

Robust GMM Tests for Structural Breaks*

Abstract

We propose a class of new robust GMM tests for endogenous structural breaks. The tests are based on supremum, average and exponential functionals derived from robust GMM estimators with bounded influence function. We study the theoretical local robustness properties of the new tests and show that they imply a uniformly bounded asymptotic sensitivity of size and power under general local deviations from a reference model. We then analyze the finite sample performance of the new robust tests in some Monte Carlo simulations, and compare it with that of classical GMM tests for structural breaks. In large samples, we find that the performance of classical asymptotic GMM tests can be quite unstable already under slight departures from some given reference distribution. In particular, the loss in power can be substantial in some models. Robust asymptotic tests for structural breaks yield important power improvements already under slight local departures from the reference model. This holds both in exactly identified and overidentified model settings. In small samples, bootstrapped versions of both the classical and the robust GMM tests provide a very accurate and very stable empirical size also for quite small sample sizes. However, bootstrapped robust GMM tests are found to provide again a higher finite sample efficiency.

Keywords: Robust Tests, Generalized Method of Moment, Structural Breaks, Monte Carlo, Bootstrap
J.E.L. Classification Numbers: C10, C12, C13, C15.

Patrick Gagliardini, Institute of Finance, University of Lugano, Via Buffi 13,

CH-6900 Lugano, e-mail: Patrick.Gagliardini@lu.unisi.ch.

Fabio Trojani, Institute of Finance, University of Lugano, Via Buffi 13,

CH-6900 Lugano, e-mail: Fabio.Trojani@lu.unisi.ch.

Giovanni Urga, Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (U.K.).

Tel.: +/44/20/70408696, Fax.: +/44/20/70408885, e-mail: g.urga@city.ac.uk

<http://www.staff.city.ac.uk/~giourga>.

*We wish to thank three anonymous referees, Pietro Balestra, Lorian Mancini, Claudio Ortelli, Paolo Porchia and participants to the International Conference on MODELLING STRUCTURAL BREAKS, LONG MEMORY AND STOCK MARKET VOLATILITY (held in London at Cass Business School, 6-7 December 2002) for discussions and very helpful comments. G. Urga wishes to thank ESRC under Grant number R000238145 ("*New methods for detecting multiple structural breaks in time series*") for funding this research. Patrick Gagliardini and Fabio Trojani gratefully acknowledge the financial support of the Swiss National Science Foundation (grant 12-65196.01 and NCCR FINRISK).

1 Introduction

We propose a class of new Generalized Method of Moments (GMM, Hansen (1982)) tests for endogenous structural breaks that ensure a uniformly bounded asymptotic sensitivity of level and power under general local departures from a reference model.

GMM based test statistics defining tests for structural breaks are typically obtained as the supremum, the average or some related functional of sequences of quadratic GMM statistics, each being asymptotically chi-square distributed under the null of no break¹ (see e.g. Andrews and Fair (1988), Ghysels and Hall (1990), Andrews (1993), Andrews and Ploberger (1994), Ghysels, Guay and Hall (1997)). Such GMM functionals are evaluated at some GMM model parameter estimates, conditionally on a given break date. A general GMM statistical functional (as for instance a GMM estimator or the level/power of a GMM test) has a bounded asymptotic sensitivity under local model perturbations if and only if it is based on a GMM model with a bounded orthogonality function. Moreover, GMM statistics with unbounded asymptotic sensitivity can be robustified by applying a weighted orthogonality function that bounds the influence of general local departures from a given reference model (see Ronchetti and Trojani (2001)). This defines locally robust GMM (RGMM) estimation and testing procedures of simple parametric hypotheses in a fairly general GMM setting. In this paper, we propose a class of RGMM tests for structural breaks, which are defined as functionals of sequences of quadratic RGMM statistics based on a bounded orthogonality function.

The need for robust statistical procedures in estimation and testing has been stressed by many authors and is now widely recognized; see for instance, Hampel (1974), Koenker and Bassett (1978), Huber (1981), Peracchi (1990, 1991), Heritier and Ronchetti (1994), Krishnakumar and Ronchetti (1997), Ronchetti and Trojani (2001), Genton and Ronchetti (2003). From a general viewpoint, the goal of robust testing procedures is to construct test statistics that maintain a

¹ Asymptotic critical values for GMM tests of endogeneous structural breaks are derived by the Functional Central Limit Theorem and have typically to be computed by simulation; see for instance Andrews (1993) and Andrews (2003). Analytical approximations have been proposed in Hansen (1997).

uniformly satisfactory level and power behavior under general local distributional departures from some given reference model. This is achieved by working with smooth test functionals that are asymptotically stable under departures from the given reference point. In particular, in the GMM setting, a necessary asymptotic robustness/stability requirement for a GMM test based on an asymptotically chi-square distributed statistic is a bounded influence function (Hampel (1974)) of the GMM estimator defining the statistic. Therefore, we define a new class of tests for breaks using sequences of RGMM statistics with bounded influence function. A bounded influence function of a GMM test statistic is equivalent to the boundedness of the given orthogonality function. This is why RGMM statistics can be obtained by truncating appropriately the unbounded orthogonality function of a nonrobust GMM model setting.

We study the theoretical properties of our robust testing procedures for structural breaks and analyze their empirical performance in some Monte Carlo experiments of a few GMM model settings. To our knowledge, this issue has been so far largely unexplored in the literature. For the estimation problem, Fiteni (2002) derived the asymptotic properties of a robust break date estimator defined through the supremum functional over a sequence of robustified loss functions. These results apply to a standard linear regression model setting. We propose a class of general RGMM tests for breaks that apply to linear and nonlinear model settings.

We first show theoretically that RGMM tests for breaks imply a uniformly bounded asymptotic sensitivity of level and power under general local deviations from a reference model. This ensures a uniform quality of the asymptotic approximation to the finite sample distribution of a RGMM statistic over a relevant neighborhood of slightly different model distributions. We then compare via Monte Carlo simulation the performance of GMM and RGMM tests for breaks, using both standard asymptotic critical values and bootstrapped versions of the tests.

In large samples, we find that the performance of classical asymptotic GMM tests can be quite unstable already under slight departures from some given reference distribution. In particular, the power under departure from conditional normality can be quite low in some models. Robust

asymptotic tests for structural breaks yield important power improvements already under slight local departures from the reference model. This holds both in exactly identified and overidentified model settings.

In small samples, bootstrapped versions of both the classical and the robust GMM tests provide a very accurate and very stable empirical size also for quite small sample sizes. However, bootstrapped robust GMM tests are found to provide again a higher finite sample power.

The remaining of the paper is organized as follows. Section 2 reviews GMM tests for structural breaks. Section 3 introduces robust GMM tests for structural breaks and studies formally their local stability properties in neighborhoods of a reference model. Section 4 analyzes by Monte Carlo simulations the empirical properties of the new robust tests in linear and nonlinear GMM testing settings, while Section 5 concludes and gives suggestions for further developments.

2 GMM tests for structural breaks

We briefly review GMM tests for structural breaks - by focusing on Andrews (1993) setting - and write the relevant statistics as functionals on a suitable set of probability distributions. This formalism will allow us to analyze in Section 3 the asymptotic local stability properties of GMM tests for structural breaks. We first discuss in Section 2.1 the different hypotheses of structural change, and then introduce GMM estimators in Section 2.2. In Section 2.3 we define the GMM test statistics for structural breaks which are relevant for our exposition.

In the following we adopt the symbol \implies to denote weak convergence in the sense of Pollard (1984, pp. 64-66) for sequences of random elements of a space of bounded Euclidean valued cadlag functions on $\Pi \subset [0, 1]$, equipped by the supremum norm topology and by the corresponding Borel sigma algebra. The symbol \rightarrow_d denotes convergence in distribution, ∇ denotes the gradient operator, $\mathcal{B}(\mathbb{R}^k)$ is the Borel sigma algebra on \mathbb{R}^k , $\|\cdot\|$ is the Euclidean norm.

2.1 Hypotheses of structural changes

We consider a parametric model indexed by parameters $(\beta_t, \delta_0) \in \Theta = \mathcal{B} \times \Delta \subset \mathbb{R}^p \times \mathbb{R}^q$, for $t = 1, 2, \dots$ and test the null hypothesis of parameter stability:

$$H_0 : \beta_t = \beta_0 \quad \text{for all } t \geq 1 \text{ and some } \beta_0 \in \mathcal{B} \subset \mathbb{R}^p \quad . \quad (1)$$

Several alternative hypotheses may be of interest in the present setting. The simple one time change alternative with known change point² $\pi \in \Pi \subset (0, 1)$ is given by:

$$H_{1T}(\pi) : \beta_t = \begin{cases} \beta_1(\pi) & \text{for } t = 1, \dots, T\pi \\ \beta_2(\pi) & \text{for } t = T\pi + 1, \dots, T \end{cases} \quad , \quad (2)$$

for some constant vectors $\beta_1(\pi), \beta_2(\pi) \in \mathcal{B}$. A natural alternative where the change point $\pi \in \Pi \subset (0, 1)$ is unknown is:

$$H_A(\Pi) = \bigcup_{\pi \in \Pi} H_{1T}(\pi) \quad .$$

In this case one tests for the presence of a break in the known interval Π . Finally, when applying tests for structural breaks as general diagnostic tools, a natural alternative may be

$$H_1 : \beta_s \neq \beta_t \quad \text{for some } s, t \geq 1 \quad .$$

Although this hypothesis is more general than $\cup_{\Pi \subset (0,1)} H_A(\Pi)$ the robust GMM tests for structural breaks considered in this paper have power also against H_1 .

2.2 GMM estimators

Let $\mathcal{W} = \{W_t : t \geq 1\}$ be a stochastic process with values in $\mathbf{W} \subset \mathbb{R}^k$, defined on a measurable space (Ω, \mathcal{F}) , and let $m : \mathbb{R}^k \times \Theta \rightarrow \mathbb{R}^v$ be an orthogonality function. A *GMM estimator* $\tilde{\theta} = (\tilde{\beta}', \tilde{\delta}')'$ is the asymptotic functional solution of a quadratic minimization problem:

$$\tilde{\theta}(P) = \arg \inf_{\theta \in \Theta} Q(P, \theta) \quad ,$$

² For technical reasons Π is assumed to be a closed set.

where P is a probability measure on (Ω, \mathcal{F}) and

$$Q(P, \theta) = \left(\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E_P [m(W_t, \theta)] \right)' \Omega(P) \left(\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E_P [m(W_t, \theta)] \right) \quad , \quad (3)$$

for some positive definite deterministic $v \times v$ matrix $\Omega = \Omega(P)$ that can depend on P .

In the following it will be convenient to work with the finite dimensional distributions P_t of W_t , defined by $P_t(A) := P(W_t \in A)$, for any $A \in \mathcal{B}(\mathbb{R}^k)$ and $t \geq 1$. Defining $\bar{P}_T = \frac{1}{T} \sum_{t=1}^T P_t$, we ensure existence of a limit \bar{P}_∞ for the sequence $\{\bar{P}_T : T \geq 1\}$ by means of the following assumption.

Assumption 1 *There exists a probability measure \bar{P}_∞ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ such that \bar{P}_∞ is the weak limit of $\{\bar{P}_T : T \geq 1\}$: $\bar{P}_T \rightarrow \bar{P}_\infty$, weakly as $T \rightarrow \infty$.*

If the functional $Q \mapsto E_Q [m(W, \beta, \delta)]$ is weakly continuous³ for any $(\beta, \delta) \in \Theta$, then

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E_P [m(W_t, \beta, \delta)] = E_{\bar{P}_\infty} [m(W, \beta, \delta)] \quad .$$

In addition, let us assume that matrix Ω is a functional of \bar{P}_∞ , $\Omega = \Omega(\bar{P}_\infty)$. Then the GMM estimator itself can be written as a functional of \bar{P}_∞ :

$$\tilde{\theta}(\bar{P}_\infty) = \arg \inf_{\theta \in \Theta} E_{\bar{P}_\infty} [m(W, \theta)]' \Omega(\bar{P}_\infty) E_{\bar{P}_\infty} [m(W, \theta)] \quad , \quad (4)$$

for any suitable \bar{P}_∞ . If the GMM model is correctly specified and identified under P , i.e. if

$$E_{\bar{P}_\infty} [m(W, \theta)] = 0 \quad , \quad (5)$$

for a unique $\theta \in \Theta$, then the solution of (5) and the GMM estimator (4) coincide. More generally, (5) can have several solutions or no solution under P . In this case, only the solution of the minimization problem in (4) defines the asymptotic functional structure of $\tilde{\theta}$.

