a] *University of North Carolina and CIRANO, USA*
[b] *Université du Québec à Montreal and CIRPEE, USA*

Accepted 22 January 2003

Abstract

Simulation-based estimation methods have become more widely used in recent years. We propose a set of tests for structural change in models estimated via simulated method of moments (see Duffe and Singleton (Econometrica 61 (1993) 929). These tests extend the work of Andrews (Econometrica 61 (1993) 821), Ghysels et al. (J. Econom. 82 (1997) 209) and Sowell (Econometrica 64 (1996) 1085; Tests for Violation of Moment Conditions, Manuscript, Graduate School of Industrial Administration, Carnegie mellon University) which covered generalized method of moments estimators not involving simulation. We derive the asymptotic distributions of various tests. We show that the number of simulations does not affect the asymptotic distribution under the null but adversely influences local asymptotic power. A Monte-Carlo investigation of the finite sample size and power reveals that a relatively small number of simulations suffices to obtain tests with desirable small sample size and power properties.

JEL classification: C13; C30

Keywords: Structural change tests; Simulated method of moments; Small sample properties

1. Introduction

The steady increase in computational speed of computers has enhanced the practical use of simulation-based estimators in econometrics. There is now a well-established

* Corresponding author.
E-mail addresses: eghysels@unc.edu (E. Ghysels), guay.alain@uqam.ca (A. Guay).

0304-4076/03/$ - see front matter
doi:10.1016/S0304-4076(03)00094-0

asymptotic distribution theory for a large variety of procedures, including the simulated method of moments estimator (henceforth SMM) which is the focus of our paper. Duffie and Singleton (1993) extended the generalized method of moments procedure (henceforth GMM) of Hansen (1982) and developed the asymptotic properties of the SMM estimation procedure in the context of dynamic econometric time-series models. They show that $\sqrt{T}(\hat{\theta}_T^S - \theta_0)$ is asymptotically normal with covariance matrix identical to the GMM estimator but scaled by $(1 + 1/S)$ where S is the number of simulations and θ_0 the parameter vector of interest.[1] Hence, the asymptotic distributions of SMM and GMM estimators coincide when the number of simulations tends to infinity, since all simulation uncertainty evaporates.

The scope of this paper is to shed light on distribution as $T \to \infty$ and S fixed, for structural change tests. Several tests for structural change with presumed breakpoint unknown already exist for GMM, see in particular Andrews (1993), Andrews and Ploberger (1994), Sowell (1996a, b), Guay (2003), Ghysels et al. (1997) and Hall and Sen (1999). We study several tests for structural change with unknown breakpoint when SMM instead of GMM estimators are involved. The design of the tests is based on the optimality principles of local asymptotic power discussed by Andrews and Ploberger (1994), Sowell (1996a), Guay (2003) and Hall and Sen (1999). While the asymptotic distributions of the partial-sample SMM estimators depend on S, we show that the asymptotic distributions of tests for structural change, which are functions of the partial-sample SMM estimators, are nuisance parameter free under the null. Hence, all the tests we propose have the same asymptotic distributions under the null of no structural change, with S finite, as their GMM counterpart with critical values tabulated in Andrews (1993), Andrews and Ploberger (1994), Sowell (1996a), Guay (2003) and Ghysels et al. (1997). Under local alternatives, the asymptotic distribution depends on the number of simulations S. We study the impact of the number of simulations on the small sample properties of tests for structural change. A Monte-Carlo investigation of the finite sample size and power reveals that simulation uncertainty does affect the properties of tests. Nevertheless, a relatively small number of simulations is sufficient to obtain tests with desirable small sample size and power properties.

In Section 2 we fix notation and discuss the regularity conditions required to establish several large sample properties of SMM estimators which are used in the derivations of the asymptotic distribution of the tests. Section 3 is devoted to testing for structural change. The null hypothesis and the test statistics are formally defined and the main results of the paper, namely the asymptotic distributions of the tests, are presented. In Section 4, we report the results of a Monte-Carlo study of the finite sample properties of the tests. Section 5 concludes the paper.

[1] The relatively simple scaling factor reflecting simulation uncertainty, is not shared by all asymptotic results. Complex specification tests may result in nontrivial influences of simulation uncertainty (see for example, the consistency tests appearing in Appendix 3 of Gouriéroux et al., 1993).

2. Models and simulations

Laroque and Selanie (1989), Ingram and Lee (1991), Duffie and Singleton (1993), Gouriéroux et al. (1993), Smith (1993), Gallant and Tauchen (1996), Gouriéroux and Monfort (1996) and Dridi and Renault (2000) provide extensive coverage of simulation-based methods. Simulation-based procedures are often the key to feasible estimation in various nonlinear dynamic models. In such applications one often wants to test whether the parametric econometric model is invariant through time. These tests either emerge naturally in the context of the model or else can be used as overall diagnostic tests. An example of the former is the estimation of stochastic volatility models, which involve simulation-based estimation procedures (see Ghysels et al., 1996 for a review). van der Sluis (1998, 1999) discusses several applications involving financial market data. He studies foreign exchange markets and examines major events like the British pound leaving the European exchange rate mechanism, in addition he also examines stock return data and considers the stability of stochastic volatility models during the October 1987 stock market crash era. It is of interest to note that several papers examining exchange rate target zone models involve SMM estimators and could also be considered, examples include de Jong (1994), Jorgensen and Mikkelsen (1991) and Iannizzotto and Taylor (1999). Many other examples can be found in the literature on estimating continuous time asset pricing models. Tests for structural change are obviously also important for the estimation of structural models, typically found in macro and other areas. Prominent examples of such models which require a simulation-based approach include Foster and Viswanathan (1995), who derive a speculative trading model with endogenous informed trading that yields a conditionally heteroskedastic time series for trading volume and the squared price changes. On a very different topic Asea and Turnovsky (1998) study how capital income taxes affect household portfolio choice and growth. The authors approach this question within the context of a structural stochastic model of a small open economy. Similarly Hairault et al. (1997) study a structural model of time to implement and aggregate fluctuations. These are examples of model building inspired by the Lucas critique which prompts questions about testing for structural change pertaining to policy invariance.

To establish the asymptotic distribution theory of tests for structural change we need to define first the class of data generating processes we can simulate, how to simulate them and how to define SMM estimators on (sub)samples of data. In the first section we present the class of data generating processes. The next section covers assumptions and definitions and the final section establishes asymptotic properties of partial-sample SMM estimators. Before dealing with these issues we need to elaborate briefly on the specification of the parameter vector in our generic setup. We will consider parametric models indexed by parameters (β, δ) where $\beta \in B$, with $B \subset R^r$ and $\delta \in \Delta \subset R^s$. Following Andrews (1993) we make a distinction between pure structural change when no subvector δ appears and the entire parameter vector is subject to structural change under the alternative and partial structural change which corresponds to cases where only a subvector β is subject to structural change under the alternative. The generic null can be written as follows:

$$H_0: \beta_t = \beta_0 \quad \forall t = 1, \ldots, T. \tag{2.1}$$

The majority of tests we will consider assume as alternative that at some point in the sample there is a single structural break, like for instance,

$$\beta_t = \begin{cases} \beta_{10} & t = 1, \ldots, [\pi T], \\ \beta_{20} & t = [\pi T] + 1, \ldots, T, \end{cases}$$

where π determines the fraction of the sample before and after the assumed break point and [.] denotes the greatest integer function. The separation $[T\pi]$ represents a possible breakpoint which is governed by an unknown parameter π. Hence, we will consider a setup with a parameter vector which encompasses any kind of partial or pure structural change involving a single breakpoint. In particular, we consider a p-dimensional parameter vector $\theta = (\beta_1', \beta_2', \delta')'$ where β_1 and $\beta_2 \in B \subset R^r$ and $\theta \in \Theta = B \times B \times \Delta \subset R^p$ where $p = 2r + v$. The parameters β_1 and β_2 apply to the samples before and after the presumed breakpoint and the null implies that

$$H_0: \beta_{10} = \beta_{20} = \beta_0. \tag{2.2}$$

We will formulate all our models in terms of θ. Special cases could be considered whenever restrictions are imposed in the general parametric formulation. One such restriction would be that $\theta_0 = (\beta_0', \beta_0')'$, which would correspond to the null of a pure structural change hypothesis. Once we have defined the moment conditions we will also translate this into over-identifying restrictions and relate it to structural change tests, following the analysis of Sowell (1996b) and Hall and Sen (1999).

2.1. The data generating process

We consider a process of endogenous observable variables y_t which is generated by the following dynamic model:

$$y_t = H(y_{t-1}, u_t, \theta_0),$$

$$u_t = G(u_{t-1}, \varepsilon_t, \theta_0), \tag{2.3}$$

where θ_0 is the presumed value of the $p \times 1$ parameter vector and u_t and ε_t are unobservable processes.[2] Since we consider simulation-based inference it is assumed that ε_t is a process of disturbances with a *known* density. The processes y_t, u_t and ε_t can either be univariate or multivariate. For the moment we will not be very specific about the conditions we need to impose on y_t and u_t. We assume however that we have a sample of observations $t = 1, \ldots, T$ for y_t. Under (2.3), one can simulate values

[2] It should be noted that in Eq. (2.3) we assume that no exogenous variable as in Duffie and Singleton (1993). The proofs of the theorems in this paper would need to be modified to accommodate the fact that $z_t = (x_t, y_t)$, where x_t are exogenous variables, and its simulated counterpart $z_t^s = (x_t, y_t^s)$ are no longer independent as they share the common exogenous process. The results in this paper can be extended to deal with DGPs that include exogenous processes at a cost of complications that do not change the main asymptotic results we present.

of y_t given initial values for y_0 and u_0 and the parameter vector θ_0. This is done by drawing simulated values ε_t^s from the known density. Simulating processes in the presence of breaks needs to be discussed more carefully as it involves some issues which are not typically addressed in the literature. We noted that $\theta=(\beta_1', \beta_2', \delta')'$ where the parameters β_1 and β_2 apply to the samples before and after the presumed breakpoint. It might therefore be more convenient to rewrite (2.3) as

$$y_t = H(y_{t-1}, u_t, \beta_1, \delta) \quad t = 1, \ldots, [\pi T], \tag{2.4}$$

$$u_t = G(u_{t-1}, \varepsilon_t, \beta_1, \delta) \quad t = 1, \ldots, [\pi T] \tag{2.5}$$

and

$$y_t = H(y_{t-1}, u_t, \beta_2, \delta) \quad t = [\pi T] + 1, \ldots, T, \tag{2.6}$$

$$u_t = G(u_{t-1}, \varepsilon_t, \beta_2, \delta) \quad t = [\pi T] + 1, \ldots, T. \tag{2.7}$$

For a DGP without exogenous variables two simulation schemes are possible. Given S draws of $\{\varepsilon_t^s\}_{t=1}^T$, with $s = 1, \ldots, S$, and a given θ we can obtain $\{u_t^s\}_{t=1}^T$ and $\{y_t^s\}_{t=1}^T$ by (2.4) and (2.6). Note that the starting value for the post-break data is $y_{[\pi T]}$ and is therefore also generated endogenously (under the pre-break regime). The SMM estimator involves moment conditions which are function of $z_t = \{y_t, \varepsilon_t, u_t\}_{t=1}^T$ and of the simulated process $z_t^s(\theta) = \{y_t^s(\theta), \varepsilon_t^s, u_t^s(\theta)\}_{t=1}^T$.[3] Alternatively by drawing $\{\varepsilon_t^s\}_{t=1}^{TS}$, for a fixed S, we can obtain a simulated path of $\{u_t^s(\theta)\}_{t=1}^{[TS\pi]}$ and $\{y_t^s(\theta)\}_{t=1}^{[TS\pi]}$ before the considered breakpoint for a given β_1, δ and a simulated path of $\{u_t^s(\theta)\}_{[TS\pi]+1}^{TS}$ and $\{y_t^s(\theta)\}_{[TS\pi]+1}^{TS}$ after the breakpoint for a given β_2, δ.[4] In the remainder of the paper we consider the first scheme, i.e. drawing S simulated paths, while the asymptotic results derived in the paper also hold for a draw of a single simulated path of TS values.