To define the finite sample GMM estimator associated with a sample $\mathcal{W}_T := \{W_t : 1 \leq t \leq T\}$ let $\bar{P}_{\mathcal{W}_T} := \frac{1}{T} \sum_{t=1}^T \delta_{W_t}$ be the empirical distribution of \mathcal{W}_T , where δ_{W_t} is the measure with point

³ A sufficient condition for the weak continuity of $Q \mapsto E_Q [m(W, \beta, \delta)]$ is the boundedness of the orthogonality function m . Boundedness of m is the condition required to ensure the local robustness of a general GMM statistic; see Section 3 below.

mass at W_t . Under standard regularity conditions one has $\bar{P}_{\mathcal{W}_T} \rightarrow \bar{P}_\infty$ weakly, P -almost surely, as $T \rightarrow \infty$. The finite sample GMM estimator is $\hat{\theta}_T := \left(\hat{\beta}'_T, \hat{\delta}'_T \right)' := \left(\tilde{\beta}'(\bar{P}_{\mathcal{W}_T}), \tilde{\delta}'(\bar{P}_{\mathcal{W}_T}) \right)'$, i.e. the solution of the minimization problem in (4) for $\bar{P}_{\mathcal{W}_T}$. Under the correct specification and identification hypothesis (5), and standard regularity conditions (see for instance Hansen (1982)), the finite sample GMM estimator $\hat{\theta}_T$ converges a.s. as $T \rightarrow \infty$ to the unique solution $\tilde{\theta}(\bar{P}_\infty)$ in (5), and is asymptotically normally distributed. When $\left\{ m(W_t, \tilde{\theta}(\bar{P}_\infty)) : t \geq 1 \right\}$ is a martingale difference sequence under P , the optimal weighting matrix Ω is:

$$\Omega(\bar{P}_\infty) = S(\bar{P}_\infty)^{-1} := \left(E_{\bar{P}_\infty} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) m(W, \tilde{\theta}(\bar{P}_\infty))' \right] \right)^{-1}. \quad (6)$$

The next section introduces GMM tests for parameter stability. They are obtained from the above GMM estimators.

2.3 Test statistics

Some consistent, asymptotically equivalent, GMM test statistics for testing H_0 against $H_{1T}(\pi)$ are Wald-type, Lagrange Multiplier-type (LM) or Likelihood ratio-type statistics. Without loss of generality we focus on LM test functionals⁴ and assume a choice of the weighting matrix as in (6). A LM test can be defined by means of the statistic

$$\widehat{LM}_T(\pi) = \frac{T}{\pi(1-\pi)} \cdot LM_T(\pi) = \frac{T}{\pi(1-\pi)} U_T(\pi)' U_T(\pi),$$

where

$$U_T(\pi) = \pi H(\bar{P}_\infty)^{1/2} \frac{1}{\pi T} \sum_{t=1}^{T\pi} m(W_t, \hat{\theta}_T),$$

with

$$\begin{aligned} H(\bar{P}_\infty) &= S(\bar{P}_\infty)^{-1} M(\bar{P}_\infty) \Sigma(\bar{P}_\infty) M'(\bar{P}_\infty) S(\bar{P}_\infty)^{-1}, \\ \Sigma(\bar{P}_\infty) &= \left[M'(\bar{P}_\infty) S(\bar{P}_\infty)^{-1} M(\bar{P}_\infty) \right]^{-1}, \quad M(\bar{P}_\infty) = E_{\bar{P}_\infty} \left[\nabla_{\beta'} m(W, \tilde{\theta}(\bar{P}_\infty)) \right]. \end{aligned}$$

⁴ An alternative way that could be also pursued in a RGMM testing approach for breaks is to use quadratic GMM statistics as proposed in Ghysels, Guay and Hall (1997). It can be seen from the exposition in Section 3 how such a RGMM inference approach can be applied to that setting.

Consistent estimators of $S(\bar{P}_\infty)$ and $M(\bar{P}_\infty)$ are given by $S(\bar{P}_{\mathcal{W}_T})$ and $M(\bar{P}_{\mathcal{W}_T})$, respectively. The LM statistic $\widehat{LM}_T(\pi)$ is particularly simple to compute, since it requires only the computation of a single GMM estimator. This is a clear advantage when working with RGMM statistics, because RGMM estimators for time series are typically more computationally intensive than classical GMM estimators, as can be seen from the description of the RGMM algorithms in the Appendices.

In order to discuss the asymptotic functional structure of $U_T(\cdot)$, we introduce a stronger version of Assumption 1.

Assumption 2 *For any $\pi \in \Pi$, there exists a probability measure $\bar{P}(\pi)$ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ such that: $\frac{1}{T\pi} \sum_{t=1}^{\pi T} P_t \rightarrow \bar{P}(\pi)$, weakly as $T \rightarrow \infty$, uniformly in $\pi \in \Pi$.*

In particular, with the above notation we have $\bar{P}_\infty = \bar{P}(1)$. Moreover, $\frac{1}{T\pi} \sum_{t=1}^{\pi T} \delta_{W_t} \rightarrow \bar{P}(\pi)$ weakly as $T \rightarrow \infty$, P almost surely. If the functional $Q \mapsto E_Q[m(W, \beta, \delta)]$ is weakly continuous, this implies an asymptotic functional structure U of U_T of the form:

$$U(\pi, P) = \pi H(\bar{P}_\infty)^{1/2} E_{\bar{P}(\pi)} \left[m \left(W, \tilde{\theta}(\bar{P}_\infty) \right) \right] \quad . \quad (7)$$

In particular, functional $U(\pi, P)$ depends on P through the finite dimensional measures $\bar{P}(\pi)$ and \bar{P}_∞ defined on $\mathcal{B}(\mathbb{R}^k)$. Hence, the asymptotic functional structure $LM(\pi, \cdot)$ of $LM_T(\pi)$ is given by

$$LM(\pi, P) = U(\pi, P)' U(\pi, P) \quad . \quad (8)$$

In this paper we focus on a class of supremum ($\widehat{\xi}_T^{\text{sup}}$), average ($\widehat{\xi}_T^{\text{ave}}$) and exponential ($\widehat{\xi}_T^{\text{exp}}$) statistics to test H_0 against alternatives of the form $H_A(\Pi)$ or H_1 . The test statistic $\widehat{\xi}_T^{\text{sup}}$ is defined by

$$\widehat{\xi}_T^{\text{sup}} := \sup_{\pi \in \Pi} \widehat{LM}_T(\pi) \quad . \quad (9)$$

The test statistic $\widehat{\xi}_T^{\text{ave}}$ is defined by

$$\widehat{\xi}_T^{\text{ave}} := \int_{\Pi} \widehat{LM}_T(\pi) d\lambda(\pi) \quad , \quad (10)$$

where λ is the Lebesgue measure on Π . Similarly, the test statistic $\widehat{\xi}_T^{\text{exp}}$ is defined by

$$\widehat{\xi}_T^{\text{exp}} := \log \int_{\Pi} \exp \left(\frac{1}{2} \widehat{LM}_T(\pi) \right) d\lambda(\pi) \quad . \quad (11)$$

The asymptotic functional structure of $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$ and $\widehat{\xi}_T^{\text{exp}}$ is completely determined by functional LM in (8). In particular, (9), (10) and (11) imply that the asymptotic stability properties of $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$ and $\widehat{\xi}_T^{\text{exp}}$ are determined by those of the functional U in (7). Therefore, one can expect to obtain a class of tests for structural breaks with better local stability properties when working with quadratic functionals based on a robust functional U . This in turn requires working with GMM test statistics and estimators based on a GMM setting with a bounded orthogonality function m . Section 3 below provides a more detailed discussion of these issues.

In a likelihood setting, statistics of the form (10) and (11) define an optimal test in terms of a weighted average power criterion based on a uniform prior for the break date $\pi \in \Pi$. Specifically, average type tests can be interpreted as the optimal test for structural breaks in the case of alternative hypotheses very near to the null. Similarly, the exponential test is the optimal test for testing more distant alternatives (see Andrews and Ploberger (1994)). When constructing robust tests for structural breaks in a likelihood setting we can therefore expect robust versions of the $\widehat{\xi}_T^{\text{ave}}$, $\widehat{\xi}_T^{\text{exp}}$ statistics to produce a higher power, when compared with robust versions of $\widehat{\xi}_T^{\text{sup}}$.

The asymptotic distribution of $\widehat{LM}_T(\cdot)$ as a process indexed by π , which implies the distribution of test statistics $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$ and $\widehat{\xi}_T^{\text{exp}}$ in (9), (10) and (11) by means of the Functional Central Limit theorem, can be studied under the following general assumption.

Assumption 3 *The model probability P satisfies the following condition:*

$$\sup_{\pi \in \Pi} \left\| \begin{pmatrix} \sqrt{T} \frac{1}{T\pi} \sum_{t=1}^{\pi T} E_P \left[m \left(W_t, \tilde{\theta}(\overline{P}_\infty) \right) \right] - \mu_1(\pi) \\ \sqrt{T} \frac{1}{T(1-\pi)} \sum_{t=\pi T+1}^T E_P \left[m \left(W_t, \tilde{\theta}(\overline{P}_\infty) \right) \right] - \mu_2(\pi) \end{pmatrix} \right\| = o(1), \quad (12)$$

for some bounded \mathbb{R}^v -valued functions μ_1, μ_2 defined on Π .

We may distinguish two cases for the interpretation of this assumption. When model P satisfies the null hypothesis H_0 of parameter stability in (1):

$$E_P \left(m \left(W_t, \tilde{\theta}(\overline{P}_\infty) \right) \right) = 0, \quad \text{for all } t,$$

then condition (12) is satisfied with $\mu_1(\pi) = \mu_2(\pi) = 0$, for all $\pi \in \Pi$. When instead either function μ_1 or function μ_2 is different from zero, then model P satisfies a local alternative hypoth-

esis, which is equivalent to Assumption 1-LP in Andrews (1993, p. 841). In addition, condition (12) implies the correct specification hypothesis $E_{\overline{P}_\infty} \left(m \left(W_t, \tilde{\theta}(\overline{P}_\infty) \right) \right) = 0$. The asymptotic distribution of $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$ and $\widehat{\xi}_T^{\text{exp}}$ is characterized in the next theorem.

Theorem 1 *Under Assumption 3 and regularity conditions on (Θ, m, P) (see Andrews (1993), Assumption 1 p. 830 and Assumption 3 p. 835) it follows:*

1. $\widehat{LM}_T(\cdot) \implies Q_p(\cdot)$ as a process indexed by $\pi \in \Pi$, where

$$Q_p(\pi) := \frac{1}{\pi(1-\pi)} (J_p(\pi) + b(\pi))' (J_p(\pi) + b(\pi)) \quad , \pi \in \Pi,$$

$J_p(\cdot)$ is a Brownian Bridge process, that is $J_p(\pi) = B_p(\pi) - \pi B_p(1)$, with $B_p(\cdot)$ a p -dimensional standard Brownian motion on $[0, 1]$, and vector b is given by:

$$b(\pi) = \pi(1-\pi) H(\overline{P}_\infty)^{1/2} [\mu_1(\pi) - \mu_2(\pi)] \quad , \quad \pi \in \Pi.$$

2. $\widehat{\xi}_T^{\text{sup}} \rightarrow_d \sup_{\pi \in \Pi} Q_p(\pi)$ under P .
3. $\widehat{\xi}_T^{\text{ave}} \rightarrow_d \int_{\Pi} Q_p(\pi) d\lambda(\pi)$ under P .
4. $\widehat{\xi}_T^{\text{exp}} \rightarrow_d \log \int_{\Pi} \exp(\frac{1}{2} Q_p(\pi)) d\lambda(\pi)$ under P .

Based on this result, critical values for test statistics $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$, $\widehat{\xi}_T^{\text{exp}}$ can be computed by simulation of process $J_p(\pi)' J_p(\pi) / \pi(1-\pi)$, $\pi \in \Pi$. Under local alternatives, the power of the test is characterized by the noncentrality vector $b(\pi)$, $\pi \in \Pi$.

3 Robust GMM tests for structural breaks

The goal of robust statistics is to provide estimation and inference procedures which are locally stable in a nonparametric neighbourhood of relevant distributions around a given reference model. In other words, those procedures are not excessively sensitive to small deviations from a reference model. Therefore, statistical robustness deals with inference procedures that are based on smooth statistical functionals. A minimal robustness requirement is continuity of such functionals. A second stronger requirement is their Fréchet differentiability⁵ (see for instance Bednarski (1993),

⁵ Let \mathcal{M} be the linear space of finite measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, equipped with a norm $\|\cdot\|$. A functional $U(P)$ defined on \mathcal{M} is Fréchet differentiable at P if there exists a bounded linear operator $DU(P, \cdot)$ such that:

$$U(Q) - U(P) = DU(P, Q - P) + o(\|Q - P\|).$$

$DU(P, \cdot)$ is called Fréchet derivative of U at P .

p. 27). From (9), (10) and (11) we can expect the power and level functionals of robust tests for breaks to satisfy the first or the second requirement if the statistical functional U in (7) does it. Therefore, a first focus is on GMM settings where such statistical functionals are Fréchet differentiable.