Let us define

$$k(z_t, \theta) = m(z_t) - E_\theta m(z_t), \tag{2.8}$$

where m is an R^q-valued function of moment conditions and θ is an element of the parameter space $\Theta \subset R^p$ where $q \geqslant p$. For simplicity we denote $k(z_t, \theta)$ as $k_t(\theta)$. The method of moments estimator of θ_0 is based on the following argument:

$$Ek_t(\theta_0) = 0. \tag{2.9}$$

A sample equivalent in simulated context can be written as follows:

$$f_T^S(\theta) = \frac{1}{T} \sum_{t=1}^{T} \left(m(z_t) - \frac{1}{S} \sum_{s=1}^{S} m(z_t^s(\theta), \theta) \right), \tag{2.10}$$

where S is the number of simulations.

[3] It will be assumed that the observed and the simulated processes are given by two triangular arrays of random vectors z_{Tt} and z_{Tt}^s (see Assumption A.2 in Appendix A and Section 2.2). For notational simplicity z_t and z_t^s refer to z_{Tt} and z_{Tt}^s, respectively.

[4] It important to note that this alternative is not possible for a DGP containing exogenous variables.

2.2. Assumptions and definitions

We need to impose restrictions on the admissible class of functions and processes involved in estimation to guarantee well-behaved asymptotic properties of SMM estimators using the entire data sample or subsamples of observations. We discuss a set of regularity conditions required to obtain weak convergence of partial-sample SMM estimators to a function of Brownian motions. To streamline the presentation we only summarize the assumptions and provide a detailed description of them in Appendix A. We first define the standard SMM estimator introduced by Duffie and Singleton (1993) using the full sample data.

Definition 2.1. The full sample SMM estimator $\{\tilde{\theta}_T^S\}$ is a sequence of random vectors such that

$$\tilde{\theta}_T^S = \text{Argmin}_\theta\, f_T^S(\theta)' \hat{W}_T f_T^S(\theta),$$

where $\hat{W}_T$ is a random positive definite symmetric $q \times q$ matrix.

The optimal weighting matrix W is defined to be the inverse of $(1 + 1/S)\Omega$ where

$$\Omega = \lim_{T\to\infty} \text{Var}\left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} [m(z_t) - Em(z_t)] \right).$$

As mentioned by Duffie and Singleton (1993), the matrix Ω depends neither on θ and on the simulated moments. The optimal weighting matrix can be estimated consistently using methods developed by Gallant (1987), Andrews and Monahan (1992), Newey and West (1994), among several others (see also Duffie and Singleton, 1993; Gouriéroux and Monfort, 1996 for a discussion in the context of SMM).

Several tests for structural change also involve partial-sample SMM estimators similar to the partial-sample GMM estimators defined by Andrews (1993). We consider again the two subsamples, the first based on observations $t = 1, \ldots, [T\pi]$ and the second covering $t = [T\pi] + 1, \ldots, T$ where $\pi \in \Pi \subset (0, 1)$. The partial-sample SMM estimators for $\pi \in \Pi$ based on the first and the second subsamples are formally defined as

Definition 2.2. A partial-sample SMM estimator $\{\hat{\theta}_T^S(\pi)\}$ is a sequence of random vectors such that

$$\hat{\theta}_T^S(\pi) = \text{Argmin}_\theta\, \bar{f}_T^S(\theta, \pi)' \hat{W}_T(\pi) \bar{f}_T^S(\theta, \pi)$$

for all $\pi \in \Pi$, where

$$\bar{f}_T^S(\theta, \pi) = \begin{bmatrix} f_{1T}^S(\theta, \pi) \\ f_{2T}^S(\theta, \pi) \end{bmatrix},$$

$$f_{1T}^S(\theta, \pi) = \frac{1}{T} \sum_{t=1}^{[T\pi]} \left(m(z_t) - \frac{1}{S} \sum_{s=1}^{S} m(z_t^s(\beta_1, \delta), \beta_1, \delta) \right),$$

$$f_{2T}^{S}(\theta,\pi)=\frac{1}{T}\sum_{t=[T\pi]+1}^{T}\left(m(z_t)-\frac{1}{S}\sum_{s=1}^{S}m(z_t^s(\beta_2,\delta),\beta_2,\delta)\right)$$

and $\hat{W}_T(\pi)$ is a random positive definite symmetric $2q\times 2q$ matrix.

The partial-sample optimal weighting matrix is defined as the inverse of $(1+1/S)\Omega(\pi)$ where

$$\Omega(\pi)=\lim_{T\to\infty}\operatorname{Var}\left(\frac{1}{\sqrt{T}}\begin{bmatrix}\sum_{t=1}^{[T\pi]}(m(z_t)-Em(z_t))\\ \sum_{t=[T\pi]+1}^{T}(m(z_t)-Em(z_t))\end{bmatrix}\right),$$

which under null (2.2) is asymptotically equal to

$$\Omega(\pi)=\begin{bmatrix}\pi\Omega & 0\\ 0 & (1-\pi)\Omega\end{bmatrix}.$$

Duffie and Singleton (1993) mention two reasons why the GMM regularity conditions cannot be used to show the convergence of SMM estimators. First, the initial conditions of the time-series processes are in general not drawn from their stationary distribution which results in a local nonstationarity of the simulated process. Second, the first moment continuity assumption used by Hansen (1982) and Andrews (1987) is not sufficient to establish the uniform convergence of the sample criterion function to its population equivalent. Indeed, this continuity is not valid for data generated by simulations which depend on the unknown parameter vector. Duffie and Singleton (1993), and more recently Carrasco et al. (1999), rely on a notion of geometric ergodicity to establish strong laws of large numbers. In the context of nonlinear models, Gallant (1987), Gallant and White (1988), Andrews (1993) and many others adopt the notion of near epoch dependence (on a strong mixing base as defined more rigorously in Appendix A).[5] More recent papers involving simulation-based estimation have maintained the notion of near epoch dependence (see Gallant and Tauchen, 1996 for a discussion on this). Therefore, we deviate from Duffie and Singleton and assume that observed and simulated series are near epoch dependent, a condition also used by Andrews (1993) in the context of tests for structural change which can accommodate local nonstationarity. This is covered by Assumption A.2. Note that under the null of no structural change, geometric ergodicity as supposed by Duffie and Singleton (1993), implies near epoch dependence.

We also need to impose a number of regularity conditions which do not appear in Andrews (1993) or the more recent work on structural change tests for GMM

[5] Related forms of temporal dependence appear in the work of Billingsley (1968), McLeish (1975a,b), Bierens (1981).

estimators. In the simulated case, the moment conditions $m(z_t^s(\theta),\theta)$ depends on θ not only directly, but indirectly through the dependence of $z_t^s(\theta)$. To circumvent this difficulty, we need to impose a global Lipschitz condition on moment conditions and their total derivatives with respect to the parameter vector in order to obtain uniform convergence. The global Lipschitz condition was also used by Duffie and Singleton (1993), and results in modifications of Andrews (1993) to establish the asymptotic distribution of structural change tests. This condition appears in Assumptions A.3 and A.5 and is a sufficient condition for stochastic equicontinuity of a triangular array of a random vector.[6] Finally, we also impose standard identification assumptions and restrict the parameter space to be compact.

2.3. Asymptotic properties of partial sample SMM estimators

We present two theorems in this section which establish the large sample properties of the partial sample SMM estimators under the null hypothesis. The first theorem establishes consistency whereas the second characterizes the asymptotic distribution.

Theorem 1. *Under Assumptions* A.1–A.3, *the partial sample SMM estimators* $\hat{\theta}_T^S(\pi)$ *satisfies* $sup_{\pi\in\Pi}\|\hat{\theta}_T^S(\pi)-\theta_0\|\xrightarrow{p}0$ *for some* θ_0 *in the interior of* Θ.

Proof. See Appendix B. □

To characterize the asymptotic distribution we define the following matrices:

$$F^{\beta}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\partial m(z_t^s(\beta_0,\delta_0),\beta_0,\delta_0)/\partial\beta'\in R^{q\times r},$$

$$F^{\delta}=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\partial m(z_t^s(\beta_0,\delta_0),\beta_0,\delta_0)/\partial\delta'\in R^{q\times v},$$

$$F(\pi)=\begin{bmatrix}\pi F^{\beta} & 0 & \pi F^{\delta}\\ 0 & (1-\pi)F^{\beta} & (1-\pi)F^{\delta}\end{bmatrix}\in R^{2q\times(2r+v)}.$$

We denote $\{B(\pi)\colon \pi\in[0,1]\}$ as a q-dimensional vector of standard Brownian motion on $[0,1]$ and define

$$G(S,\pi)=\left(1+\frac{1}{S}\right)^{1/2}\begin{bmatrix}\Omega^{1/2}B(\pi)\\ \Omega^{1/2}(B(1)-B(\pi))\end{bmatrix},$$

where $B(\pi)$ is a q-dimensional vector of standard Brownian motion.

[6] As noted by Andrews (1992), stochastic equicontinuity (strong or in probability) is a necessary and sufficient condition to go from pointwise to uniform convergence (strong or in probability).

Theorem 2. *Under null* (2.2) *of no structural change and Assumptions* A.1–A.5, *for a fixed* S, *any sequence of partial sample SMM estimators* $\{\hat{\theta}_T^S(\cdot)\colon T \geqslant 1\}$ *satisfies*

$$\sqrt{T}(\hat{\theta}_T^S(\cdot) - \theta_0) \Rightarrow (F(\cdot)'W(\cdot)F(\cdot))^{-1}F(\cdot)'W(\cdot)G(S,\cdot)$$

as a process indexed by $\pi \in \Pi$, *provided* Π *has a closure in* $(0,1)$.

Proof. See Appendix B. □

It should be noted that the asymptotic distribution of the partial sample SMM estimators depends on a nuisance parameter, namely the number of simulations S and yields a process indexed by $\pi \in \Pi$, the latter being a nuisance parameter encountered in the context of GMM partial sample estimators as well. As $S \to \infty$, the asymptotic distribution coincides with the GMM partial sample estimators discussed in Andrews (1993). However, in practice, the number of simulations S is fixed and usually low. Fortunately, although the asymptotic distribution of the partial sample SMM estimators depends on S, we will see that the asymptotic distribution of the test statistics for structural change is independent of S under the null, and hence is nuisance parameter free once we deal with π.

3. Structural change test statistics

The purpose of this section is to refine the null hypothesis. Such a refinement will enable us to construct various tests for structural change in the spirit of Sowell (1993, 1996a) and Hall and Sen (1999). In a first subsection we refine the null hypothesis, a second section is devoted to tests for parameter constancy, a final section covers tests for stability of over-identifying restrictions.

3.1. Refining the null hypothesis

The moment conditions for the full sample under the null can be written as $Ek_t(\theta_0)=0$, $\forall t=1,\dots,T$. Following Sowell (1993, 1996a), we can project the moment conditions on the subspace identifying the parameters and the subspace of over-identifying restrictions. In particular, considering the (standardized) moment conditions for the full sample SMM estimator, such a decomposition corresponds to

$$\Omega^{-1/2}Ek_t(\theta_0) = P_F\Omega^{-1/2}Ek_t(\theta_0) + (I_q - P_F)\Omega^{-1/2}Ek_t(\theta_0),$$

where $P_F = \Omega^{-1/2}F[F'\Omega^{-1}F]^{-1}F'\Omega^{-1/2}$. The first term is the projection identifying the parameter vector and the second term is the projection for the over-identifying restrictions. The projection argument enables us to refine the null hypothesis (2.2). For instance, following Hall and Sen (1999) we can consider the null, for the case of a possible single breakpoint, which separates the identifying restrictions across the two subsamples:

$$\mathrm{H}_0^I(\pi) = \begin{cases} P_F \Omega^{-1/2} E[k_t(\theta_0)] = 0 & \forall t = 1, \ldots, [T\pi], \\ P_F \Omega^{-1/2} E[k_t(\theta_0)] = 0 & \forall t = [T\pi] + 1, \ldots, T. \end{cases}$$

Moreover, the overidentifying restrictions are stable if they hold before and after the possible breakpoint. This is formally stated as $\mathrm{H}_0^O(\pi) = \mathrm{H}_0^{O1}(\pi) \cap \mathrm{H}_0^{O2}(\pi)$ with

$$\mathrm{H}_0^{O1}(\pi) : (I_q - P_F)\Omega^{-1/2} E[k_t(\theta_0)] = 0 \quad \forall t = 1, \ldots, [T\pi],$$

$$\mathrm{H}_0^{O2}(\pi) : (I_q - P_F)\Omega^{-1/2} E[k_t(\theta_0)] = 0 \quad \forall t = [T\pi] + 1, \ldots, T.$$

The projection reveals that instability must be a result of a violation of at least one of the three hypotheses: $\mathrm{H}_0^I(\pi), \mathrm{H}_0^{O1}(\pi)$ or $\mathrm{H}_0^{O2}(\pi)$. Various tests can be constructed with local power properties against any particular one of these three null hypotheses (and typically no power against the others). To elaborate further on this we consider a sequence of local alternatives based on the moment conditions:

Assumption 3.1. A sequence of local alternatives is specified as

$$Ek_t(\theta_0) = h\left(\eta, \tau, \frac{t}{T}\right) \Big/ \sqrt{T}, \tag{3.1}$$

where $h(\eta, \tau, \pi)$, for $\pi \in [0, 1]$, is a q-dimensional function. The parameter τ locates structural changes as a fraction of the sample size and the vector η defines the local alternatives.[7] These local alternatives are chosen to show that the structural change tests presented in this paper have nontrivial power against a large class of alternatives. Also, our asymptotic results can be compared with Sowell's results for GMM framework. The following theorem, which extends Theorem 1 of Sowell (1996a), provides the asymptotic distribution of the moment condition for the partial-sample SMM under the general sequence of local alternatives appearing in Assumption 3.1.

Theorem 3. *Let Assumptions* A.1–A.5 *and Assumption* 3.1 *hold, then*

$$\sqrt{T} \bar{f}_T^S(\theta_0, \pi) \Rightarrow \begin{bmatrix} \left(1 + \dfrac{1}{S}\right)^{1/2} \Omega^{1/2} B(\pi) + H(\pi) \\ \left(1 + \dfrac{1}{S}\right)^{1/2} \Omega^{1/2} (B(1) - B(\pi)) + (H(1) - H(\pi)) \end{bmatrix},$$

where $H(\pi) = \int_0^\pi h(\eta, \tau, u)\,\mathrm{d}u$ and $B(\pi)$ is a q-dimensional vectors of standard Brownian motion and $\bar{f}_T^S(\theta_0, \pi)$ is defined in Definition 2.2.

[7] The function $h(\cdot)$ allows for a wide range of alternative hypotheses (see Sowell, 1996a). In its generic form it can be expressed as the uniform limit of step functions, $\eta \in R^i$, $\tau \in R^j$ such that $0 < \tau_1 < \tau_2 < \cdots < \tau_j < 1$ and θ^* is in the interior of Θ. Therefore, it can accommodate multiple breaks.

Proof. See Appendix B. □

3.2. Tests for parameter constancy

In this section, we introduce several tests for structural change for parameter stability and establish their asymptotic distribution. The null hypothesis is (2.2), or more precisely $H_0^I(\pi)$. We present Wald, Lagrange multiplier and likelihood ratio-type statistics. Predictive tests will be discussed in the next section. For brevity, we consider only test statistics based on the optimal weighting matrix. The results obtained are valid under an arbitrary weighting matrix. The first is the Wald statistic which is given by

$$\text{Wald}_T^S(\pi) = T(\hat{\beta}_{1T}^S(\pi) - \hat{\beta}_{2T}^S(\pi))'(\hat{V}_\Omega(\pi))^{-1}(\hat{\beta}_{1T}^S(\pi) - \hat{\beta}_{2T}^S(\pi)),$$

where $(\hat{V}_\Omega(\pi))=(\hat{V}_1(\pi)/\pi+\hat{V}_2(\pi)/(1-\pi))$ and $\hat{V}_j(\pi)=(1+1/S)(\hat{F}_j^\beta(\pi)'\hat{\Omega}_j^{-1}(\pi)\hat{F}_j^\beta(\pi))^{-1}$ for $j=1,2$. When $j=1$, the estimators of F^β and Ω are obtained with data from the first part of the sample $t=1,\dots,[T\pi]$ while for $j=2$, the estimators are obtained with data from the remainder of the sample $t=[T\pi]+1,\dots,T$. As discussed in Section 2, we proceed with simulations for each subsample, a process which may be prohibitively expensive in terms of computer time, a reason why the Wald-type test may not be the most appealing in this context.

The Lagrange multiplier statistic does not involve estimators obtained from subsamples, rather it relies on full sample parameter estimates. More precisely, the $\text{LM}_T^S(\pi)$ statistic is based on the first-order conditions for the partial sample SMM estimators evaluated at the full sample estimator:

$$\text{LM}_T^S(\pi) = c_T^S(\pi)'(\hat{V}_1(\pi)/\pi + \hat{V}_2(\pi)/(1-\pi))^{-1}c_T^S(\pi),$$

where

$$c_T^S(\pi) = L\begin{bmatrix} \frac{1}{\pi}\hat{G}_1(\pi) & 0 \\ 0 & \frac{1}{(1-\pi)}\hat{G}_2(\pi) \end{bmatrix}\sqrt{T}\bar{f}_T^S(\tilde{\theta}_T^S,\pi),$$

$L=[I_p,-I_p]$, and $\hat{G}_j(\pi)=((\hat{F}_j^\beta(\pi))'\hat{\Omega}_j^{-1}(\pi)\hat{F}_j^\beta(\pi))^{-1}(\hat{F}_j^\beta(\pi))'\hat{\Omega}_j^{-1}(\pi)$. Note that $\hat{F}_j^\beta(\pi)$ and $\hat{\Omega}_j^{-1}(\pi)$ are now matrices evaluated at the restricted full sample SMM estimator computed over the first and the second subsamples. Andrews (1993) shows that the $\text{LM}_T^S(\pi)$ simplifies to

$$\frac{T}{\pi(1-\pi)}\left(1+\frac{1}{S}\right)^{-1} f_{1T}^S(\tilde{\theta}_T^S)'\hat{\Omega}^{-1}\hat{F}^\beta[(\hat{F}^\beta)'\hat{\Omega}^{-1}\hat{F}^\beta]^{-1}(\hat{F}^\beta)'\hat{\Omega}^{-1}f_{1T}^S(\tilde{\theta}_T^S),$$

where

$$f_{1T}^S(\tilde{\theta}_T^S) = \frac{1}{T}\sum_{t=1}^{[T\pi]}\left(m(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} m(z_{Tt}^s(\tilde{\beta},\tilde{\delta}),\tilde{\beta},\tilde{\delta})\right).$$

Thus, the $\mathrm{LM}_T^S(\pi)$ corresponds to the projection of the moment conditions evaluated at the full sample estimator on the subspace identifying the parameter vector β.

The LR-type statistic is defined as the difference between the objective function for the partial sample SMM evaluated at the full sample estimator and at the partial sample estimators:

$$\mathrm{LR}_T^S(\pi) = T\left(1+\frac{1}{S}\right)^{-1}\left(\bar{f}_T^S(\tilde{\theta}_T^S,\pi)'\hat{\Omega}_T^{-1}(\pi)\bar{f}_T^S(\tilde{\theta}_T^S,\pi) - \bar{f}_T^S(\hat{\theta}_T^S(\pi),\pi)'\hat{\Omega}_T^{-1}(\pi)\bar{f}_T^S(\hat{\theta}_T^S(\pi),\pi)\right).$$

We state now the main theorem which establishes the asymptotic distribution of the Wald, LM and LR-type test statistics under the local alternative (3.1).

Theorem 4. *Under the null hypothesis* H_0 *in* (2.2) *and Assumptions* A.1–A.5, *the following processes indexed by* π *for a given set* Π *whose closure lies in* $(0,1)$ *satisfy*

$$\mathrm{Wald}_T^S(\pi) \Rightarrow Q_r(\pi), \quad \mathrm{LM}_T^S(\pi) \Rightarrow Q_r(\pi), \quad \mathrm{LR}_T^S(\pi) \Rightarrow Q_r(\pi)$$

with

$$Q_r(\pi) = \frac{\mathrm{BB}_r(\pi)'\mathrm{BB}_r(\pi)}{\pi(1-\pi)}$$

and under alternative (3.1)

$$Q_r(\pi) = \frac{\mathrm{BB}_r(\pi)'\mathrm{BB}_r(\pi)}{\pi(1-\pi)} + \left(1+\frac{1}{S}\right)^{-1}\frac{(H(\pi)-\pi H(1))'\Omega^{-1/2}P_{F^\beta}\Omega^{-1/2}(H(\pi)-\pi H(1))}{\pi(1-\pi)},$$

where $\mathrm{BB}_r(\pi) = B_r(\pi) - \pi B_r(1)$ *is a Brownian bridge,* B_r *is r-vector of independent Brownian motions and* $P_{F^\beta} = \Omega^{-1/2}F^\beta[(F^\beta)'\Omega^{-1}F^\beta]^{-1}(F^\beta)'\Omega^{-1/2}$.

Proof. See Appendix B. □

The result in Theorem 4 tells us that the asymptotic distributions under the null of the Wald, LR-type and LM statistics are the same as those obtained by Andrews (1993) for the GMM estimator. Note that this implies that the asymptotic distribution of these statistics does not depend on S. Under the alternative, the asymptotic power depends on the number of simulations. A small number of simulation paths implies less asymptotic power reflected by the term $(1+1/S)^{-1}$.

When π is unknown, i.e. the case of unknown breakpoint, we can use the above result to construct statistics across $\pi \in \Pi$. In the context of maximum likelihood estimation, Andrews and Ploberger (1994) derive asymptotic optimal tests for a Gaussian a priori of the amplitude of the structural change based on the Neyman–Pearson approach which are characterized by an average exponential form. The Sowell (1996a) optimal tests are a generalization of the Andrews and Ploberger approach to the case of two measures that do not admit densities. The most powerful test is given by the Radon–Nikodym derivative of the probability measure implied by the local alternative with respect to the probability measure implied by the null hypothesis.

The optimal average exponential form is the following:

$$\mathrm{Exp}-Q_T^S=(1+c)^{-r/2}\int \exp\left(\frac{1}{2}\,\frac{c}{1+c}Q_T^S(\pi)\right)\,\mathrm{d}J(\pi),$$

where various choices of c determine power against close or more distant alternatives and $J(\cdot)$ is the weight function over the value of $\pi\in\Pi$. In the case of close alternatives $(c=0)$, the optimal test statistic takes the average form

$$\mathrm{ave}\,Q_T^S=\int_\Pi Q_T^S(\pi)\,\mathrm{d}\,J(\pi).$$

For a distant alternative $(c=\infty)$, the optimal test statistics takes the exponential form

$$\exp Q_T^S=\log\left(\int_\Pi \exp\left[\frac{1}{2}Q_T^S(\pi)\right]\,\mathrm{d}\,J(\pi)\right).$$

The supremum form often used in the literature corresponds to the case where $c/(1+c)\to\infty$. The sup test is

$$\sup Q_T^S=\sup_{\pi\in\Pi} Q_T^S(\pi).$$

The following theorem gives the asymptotic distribution for the exponential mapping for Q_T^S when Q_T^S corresponds to the Wald, LM and LR ratio-type tests under the null.

Theorem 5. *Under the null hypothesis* H_0 *in* (2.2) *and Assumptions* A.1–A.5, *the following processes indexed by* π *for a given set* Π *whose closure lies in* (0,1) *satisfy*

$$\sup Q_T^S\Rightarrow\sup_{\pi\in\Pi} Q_r(\pi),\qquad \mathrm{ave}Q_T^S(\pi)\Rightarrow\int_\Pi Q_r(\pi)\,\mathrm{d}J(\pi),$$

$$\exp Q_T^S\Rightarrow\log\left(\int_\Pi \exp[\tfrac{1}{2}Q_r(\pi)]\,\mathrm{d}J(\pi)\right)$$

with

$$Q_r(\pi)=\frac{\mathrm{BB}_r(\pi)'\mathrm{BB}_r(\pi)}{\pi(1-\pi)}.$$

This result is obtained through the application of the continuous mapping theorem (see Pollard, 1984).

This implies that we can rely on the critical values tabulated for the case of GMM-based tests in the context of SMM estimators. For example the critical values for the statistics defined by the supremum over all breakpoints $\pi \in \Pi$ of $\mathrm{Wald}_T^S(\pi)$, $\mathrm{LM}_T^S(\pi)$ or $\mathrm{LR}_T^S(\pi)$ can be found in the original paper by Andrews (1993). The same is true for the Sowell (1996a) and Andrews and Ploberger (1994) asymptotic optimal tests.