3.1 Fréchet differentiability

Boundedness of the orthogonality function m is a necessary condition for a general GMM statistic like the GMM estimator $\tilde{\theta}$ in (4) or the functional U in (7) to be Fréchet differentiable. More specifically, an orthogonality function m is unbounded if and only if the influence function (Hampel (1974)) of a GMM statistic is unbounded. Unboundedness of the influence function in turn implies an unbounded asymptotic sensitivity of a GMM statistic in a neighborhood of \bar{P}_∞ , a fact that is not compatible with Fréchet differentiability (see for instance Heritier and Ronchetti (1994)). Therefore, for the rest of the paper we consider a GMM setting based on a bounded orthogonality function m .

Assumption 4 *The orthogonality function m is such that*

$$\|m\|_\infty := \sup_{(w,\theta) \in \mathbf{W} \times \Theta} \|m(w, \theta)\| < \infty \quad .$$

Under Assumption 4 and further regularity conditions, Fréchet differentiability of the GMM functionals $\tilde{\theta}$ and U can be ensured (see for instance Clarke (1986), Bednarski (1993), Heritier and Ronchetti (1994)). We assume in the sequel the Fréchet differentiability of such functionals⁶ .

Assumption 5 *The functionals $\tilde{\theta}$ and U are Fréchet differentiable.*

The important property of Fréchet differentiable testing functionals for robust inference purposes is their uniform convergence in distribution over asymptotic neighborhoods of the reference model. This feature provides a way to compute uniform asymptotic expansions where the linearized asymptotic sensitivity of the level and the power of the test can be uniformly bounded over neighborhoods of the reference model. In the next section we define asymptotic neighbourhoods

⁶ The Fréchet derivative of functionals $\tilde{\theta}$ and U are computed in Appendix 1 in the proof of Theorem 2.

of the reference model. Then, we address (Sections 3.3 and 3.4) the issue of uniform convergence of robust tests for structural breaks and the uniform expansion of their level and power functionals.

3.2 Asymptotic neighbourhoods

Let P be a probability measure on (Ω, \mathcal{F}) . This will be the reference model in our robust setting.

The next assumption is imposed on the reference model.

Assumption 6 *Under the reference model P , condition (12) is satisfied.*

In particular, recall that under Assumption 6 the reference model satisfies the correct specification hypothesis (5).

In order to study the local stability of GMM tests for structural breaks we now introduce asymptotic neighbourhoods of the reference model P . Without loss of generality we work in the sequel with asymptotic ε -contaminated neighbourhoods. Let \mathcal{M}_∞ be a set of measures satisfying Assumption 2:

$$\mathcal{M}_\infty = \left\{ Q : Q \text{ is a probability measure and } \frac{1}{\pi T} \sum_{t=1}^{T\pi} Q_t \rightarrow \overline{Q}(\pi) \text{ weakly as } T \rightarrow \infty, \right. \\ \left. \text{uniformly in } \pi \in \Pi \right\} .$$

For such measures the uniform weak limit of $\frac{1}{(1-\pi)T} \sum_{t=T\pi+1}^T Q_t$ as $T \rightarrow \infty$ also exists, and is denoted by $\underline{Q}(\pi)$, $\pi \in \Pi$. An ε -contaminated local neighbourhood $\mathcal{U}_{\varepsilon,T}$ of P is defined by:

$$\mathcal{U}_{\varepsilon,T} = \left\{ Q^{\eta,T} = \left(1 - \frac{\eta}{\sqrt{T}} \right) P + \frac{\eta}{\sqrt{T}} Q : Q \in \mathcal{M}_\infty \text{ and } \eta < \varepsilon \right\} . \quad (13)$$

Neighbourhood $\mathcal{U}_{\varepsilon,T}$ represents a set of relevant process distributions close to the reference model P , over which the econometrician desires a smooth behaviour of test statistics for structural breaks. Depending on whether the reference model P satisfies condition (12) with zero or non-zero μ_1, μ_2 functions, set $\mathcal{U}_{\varepsilon,T}$ represents an asymptotic neighbourhood of a model satisfying the null hypothesis of structural stability, or a neighborhood of a local alternative model, respectively. We emphasize that $\mathcal{U}_{\varepsilon,T}$ is a nonparametric neighbourhood of distributions, since virtually no

parametric assumption is imposed on the local deviation directions Q . The only restriction is that any $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$ is a mixture of distributions⁷.

It is important to notice that, in an overidentified setting, for any finite sample size T , local deviations $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$ may or may not admit the existence of a solution for the corresponding finite sample population moment conditions. In the second case, a GMM local misspecification in the sense of Hall and Inoue (2003) arises. In particular, local robustness of our RGMM test statistics for structural breaks ensures automatically stability of level and power under possible local GMM misspecifications.

We conclude this section by presenting a restricted class of local contamination directions Q , in order to provide some more insight into the structure of partial sample asymptotic measures $\bar{Q}(\pi)$, $\pi \in \Pi$, in applied examples. This class will be useful later on to illustrate in a simple setting some of our results. It is characterized by the following Assumption.

Assumption 7 *Measure $Q \in \mathcal{M}_\infty$ satisfies:*

$$E_{\bar{Q}(\pi)} \left[m \left(W, \tilde{\theta}(\bar{P}_\infty) \right) \right] = \gamma(\pi) E_{\bar{Q}_\infty} \left[m \left(W, \tilde{\theta}(\bar{P}_\infty) \right) \right], \quad \pi \in \Pi,$$

for some continuous function $\gamma(\cdot)$ defined on Π .

If Assumption 7 holds, we will always assume that local deviation directions Q spanning neighbourhood $\mathcal{U}_{\varepsilon,T}$ satisfy such assumption.

The class of measures Q satisfying Assumption 7 includes several relevant cases of practical interest. For instance, time homogeneous local deviations Q , such that $\bar{Q}(\pi) = \bar{Q}_\infty$, for all $\pi \in \Pi$, correspond to the case $\gamma(\pi) = 1$, for all $\pi \in \Pi$. Assumption 7 is also satisfied for time non-homogeneous local deviations Q for which only a fraction of the sample is contaminated, as illustrated in the following example.

Example 1 *In this example we consider local deviations involving replacement outliers. Let draws from measure Q be obtained by replacing with probability η^* coordinates W_t of draws from P with independent draws from a distribution $P^* \in \mathcal{B}(\mathbb{R}^k)$, for $\pi = t/T < \pi_0$, where $\pi_0 \in (0, 1)$. Then*

⁷ In particular, when we have a setting with a $(1 - \epsilon)$ percentage of clean data, it is possible to show that such local deviations can be represented as in (13); see Künsch (1984), p. 486.

under Q :

$$Q : \begin{cases} W_t \sim P_t, & \pi \leq \pi_0, \\ W_t \sim (1 - \eta^*) P_t + \eta^* P^*, & \pi > \pi_0. \end{cases}$$

It can be verified that:

$$E_{\overline{Q}(\pi)} \left[m \left(W, \tilde{\theta} \left(\overline{P}_\infty \right) \right) \right] = \begin{cases} 0 & \pi \leq \pi_0, \\ \frac{1 - \pi_0 / \pi}{1 - \pi_0} E_{\overline{Q}_\infty} \left[m \left(W, \tilde{\theta} \left(\overline{P}_\infty \right) \right) \right] & \pi > \pi_0. \end{cases}$$

Hence, this example satisfies Assumption 7 with $\gamma(\pi) = 0$ for $\pi \leq \pi_0$ and $\gamma(\pi) = \frac{1 - \pi_0 / \pi}{1 - \pi_0}$ for $\pi > \pi_0$. Similarly, convex combinations of measures of this kind with different break dates π_0 , or measures where only the first portion π_0 of the sample is contaminated, also satisfy Assumption 7. Finally, the limit case $\pi_0 = 0$ corresponds to a time homogeneous local deviation with replacement outliers.

Local contaminations of the form in Example 1 are relevant in applications. As an illustration, one can consider the situation where a structural break in the parametric part of the model - due for instance to a change in economic policy or a change in the institutional context - is associated with isolated, abrupt movements in the series, caused by some instability in financial markets. These movements correspond to general changes in the distribution of the process, which typically cannot be fully incorporated in the parametric part of the model. In such a situation, it is important to ensure that tests for structural breaks still maintain their power against breaks in the parametric (structural) part of the model.

3.3 Uniform convergence

In this section we provide a uniform convergence result for robust GMM test statistics for structural breaks. The motivation for this result is that uniform convergence ensures a uniform quality of the asymptotic approximation over a relevant set of slightly different model distributions. In particular, uniform convergence gives us a way to control uniformly the stability of the asymptotic level and power functionals of $\hat{\xi}_T^{\text{sup}}$, $\hat{\xi}_T^{\text{ave}}$ and $\hat{\xi}_T^{\text{exp}}$ under sequences of local departures $Q^{\eta, T}$, $T \geq 1$, from the reference model. We first define uniform weak convergence of process $U_T(\cdot)$ as a process indexed by π .

Definition 1 *The sequence $\{U_T(\cdot) : T \geq 1\}$ converges weakly as a process indexed by π to $J_p(\cdot)$, uniformly over the asymptotic neighborhood $\mathcal{U}_\varepsilon = \{U_{\varepsilon, T} : T \geq 1\}$, if*

$$\mathcal{L}^{\eta, T} \left(\sqrt{T} (U_T(\cdot) - U(\cdot, Q^{\eta, T})) \right) \Longrightarrow J_p(\cdot) \quad (T \rightarrow \infty)$$

uniformly in $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$, where $\mathcal{L}^{\eta,T}$ is the process distribution under $Q^{\eta,T}$ whereas $J_p(\cdot)$ is the Brownian Bridge process in Theorem 1.

Note that Definition 1 applies independently of whether P satisfies the null hypothesis H_0 of no break or the alternative hypothesis $H_A(\Pi)$.

Given the Fréchet differentiability Assumption 5, we can assume uniform convergence in distribution of the sequence $\{U_T(\cdot) : T \geq 1\}$ (see Clarke (1986) and Heritier and Ronchetti (1994) for more details on the relation between Fréchet differentiability and uniform convergence in distribution).

Assumption 8 *The sequence $\{U_T(\cdot) : T \geq 1\}$ converges weakly as a process indexed by π to $J_p(\cdot)$, uniformly over the asymptotic neighborhood $\mathcal{U}_{\varepsilon}$.*

The Fréchet differentiability Assumption 5 and the uniform convergence Assumption 8 imply the following uniform convergence of the RGMM test statistics for breaks over asymptotic neighborhoods of the reference model.

Theorem 2 *Under Assumptions 4, 5, 6, 8 it follows that:*

1. $\widehat{LM}_T(\cdot) \implies Q_p^*(\cdot)$ as a process indexed by $\pi \in \Pi$, uniformly in $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$, where

$$Q_p^*(\pi) = \frac{1}{\pi(1-\pi)} (J_p(\pi) + b^*(\pi))' (J_p(\pi) + b^*(\pi)),$$

with:

$$b^*(\pi) = \pi(1-\pi) H(\overline{P}_{\infty})^{1/2} (\mu_1^*(\pi) - \mu_2^*(\pi)) \quad ,$$

and:

$$\mu_1^*(\pi) = \mu_1(\pi) + \eta E_{\overline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right), \quad \mu_2^*(\pi) = \mu_2(\pi) + \eta E_{\underline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right).$$

2. $\widehat{\xi}_T^{\sup} \rightarrow_d \sup_{\pi \in \Pi} Q_p^*(\pi)$, uniformly in $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$.
3. $\widehat{\xi}_T^{\text{ave}} \rightarrow_d \int_{\Pi} Q_p^*(\pi) d\lambda(\pi)$, uniformly in $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$.
4. $\widehat{\xi}_T^{\text{exp}} \rightarrow_d \log \int_{\Pi} \exp \left(\frac{1}{2} Q_p^*(\pi) \right) d\lambda(\pi)$, uniformly in $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$.