3.3. Tests for stability of over-identifying restrictions

Tests presented in the preceeding section are based on the projection of the moment conditions on the subspace of identifying restrictions. In this section, we are interested by test against violations of $\mathrm{H}_0^{O1}(\pi)$ or $\mathrm{H}_0^{O2}(\pi)$. The local alternatives are given by the projection of the moment condition on the subspace orthogonal to the identifying restrictions. For instance, in the case of a single breakpoint, the local alternatives by Assumption 3.1 correspond to: [8]

$$\mathrm{H}_A^{O1}(\pi) : (I_q - P_F)\Omega^{-1/2}E[k_t(\theta_0)] = (I_q - P_F)\Omega^{-1/2}\frac{\eta_1}{\sqrt{T}} \qquad t = 1,\dots,[T\pi],$$

$$\mathrm{H}_A^{O2}(\pi) : (I_q - P_F)\Omega^{-1/2}E[k_t(\theta_0)] = (I_q - P_F)\Omega^{-1/2}\frac{\eta_2}{\sqrt{T}} \qquad t = [T\pi]+1,\dots,T.$$

Sowell (1996b) introduces optimal tests for the violation of the over-identifying restrictions when the violation occurs before the breakpoint corresponding to the alternative H_A^{O1}. The statistic is based on the projection of the partial sum of the full sample estimator on the appropriated subspace. Hall and Sen (1999) introduce optimal test for the case where the violation occurs after the breakpoint i.e. H_A^{O1}. The statistic is based on the projection of the backward partial sum of the full sample estimator on the appropriated subspace.

The next theorem gives the asymptotic distribution of the partial sum of the full sample estimator for the general local alternatives (Assumption 3.1). With this result, we can obtain the asymptotic distribution for optimal tests for the stability of over-identifying restrictions. The asymptotic distributions are the same than derived by Sowell and Hall and Sen under the null but depend on the number of simulated paths under the alternative.

Theorem 6. *Under Assumptions* A.1–A.5 *and Assumption* 3.1, *then*

$$\sqrt{T} f_{1T}^S(\tilde{\theta}_T^S) \Rightarrow \left(1+\frac{1}{S}\right)^{1/2}\Omega^{1/2}B(\pi) + H(\pi)$$
$$-\pi F(F'WF)^{-1}F'W\left[\left(1+\frac{1}{S}\right)^{1/2}\Omega^{1/2}B(1) + H(1)\right],$$

where $B(\pi)$ is a q-dimensional vectors of mutually independent Brownian motion and $f_{1T}^S(\cdot)$ and Ω are defined in Section 2.

[8] In the notation of Assumption 3.1 this case corresponds to $\tau = \pi$.

The statistic for the violation of over-identifying restriction before the breakpoint is

$$\mathrm{Sow}_T^S(\tilde{\theta}_T^S,\pi)=f_{1T}^S(\tilde{\theta}_T^S,\pi),$$

where $f_{1T}^S(\cdot)$ is defined in Definition 2.2 and $\tilde{\theta}_T^S$ is the full sample estimator. The statistic for the violation of over-identification after the breakpoint is

$$HS_{t,T}^S(\tilde{\theta}_T^S,\pi)=f_{2T}^S(\tilde{\theta}_T^S,\pi),$$

where $f_{2T}^S(\cdot)$ is defined in Definition 2.2. The statistic test for the stability of over-identifying restrictions consists in the projection of $\Omega^{-1/2}\mathrm{Sow}_{t,T}^S$ and $\Omega^{-1/2}HS_{t,T}^S$ on the subspace orthogonal to the subspace identifying the parameter vector: (I_q-P_F). Thus, these statistic tests are

$$T\left(1+\frac{1}{S}\right)^{-1}\mathrm{Sow}_T^S(\tilde{\theta}_T^S,\pi)'\Omega^{-1/2}(I-P_F)\Omega^{-1/2}\mathrm{Sow}_T^S(\tilde{\theta}_T^S,\pi), \tag{3.2}$$

$$T\left(1+\frac{1}{S}\right)^{-1}HS_{t,T}^S(\tilde{\theta}_T^S,\pi)'\Omega^{-1/2}(I-P_F)\Omega^{-1/2}HS_{t,T}^S(\tilde{\theta}_T^S,\pi). \tag{3.3}$$

As we show above, these statistics have to be multiplied by the scalar $(1+1/S)^{-1}$ to obtained an asymptotic distribution free of the nuisance parameter S under the null.

In the case of close alternative, the optimal test takes the average form

$$\mathrm{ave}\, Q_T^{S,O}=\int_\Pi Q_{t,T}^{S,O}\,\mathrm{d}J(\pi),$$

where $Q_{t,T}^{S,O}$ corresponds to Eqs. (3.2) and (3.3). For a distant alternative, we have the following exponential form:

$$\exp Q_T^{S,O}=\log\left(\int_\Pi \exp\left[\frac{1}{2}Q_{t,T}^{S,O}\right]\,\mathrm{d}J(\pi)\right).$$

The supremum form is

$$\sup Q_T^{S,O}=\sup_{\pi\in\Pi} Q_{t,T}^{S,O}(\pi).$$

The following theorem provides the asymptotic distribution of those optimal test statistics.

Theorem 7. *Under the null of no structural change and Assumptions* A.1–A.5, *the following processes indexed by* π *for a given set* Π *whose closure lies in* (0,1) *satisfy*

$$\sup Q_T^{S,O}\Rightarrow \sup_{\pi\in\Pi} Q_{q-p}(\pi),\qquad \mathrm{ave}\, Q_T^{S,O}\Rightarrow\int_\Pi Q_{q-p}(\pi)\,\mathrm{d}J(\pi),$$

$$\exp Q_T^{S,O}\Rightarrow\log\left(\int_\Pi \exp\left[\frac{1}{2}Q_{q-p}(\pi)\right]\,\mathrm{d}J(\pi)\right)$$

with

$$Q_{q-p}(\pi) = B_{q-p}(\pi)' B_{q-p}(\pi)$$

and under alternative (3.1)

$$Q_{q-p}(\pi) = B_{q-p}(\pi)' B_{q-p}(\pi) + \left(1 + \frac{1}{S}\right)^{-1} (H(1) - H(\pi))' \Omega^{-1/2} (I - P_F)$$
$$\times \Omega^{-1/2} (H(1) - H(\pi)),$$

where $B_{q-p}(\pi)$ *is a* $q-p$*-dimensional vector of independent Brownian motion and* $P_F = \Omega^{-1/2} F (F' \Omega^{-1} F)^{-1} F' \Omega^{-1/2}$.

Proof. See Appendix B. □

As with previous cases, the asymptotic distributions of optimal tests again do not depend on the number of simulation paths under the null but the asymptotic power is affected by S.

To conclude we should note that the sequence of local alternatives in (3.1) is expressed in terms of violations of moment conditions instead of parameters as in (2.2). This brings us to the subject of predictive tests for structural change considered by Ghysels et al. (1997) for the case of GMM estimators. They consider a single breakpoint, which amounts to the following null:

$$\mathrm{H}_0 : Ek_t(\theta_0) = 0 \quad \forall t = 1, \ldots, T$$

and alternative

$$Ek_t(\theta_0) = \begin{cases} 0 & \forall t = 1, \ldots, [\pi T], \\ T^{-\frac{1}{2}} \mu_2 & \forall t = [\pi T] + 1, \ldots, T \end{cases}$$

with $\mu_2 \neq 0$. The predictive test is based on evaluating the sample moment conditions for the subsample $t = [\pi T] + 1, \ldots, T$ using $\hat{\theta}^S_{1T}(\pi)$, i.e. the parameter estimates from the first subsample. The test statistic is defined as

$$\mathrm{Pred}_T(S, \pi) = (T - [T\pi]) f^S_{2T}(\hat{\theta}^S_{1T}(\pi))' \hat{V}^{-1}_{PR} f^S_{2T}(\hat{\theta}^S_{1T}(\pi)),$$

where $\hat{V}_{\mathrm{PR}}$ is a covariance matrix defined in Ghysels and Hall (1990) and

$$f^S_{2T}(\hat{\theta}^S_{1T}(\pi)) = \frac{1}{T - [T\pi]} \sum_{t=[T\pi]+1}^{T} \left(m(z_{Tt}) - \frac{1}{S} \sum_{s=1}^{S} m(z^s_{Tt}(\hat{\theta}^S_{1T}(\pi)), \hat{\theta}^S_{1T}(\pi)) \right).$$

One can apply again arguments similar to those in Theorem 7 to show that the critical values appearing in Ghysels et al. (1999, corrigendum of 1997) apply.

4. Finite sample properties

The results in Sections 3 and 4 imply that the asymptotic distribution of simulation-based tests for structural change under the null as well as under a sequence of local alternatives is independent of S, the number of simulations. Yet, one might expect that in finite samples both the size and power are affected by the number of simulations. We conduct a Monte-Carlo study to appraise the extend to which the finite sample size and power depend on S. The setup we consider is the following AR(1) model:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t, \tag{4.1}$$

where ε_t is i.i.d. N(0, 1). We consider both GMM and SMM estimators. Comparing the two will allow us to appraise the impact of simulation uncertainty in finite samples. The GMM estimator of the parameter vector $\theta \equiv (\alpha_0, \alpha_1, \sigma_\varepsilon)$ is based on the following set of moment conditions:

$$\alpha_0/(1-\alpha_1) - 1/T\sum_{\tau=1}^{T} y_\tau,$$

$$\sigma_\varepsilon^2/(1-\alpha_1^2) - 1/T\sum_{\tau=1}^{T}(y_\tau - \bar{y})^2,$$

$$(\alpha_1\sigma_\varepsilon^2)/(1-\alpha_1^2) - 1/T\sum_{\tau=2}^{T}(y_\tau - \bar{y})(y_{\tau-1} - \bar{y}),$$

whereas the SMM estimator is based on the same moment conditions, though computed via simulations, that is

$$1/T\sum_{\tau=1}^{T}\left(y_\tau - 1/S\sum_{s=1}^{S} y_\tau^s(\theta)\right),$$

$$1/T\sum_{\tau=1}^{T}\left((y_\tau - \bar{y})^2 - 1/S\sum_{s=1}^{S}(y_\tau^s(\theta) - \bar{y}^{\,s}(\theta))^2\right),$$

$$1/T\sum_{\tau=2}^{T}\left((y_\tau - \bar{y})(y_{\tau-1} - \bar{y}) - 1/S\sum_{s=1}^{S}(y_\tau^s(\theta) - \bar{y}^{\,s}(\theta))(y_{\tau-1}^s(\theta) - \bar{y}^{\,s}(\theta))\right).$$

Under the null hypothesis we consider the following parameter values: $\alpha_0 = 0$ and 1, $\alpha_1 = 0.5$ and 0.9. Pairwise combinations yields four process specifications. Under the alternative we consider either a shift in the intercept or in the slope. For the former we take

$$\alpha_{0t} = \begin{cases} 0 & t = 1, \ldots, [\pi T], \\ 1 & t = [\pi T] + 1, \ldots, T. \end{cases} \tag{4.2}$$

For the breaks in the slope we take as alternative

$$\alpha_{1t} = \begin{cases} 0.5 & t = 1, \dots, [\pi T], \\ 0.9 & t = [\pi T] + 1, \dots, T. \end{cases} \tag{4.3}$$

The sample sizes are $T = 100$ and 200. Under the alternative breaks occur at $\pi = 0.25, 0.50$ and 0.75. The AveLM, ExpLM and SupLM statistics are considered with $S = 1, 2$ and 10. We examine LM-type tests as they are the most appealing from a computational point of view, only involving full sample estimators. The three variations on the LM-type statistics allow us to explore the effect of local asymptotic power schemes on finite sample power properties. The Monte-Carlo study is based on 1000 replications. Table 1 reports the size of the various tests, Tables 2 and 3 cover power properties for intercept and slope breaks. The nominal level of the tests is 5%.

Table 1, as well as the other tables, reports the results for AveLM tests in the top panel, ExpLM in the middle panel and SupLM in the lower panel. Each panel contains three SMM estimators with $S = 1, 2$ and 10 and the GMM estimator involving the same moment conditions appear for the four parameter combinations considered and the two sample sizes. Regardless of the test, sample size or parameter settings, we note from Table 1 that there are serious size distortions for all the SMM estimators involving only one simulation, i.e. $S = 1$. It takes only a second simulation, i.e. $S = 2$ to completely eliminate any size distortion due to simulation error, that is the size of the tests are roughly the same for the GMM and the SMM estimator-based tests with $S = 2$. Considering a larger number of simulations, i.e. increase S to 10, adds very little improvement. These results tell us that simulation uncertainty in the computation of moments matters, but it is fairly straightforward to accommodate as only a very small number of simulations is required.

In Table 2 we examine the power of the tests when a break occurs in the intercept of the AR(1) model, with breaks occuring at fractions $\pi = 0.25, 0.50$ and 0.75 of samples of size $T = 100$ and 200. Tests involving $S = 1$ should be largely ignored as they suffer from serious size distortions. We focus therefore our analysis on the SMM estimators with $S = 2$ and 10. Typically, the SMM-based tests are less powerful than their GMM-based counterpart. In some cases the SMM-based tests appear more powerful, which is to be attributed to Monte-Carlo simulation error. The differences between tests with two or ten simulations are neglible, similar to the findings reported in Table 1 regarding size properties. The loss of power compared to the GMM estimators does not seem to be affected by the persistence in the series (i.e. the value of α_1) nor by the location of the break (i.e. the value of π). Moreover, the SMM-based tests inherit the power profile of GMM-based tests, as is expected. In particular, as the sample size increases so does the power and the power is higher for mid-sample breaks and roughly equivalent (i.e. symmetric) for $\pi = 0.25$ and $\pi = 0.75$. when S increases. Symmetry of power properties with regard to the location of the break was also reported by Ghysels et al. (1997) for GMM-based tests. For breaks in the slope, reported in Table 3, many of the aforementioned conclusions hold as well, except that the power of the GMM and SMM tests is no longer symmetric around a mid-sample break, but instead declines as π increases. Finally, the expLM tests are usually the most powerful ones,

Table 1
Size properties of SMM and GMM-based LM-type tests

T	α_0	α_1	SMM		GMM	
			$S=1$	$S=2$	$S=10$	$S=\infty$
AveLM tests						
100	0	0.5	0.411	0.027	0.033	0.027
	0	0.9	0.587	0.167	0.152	0.094
	1	0.5	0.398	0.041	0.028	0.019
	1	0.9	0.592	0.150	0.140	0.115
200	0	0.5	0.365	0.035	0.035	0.036
	0	0.9	0.510	0.087	0.082	0.090
	1	0.5	0.334	0.041	0.049	0.033
	1	0.9	0.515	0.098	0.100	0.087
ExpLM tests						
100	0	0.5	0.521	0.034	0.037	0.024
	0	0.9	0.714	0.167	0.251	0.163
	1	0.5	0.517	0.040	0.029	0.024
	1	0.9	0.721	0.240	0.237	0.192
200	0	0.5	0.453	0.033	0.035	0.037
	0	0.9	0.645	0.138	0.128	0.116
	1	0.5	0.427	0.033	0.040	0.032
	1	0.9	0.649	0.131	0.154	0.131
ExpLM tests						
100	0	0.5	0.494	0.022	0.031	0.013
	0	0.9	0.707	0.233	0.252	0.170
	1	0.5	0.493	0.028	0.034	0.018
	1	0.9	0.718	0.232	0.246	0.191
200	0	0.5	0.451	0.019	0.041	0.030
	0	0.9	0.663	0.151	0.151	0.129
	1	0.5	0.421	0.025	0.040	0.022
	1	0.9	0.651	0.150	0.166	0.143

Notes: Entries to the Table are Monte-Carlo simulations (1000 replications) of p-values for 5% nominal tests of the LM-type for the model appearing in (4.1).

not surprisingly as they are designed to have power again "distant" alternatives. The differences in power across the three types of tests are usually small, however.

5. Conclusion

In this paper, we examined tests for structural change in the context of simulated method of moments estimators. We found that the asymptotic distribution for such tests coincide with their GMM counterpart under the null but depends on the number

Table 2
Power properties of SMM- and GMM-based LM-type tests

T	α_0^1	α_0^2	α_1	π	SMM		GMM	
					$S=1$	$S=2$	$S=10$	$S=\infty$
AveLM tests								
100	0	1	0.5	0.25	0.404	0.158	0.159	0.279
				0.50	0.651	0.470	0.631	0.845
				0.75	0.723	0.423	0.639	0.332
100	0	1	0.9	0.25	0.652	0.233	0.193	0.696
				0.50	0.696	0.701	0.527	0.715
				0.75	0.844	0.854	0.606	0.681
200	0	1	0.5	0.25	0.568	0.386	0.406	0.686
				0.50	0.861	0.862	0.969	0.998
				0.75	0.909	0.813	0.964	0.739
200	0	1	0.9	0.25	0.662	0.204	0.182	0.683
				0.50	0.699	0.429	0.439	0.567
				0.75	0.882	0.554	0.700	0.557
ExpLM tests								
100	0	1	0.5	0.25	0.484	0.145	0.127	0.228
				0.50	0.675	0.408	0.562	0.782
				0.75	0.798	0.493	0.723	0.304
100	0	1	0.9	0.25	0.719	0.292	0.240	0.740
				0.50	0.766	0.433	0.608	0.769
				0.75	0.890	0.637	0.660	0.719
200	0	1	0.5	0.25	0.615	0.369	0.379	0.729
				0.50	0.876	0.831	0.947	0.990
				0.75	0.949	0.883	0.990	0.789
200	0	1	0.9	0.25	0.735	0.277	0.254	0.744
				0.50	0.659	0.368	0.497	0.621
				0.75	0.872	0.481	0.770	0.599
SupLM tests								
100	0	1	0.5	0.25	0.435	0.082	0.067	0.110
				0.50	0.633	0.260	0.409	0.770
				0.75	0.783	0.400	0.642	0.317
100	0	1	0.9	0.25	0.717	0.290	0.230	0.772
				0.50	0.757	0.442	0.607	0.817
				0.75	0.864	0.622	0.629	0.750
200	0	1	0.5	0.25	0.598	0.262	0.252	0.573
				0.50	0.839	0.675	0.850	0.992
				0.75	0.945	0.833	0.972	0.644
200	0	1	0.9	0.25	0.738	0.296	0.255	0.724
				0.50	0.756	0.483	0.476	0.601
				0.75	0.921	0.603	0.761	0.582

Notes: Entries to the Table are Monte-Carlo simulations (1000 replications) of p-values for 5% nominal tests of the LM-type for the model appearing in (4.1) under alternative (4.2).

Table 3
Power properties of SMM- and GMM-based LM-type tests

T	α_0	α_1^1	α_1^2	π	SMM		GMM	
					$S=1$	$S=2$	$S=10$	$S=\infty$
AveLM tests								
100	0	0.5	0.9	0.25	0.422	0.202	0.231	0.174
				0.50	0.470	0.208	0.247	0.190
				0.75	0.428	0.148	0.144	0.096
100	1	0.5	0.9	0.25	0.801	0.394	0.379	0.872
				0.50	0.797	0.559	0.570	0.670
				0.75	0.823	0.505	0.460	0.389
200	0	0.5	0.9	0.25	0.544	0.388	0.481	0.419
				0.50	0.578	0.393	0.503	0.494
				0.75	0.464	0.188	0.219	0.193
200	1	0.5	0.9	0.25	0.875	0.459	0.572	0.971
				0.50	0.856	0.619	0.709	0.472
				0.75	0.878	0.578	0.748	0.147
ExpLM tests								
100	0	0.5	0.9	0.25	0.357	0.240	0.270	0.213
				0.50	0.552	0.233	0.305	0.200
				0.75	0.563	0.211	0.221	0.123
100	1	0.5	0.9	0.25	0.860	0.442	0.399	0.911
				0.50	0.853	0.653	0.713	0.801
				0.75	0.878	0.573	0.550	0.461
200	0	0.5	0.9	0.25	0.620	0.425	0.547	0.475
				0.50	0.631	0.399	0.503	0.489
				0.75	0.530	0.232	0.249	0.211
200	1	0.5	0.9	0.25	0.913	0.532	0.596	0.992
				0.50	0.907	0.685	0.790	0.724
				0.75	0.912	0.677	0.818	0.206
SupLM tests								
100	0	0.5	0.9	0.25	0.525	0.241	0.280	0.204
				0.50	0.552	0.216	0.289	0.175
				0.75	0.567	0.211	0.221	0.131
100	1	0.5	0.9	0.25	0.849	0.401	0.300	0.867
				0.50	0.859	0.621	0.690	0.799
				0.75	0.863	0.568	0.550	0.459
200	0	0.5	0.9	0.25	0.641	0.431	0.534	0.475
				0.50	0.621	0.352	0.438	0.401
				0.75	0.571	0.247	0.265	0.203
200	1	0.5	0.9	0.25	0.920	0.530	0.517	0.989
				0.50	0.917	0.687	0.792	0.778
				0.75	0.919	0.669	0.782	0.214

Notes: Entries to the Table are Monte-Carlo simulations (1000 replications) of p-values for 5% nominal tests of the LM-type for the model appearing in (4.1) under alternative (4.3).

of simulated paths under the alternative. Obviously, there are limitations to our results as well as unresolved challenges. Regarding the limitations we should mention the finite sample performance of the tests. There is a fair amount of Monte-Carlo evidence regarding the finite sample performance of GMM tests for structural change. In finite samples the simulation uncertainty is one more factor that may deteriorate the finite sample performance. As noted in Section 4, the simulation uncertainty matters in finite samples, yet it is easy to account for with only a small number of simulations. It appears from our Monte-Carlo study that $S=2$ suffices, i.e. tremendous improvements occur by simply adding a second simulation.

In general, power properties of GMM and SMM-based tests deteriorate against structural change at the very beginning or the very end of a sample. Dufour et al. (1994) proposed tests, with null and alternative similar to predictive tests, designed to handle situations where one sample is large (the estimation sample) and a second sample is small (the prediction sample) even containing only one observation. Unfortunately, the results obtained for SMM-based tests we discussed in this paper do not extend to the setup considered by Dufour et al. Their tests will depend on the nuisance parameter S, though one could let $S \to \infty$, and obtain the same results as in Dufour et al. Since the second sample is small generating a large number of simulations would be particularly difficult.

The next generation of estimators are simulation-based procedures involving two models, an auxiliary model and a model of interest. Such procedures, discussed by Gouriéroux et al. (1993) and Gallant and Tauchen (1996) add some nontrivial complications regarding testing since they involve parameters of two models. We leave testing for structural change in such settings for further research.

6. For further reading

The following references may also be of interest to the reader: Andrews, 1994 and Chow, 1960.

Acknowledgements

We benefited from the suggestions of an Associate Editor and two Referees whose comments helped us to greatly improve our paper. We would also like to thank Herman Bierens, Ron Gallant and Alastair Hall for comments on an earlier draft of the paper as well as participants at the North American Econometric Society Meetings in Montreal, June 1998. We would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada through a grant to NCM2 (Network for Computing and Mathematical Modeling). The second author would like to thank CREST (Paris) for their outstanding hospitality where parts of this work was done.

Appendix A. Detailed description of regularity conditions

Assumption A.1. The parameter space Θ is a compact subset of R^p and Θ_0 is a dense subset of Θ.