Proof. See Appendix 1. ■

The local contamination direction Q affects the asymptotic distribution of statistics $\widehat{\xi}_T^{\sup}$, $\widehat{\xi}_T^{\text{ave}}$, $\widehat{\xi}_T^{\text{exp}}$, in particular their asymptotic level and power, through the additional term

$$\begin{aligned} & \eta\pi(1-\pi) H(\overline{P}_{\infty})^{1/2} \left\{ E_{\overline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right) - E_{\underline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right) \right\} \\ = & \eta\pi H(\overline{P}_{\infty})^{1/2} \left\{ E_{\overline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right) - E_{\underline{Q}(\pi)} \left(m \left(W, \tilde{\theta}(\overline{P}_{\infty}) \right) \right) \right\} \quad , \end{aligned}$$

in the non-centrality vector $b^*(\pi)$. This term involves the difference between the moment conditions computed on a portion of the sample and the moment conditions computed on the full sample, evaluated at the reference model parameter $\tilde{\theta}(\bar{P}_\infty)$. In general, time non-homogeneous local deviation directions Q such that $E_{\bar{Q}(\pi)}\left(m\left(W, \tilde{\theta}(\bar{P}_\infty)\right)\right)$ varies with π affect asymptotically the level and the power of test statistics for structural breaks. Conversely, time homogeneous local deviations Q such that $\bar{Q}(\pi) = \bar{Q}_\infty$, $\pi \in \Pi$, have no asymptotic impact. This is a consequence of the fact that functional $U(\cdot, Q)$ in (7) is equal to zero for all measures such that $\bar{Q}(\pi) = \bar{Q}_\infty$, $\pi \in \Pi$. We emphasize that Theorem 2 provides *uniform* convergence results over neighbourhoods of the reference model, which guarantee the stability of level and power of the RGMM test statistics for structural breaks uniformly over small local deviations from the reference model. For instance, Theorem 2 implies that time homogeneous local deviations do not affect asymptotically size and power of test statistics, *uniformly* in a neighbourhood of P . It is important to stress that this result only holds under Assumption 4 of a bounded orthogonality function m . With an unbounded m function, these convergence results can hold only pointwise with respect to local deviation directions $Q^{\eta, T}$. This implies that, for any fixed sample size T , the distortion of the level and power of test statistics based on unbounded orthogonality functions m may become arbitrary large for some local deviation $Q^{\eta, T}$ very close (in distribution) to the reference model P . Therefore, in finite samples we expect more stable level and power properties across different local deviations for GMM test statistics based on a bounded orthogonality function m .

When the local deviation direction Q satisfies Assumption 7, the non-centrality vector b^* gets the more explicit representation:

$$b^*(\pi) = \pi(1 - \pi) H(\bar{P}_\infty)^{1/2} [\mu_1(\pi) - \mu_2(\pi)] + \eta\pi [\gamma(\pi) - 1] d(\bar{P}_\infty, \bar{Q}_\infty) \quad ,$$

where

$$d(\bar{P}_\infty, \bar{Q}_\infty) = H(\bar{P}_\infty)^{1/2} E_{\bar{Q}_\infty} \left(m \left(W, \tilde{\theta}(\bar{P}_\infty) \right) \right).$$

In this particular case, a time non-homogeneous local deviation Q can affect asymptotically the level or the power of statistics for structural breaks if and only if $E_{\overline{Q}_\infty} \left(m \left(W, \tilde{\theta} \left(\overline{P}_\infty \right) \right) \right) \neq 0$, that is if the orthogonality conditions at $\tilde{\theta} \left(\overline{P}_\infty \right)$ are not satisfied under \overline{Q}_∞ .

3.4 Asymptotic expansions of level and power functionals

In this section we provide a uniform expansion for small contamination amounts ε of the asymptotic level and power functionals of robust GMM tests for structural breaks, over asymptotic neighborhoods of the reference model P .

Let us first consider the level functional. In this case the reference model P satisfies Assumption 6 with $\mu_1 = \mu_2 = 0$ in (12). The asymptotic level functional is defined by:

$$\lim_{T \rightarrow \infty} \alpha \left(Q^{\eta, T} \right) = \lim_{T \rightarrow \infty} Q^{\eta, T} \left(\widehat{\xi}_T > \xi_0 \right),$$

where $\widehat{\xi}_T$ is any of the supremum, average or exponential statistics (9), (10) or (11), respectively, and ξ_0 is the corresponding critical value for a given nominal size α_0 , computed from Theorem 1 (or Theorem 2) with $b = 0$ ($b^* = 0$, respectively). From Theorem 2 it follows that the level converges to $\mathcal{L}^* \left(\xi_0 \right)$, where $\mathcal{L}^* \left(\cdot \right)$ is the cumulative distribution function of the limit variables in points 2., 3., or 4. of Theorem 2, uniformly in $Q^{\eta, T} \in \mathcal{U}_{\varepsilon, T}$. The general analysis of level distortions induced by local deviations follows the line of the discussion of Theorem 2. In some particular settings, it is possible to give a more precise characterizations of such level distortions. This is the case for instance when Assumption 7 is satisfied. Since the direction of the noncentrality parameter $b^* \left(\pi \right)$ does not depend on π (Theorem 2), and:

$$\|b^* \left(\pi \right)\|^2 = \eta^2 \pi^2 \left[1 - \gamma \left(\pi \right) \right]^2 \|d \left(\overline{P}_\infty, \overline{Q}_\infty \right)\|^2,$$

the distribution of stochastic process $Q_p^* \left(\cdot \right)$ depends on local deviation direction \overline{Q}_∞ only through⁸ $\eta^2 \|d \left(\overline{P}_\infty, \overline{Q}_\infty \right)\|^2$. Denote by $Q_p^* \left(\pi, \eta^2 \|d \left(\overline{P}_\infty, \overline{Q}_\infty \right)\|^2 \right)$, $\pi \in \Pi$, a process with such a distribution. We provide a theorem for the asymptotic level expansion of the robust GMM statistics

⁸ As in the standard argument for the distribution of a chi-square variable, we use the fact that there exists a rotation matrix R such that $Q_p^* \left(\pi \right) = \left(R J_p \left(\pi \right) + R b^* \left(\pi \right) \right)' \left(R J_p \left(\pi \right) + R b^* \left(\pi \right) \right)$, $R b^* \left(\pi \right) = \|b^* \left(\pi \right)\| \left(1, 0, \dots, 0 \right)'$ and $R \cdot J_p$ is a Brownian Bridge.

for structural breaks under small contamination amounts ε . We focus in the exposition on the average statistic, the case of supremum or exponential statistics being completely similar.

Theorem 3 *Under Assumptions 4-8 it follows that, for any $\eta < \varepsilon$:*

$$\lim_{T \rightarrow \infty} \alpha(Q^{\eta, T}) = \alpha_0 + \eta^2 \mu \cdot \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 + o(\eta^2),$$

uniformly in $Q^{\eta, T} \in \mathcal{U}_{\varepsilon, T}$, where $\mu = -\partial \mathcal{L}(\xi_0, y) / \partial y|_{y=0}$, $\mathcal{L}(\cdot, y)$ is the cumulative distribution function of the random variable

$$\int_{\Pi} Q_p^*(\pi, y) d\lambda(\pi),$$

and:

$$d(\bar{P}_\infty, \bar{Q}_\infty) = H(\bar{P}_\infty)^{1/2} E_{\bar{Q}_\infty} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right]. \quad (14)$$

Proof. See Appendix 2. ■

It is possible to provide the corresponding theorem for the asymptotic expansion of the power functional, defined by:

$$\lim_{T \rightarrow \infty} \pi(Q^{\eta, T}) = \lim_{T \rightarrow \infty} Q^{\eta, T}(\hat{\xi}_T > \xi_0),$$

where now the reference model P satisfies Assumption 6 with some given non-zero functions μ_1, μ_2 .

Let $Q_p^\#(\pi, y)$ denote the random variable $Q_p^\#(\pi, y) = (J_p(\pi) + b(\pi, y))' (J_p(\pi) + b(\pi, y))$, where

$$b(\pi, y) = \pi(1 - \pi) H(\bar{P}_\infty)^{1/2} [\mu_1(\pi) - \mu_2(\pi)] + \pi[\gamma(\pi) - 1]y, \quad y \in \mathbb{R}^v.$$

Then, we have the following power counterpart of Theorem 3.

Theorem 4 *Under Assumptions 4-8 it follows that, for any $\eta < \varepsilon$:*

$$\lim_{T \rightarrow \infty} \pi(Q^{\eta, T}) = \pi(\bar{P}_\infty) + \eta \mu' \cdot d(\bar{P}_\infty, \bar{Q}_\infty) + o(\eta),$$

uniformly in $Q^{\eta, T} \in \mathcal{U}_{\varepsilon, T}$, where $\mu = -\partial \mathcal{L}(\xi_0, y) / \partial y|_{y=0}$, $\mathcal{L}(\cdot, y)$ is the cumulative distribution function of the random variable

$$\int_{\Pi} Q_p^\#(\pi, y) d\lambda(\pi),$$

and $d(\bar{P}_\infty, \bar{Q}_\infty)$ is given in (14).

Proof. See Appendix 2. ■

Versions of Theorems 3 and 4 for the supremum and exponential statistics are completely similar,

with $\mathcal{L}(\cdot, y)$ the cumulative distribution function of the random variables $\sup_{\pi \in \Pi} Q_p^a(\pi, y)$ and

$\log \int_{\Pi} \exp(\frac{1}{2} Q_p^a(\pi, y)) d\lambda(\pi)$, $a = *, \#$, respectively.

Theorems 3 and 4 show that the asymptotic linearized distortion in the level and power of tests for structural breaks is proportional to $\|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2$ and $d(\bar{P}_\infty, \bar{Q}_\infty)$, respectively. In particular, if m is bounded, for any given function γ in Assumption 7 the distortion in the level or power of supremum, average and exponential tests for breaks is uniformly bounded over asymptotic neighborhoods of the reference model. In this setting, this implies the robustness of GMM tests for breaks based on bounded orthogonality functions.

Theorem 3 can be used to give uniform asymptotic bounds on the maximal sensitivity in the level of tests based on a bounded orthogonality function. In particular, an orthogonality function such that

$$\sup_{(w, \theta) \in \mathbf{W} \times \Theta} |m(w, \theta)' H(\bar{P}_\infty) m(w, \theta)| < c^2 \quad , \quad (15)$$

for some constant $c > \sqrt{v}$, implies up to terms of uniform order $o(\eta^2)$:

$$\left| \lim_{T \rightarrow \infty} \alpha(Q^{\eta, T}) - \alpha_0 \right| \leq \eta^2 \mu \cdot c^2 \quad . \quad (16)$$

The tuning constant c of our RGMM estimators determines their degree of robustness. In applications c has to be determined by the econometrician, for instance on the basis of some prior information about a maximal realistic extent η of deviation from the reference model which can be expected in the data. A lower constant c implies a higher robustness under a departure from the reference model. For testing purposes, the bounds (16) can be used to choose the constant c in dependence of the maximal amount of contamination expected (ε) and the maximal distortion in the asymptotic level which a researcher is willing to accept. In this case, the derivatives μ will have to be computed numerically, by simulating the distribution of

$$\sup_{\pi \in \Pi} Q_p^*(\pi, y) \quad , \quad \int_{\Pi} Q_p^*(\pi, y) d\lambda(\pi) \quad , \quad \log \int_{\Pi} \exp\left(\frac{1}{2} Q_p^*(\pi, y)\right) d\lambda(\pi) \quad ,$$

for several values of y in a neighborhood of 0. For instance, the local robustness of the level of tests for structural breaks could be studied in dependence of c , thereby producing information about the degree of asymptotic local stability in the level required for a particular model setting. Notice

that in the case where $\Pi = \{\pi_0\}$ and $\lambda = \delta_{\pi_0}$ (the Dirac distribution at π_0) the above tests collapse to a test for a break at a known break date. In this case Theorem 3 and Theorem 4 coincide with Theorem 1 and Theorem 3 in Ronchetti and Trojani (2001), where the local robustness of the level and power functionals for standard Maximum Likelihood-type GMM tests has been characterized. Since in this case the distribution of the random variable $Q_p^*(\pi_0, y)$ is noncentral chi-square with noncentrality parameter y , full analytical expressions for μ become available.

The bound (15) is satisfied by RGMM estimators with bounded self-standardized sensitivity, i.e. such that⁹

$$\sup_{(w, \theta) \in \mathbf{W} \times \Theta} \left| m(w, \theta)' S(\bar{P}_\infty)^{-1} m(w, \theta) \right| < c^2 \quad . \quad (17)$$

An analogous result applies for the power of RGMM tests for structural breaks based on a bounded orthogonality function m satisfying (17). Therefore, we consider supremum, average and exponential tests for breaks based on such RGMM estimators and their orthogonality functions. Details on the definition and the computation of such RGMM estimators and their orthogonality functions in a general GMM setting are provided in Ronchetti and Trojani (2001), p. 45-48.