Assumption A.2. For all $\theta \in \Theta$, $\{z^s_{Tt}(\theta): t \leqslant T, T \geqslant 1\}$ is a triangular array of Z-values random vectors that is L^0-near epoch dependent on a strong mixing base $\{V_{Tt}: t = \ldots, 0, 1, \ldots; T \geqslant 1\}$, where Z is a Borel subset of R^k.[9]

Assumption A.3. The set of moment conditions satisfies:

- For some $r > 2$ and for all $\theta \in \Theta$, $\{m(z^s_{Tt}(\theta), \theta): t \leqslant T, T \geqslant 1\}$ is a triangular array of R^q-valued random vector that is L^2-near epoch dependent of size—$\frac{1}{2}$ on a strong mixing base $\{V_{Tt}: t = \ldots, 0, 1, \ldots; T \geqslant 1\}$ of size $-r/(r-2)$, $\sup_{t \leqslant T, T \geqslant 1} E\|m(z^s_{Tt}(\theta), \theta)\|^r < \infty$.
- $\overline{\lim}_{T\to\infty}(1/T)\sum_{t=1}^{T} E \sup_{\theta\in\Theta}|m(z^s_{Tt}(\theta), \theta)|^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$.
- For all $\theta, \theta' \in \Theta$, there is a sequence $\{B_{Tt}: t \leqslant T, T \geqslant 1\}$ not depending on θ with $(1/T)\sum_{t=1}^{T} EB_{Tt} = O_p(1)$ such that $\|m(z^s_{Tt}(\theta), \theta) - m(z^s_{Tt}(\theta'), \theta')\| \leqslant B_{Tt}\|\theta - \theta'\|$.
- $\sup_{\pi\in\Pi}\|\hat{W}_T(\pi) - W(\pi)\| \xrightarrow{p} 0$ for $2q \times 2q$ matrices $W(\pi)$ for which $\sup_{\pi\in\Pi}\|W(\pi)\| < \infty$ (including the optimal one).
- $\lim_{T\to\infty}(1/T)\sum_{t=1}^{[T\pi]} Em(z^s_{Tt}(\beta, \delta), \beta, \delta)$ exists uniformly over $(\beta, \delta, \pi) \in B \times \Delta \times \Pi$ and equals $\pi \lim_{T\to\infty}(1/T)\sum_{t=1}^{T} Em(z^s_{Tt}(\beta, \delta), \beta, \delta)$.
- $\mathrm{Var}(\frac{1}{\sqrt{T}}\sum_{t=1}^{T\pi}[(m(z_{Tt}) - Em(z^s_{Tt}))]) \xrightarrow{p} \pi\Omega$, $\forall \pi \in [0, 1]$ for positive $q \times q$ matrix Ω.
- $\tilde{f}^S(\beta_0, \delta_0) = 0$, where

$$\tilde{f}^S(\beta, \delta) = \tilde{m} - \tfrac{1}{S}\sum_{s=1}^{S}\tilde{m}^s(\beta, \delta)$$

$$= \lim_{T\to\infty}(1/T)\sum_{t=1}^{T}(m(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} m(z^s_{Tt}(\beta, \delta), \beta, \delta))$$

  and for every neighborhood $\Theta_0(\subset \Theta)$ of θ_0, $\inf_{\pi\in\Pi}\inf_{\theta\in\Theta/\Theta_0} f^S(\theta, \pi)' W(\pi) f^S(\theta, \pi) > 0$, where $f^S(\theta, \pi) = (\pi \tilde{f}^S(\beta_1, \delta))', (1 - \pi)(\tilde{f}^S(\beta_2, \delta))')'$.

Assumption A.4. $F(\pi)'W(\pi)F(\pi)$ is nonsingular $\forall \pi \in \Pi$ and has eigenvalues bounded away from zero.

We define the total derivative as

$$D_{\theta'} m(z^s_{Tt}(\theta), \theta) = \frac{\mathrm{d}}{\mathrm{d}\theta'} m(z^s_{Tt}(\theta), \theta). \tag{A.1}$$

Assumption A.5. The total derivative satisfies:

- For some $r > 2$ and for all $\theta \in \Theta$, $\{D_{\theta'} m(z^s_{Tt}(\theta), \theta): t \leqslant T, T \geqslant 1\}$ is a triangular array of R^q-valued random vector that is L^2-near epoch dependent of size $-1/2$ on a

[9] For a definition of L^p-near epoch dependence see Andrews (1993, p. 830), Davidson (1994) or Gallant and White (1988).

strong mixing base $\{V_{Tt}\colon t=\ldots,0,1,\ldots;T\geqslant 1\}$ of size $-r/(r-2)$ and $\sup_{t\leqslant T,T\geqslant 1}E\|D_{\theta'}m(z^s_{Tt}(\theta),\theta)\|^r<\infty$.

- $m(z^s_{Tt}(\theta),\theta)$ is differentiable in $(\beta,\delta)\in B_0\times\Delta_0$ $\forall z\in Z$, where B_0 and Δ_0 are some neighborhood of β_0 and δ_0.
- For all $\theta,\theta'\in\Theta$, there is a sequence $\{C_{Tt}\colon t\leqslant T,T\geqslant 1\}$ not depending on θ with $\frac{1}{T}\sum_{t=1}^{T}EC_{Tt}=\mathrm{O_p}(1)$ s.t. $\|D_{\theta'}m(z^s_{Tt}(\theta),\theta)-D_{\theta'}m(z^s_{Tt}(\theta),\theta')\|\leqslant C_{Tt}\|\theta-\theta'\|$ and $\sup_{t\leqslant T,T\geqslant 1}E\sup_{\theta\in\Theta_0}\|D_{\theta'}m(z^s_{Tt}(\theta),\theta)\|^{1+\varepsilon}<\infty$ for some $\varepsilon>0$.
- $\lim_{T\to\infty}(1/T)\sum_{t=1}^{[T\pi]}((1/S)\sum_{s=1}^{S}ED_{\theta'}m(z^s_{Tt}(\theta),\theta))$ exists uniformly over $\pi\in\Pi$ equals πF, $\forall\pi\in\Pi$ where

$$F=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\partial m(z^s_t(\theta_0),\theta_0)/\partial\theta'\in R^{q\times p}.$$

Appendix B. Proofs of theorems

To proof Theorems 1 and 2 we establish first a lemma for a generic function $Q(.)$. Different applications of the lemma will involve different specific functions.

Lemma A.1. *Suppose* (*a*) *Assumptions* A.1 *holds*, (*b*) *Assumptions* A.2 *holds*, (*c*) *for some* $r>2$, $Q(z^s_{Tt}(\theta),\theta)$ *is a triangular array of* R^q*-valued random vectors that is* L^2*-near epoch dependent of size* $-1/2$ *on a strong mixing base* $\{V_{Tt}\colon t=\ldots,0,1,\ldots;T\geqslant 1\}$ *of size* $-r/(r-2)$ *and* $\sup_{t\leqslant T,T\geqslant 1}E\|Q(z^s_{Tt}(\theta),\theta)\|^r<\infty$, (*d*) *the q-vector* $Q(z^s_{Tt}(\theta),\theta)$ *follows a global Lipschitz condition in* θ *and* (*e*) $\limsup_{T\to\infty}(1/T)\sum_{t=1}^{T}E\sup_{\theta\in\Theta}|Q(z^s_{Tt}(\theta),\theta)|^{1+\varepsilon}<\infty$ *for some* $\varepsilon>0$. *Then*

$$\sup_{\theta\in\Theta}\sup_{R\leqslant T}\left|\frac{1}{T}\sum_{t=1}^{R}[Q((z^s_{Tt}(\theta),\theta)-EQ((z^s_{Tt}(\theta),\theta)]\right|\xrightarrow{p}0.$$

Proof. We define $G_{Tt}(\theta)=\sup_{R\leqslant T}|(1/T)\sum_{t=1}^{R}(Q(z^s_{Tt}(\theta),\theta)-EQ(z^s_{Tt}(\theta),\theta))|$. By Theorem 21.9 of Davidson (1994), $\sup_{\theta\in\Theta}G_{Tt}(\theta)\xrightarrow{p}0$ if and only if (i) the pointwise convergence of $G_{Tt}(\theta)$ for $\theta\in\Theta_0$, where Θ_0 is a dense subset of Θ and (ii) $\{G_{Tt}\}$ is stochastically equicontinuous.[10] For (i), by Assumption (a), Θ_0 is a dense subset of Θ. By Assumption (c), $Q(z^s_{Tt}(\cdot),\cdot)$ is L^2-near epoch dependent, which implies that this process is L^2-approximable.[11] An L^2-approximable process is L^0-approximable. By Assumption (e), the approximators can be taken to be conditional means $\{EQ(z^s_{Tt}(\theta),\theta)|V_{Tt-r},\ldots,V_{Tt+r}\colon t\leqslant T,T\geqslant 1,r\geqslant 1\}$. Thus $Q(z^s_{Tt}(\cdot),\cdot)$ is L^0-near epoch dependent. Using Lemma A.2 of Andrews (1993) with X_{Tt} equal to an element of the q-dimensional vector $Q(z^s_{Tt}(\theta),\theta)-EQ(z^s_{Tt}(\theta),\theta)$, we obtain (i). For (ii), by Theorem 21.11 from Davidson (1994), global Lipschitz condition (Assumption (d)) is sufficient to obtain that $\{G_{Tt}\}$ is stochastically equicontinuous. □

[10] The presentation of Davidson is drawn mainly from Andrews (1992). Newey (1991) provides also conditions for uniform convergence based on stochastic equicontinuity.

[11] The concept of L^p-approximable process is due to Pötscher and Prucha (1991). See also Davidson (1994) for a presentation of L^p-approximable process.

Proof of Theorem 1. First, we need to show that

$$\sup_{\pi\in\Pi,\theta\in\Theta} |\bar{f}_T^S(\theta,\pi)'\hat{W}_T(\pi)\bar{f}_T^S(\theta,\pi) - f^S(\theta,\pi)'W(\pi)f^S(\theta,\pi)| \xrightarrow{p} 0.$$

Using Assumption A.3, the expression above holds if

$$\sup_{\pi\in\Pi}\sup_{\theta\in\Theta} |\bar{f}_T^S(\theta,\pi) - f^S(\theta,\pi)| \xrightarrow{p} 0.$$

Using $\sum_{[T\pi]+1}^{T} = \sum_{t=1}^{T} - \sum_{t=1}^{[T\pi]}$, the expression above holds if

$$\sup_{\theta\in\Theta}\sup_{[T\pi_1]\leqslant R\leqslant T}\left|\frac{1}{T}\sum_{t=1}^{R}\left[\left(m(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} m((z_{Tt}^s(\theta),\theta)\right)\right.\right.$$
$$\left.\left. - \left(Em(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} Em((z_{Tt}^s(\theta),\theta))\right)\right]\right| \xrightarrow{p} 0, \tag{B.1}$$

where $\pi_1 = \inf\{\pi : \pi\in\Pi\} > 0$ and

$$\sup_{\theta\in\Theta}\sup_{\pi\in\Pi}\left|\frac{1}{T}\sum_{t=1}^{[T\pi]}\left[\left(Em(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} Em((z_{Tt}^s(\theta),\theta)\right)\right.\right.$$
$$\left.\left. - \left(\tilde{m} - \frac{1}{S}\sum_{s=1}^{S}\tilde{m}^s(\theta)\right)\right]\right| \xrightarrow{p} 0. \tag{B.2}$$

For B.1, by the triangle inequality

$$\sup_{R\leqslant T}\left|\frac{1}{T}\sum_{t=1}^{R}\left[\left(m(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} m((z_{Tt}^s(\theta),\theta)\right)\right.\right.$$
$$\left.\left. - \left(Em(z_{Tt}) - \frac{1}{S}\sum_{s=1}^{S} Em((z_{Tt}^s(\theta),\theta)\right)\right]\right|$$
$$\leqslant \sup_{R\leqslant T}\left|\frac{1}{T}\sum_{t=1}^{R}(m(z_{Tt}) - Em(z_{Tt}))\right|$$
$$+ \left|\frac{1}{T}\sum_{t=1}^{R}\left(\frac{1}{S}\sum_{s=1}^{S} Em((z_{Tt}^s(\theta),\theta) - m((z_{Tt}^s(\theta),\theta)\right)\right|.$$

By Assumption A.2, the first term on the right-hand side of the expression above converges to zero in probability. By using Lemma 1 for $m(z^s_{Tt}(\cdot),\cdot)$ with Assumptions A.1–A.3, we establish the uniform WLLN for $m(z^s_{Tt}(\cdot),\cdot)$ for $s=1,\dots,S$. Thus, the second term of the expression above also converges to zero in probability. Hence, Eq. (B.1) holds, whereas Eq. (B.2) holds by Assumption A.3 and the triangle inequality. Finally, we now apply Lemma A.1 of Andrews (1993) under Assumption A.3 with $Q_T(\theta,\pi)=\bar f^S_T(\theta,\pi)'\hat W_T(\pi)\bar f^S_T(\theta,\pi)$, $Q(\theta,\pi)=f^S(\theta,\pi)'W(\pi)f^S(\theta,\pi)$, we then obtain that $\sup_{\pi\in\Pi}\|\hat\theta(\pi)-\theta_0\|\xrightarrow{p}0$.