4 Monte Carlo simulations

Having presented in the last section the theoretical background of RGMM tests for structural breaks, in this section we report a series of Monte Carlo simulations in order to evaluate their finite sample level and power properties across different model settings. We compare the performance of RGMM tests with the one of classical GMM tests by focusing on the stability of power and level under local departures from a given reference model. We first provide results for relatively large samples sizes using standard asymptotic critical values. In a second step, we also present results for bootstrapped versions of the tests in small samples. It is known that bootstrap procedures can provide very accurate refinements of the finite sample distribution of classical tests for structural

⁹ This follows from the orthogonal projection property of the matrix

$$S(\bar{P}_\infty)^{\frac{1}{2}} H(\bar{P}_\infty) S(\bar{P}_\infty)^{\frac{1}{2}} \quad .$$

breaks; see for instance Diebold and Chen (1996) for such an evidence in a simple linear model setting. Our simulations investigate how far bootstrapping RGMM statistics for structural breaks can help providing a uniform bootstrap performance over a relevant model neighbourhood.

4.1 Testing for structural breaks in a linear regression model

We first consider tests for a break in the slope coefficient of a linear regression model with an autoregressive regressor. The model is given by:

$$\begin{aligned} y_t &= \gamma + \beta_t x_t + \sigma u_t \\ x_t &= \alpha + \rho x_{t-1} + \sigma_\varepsilon \varepsilon_t \end{aligned}, \tag{18}$$

where:

$$\beta_t = \begin{cases} \beta_1, & \text{for } t = 1, \dots, T\pi_0 \\ \beta_2, & \text{for } t = T\pi_0 + 1, \dots, T \end{cases},$$

for some $\pi_0 \in \Pi$. The error term ε_t in the process x_t is *i.i.d.* $N(0, 1)$ distributed. For the error term u_t in the linear regression model (18) we simulate a set of different distributions according to Model 1a-1e below. Specifically, we set:

- **Model 1a:** $u_t \sim i.i.d. N(0, 1)$,
- **Model 1b:** $u_t \sim i.i.d. t_5/\sqrt{5/3}$,
- **Model 1c:** $u_t \sim i.i.d. t_3/\sqrt{3}$,
- **Model 1d:** $u_t \sim i.i.d. CN(0.05, 0, 3)/\sqrt{1.4}$,
- **Model 1e:** $u_t \sim i.i.d. CN(0.1, 0, 3)/\sqrt{1.8}$,

where t_n is a Student distribution with n degrees of freedom and $CN(x, 0, 3)$ is a standard normal distribution contaminated with probability x by a further zero mean normal distribution having standard deviation 3. All error distributions have been standardized. The standard orthogonality conditions for a least squares estimation of the linear regression model (18) are based on an

orthogonality function given by

$$\psi(W_t, \beta, \delta) = \begin{pmatrix} \psi_1(W_t, \beta, \delta) \\ \psi_2(W_t, \beta, \delta) \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 1 \\ x_t \end{pmatrix} (y_t - \gamma - \beta x_t) \\ (y_t - \gamma - \beta x_t)^2 / \sigma^2 - 1 \end{pmatrix},$$

where $W_t = (y_t, x_t)'$, $\delta = (\gamma, \sigma^2)'$. Classical GMM estimators are obtained by using such orthogonality conditions. In particular, the model is exactly identified and under a Gaussian error distribution the estimating function ψ defines the maximum likelihood estimator of θ . Therefore, under Gaussianity of the error distributions we can expect classical tests based on the average or the exponential statistics to provide the highest power (see again Andrews and Ploberger (1994)). However, since ψ is unbounded, maximum likelihood estimators and tests based on such an estimating function are not robust. RGMM estimators of (18) can be constructed by applying orthogonality conditions based on a truncated orthogonality function given by

$$m(W_t, \beta, \delta) = \begin{pmatrix} m_1(W_t, \beta, \delta) \\ m_2(W_t, \beta, \delta) \end{pmatrix} = \begin{pmatrix} A_1 \psi_1(W_t, \beta, \delta) w_{c_1} (A_1 [\psi_1(W_t, \beta, \delta)]) \\ A_2 [\psi_2(W_t, \beta, \delta) - \tau_2] w_{c_2} (A_2 [\psi_2(W_t, \beta, \delta) - \tau_2]) \end{pmatrix}, \quad (19)$$

where $w_c(z) := \min(1, c/\|z\|)$ defines a set of Huber's weights that downweight observations which are influential (in terms of asymptotic bias) for a classical least squares estimation of the model; see also Hampel et al. (1986), Section 4.4., for more details. The constants $c_1 > \sqrt{2}$, $c_2 > 1$ are tuning constants that control the amount of robustness in the estimation of $(\gamma, \beta)'$ and σ^2 , respectively. The matrix $A_1 \in \mathbb{R}^{2 \times 2}$ and the scalars A_2, τ_2 are determined as the solution of the implicit equations¹⁰ :

$$0 = E_{P_0} [m(W_t, \beta_0, \delta_0)] \quad , \quad (20)$$

$$I = E_{P_0} \left[m(W_t, \beta_0, \delta_0) m(W_t, \beta_0, \delta_0)' \right] \quad , \quad (21)$$

¹⁰ Shift factor τ_2 is introduced to ensure Fisher consistency at the reference model P_0 , see equation (20). In particular, no such shift factor is needed for orthogonality function m_1 , since ψ_1 is a symmetric random variable under P_0 . Matrix A_1 and scalar A_2 are normalization factors, which ensure that the self-standardized sensitivity of the RGMM estimator is bounded by $c := \sqrt{c_1^2 + c_2^2}$; see also equation (17).

where P_0 is the reference model distribution of a linear regression model (18) with normally distributed¹¹ error terms u_t , that is Model 1a, and $\theta_0 = \tilde{\theta}(P_0)$.

We emphasize that the reference model distribution P_0 is used in the definition of RGMM statistics only to define a truncated GMM orthogonality condition which, by virtue of the uniform asymptotic results in Theorem 12, 13 and 14, ensures a uniform asymptotic behavior of the statistics $\widehat{\xi}_T^{\text{sup}}$, $\widehat{\xi}_T^{\text{ave}}$, $\widehat{\xi}_T^{\text{exp}}$ under local departure from model P_0 . In particular, when we simulate - for instance - under a student t_7 or t_5 distribution in Model 1b or 1c, the student t distribution is by no way used to define the RGMM moment conditions in such model setting. Instead, the robustified bounded orthogonality function m is still computed using only the structure of the given reference P_0 , according to (20), (21).

In addition to Models 1a-1e, we also consider a further model of local contamination, where replacement outliers invalidate for any finite sample size T the population moment conditions evaluated at the solution $\tilde{\theta}(P_0)$ implied by the reference model P_0 . Specifically, we study the effects of the following time homogeneous outliers replacement model:

- **Model 1f:** *the observations of a sample y_t , $t = 1, \dots, T$ from model 1a are replaced with probability 0.05 with an outlier $y_t \sim 3N(0, 1)$.*

Model 1f corresponds to a time homogeneous local deviation as in Example 1 with $\pi_0 = 0$. Since the GMM moment conditions are exactly identified, no GMM misspecification is induced.

In the given setting, RGMM tests based on the bounded orthogonality function m will suffer an efficiency loss under an exact Gaussian distribution for u_t . However, already under slight departures from Gaussianity they can provide a higher power of inferences on structural breaks. In the next subsections, we study these issues by simulation for the above Models 1a-1f of local departure from a conditionally Gaussian linear regression model.

¹¹ Specifically, since under P_0 the random variable $(y_t - \gamma_0 - \beta_0 x_t)^2 / \sigma_0^2$ is χ_1^2 distributed, irrespectively of (β_0, δ_0) , the correction constants A_2, τ_2 have to be computed only once, before starting the robust estimation algorithm.

4.1.1 Asymptotic tests

Tables 1a-1f report the results of our Monte Carlo simulations for Models 1a-1f in the given linear regression setting. The break date is fixed at $t_0 = 0.5 \cdot T$, where $T = 100, 200, 300$, respectively. In the simulations for $T = 200$ we have set $\alpha = \gamma = 0$, $\beta_1 = 1$, $\beta_2 = 1, 1.1, 1.2, 1.3$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$ ¹². In the simulations for the other sample sizes the values of $\beta_2 - \beta_1$ applied for $T = 200$ have been multiplied by a factor $\sqrt{200/T}$ in order to obtain comparable local alternatives across the different sample sizes. We also fixed $\Pi = [p_0, 1 - p_0]$, where $p_0 = 0.25$. We provide the results for the supremum and average classical GMM statistics, and for the corresponding RGMM statistics using the bounded orthogonality function (19). Though available on request, we omit the results for the exponential statistics, since they are very close to those of the average statistics across all designs. Finally, the tuning constants for the RGMM test have been set at $c_1 = 3, c_2 = 3$.

Tables 1a-1f about here

Table 1a shows that the power loss under normality of supremum and average RGMM tests for breaks is moderate, with losses relatively to the classical GMM tests that are typically below 10%. As expected, the power of classical and robust average tests is above the one of tests based on the supremum functional. Table 1b shows that already under a t_5 error distribution the power curves of classical and robust GMM tests are very similar. In Table 1c, under a t_3 error distribution, these issues are more pronounced with a clearly higher power of robust GMM tests relatively to standard procedures. In this model setting some oversize of classical average tests arises. For example, the empirical sizes of the classical GMM tests are above 8.5% for $T = 100$. Table 1d shows similar patterns as for a t_5 distribution when simulating under a $CN(0.05, 0, 3)$ error distribution: the power curves of classical and robust GMM tests under such a setting are very similar and the classical average test shows a tendency to a slight oversize when $T = 100$. Under the $CN(0.1, 0, 3)$ setting in Table 1e this last pattern is more pronounced, and RGMM tests improve on the power

¹² Simulations for more extreme autocorrelation coefficients $\rho = 0.2, 0.8$ give similar findings as for $\rho = 0.5$. These results are available from the authors on request.

of classical GMM procedures.

Finally Table 1f shows that the empirical size of GMM and RGMM statistics under replacement outliers is for $T = 300$ quite accurate. At the same time, RGMM tests provide for $T = 300$ a higher power, both for the supremum and the average statistics. For moderate sample sizes $T = 100, 200$, RGMM tests based on the average statistics still perform satisfactorily. On the other hand, classical GMM tests based on the average tend to produce a slight oversize which induces an "artificial" power increase. This pattern is confirmed by our bootstrap results for sample size $T = 50$, presented in the next section. GMM and RGMM results for the supremum tests and sample sizes $T = 100, 200$, show that all statistics produce an undersize which reduces uniformly the finite sample power of the tests. This pattern is slightly more pronounced for the robust tests. Moreover, it seems to be due to a pure finite sample effect which disappears when applying bootstrap resampling methods that provide a more accurate finite sample inference. This is confirmed by our bootstrap results for sample size $T = 50$, presented in the next section.

The observed stability of level and power of RGMM statistics - in particular those based on the average functional - across different local deviations from the Gaussian reference model is a consequence of their uniform convergence (see Theorem 2 and the discussion thereafter). Uniform convergence implies that, for any finite sample size T , the distortion of level and power is bounded uniformly across local deviations from the reference model, whereas they can be arbitrarily large in some direction for the classical GMM statistics.

4.1.2 Bootstrap tests

The two last panels of Tables 1a-1f present for a sample size $T = 50$ the results implied by GMM and RGMM tests for breaks based on (i) standard asymptotic critical values and (ii) bootstrapped versions of the relevant GMM statistics. In all simulations we used 1000 bootstrap samples. Generally, from these results we see that the finite sample size of GMM and RGMM asymptotic tests based on the supremum functional can be quite biased relatively to the given nominal level.

Indeed, for basically all Models 1a-1f we observe an empirical size of such tests which is systematically below the correct nominal level of 5%. Similarly, classical asymptotic GMM tests based on the average tend to produce a clear oversize relatively to the correct nominal level. By contrast, RGMM tests based on the average seem to control quite well the empirical sizes across all Models 1a-1f. Bootstrapped versions of GMM and RGMM tests for $T = 50$ are found to provide very accurate finite sample sizes, thus correcting the distortion observed for the asymptotic tests discussed previously. This is consistent with the findings of Diebold and Chen (1996). However, more strikingly, it appears that bootstrapped versions of the RGMM tests tend to produce a uniformly higher power across all Models 1a-1f. For instance, the power of bootstrapped RGMM tests under Model 1a for $\beta_2 = 1.3$ is not smaller than the one of bootstrapped classical GMM tests for both the supremum and the average statistics. Under local deviations from Gaussianity of the errors in the regression model (Table 1b-1f) such power increases appear to be often quite substantial, as for instance under a student t_3 error distribution in Table 1c.

4.2 Testing for structural breaks in an ARCH model

In order to investigate the properties of classical and robust GMM tests for structural breaks in a nonlinear model, we now consider an ARCH model setting (Engle (1982)). Specifically we analyze tests for breaks in the autoregressive coefficient of the conditional variance equation in an ARCH(1) model.

The model specification is given by:

$$\begin{aligned} y_t &= \sigma_t u_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_{1,t} y_{t-1}^2, \end{aligned}$$

where

$$\alpha_{1,t} = \begin{cases} \alpha_1, & \text{for } t = 1, \dots, T\pi_0 \\ \alpha_2, & \text{for } t = T\pi_0 + 1, \dots, T \end{cases},$$

for some $\pi_0 \in \Pi$. For u_t we simulate again a set of distributions near to a standard normal one,

according to Models 2a-2e below:

- **Model 2a:** $u_t \sim i.i.d. N(0, 1)$,
- **Model 2b:** $u_t \sim i.i.d. t_7/\sqrt{7/5}$
- **Model 2c:** $u_t \sim i.i.d. t_5/\sqrt{5/3}$
- **Model 2d:** $u_t \sim i.i.d. CN(0.05, 0, 3)/\sqrt{1.4}$
- **Model 2e:** $u_t \sim i.i.d. CN(0.1, 0, 3)/\sqrt{1.8}$

All error distributions have been standardized. As for the previous section, we also consider replacement outliers which destroy the structure of the model by distorting the GMM parameter estimate which satisfies the population orthogonality conditions for a finite sample size T :

- **Model 2f:** *the observations from a sample y_t , $t = 1, \dots, T$, of model 2a are replaced with probability 0.025 with an outlier $y_t \sim N(0, 4\sqrt{\alpha_0/(1 - \alpha_1)})$*

The orthogonality conditions for a classical GMM estimation of the model are defined by an orthogonality function given by:

$$\psi(W_t, \beta, \delta) = \begin{pmatrix} 1 \\ y_{t-1}^2 \end{pmatrix} \frac{1}{\sigma_t^2} \left(\frac{y_t^2}{\sigma_t^2} - 1 \right) \quad ,$$

where $\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2$, $W_t = (y_t, y_{t-1})$, $\beta = \alpha_1$, $\delta = \alpha_0$. Similarly to the previous simulation setting, under a Gaussian error distribution ψ defines a maximum-likelihood estimator of θ . However, this orthogonality function is unbounded, so that the implied GMM estimators are not robust. The orthogonality function for a robust GMM estimation of the model is given by:

$$m(W_t, \beta, \delta) = A [\psi(W_t, \beta, \delta) - \tau(y_{t-1})] w_c(A [\psi(W_t, \beta, \delta) - \tau(y_{t-1})]) \quad , \quad (22)$$

for some tuning constant $c > \sqrt{2}$. The matrix $A \in \mathbb{R}^{2 \times 2}$ and the vector function τ are defined by

the implicit equations¹³ :

$$\begin{aligned} 0 &= E_{P_0} \left[m(W_t, \beta_0, \delta_0) \mid \underline{y_{t-1}} \right] \quad , \\ I &= E_{P_0} \left[m(W_t, \beta_0, \delta_0) m(W_t, \beta_0, \delta_0)' \right] \quad , \end{aligned} \tag{23}$$

where P_0 is the reference model distribution of an ARCH(1) model with conditionally normally distributed error terms u_t (Model 2a). The shift factors $\tau(y_{t-1})$, $t = 1, \dots, T$, can be computed by using an analytical Laplace approximation of the integrals involved in the solution of (23), as proposed in Mancini, Ronchetti and Trojani (2003). This avoids the numerical computation of such integrals and largely reduces the computation time of robust GMM estimators in the present and related settings. Details on the computation of $\tau(y_{t-1})$ and the corresponding robust GMM estimator for the moment conditions (23) are given in Appendix 3.

4.2.1 Asymptotic tests

Tables 2a-2f present the results of our Monte Carlo simulations for Models 2a-2f in the given ARCH(1) model setting. The break date is fixed at $t_0 = 0.5 \cdot T$, where $T = 250, 500, 1000$. In the simulations for $T = 1000$ we have set $\alpha_0 = 0.01, \alpha_1 = 0.6, \alpha_2 = 0.6, 0.7, 0.8, 0.9$. In the simulations for the other sample sizes the values of $\alpha_2 - \alpha_1$ applied for $T = 1000$ have been multiplied by a factor $\sqrt{1000/T}$ in order to obtain comparable local alternatives across the different sample sizes. The tuning constants for the RGMM test have been set at $c = 6.18$. We also fixed $\Pi = [p_0, 1 - p_0]$, where $p_0 = 0.45$. The nominal level of the test is 5%.

Tables 2a-2f about here

Table 2a shows that the power loss under normality of average and exponential RGMM test for breaks is moderate and always below 10%. Moreover, even under normality, the power of RGMM supremum tests is above the one of classical GMM procedures. As expected, the power of classical and robust average and exponential tests is above the one of tests based on the supremum

¹³ Similarly to the previous simulation setting, shift vectors $\tau(y_{t-1})$, $t = 2, \dots, T$, ensure conditional Fisher consistency of RGMM estimators for ARCH models at the reference model P_0 . The normalization matrix A ensures that the self-standardized sensitivity of the RGMM estimator is bounded by c ; see also equation (17).

functional. Table 2b shows that under a t_7 error distribution the gain in power of RGMM tests based on the supremum is very large, with relative increases that are sometimes around 80%-100%. Also in the case of the average and exponential statistics RGMM procedures do produce clear power increases in this setting. Such patterns are even more apparent in Table 2c, under a t_5 distribution, where average and supremum RGMM statistics yield very large power improvements. The results in Tables 2d and 2e (for a $CN(0.05, 0, 3)$ and a $CN(0.1, 0, 3)$ error distribution, respectively) are qualitatively similar to those of Table 2b and 2c, with effects that are however quantitatively even larger than in the case of a Student t error distribution. Finally, Table 2f shows the results under the outlier replacement Model 2f. These findings further confirm the large power improvement of RGMM tests.

4.2.2 Bootstrap tests

The last two panels of Tables 2a-2f present for a sample size $T = 125$ the results implied by GMM and RGMM tests for breaks based on (i) standard asymptotic critical values and (ii) bootstrapped versions of the relevant GMM statistics. Generally, from these results we see that the finite sample size of GMM and RGMM asymptotic tests for all Model 2a-2f can be quite biased downwards relatively to the given nominal level.

Bootstrapped versions of GMM and RGMM tests for $T = 125$ provide very accurate finite sample sizes, thus correcting very well the bias observed for all asymptotic tests discussed previously. Similarly to the previous section, we observe that bootstrapped versions of the RGMM tests tend to produce a uniformly higher power across all Models 2a-2f. Under local deviations from conditional Gaussianity in the ARCH(1) model such power increases appear to be often quite substantial, as for instance under the contaminated normal model in Table 2e.

4.3 Testing for structural breaks in overidentified models

In this last section we investigate the properties of classical and robust GMM tests for structural breaks in an overidentified model. Such models are important for applications, since they are

the basis of many structural economic specifications, such as for instance regression models with more instruments than endogenous variables or intertemporal models with more Euler equations than structural parameters. As it is well-known, a major concern with GMM inference in overidentified models is the finite sample performance of GMM estimators and tests, especially when the number of orthogonality conditions is large (see for instance Altonji and Segal (1996), Burnside and Eichenbaum (1996) and Hansen, Heaton and Yaron (1996)). In order to address this point, we consider testing for structural breaks in a standard specification from the literature on finite sample properties of GMM, proposed by Burnside and Eichenbaum (1996). The model is characterized by the orthogonality conditions:

$$E [W_t - \beta_t \iota] = 0,$$

where $W_t = (W_{1,t}, \dots, W_{d,t})'$ is a d -dimensional vector of independent random variables, $d = 10$, $\iota = (1, \dots, 1) \in \mathbb{R}^d$, and β_t is a scalar parameter such that:

$$\beta_t = \begin{cases} \beta_1, & \text{for } t = 1, \dots, T\pi_0 \\ \beta_2, & \text{for } t = T\pi_0 + 1, \dots, T \end{cases},$$

for some $\pi_0 \in \Pi$. We test therefore for a break in the common mean β of a set of variables $W_{1,t}, \dots, W_{d,t}$. For each component $W_{i,t}$ we simulate a set of distributions close to the chi-square distribution with one degree of freedom:

$$W_{i,t} = \beta_t u_{i,t}^2,$$

according to the following models:

- **Model 3a:** $u_{i,t} \sim i.i.d. N(0, 1)$,
- **Model 3b:** $u_{i,t} \sim i.i.d. t_7 / \sqrt{7/5}$,
- **Model 3c:** $u_{i,t} \sim i.i.d. t_5 / \sqrt{5/3}$,
- **Model 3d:** $u_{i,t} \sim i.i.d. CN(0.05, 0, 3) / \sqrt{1.4}$,

- **Model 3e:** $u_{i,t} \sim i.i.d. CN(0.1, 0, 3) / \sqrt{1.8}$.

The orthogonality function in Burnside and Eichenbaum (1996) model is given by:

$$\psi(W_t, \beta) = W_t - \beta\iota,$$

and is unbounded, implying nonrobust GMM statistics. Moreover, even under a Gaussian distribution for $u_{i,t}$, the GMM estimator implied by ψ is not a maximum likelihood estimator. A RGMM estimator can be constructed by truncating the orthogonality conditions:

$$m(W_t, \beta) = A(W_t - \beta\iota - \tau) w_c(A[W_t - \beta\iota - \tau]),$$

for some tuning constant c . Vector τ and matrix A are defined by the implicit equations:

$$\begin{aligned} 0 &= E_{P_0} [m(W_t, \beta_0)] \quad , \\ I &= E_{P_0} [m(W_t, \beta_0)m(W_t, \beta_0)'] \quad , \end{aligned}$$

where P_0 is the reference model distribution¹⁴ corresponding to a chi-square distribution $\chi^2(1)$ for the variables $W_{i,t}/\beta_0$. All above models of departure from a Gaussian distribution for u_t satisfy the moment conditions implied by the orthogonality function ψ .

Finally, we consider local models of contamination which invalidate for any finite sample size T the population moment conditions implied by the orthogonality function ψ . Specifically, we study the effect of the two following outlier replacement models:

- **Model 3f:** *the observations of components $i = 1, \dots, 5$ of a sample $W_{i,t}$, $t = 1, \dots, T$ from model 3a are replaced with probability 0.05 by an outlier $W_{i,t} \sim (3N(0, 1))^2$.*
- **Model 3g:** *the observations of components $i = 1, \dots, 5$ of a sample $W_{i,t}$, $t = 1, \dots, T$ from model 3a are replaced with probability η by an outlier $W_{i,t} \sim (3N(0, 1))^2$ for $t = 1, \dots, 0.5T$.*

¹⁴ As in section 4.1 vector τ ensures Fisher consistency at the reference model and has to be computed only once, before starting the robust estimation algorithm. Matrix A is computed in the robust estimation algorithm, using an empirical version of the second implicit equation, and ensures a self-standardized sensitivity of the RGMM estimator bounded by c ; see again equation (17).

Models 3f and 3g represent local deviations which induce a local GMM model misspecification in the sense of Hall and Inoue (2003). More specifically, Model 3f implies a time homogeneous local departure from the reference model P_0 . Model 3g, instead, is a time non-homogeneous local deviation. It represents a shift in the distribution of $W_{i,t}$, which cannot be exhausted by a break in the structural parameter β only. In particular, local deviations in Model 3f and 3g are of the type considered in Example 1 and satisfy Assumption 7.

4.3.1 Asymptotic tests

Tables 3a-3g present the results of our Monte Carlo simulations for Models 3a-3g in Burnside and Eichenbaum's (1996) model. The break date is fixed at $t_0 = 0.5 \cdot T$, where $T = 100, 200$ ¹⁵. In the simulations for $T = 200$ we have set $\beta_1 = 1, \beta_2 = 1, 1.1, 1.2, 1.3$. In the simulations for the other sample sizes the values of $\beta_2 - \beta_1$ applied for $T = 200$ have been multiplied by a factor $\sqrt{200/T}$ in order to obtain comparable local alternatives across the different sample sizes. The tuning constants for the RGMM test have been set at $c = 4.18$. We also fixed $\Pi = [p_0, 1 - p_0]$, where $p_0 = 0.25$. The nominal level of the tests is 5%.

Tables 3a-3g about here

Table 3a shows that under the given reference model classical and robust GMM tests for breaks perform very similarly. Finite sample sizes of all tests are quite near to their nominal levels. Moreover, virtually no loss in power of RGMM tests is observed, relatively to their classical counterparts. Under a student t_7 or t_5 distribution, in Tables 3b and 3c, RGMM tests provide a clearly higher power than their classical counterparts. For instance, in Table 3c for $T = 200$ the power of RGMM tests against the alternative $\beta_2 = 1.3$ is about 46%, 53% and 52% for the supremum, the average and the exponential statistics. The power obtained when applying classical GMM tests is instead only about 34%, 41% and 40%, for the supremum, the average and the exponential statistics.