Proof of Theorem 2. We start with the following property for the sample score function:

$$\left(\frac{\partial}{\partial\theta'}\bar f^S_T(\hat\theta^S_T(\pi),\pi)\right)'\hat W_T(\pi)\sqrt{T}\bar f^S_T(\hat\theta^S_T(\pi),\pi)=\mathrm{o_p}(1). \tag{B.3}$$

Using Taylor's theorem, we have

$$\sqrt{T}\bar f^S_T(\hat\theta^S_T(\pi),\pi)=\sqrt{T}\bar f^S_T(\theta_0,\pi)+\left(\frac{\partial}{\partial\theta'}\bar f^S_T(\bar\theta^S_T(\pi),\pi)\right)\sqrt{T}(\hat\theta^S_T(\pi)-\theta_0), \tag{B.4}$$

where $\bar\theta^S_T(\pi)'=[\bar\theta^{S,(1)}_T(\pi)\dots\bar\theta^{S,(p)}_T(\pi)]$ and $\bar\theta^{S,(k)}_T(\pi)=\lambda^{(k)}\theta^{(k)}_0+(1-\lambda^{(k)})\hat\theta^{S,(k)}_T(\pi)$ for some $0\leqslant\lambda^{(k)}\leqslant 1$ and $k=1,\dots,p$. Since $\hat\theta^S_T(\pi)$ is consistent for θ_0 (see Theorem 1 and by the consistency of the full sample estimator), $\bar\theta^S_T(\pi)\xrightarrow{p}\theta_0$. We now have to show that

$$\sup_{\pi\in\Pi}\left\|\frac{\partial}{\partial\theta'}\bar f^S_T(\bar\theta^S_T(\pi),\pi)-F(\pi)\right\|\xrightarrow{p}0, \tag{B.5}$$

whenever $\bar\theta^S_T(\pi)$ satisfies $\sup_{\pi\in\Pi}\|\bar\theta^S_T(\pi)-\theta_0\|\xrightarrow{p}0$. To establish this, we can write

$$\begin{aligned}
&\sup_{\pi\in\Pi}\left\|\frac{\partial}{\partial\theta'}\bar f^S_T(\bar\theta^S_T(\pi),\pi)-F(\pi)\right\|\\
&\quad\leqslant\sup_{\pi\in\Pi}\left\|\frac{\partial}{\partial\theta'}\bar f^S_T(\bar\theta^S_T(\pi),\pi)-E\frac{\partial}{\partial\theta'}\bar f^S_T(\theta,\pi)\Big|_{\theta=\bar\theta^S_T(\pi)}\right\|\\
&\qquad+\sup_{\pi\in\Pi}\left\|E\frac{\partial}{\partial\theta'}\bar f^S_T(\theta,\pi)\Big|_{\theta=\bar\theta^S_T(\pi)}-E\frac{\partial}{\partial\theta'}\bar f^S_T(\theta_0,\pi)\right\|\\
&\qquad+\sup_{\pi\in\Pi}\left\|E\frac{\partial}{\partial\theta'}\bar f^S_T(\theta_0,\pi)-F(\pi)\right\|.
\end{aligned} \tag{B.6}$$

For the first term of Eq. (B.6), we need to show a WLLN for $\partial\bar f^S_T(\theta,\pi)/\partial\theta'$. Using Lemma 1 for $D_{\theta'}m(z^s_{Tt},\theta)$ combined with Assumptions A.1, A.2, A.5, we obtain a uniform WLLN for $D_{\theta'}m(z^s_{Tt},\theta)$. Hence, the first term of the expression above converges to zero in probability.

For the second term of Eq. (B.6), we have

$$\sup_{\pi\in\Pi}\left|\left|E\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta,\pi)\Big|_{\theta=\bar{\theta}^{S}_{T}(\pi)}-E\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta_0,\pi)\right|\right|$$

$$\leqslant E\sup_{\pi\in\Pi}\left|\left|\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta,\pi)\Big|_{\theta=\bar{\theta}^{S}_{T}(\pi)}-\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta_0,\pi)\right|\right|.$$

By the global Lipschitz condition,

$$\left|\left|\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta,\pi)\Big|_{\theta=\bar{\theta}^{S}_{T}(\pi)}-\frac{\partial}{\partial\theta'}\bar{f}^{S}_{T}(\theta_0,\pi)\right|\right|\leqslant C_{Tt}\left|\left|\bar{\theta}^{S}_{T}(\pi)-\theta_0\right|\right|.$$

Since $\sup_{\pi\in\Pi}\|\bar{\theta}^{S}_{T}(\pi)-\theta_0\|\xrightarrow{p}0$, we obtain that the second term converges to zero in probability. The third term of Eq. (B.6) converges to zero in probability by Assumption A.5. Next, we show that

$$\sqrt{T}\bar{f}^{S}_{T}(\theta_0,\pi)\Rightarrow G(S,\pi). \tag{B.7}$$

We define

$$f^{S}_{1T}(\theta,\pi)=\frac{1}{T}\sum_{t=1}^{[T\pi]}\left(m((z^{s}_{Tt})-\frac{1}{S}\sum_{s=1}^{S}m((z^{s}_{Tt}(\theta_0),\theta_0)\right),$$

which corresponds to the moment conditions for the first subsample. Under the null hypothesis $f^{S}_{1T}(\theta,\pi)$ has the same asymptotic distribution as

$$\frac{1}{T}\sum_{t=1}^{[T\pi]}[m((z^{s}_{Tt})-Em((z^{s}_{Tt}(\theta_0),\theta_0)]$$

$$-\frac{1}{T}\sum_{t=1}^{[T\pi]}\left(\frac{1}{S}\sum_{s=1}^{S}[m((z^{s}_{Tt}(\theta_0),\theta_0)-Em((z^{s}_{Tt}(\theta_0),\theta_0)]\right).$$

Both terms above have asymptotically independent increments and are mutually independent. Define $v_{1,T}(\pi)=(1/\sqrt{T})\sum_{t=1}^{[T\pi]}[m(z_t)-Em(z_t)]$ and $v^{s}_{2,T}(\pi)=(1/\sqrt{T})\sum_{t=1}^{[T\pi]}[m((z^{s}_{Tt}(\theta_0),\theta_0)-Em((z^{s}_{Tt}(\theta_0),\theta_0)]$, for $s=1,\ldots,S$. We can apply Lemma A.4 of Andrews (1993), under Assumptions A.2 and A.3, to the sequences $\{v_{1,T}(\cdot)\colon T\geqslant 1\}$ and $\{v^{s}_{2,T}(\cdot)\colon T\geqslant 1\}$ for $s=1,\ldots,S$, since they have asymptotically independent increments. Therefore, $v_{1,T}(\cdot)\Rightarrow\Omega^{1/2}B_1(\cdot)$ and $v^{s}_{2,T}(\cdot)\Rightarrow\Omega^{1/2}B^{s}_{2}(\cdot)$ where $B_1(\cdot)$ and $B^{s}_{2}(\cdot)$ are q-dimensional vectors of independent Brownian motions. Since $\sqrt{T}f^{S}_{T}(\theta_0,\pi)=(\sqrt{(T)}f^{S}_{1T}(\theta_0,\pi)',\sqrt{(T)}f^{S}_{1T}(\theta_0,\pi)')'=((v_{1,T}(\pi)-\frac{1}{S}\sum_{s=1}^{S}v^{s}_{2,T}(\pi))',((v_{1,T}(1)-v_{1,T}(\pi))-\frac{1}{S}\sum_{s=1}^{S}(v^{s}_{2,T}(1)-v^{s}_{2,T}(\pi)))')'$, we have the following asymptotic distribution for the moment condition of the partial-sample SMM estimator:

$$\sqrt{T}\bar{f}^{S}_{T}(\theta_0,\pi)=\begin{bmatrix}\sqrt{T}f^{S}_{1T}(\theta_0,\pi)\\ \sqrt{T}f^{S}_{2T}(\theta_0,\pi)\end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \Omega^{1/2}[B_1(\pi) - \frac{1}{S}\sum_{s=1}^{S} B_2^s(\pi)] \\ \\ \Omega^{1/2}[(B_1(1) - B_1(\pi)) - \frac{1}{S}\sum_{s=1}^{S}(B_2^s(1) - B_2^s(\pi))] \end{bmatrix}. \tag{B.8}$$

The right expression can be rewritten as

$$\left(1 + \frac{1}{S}\right)^{1/2} \begin{bmatrix} \Omega^{1/2}B(\pi) \\ \Omega^{1/2}(B(1) - B(\pi)) \end{bmatrix}, \tag{B.9}$$

where $B(\pi)$ is a q-dimensional vector of standard Brownian motion since $((1+\frac{1}{S})^{-1/2})$ $[B_1(\pi) - \frac{1}{S}\sum_{s=1}^{S} B_2^s(\pi)]$ is a q-dimensional vector of standard Brownian motion.

Moreover, using Eqs. (B.3), (B.4), (B.5) and (B.7), Assumptions A.4 and A.5 and the continuous mapping theorem (see Pollard, 1984), we obtain the desired result. □

Proof of Theorem 3. The result is obtained by noting that

$$\sqrt{T}\bar{f}_T^S(\theta_0,\pi) = \sqrt{T}(\bar{f}_T^S(\theta_0,\pi) - E\bar{f}_T^S(\theta_0,\pi)) + \sqrt{T}E\bar{f}_T^S(\theta_0,\pi)$$

and applying the arguments from the proof of Theorem 2 under the alternative 3.1.

Proof of Theorem 4. The proof is a modification of Andrews' proof for the Wald test which takes into account the presence of simulated moments and the asymptotic distribution is derived under the alternative. The result under the null follows directly. We define $\Xi = [I_r, -I_r, 0_s] \in R^{r\times(2r+s)}$ and have

$$\sqrt{T}(\hat{\beta}_{1T}^S(\pi) - \hat{\beta}_{2T}^S(\pi)) = \Xi\sqrt{T}(\hat{\theta}_T^S(\pi) - \theta_0)$$
$$\Rightarrow \Xi(F(\pi)'W(\pi)F(\pi))^{-1}F(\pi)'W(\pi)G(S,\pi).$$

For a optimal choice of the weighting matrix, we can write

$$F(\pi)'W(\pi)F(\pi)$$
$$= \left(1+\frac{1}{S}\right)^{-1} \begin{bmatrix} \pi(F^\beta)'\Omega^{-1}F^\beta & 0 & \pi(F^\beta)'\Omega^{-1}F^\delta \\ 0 & (1-\pi)(F^\beta)'\Omega^{-1}F^\beta & (1-\pi)(F^\beta)'\Omega^{-1}F^\delta \\ \pi(F^\delta)'\Omega^{-1}F^\beta & (1-\pi)(F^\delta)'\Omega^{-1}F^\beta & (F^\delta)'\Omega^{-1}F^\delta \end{bmatrix}.$$

Therefore, by the Lemma A.5 of Andrews (1993)

$$\Xi(F(\pi)'W(\pi)F(\pi))^{-1}F(\pi)'W(\pi)G(S,\pi)$$
$$= [I_r : -I_r]\left(1+\frac{1}{S}\right)^{1/2} \begin{bmatrix} \pi(F^\beta)'\Omega^{-1}F^\beta & 0 \\ 0 & (1-\pi)(F^\beta)'\Omega^{-1}F^\beta \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} (F^{\beta})'\Omega^{-1/2}[B(\pi) + (1+\frac{1}{S})^{-1/2}\Omega^{-1/2}H(\pi)] \\ (F^{\beta})'\Omega^{-1/2}[(B(1)-B(\pi)) + (1+\frac{1}{S})^{-1/2}\Omega^{-1/2}(H(1)-H(\pi))] \end{bmatrix}$$

$$= C\left(\frac{[B(\pi) + (1+\frac{1}{S})^{-1/2}\Omega^{-1/2}H(\pi)]}{\pi} - \frac{[(B(1)-B(\pi)) + (1+\frac{1}{S})^{-1/2}\Omega^{-1/2}(H(1)-H(\pi))]}{(1-\pi)}\right),$$

where $C = (1+\frac{1}{S})^{1/2}((F^{\beta})'\Omega^{-1}F^{\beta})^{-1}(F^{\beta})'\Omega^{-1/2}$.

We have also

$$\frac{\hat{V}_1(\pi)}{\pi} + \frac{\hat{V}_2(\pi)}{(1-\pi)} \Rightarrow \frac{V}{\pi(1-\pi)} = \frac{CC'}{\pi(1-\pi)},$$

where $V = (1+\frac{1}{S})((F^{\beta})'\Omega^{-1}F^{\beta})^{-1}$. The Wald statistic then converges to

$$BH(\pi)'C'\left(\frac{CC'}{\pi(1-\pi)}\right)^{-1} CBH(\pi),$$

where $BH(\pi) = [B(\pi) - \pi B(1) + (1+\frac{1}{S})^{-1/2}\Omega^{-1/2}(H(\pi) - \pi H(1))]$.