¹⁵ Sample size $T = 100$ is used in Burnside and Eichenbaum (1996).

The results in Tables 3d and 3e for the contaminated normal models are qualitatively similar to those under a student t distribution. However, the quantitative gains in power implied by RGMM statistics are even larger, and are in some cases near to 50%, as for instance in the models with $\beta_2 = 1.3$ for $T = 200$.

Finally, in Tables 3f and 3g it is shown that also under an outlier replacement model RGMM tests for breaks provide a clearly higher efficiency of inferences on breaks than tests based on the classical GMM. For instance, in Model 3f with $\beta_2 = 1.3$ for $T = 200$ the power of RGMM tests is about 60%, 65% and 64% for the supremum, the average and the exponential statistics. This is clearly above the power implied by classical GMM testing procedures. In Table 3g we compare the power of GMM and RGMM tests under Model 3g for $(\beta_2 - \beta_1) \sqrt{\frac{200}{T}} = 0.3$ and different contamination probabilities $\eta = 0.0, 0.025, 0.05, 0.075$. It is shown that RGMM statistics maintain a higher power against structural breaks in the parametric part of the model, also in the presence of time non-homogeneous local deviations (see again the discussion after example 1).

4.3.2 Bootstrap tests

The last two panels of Tables 3a-3g present for a sample size $T = 50$ the results implied by GMM and RGMM tests for breaks based on (i) standard asymptotic critical values and (ii) bootstrapped versions of the relevant GMM statistics. Generally, from these results we see that the finite sample sizes of GMM and RGMM asymptotic tests based on the supremum functional can be quite biased downwards relatively to the given nominal level for all Model 3a-3g. By contrast, the empirical sizes of both classical and robust GMM tests based on the average and the exponential seem to be quite well controlled. However, for all statistics we observe again in all Models 3a-3g a larger power of RGMM tests when compared with the classical ones. In some cases the power increase is very substantial, as for instance in Table 3e for a contaminated normal distribution.

Bootstrapped versions of GMM and RGMM tests¹⁶ for $T = 50$ provide all accurate finite

¹⁶ They are based on the recentered bootstrap of Hall and Horowitz (1996) for overidentified models.

sample sizes, thus correcting well the bias observed for the asymptotic tests based on the supremum statistic discussed previously. Moreover, the power curves of bootstrapped GMM and RGMM tests based on the average and the exponential are very similar to those obtained for the corresponding asymptotic tests. This further confirms the good finite sample performances of such asymptotic tests in the present model setting.

Similarly to the previous section, we finally observe that bootstrapped versions of the RGMM tests tend to produce a uniformly higher power across all Models 3a-3g. Again, under local deviations from the reference model, such power increases appear to be often quite substantial, as is for instance illustrated by the results for the contaminated normal model in Table 3e.

5 Conclusions

We proposed a class of new supremum, average and exponential RGMM tests for structural breaks, which imply a bounded sensitivity of level and power under local departures from a reference model. Robustness of the new tests is obtained by computing the supremum, the average or the exponential functionals over a sequence of GMM Lagrange Multiplier statistics in a setting based on a bounded orthogonality function. Monte Carlo simulations showed that the new robust GMM tests perform well across a quite broad set of model configurations, both in terms of the efficiency and the robustness of the inference procedure, when compared with standard GMM tests for structural breaks. Due to the intrinsic difficulties in the formulation and the identification of econometric models that exactly describe the whole data distribution, it is expected that RGMM tests for breaks can help in providing some more robust and consistent evidence on the presence of breaks in the statistical analysis of economic data series.

6 Appendix 1: Proof of Theorem 2

By Assumption 8 we have:

$$\sqrt{T}U_T(\cdot) \implies J_p(\cdot) + x(\cdot),$$

uniformly over $Q^{\eta,T} \in \mathcal{U}_{\varepsilon,T}$, where:

$$x(\pi) = \lim_{T \rightarrow \infty} \sqrt{T}U(\pi, Q^{\eta,T}), \quad \pi \in \Pi.$$

Let us compute x . By Assumptions 5 and 6, a von Mises expansion up to terms $O(\eta/\sqrt{T})$ gives:

$$\begin{aligned} \sqrt{T}U(\pi, Q^{\eta,T}) &= \sqrt{T}\pi H(\bar{P}_\infty)^{\frac{1}{2}} E_{\bar{Q}^{\eta,T}(\pi)} \left[m(W, \tilde{\theta}(\bar{Q}_\infty^{\eta,T})) \right] \\ &= \sqrt{T}U(\pi, P) + \eta\pi H(\bar{P}_\infty)^{\frac{1}{2}} \left\{ E_{\bar{P}_\infty} \left[\nabla_{\theta'} m(W, \tilde{\theta}(\bar{P}_\infty)) \right] D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty) \right. \\ &\quad \left. + E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] \right\}, \end{aligned} \quad (24)$$

where $D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty)$ is the Fréchet derivative of $\tilde{\theta}$ at \bar{P}_∞ in direction $\bar{Q}_\infty - \bar{P}_\infty$. To compute $D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty)$ note first that the asymptotic estimating equation for $\tilde{\theta}(\bar{P}_\infty)$ is

$$0 = N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) E_{\bar{P}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right), \quad (25)$$

where

$$N(\bar{P}_\infty) = E_{\bar{P}_\infty} \left(\nabla_{\theta'} m(W, \tilde{\theta}(\bar{P}_\infty)) \right).$$

The directional derivative $D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty)$ can be then computed by implicitly differentiating (25) in the direction $\bar{Q}_\infty - \bar{P}_\infty$ to get

$$0 = N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) \left[N(\bar{P}_\infty) D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty) + E_{\bar{Q}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right) \right],$$

i.e.

$$D\tilde{\theta}(\bar{P}_\infty, \bar{Q}_\infty - \bar{P}_\infty) = - \left(N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) N(\bar{P}_\infty) \right)^{-1} N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) E_{\bar{Q}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right),$$

where we used that P satisfies (5). Inserting this expression into (24) gives up to uniform terms of order $o(\eta)$:

$$\begin{aligned}\sqrt{TU}(\pi, Q^{\eta, T}) &= \sqrt{TU}(\pi, P) + \eta\pi H(\bar{P}_\infty)^{\frac{1}{2}} \left\{ E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] \right. \\ &\quad \left. - R(\bar{P}_\infty) E_{\bar{Q}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right) \right\},\end{aligned}$$

where :

$$R(\bar{P}_\infty) = N(\bar{P}_\infty) (N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) N(\bar{P}_\infty))^{-1} N'(\bar{P}_\infty) \Omega(\bar{P}_\infty) .$$

Moreover, since $H(\bar{P}_\infty) R(\bar{P}_\infty) = H(\bar{P}_\infty)$, we have $H(\bar{P}_\infty)^{1/2} R(\bar{P}_\infty) = H(\bar{P}_\infty)^{1/2}$, and we get:

$$\begin{aligned}\sqrt{TU}(\pi, Q^{\eta, T}) &= \sqrt{TU}(\pi, P) \\ &\quad + \eta\pi H(\bar{P}_\infty)^{\frac{1}{2}} \left\{ E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] - E_{\bar{Q}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right) \right\}.\end{aligned}$$

From Theorem 1 we have:

$$\sqrt{TU}(\pi, P) \rightarrow \pi(1-\pi) H(\bar{P}_\infty)^{\frac{1}{2}} [\mu_1(\pi) - \mu_2(\pi)], \quad T \rightarrow \infty.$$

Thus we get:

$$\sqrt{TU}_T(\cdot) \implies J_p(\cdot) + b^*(\cdot),$$

uniformly in $Q^{\eta, T} \in \mathcal{U}_{\varepsilon, T}$, where:

$$\begin{aligned}b^*(\pi) &= \pi(1-\pi) H(\bar{P}_\infty)^{\frac{1}{2}} [\mu_1(\pi) - \mu_2(\pi)] \\ &\quad + \eta\pi H(\bar{P}_\infty)^{\frac{1}{2}} \left\{ E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] - E_{\bar{Q}_\infty} \left(m(W, \tilde{\theta}(\bar{P}_\infty)) \right) \right\} \\ &= \pi(1-\pi) H(\bar{P}_\infty)^{\frac{1}{2}} \left\{ \mu_1(\pi) + \eta E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] \right. \\ &\quad \left. - \mu_2(\pi) - \eta E_{\bar{Q}(\pi)} \left[m(W, \tilde{\theta}(\bar{P}_\infty)) \right] \right\},\end{aligned}$$

where we used that $\bar{Q}_\infty = \pi \bar{Q}(\pi) + (1-\pi) \underline{Q}(\pi)$. Points 2., 3. and 4. follow by the Functional Central Limit theorem.

7 Appendix 2: Proof of Theorems 3 and 4

Let us prove Theorem 3. Let $\mathcal{L}(\cdot, \eta^2 \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2)$, denote the cumulative distribution function of

$$\int_{\Pi} Q_p^* \left(\pi, \eta^2 \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 \right) d\lambda(\pi) .$$

We have:

$$\begin{aligned} \lim_{T \rightarrow \infty} \alpha(Q^{\eta, T}) &= \lim_{T \rightarrow \infty} Q^{\eta, T}(\widehat{\xi}_T^{ave} > \xi_0) \\ &= 1 - \mathcal{L} \left(\xi_0, \eta^2 \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 \right) \\ &= 1 - \mathcal{L}(\xi_0, 0) - \frac{\partial \mathcal{L}}{\partial y}(\xi_0, 0) \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 \eta^2 + o(\eta^2) \\ &= \alpha_0 - \frac{\partial \mathcal{L}}{\partial y}(\xi_0, 0) \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 \eta^2 + o(\eta^2), \end{aligned}$$

uniformly over the asymptotic neighborhood $U_{\varepsilon, T}$, where $\frac{\partial \mathcal{L}}{\partial y}$ denotes the derivative of \mathcal{L} with respect to the noncentrality parameter $y = \|d(\bar{P}_\infty, \bar{Q}_\infty)\|^2 \eta^2$. The proof of Theorem 4 is similar.

8 Appendix 3: Computation of the robust GMM estimator of the ARCH(1) model

In this Appendix we discuss some computational issues involved with the computation of the robust GMM estimator $(\tilde{\alpha}_0, \tilde{\alpha}_1)'$ in the ARCH(1) model of Section 4.2. The estimator $(\tilde{\alpha}_0, \tilde{\alpha}_1)'$, the shift factors $\tau(y_{t-1})$, $t = 1, \dots, T$, and the normalization matrix A have to be computed by an iterative algorithm. We first discuss the single steps necessary for computing $\tau(y_{t-1})$, A and $(\tilde{\alpha}_0, \tilde{\alpha}_1)'$, respectively, and then present the complete algorithm.

8.1 The shift factors

The first condition in (23) is satisfied for $\tau(y_{t-1})$ defined by

$$\tau(y_{t-1}) = \frac{1}{\sigma_t^2} \begin{pmatrix} \tau^*(y_{t-1}) \\ y_{t-1}^2 \tau^*(y_{t-1}) \end{pmatrix}$$

if $\tau^*(y_{t-1})$ is such that

$$E_{P_0} [(u_t^2 - 1 - \tau^*(y_{t-1})) w_t \mid y_{t-1}] = 0 \quad , \quad (26)$$

where

$$w_t := w_c \left(\frac{Ax_t}{\sigma_t^2} (u_t^2 - 1 - \tau^*(y_{t-1})) \right)$$

and $x_t = (1, y_{t-1}^2)'$. Equation (26) is equivalent to the implicit equation:

$$\tau^*(y_{t-1}) = \frac{E_{P_0} [(u_t^2 - 1) w_t \mid y_{t-1}]}{E_{P_0} [w_t \mid y_{t-1}]} \quad . \quad (27)$$

Therefore, we can focus on iterative procedures for the computation of the numerator and the denominator of $\tau^*(y_{t-1})$. Let values $(A'A)^0, \tilde{\alpha}^0, \tau^{*0}(y_{t-1}), t = 1, \dots, T$, be given on the RHS of (27). Let further $z_{1,t} \leq z_{2,t}$ denote the two solutions of the equation:

$$\left\| \frac{A^0 x_t}{\sigma_t^2} (z - 1 - \tau^{*0}(y_{t-1})) \right\| = \frac{\sqrt{x_t' (A'A)^0 x_t}}{\sigma_t^2} |z - 1 - \tau^{*0}(y_{t-1})| = c,$$

where $\sigma_t^2 = x_t' \tilde{\alpha}^0$. Then, if $z_{1,t} \leq 0 \leq z_{2,t} =: d_{2,t}^2$, as it is the case for any parameter choice we have investigated, the denominator in (27) becomes:

$$E_{P_0} [w_t \mid y_{t-1}] = 2 \left(\int_0^{d_{2,t}} \Phi(du) + \frac{c\sigma_t^2}{\sqrt{x_t' (A'A)^0 x_t}} \int_{d_{2,t}}^\infty \frac{1}{u^2 - 1 - \tau^{*0}(y_{t-1})} \Phi(du) \right).$$

When $d_{2,t}$ is large enough, the second integral can be approximated by a Laplace approximation (see for instance Jensen (1995)). In any parameter choice we considered, this approximation has shown to be very accurate and efficient. A similar analytical approximation can be used to compute the numerator of (27). Using these approximations, from the RHS of (27) we get updated values $\tau^{*1}(y_{t-1}), t = 1, \dots, T$.