We can decompose $C'(CC')^{-1}C$ as $C'(CC')^{-1/2}(CC')^{-1/2}C$. When we define the Brownian motions $(CC')^{-1/2}CB(\pi)$, the asymptotic distribution result follows. In particular, the asymptotic distribution does not depend on the number of simulations S, under the null and hence is nuisance parameter free. For the $\mathrm{LM}_T(S,\pi)$ statistic and $\mathrm{LR}_T(S,\pi)$, we can show that

$$\mathrm{LM}_T^S(\pi) = \mathrm{Wald}_T^S(\pi) + o_p(1)$$

and

$$\mathrm{LR}_T^S(\pi) = \mathrm{Wald}_T^S(\pi) + o_p(1)$$

using the same arguments as in Andrews for the GMM case. For brevity the proof is omitted. □

Proof of Theorem 6. From Theorem 3 we have that

$$\sqrt{T} f_{1T}^S(\theta_0) \Rightarrow \Omega^{1/2}\left[B_1(\pi) - \frac{1}{S}\sum_{s=1}^{S} B_2(\pi)\right] + H(\pi). \tag{B.10}$$

Let us consider an expansion of the moment conditions for the full sample evaluated with the restricted estimator

$$f_T^S(\tilde{\theta}_T^S) = f_T^S(\theta_0) + \frac{\partial}{\partial\theta'} f_T^S(\bar{\theta})(\tilde{\theta}_T^S - \theta_0),$$

where $\bar{\theta}_T^{S\prime} = [\bar{\theta}_T^{S,(1)} \dots \bar{\theta}_T^{S,(p)}]$ and $\tilde{\theta}_T^{S,(k)} = \lambda^{(k)}\theta_0^{(k)} + (1-\lambda^{(k)})\hat{\theta}_T^{S,(k)}$ for some $0 \leqslant \lambda^{(k)} \leqslant 1$ and $k = 1,\dots,p$. Multiplying both sides by $\partial f_T^S(\tilde{\theta}^S)' W_T / \partial\theta'$ yields

$$(\tilde{\theta}_T^S - \theta_0) = -\left[\frac{\partial}{\partial\theta'} f_T^S(\tilde{\theta}_T^S)' W_T \frac{\partial}{\partial\theta'} f_T^S(\bar{\theta})\right]^{-1} \frac{\partial}{\partial\theta'} f_T^S(\tilde{\theta}_T^S)' W_T f_T^S(\theta_0) \tag{B.11}$$

since $(\partial f_T^S(\tilde{\theta}_T^S)/\partial\theta')' W_T f_T^S(\tilde{\theta}_T^S) = \mathrm{o_p}(1)$.

Furthermore, expanding the moment conditions for the first subsample evaluated at the restricted full sample estimator yields

$$f_{1T}^S(\tilde{\theta}_T^S) = f_{1T}^S(\theta_0) + \frac{\partial}{\partial\theta'} f_{1T}^S(\ddot{\theta})(\tilde{\theta}_T^S - \theta_0), \tag{B.12}$$

where $\ddot{\theta}_T^{S\prime} = [\ddot{\theta}_T^{S,(1)} \dots \ddot{\theta}_T^{S,(p)}]$ and $\ddot{\theta}_T^{S,(k)} = \lambda^{(k)}\theta_0^{(k)} + (1-\lambda^{(k)})\bar{\theta}_T^{S,(k)}$ for some $0 \leqslant \lambda^{(k)} \leqslant 1$ and $k = 1,\dots,p$. Substitution of (B.9) into (B.10) results in

$$f_{1T}^S(\tilde{\theta}_T^S) = f_{1T}^S(\theta_0) - \frac{\partial}{\partial\theta'} f_{1T}^S(\ddot{\theta}) \left[\frac{\partial}{\partial\theta'} f_T^S(\tilde{\theta}_T^S)' W_T \frac{\partial}{\partial\theta'} f_T^S(\bar{\theta})\right]^{-1} \frac{\partial}{\partial\theta'} f_T^S(\tilde{\theta}_T^S)' W_T f_T^S(\theta_0).$$

Combining Eqs. (B.8) and (B.5) and Assumption (A.3) yields

$$\sqrt{T} f_{1T}^S(\tilde{\theta}_T^S) \Rightarrow \Omega^{1/2}\left(B_1(\pi) - \frac{1}{S}\sum_{s=1}^S B_2^s(\pi)\right) + H(\pi)$$
$$-\pi F(F'WF)^{-1}F'W\left[\Omega^{1/2}\left(B_1(1) - \frac{1}{S}\sum_{s=1}^S B_2^s(1)\right) + H(1)\right].$$

We need the following corollary of Theorem 6 to show the result of Theorem 7. This corollary is an extension of the Sowell's Corollary 1 (Sowell, 1996a) to the case of simulation-based estimators. □

Corollary B.1. *Under Assumptions* A.1–A.5 *and the null hypothesis, there exists an orthonormal matrix* C *such that*

$$\left(1+\frac{1}{S}\right)^{-1/2} C\sqrt{T}\Omega_T^{-1/2} f_{1T}^S(\tilde{\theta}_T^S) \Rightarrow \begin{bmatrix} \mathrm{BB}_p(\pi) \\ B_{q-p}(\pi) \end{bmatrix}$$

and under the alternative in Eq. (4.1)

$$\left(1+\frac{1}{S}\right)^{-1/2} C\sqrt{T}\Omega_T^{-1/2} f_{1T}^S(\tilde{\theta}_T^S)$$
$$\Rightarrow \begin{bmatrix} \mathrm{BB}_p(\pi) + (1+\frac{1}{S})^{-1/2} C_1 \Omega^{-1/2}(H(\pi) - \pi H(1)) \\ B_{q-p}(\pi) + (1+\frac{1}{S})^{-1/2} C_2 \Omega^{-1/2} H(\pi) \end{bmatrix},$$

where $\mathrm{BB}_p(\pi)$ *is a* p*-vector of standard Brownian bridge and* $B_{q-p}(\pi)$ *is a* $(q-p)$*-vector of standard Brownian motion and* $C'\Lambda C = \Omega^{-1/2}F(F'\Omega^{-1}F)^{-1}F'\Omega^{-1/2}$

where

$$\Lambda = \begin{bmatrix} I_p & 0_{p\times(q-p)} \\ 0_{(q-p)\times p} & 0_{(q-p)\times(q-p)} \end{bmatrix}$$

and C_1 is the matrix of the first p rows of C and C_2 the last $q-p$ rows of C.

The corollary follows from Theorem 6 and the fact that

$$\left(1+\frac{1}{S}\right)^{-1/2} C\left[B_1(\pi) - \frac{1}{S}\sum_{s=1}^{S} B_2^s(\pi)\right] \tag{B.13}$$

is a q-dimensional vector of standard Brownian motion.

Proof of Theorem 7. The statistic $Q_{t,T}^{S,O}$ consists in the quadratic form of the projection of $(1+1/S)^{-1/2}\Omega^{-1/2}\mathrm{Sow}_T^S(\tilde{\theta}_T^S,\pi)$ and $(1+1/S)^{-1/2}\Omega^{-1/2}HS_{t,T}^S(\tilde{\theta}_T^S,\pi)$ on $C'(I_q-\Lambda)C$. By Corollary B.1, this projection cancels out the p-dimensional vector of Brownian bridge. Thus, the asymptotic distribution of $Q_{t,T}^{S,O}$ is given by the quadratic form of the $q-p$-dimensional vector of Brownian motion under the null and $(1+1/S)^{-1/2}C_2\Omega^{-1/2}H(\pi)$ under the alternative. The asymptotic distribution of the supremum, the average and the exponential mappings is obtained by the continuous mapping theorem (Pollard, 1984).

References

Andrews, D.W.K., 1987. Consistency in nonlinear econometric models: a generic uniform law of large numbers. Econometrica 55, 1465–1472.

Andrews, D.W.K., 1992. Generic uniform convergence. Econometric Theory 8, 241–257.

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

Andrews, D.W.K., 1994. The large sample correspondence between classical hypothesis tests and Bayesian posterior odds tests. Econometrica 62, 1207–1232.

Andrews, D.W.K., Monahan, J.C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. Econometrica 60, 953–966.

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

Asea, P.K., Turnovsky, S.J., 1998. Capital income taxation and risk-taking in a small open economy. Journal of Public Economics 68, 55–90.

Bierens, H., 1981. Robust Methods and Asymptotic Theory. In: Lecture Notes in Economics and Mathematical Systems, Vol. 192. Springer, Berlin.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Carrasco, M., Hansen, L.P., Chen, X., 1999. Timed deformation and dependence, Mimeo, University of Chicago.

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

Davidson, J., 1994. Stochastic Limit Theory: An Introduction for Econometricians. Oxford University Press, Oxford.

de Jong, F., 1994. A univariate analysis of EMS exchange rates using a target zone model. Journal of Applied Econometrics 9, 31–45.

Dridi, R., Renault, E., 2000. Semiparametric indirect inference. Working Paper, Department of Economics, London School of Economics.

Duffie, D., Singleton, K.J., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.

Dufour, J.M., Ghysels, E., Hall, A., 1994. Generalized predictive tests and structural change analysis in econometrics. International Economic Review 35, 199–229.

Foster, F.D., Viswanathan, S., 1995. Can speculative trading explain the volume-volatility relation? Journal of Business and Economics Statistics 13, 379–396.

Gallant, A.R., 1987. Nonlinear Statistical Models. Wiley, New York.

Gallant, A.R., Tauchen, G., 1996. Which moments to match? Econometric Theory 12, 657–681.

Gallant, A.R., White, H., 1988. A unified theory of estimation and inference for nonlinear dynamic models. Basil Blackwell, New York.

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

Ghysels, E., Harvey, A., Renault, E., 1996. Stochastic Volatility. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics, Vol. 14. North-Holland, Amsterdam, pp. 119–192.

Ghysels, E., Guay, A., Hall, A., 1997. Predictive test for structural change with unknown breakpoint. Journal of Econometrics 82, 209–233. [Corrigendum Journal of Econometrics, 30, 337-343].

Gouriéroux, C., Monfort, A., 1996. Simulation-based Econometric Methods. Oxford University Press, Oxford.

Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.

Guay, A., 2003. Optimal predictive tests. Econometric Reviews, forthcoming.

Hairault, J.O., Langot, F., Portier, F., 1997. Time to implement and aggregate fluctuations. Journal of Economic Dynamics and Control 22, 109–121.

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

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Iannizzotto, M., Taylor, M., 1999. The target zone model, non-linearity and mean reversion: is the honeymoon really over?. Economic Journal 109, C96–C110.

Ingram, B., Lee, B.S., 1991. Simulation estimation of time series models. Journal of Econometrics 47, 197–205.

Jorgensen, B.N., Mikkelsen, H.O., 1991. An arbitrage free trilateral target zone model. Journal of International Money and Finance 15, 117–134.

Laroque, G., Selanie, B., 1989. Estimation of multi-market fix-price models. Econometrica 57, 831–860.

McLeish, D.L., 1975a. A maximal inequality and dependent strong laws. Annals of Probability 3, 829–839.

McLeish, D.L., 1975b. Invariance principles for dependent variables. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 32, 165–178.

Newey, W.K., 1991. Uniform convergence in probability and stochastic equicontinuity. Econometrica 59, 703–708.

Newey, W.K., West, K., 1994. Automatic lag selection in covariance matrix estimation. Review of Economic Studies 61, 631–653.

Pollard, D., 1984. Convergence of Stochastic Processes. Springer, New York.

Potscher, B.M., Prucha, I.R., 1991. Basic structure of the asymptotic theory in dynamic nonlinear econometric model. Part I: consistency and approximation concepts. Econometric Reviews 10, 125–216.

Smith, A., 1993. Estimating nonlinear time series models using simulated vector autoregressions. Journal of Applied Econometrics S8, S63–S84.

Sowell, F., 1993. A new approach to testing parameter instability in the generalized method of moments framework. Manuscript, Graduate School of Industrial Administration. Carnegie Mellon University, Pittsburgh, PA.

Sowell, F., 1996a. Optimal tests for parameter instability in the generalized method of moments framework. Econometrica 64, 1085–1107.

Sowell, F., 1996b. Tests for violations of moment conditions. Manuscript, Graduate School of Industrial Administration, Carnegie Mellon University.

van der Sluis, P.J., 1998. Computationally attractive stability tests for efficient methods of moments. Econometric Journal 1, C203–C337.

van der Sluis, P.J., 1999. Estimation and inference with the efficient method of moments—with applications to stochastic volatility and option pricing. Doctoral Dissertation, University of Amsterdam.