Similarly to above, from the second condition in (23), an updated estimate of $A'A$ is given by:

$$(A'A)^1 = \left[\frac{1}{T} \sum_{t=1}^T \frac{1}{\sigma_t^4} w_t^2 x_t x_t' (u_t^2 - 1 - \tau^{*0}(y_{t-1}))^2 \right]^{-1},$$

where $\sigma_t^2 = x_t' \tilde{\alpha}^0$, $u_t = y_t / \sigma_t$, $w_t := w_c (A^0 x_t [u_t^2 - 1 - \tau^{*0}(y_{t-1})] / \sigma_t^2)$.

8.2 Robust GMM estimator as iterated WLS

The robust GMM estimator is the solution $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1)'$ of the finite sample estimating equation

$$\frac{1}{T} \sum_{t=1}^T \frac{1}{\sigma_t^4} w_t x_t \left[y_t^2 - (1 + \tau^*(y_{t-1})) x_t' \alpha \right] = 0 \quad ,$$

where $\sigma_t^2 = x_t' \alpha = \alpha_0 + \alpha_1 y_{t-1}^2$, $\alpha = (\alpha_0, \alpha_1)'$, $w_t := w_c (Ax_t [u_t^2 - 1 - \tau^*(y_{t-1})] / \sigma_t^2)$. The solution $\hat{\alpha}$ satisfies the implicit equation:

$$\hat{\alpha} = \left[\sum_{t=1}^T \frac{1}{\sigma_t^4} w_t (1 + \tau^*(y_{t-1})) x_t x_t' \right]^{-1} \cdot \sum_{t=1}^T \frac{1}{\sigma_t^4} w_t x_t y_t^2 \quad . \quad (28)$$

Given values $\hat{\alpha}^0$, $\tau^{*0}(y_{t-1})$, $t = 1, \dots, T$, and $(A'A)^0$, they can be inserted in the RHS of (28) to get an updated estimate $\hat{\alpha}^1$, which has the form of a WLS estimator.

8.3 The robust GMM estimation algorithm

The algorithm for computing the robust GMM estimator in the ARCH(1) model is as follows:

1. Consider starting values $\tilde{\alpha}^0$, $\tau^{*0}(y_{t-1}) = 0$, $t = 1, \dots, T$, and

$$(A'A)^0 = \left[\frac{1}{T} \sum_{t=1}^T \frac{1}{\sigma_t^4} x_t x_t' (u_t^2 - 1)^2 \right]^{-1}$$

where $\sigma_t^2 = x_t' \tilde{\alpha}^0$, $u_t = y_t / \sigma_t$

2. Compute the weights:

$$w_t := w_c \left(\frac{A^0 x_t}{\sigma_t^2} [u_t^2 - 1 - \tau^{*0}(y_{t-1})] \right)$$

3. Compute the shift and normalization factors $\tau^{*1}(y_{t-1})$, $t = 1, \dots, T$, and $(A'A)^1$ by the approach discussed above:

$$\tau^{*1}(y_{t-1}) = \frac{E_{P_0} [(u_t^2 - 1) w_t \mid y_{t-1}]}{E_{P_0} [w_t \mid y_{t-1}]},$$

$$(A'A)^1 = \left[\frac{1}{T} \sum_{t=1}^T \frac{1}{\sigma_t^4} w_t^2 x_t x_t' (u_t^2 - 1 - \tau^{*0}(y_{t-1}))^2 \right]^{-1}$$

4. Compute the GMM estimator:

$$\hat{\alpha}^1 = \left[\sum_{t=1}^T \frac{1}{\sigma_t^4} w_t (1 + \tau^{*0}(y_{t-1})) x_t x_t' \right]^{-1} \cdot \sum_{t=1}^T \frac{1}{\sigma_t^4} w_t x_t y_t^2$$

5. Use $\tau^{*1}(y_{t-1})$, $t = 1, \dots, T$, $(A'A)^1$, $\hat{\alpha}^1$ as new starting values

Steps 2 to 5 are iterated until numerical convergence is obtained.

9 Tables

Table 1a: GMM LM test results for Model 1a: $u_t \sim i.i.d.N(0, 1)$. Parameters: $\alpha = \gamma = 0$, $\beta_1 = 1$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	4.02	5.34	3.94	5.12	0.0	4.16	5.94	3.80	4.80
0.1	9.12	11.7	8.96	11.1	0.1	14.0	18.9	12.0	15.7
0.2	25.7	32.4	23.8	28.8	0.2	44.4	54.6	40.4	48.3
0.3	53.4	61.6	49.1	55.4	0.3	81.0	86.3	76.0	81.5
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	3.92	5.84	3.48	5.06	0.0	2.92	6.72	2.70	4.46
0.1	8.10	12.7	6.72	9.74	0.1	5.86	12.6	4.90	8.04
0.2	21.4	31.1	18.2	25.4	0.2	15.6	29.5	13.2	22.0
0.3	44.6	58.1	39.0	50.3	0.3	34.6	54.4	29.1	42.9
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	4.60	4.84	5.02	5.06					
0.1	8.14	9.00	8.28	9.08					
0.2	19.0	22.2	18.8	21.8					
0.3	36.9	42.2	31.1	41.8					

Table 1b: GMM LM test results for Model 1b: $u_t \sim i.i.d.t_5/\sqrt{5/3}$. Parameters: $\alpha = \gamma = 0$, $\beta_1 = 1$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	3.74	5.06	4.26	5.06	0.0	4.70	6.12	4.28	5.02
0.1	9.62	12.0	10.2	12.3	0.1	9.44	13.4	9.02	12.5
0.2	26.9	33.3	29.8	36.1	0.2	27.3	35.2	27.7	34.3
0.3	54.8	62.9	59.1	66.1	0.3	53.5	61.9	55.9	63.2
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	3.16	6.26	3.14	4.90	0.0	3.24	7.62	1.92	3.64
0.1	8.36	13.8	7.74	11.5	0.1	6.46	14.2	4.96	9.66
0.2	24.8	35.5	23.9	32.2	0.2	19.2	34.8	16.5	26.8
0.3	50.2	61.7	49.5	59.8	0.3	40.1	58.9	36.1	50.3
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	5.02	5.42	4.88	5.46					
0.1	9.34	10.4	9.08	10.4					
0.2	21.4	24.3	22.7	25.8					
0.3	41.6	45.5	44.1	48.5					

Table 1c: GMM LM test results for Model 1c: $u_t \sim i.i.d.t_3/\sqrt{3}$. Parameters: $\alpha = \gamma = 0$, $\beta_1 = 1$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	5.34	6.60	4.42	5.16	0.0	4.76	8.14	3.80	4.96
0.1	11.7	15.1	12.4	16.6	0.1	13.5	18.2	12.8	16.4
0.2	33.6	40.5	41.0	47.9	0.2	35.1	43.2	39.6	47.3
0.3	63.5	69.2	76.2	81.0	0.3	63.8	71.1	73.0	79.4
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	5.04	8.68	3.18	4.26	0.0	4.36	9.78	2.54	4.64
0.1	11.7	17.5	9.64	13.6	0.1	8.88	18.4	7.26	12.3
0.2	32.9	43.2	31.7	41.3	0.2	26.3	43.9	22.8	35.8
0.3	60.8	70.4	63.2	72.9	0.3	51.8	68.9	47.3	62.4
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	4.90	4.94	5.22	5.40					
0.1	17.7	19.3	18.8	21.5					
0.2	48.7	52.2	53.9	59.8					
0.3	75.9	78.2	81.2	85.0					

Table 1d: GMM LM test results for Model 1d: $u_t \sim i.i.d.CN(0.05, 0, 3) / \sqrt{1.4}$.

Parameters: $\alpha = \gamma = 0$, $\beta_1 = 1$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$. Break date: $t_0 = 0.5 \cdot T$.

Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations.

The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	4.72	5.98	4.44	5.12	0.0	4.00	5.48	4.00	4.70
0.1	10.5	14.2	10.1	12.8	0.1	8.54	12.1	8.74	11.7
0.2	28.1	34.9	29.6	35.8	0.2	26.3	34.7	26.4	33.9
0.3	56.0	63.3	59.6	66.6	0.3	54.1	63.4	55.3	63.5
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	3.76	6.50	2.70	3.96	0.0	2.78	6.86	2.34	4.50
0.1	8.16	13.8	7.16	11.1	0.1	6.42	14.4	5.64	11.3
0.2	24.4	35.1	22.7	31.9	0.2	19.3	35.0	17.4	28.0
0.3	51.1	62.9	50.0	59.5	0.3	41.1	59.9	37.1	51.4
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	4.58	4.92	4.30	4.46					
0.1	13.4	16.1	14.4	17.0					
0.2	39.5	43.8	42.7	48.7					
0.3	68.6	72.6	71.4	76.4					

Table 1e: GMM LM test results for Model 1e: $u_t \sim i.i.d.CN(0.1, 0, 3) / \sqrt{1.8}$.

Parameters: $\alpha = \gamma = 0$, $\beta_1 = 1$, $\sigma_u = \sigma_e = 1$, $\rho = 0.5$. Break date: $t_0 = 0.5 \cdot T$.

Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations.

The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	4.56	6.22	4.72	5.24	0.0	4.08	5.40	3.76	4.56
0.1	10.6	14.1	11.1	14.1	0.1	8.96	12.8	9.62	13.0
0.2	28.7	35.1	33.4	39.4	0.2	27.4	35.4	30.3	36.9
0.3	56.0	63.3	64.7	71.0	0.3	56.1	64.9	61.3	68.5
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	3.82	7.12	2.72	4.38	0.0	3.12	7.42	2.40	4.22
0.1	8.92	15.3	7.66	11.9	0.1	7.70	15.4	6.08	11.9
0.2	26.2	36.9	26.0	35.3	0.2	22.7	37.7	19.2	30.0
0.3	52.8	64.5	54.4	64.1	0.3	44.5	61.8	40.4	55.0
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	4.58	5.48	4.40	4.74					
0.1	11.0	11.3	11.3	12.5					
0.2	26.9	29.3	30.0	33.9					
0.3	50.0	54.0	56.1	61.0					

Table 1f: GMM LM test results for Model 1f: the observations from a sample of Model 1a are replaced with probability 0.05 by an outlier $y_t \sim N(0, 3)$. Parameters: $\alpha = \gamma = 0, \beta = 1, \sigma_u = \sigma_e = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c_1 = 3, c_2 = 3$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 300$	GMM		RGMM		$T = 200$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{3/2}$	sup	ave	sup	ave	$\beta_2 - \beta_1$	sup	ave	sup	ave
0.0	4.86	6.08	4.08	5.00	0.0	4.36	6.52	3.44	4.72
0.1	7.78	10.2	7.72	9.60	0.1	7.90	11.0	6.92	9.32
0.2	19.2	24.1	20.2	25.2	0.2	18.0	23.9	18.4	24.0
0.3	37.3	43.8	41.4	49.0	0.3	36.5	43.5	39.0	47.0
$T = 100$	GMM		RGMM		$T = 50$	GMM		RGMM	
$(\beta_2 - \beta_1)\sqrt{1/2}$	sup	ave	sup	ave	$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave
0.0	3.52	6.84	2.56	4.10	0.0	3.92	9.14	2.44	4.48
0.1	7.26	11.3	5.60	8.52	0.1	5.72	12.5	3.78	6.64
0.2	16.9	24.9	14.6	22.1	0.2	11.6	23.4	9.76	16.4
0.3	34.3	45.7	33.0	44.1	0.3	21.6	36.2	19.5	31.2
$T = 50$	GMM		RGMM						
Bootstrap									
$(\beta_2 - \beta_1)\sqrt{1/4}$	sup	ave	sup	ave					
0.0	5.32	5.90	5.16	5.14					
0.1	8.06	9.32	9.18	9.86					
0.2	18.1	21.0	20.6	23.8					
0.3	33.3	37.5	39.6	43.8					

Table 2a: GMM LM test results for Model 2a: $u_t \sim i.i.d.N(0, 1)$. Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. The nominal level of the tests is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	2.60	5.30	4.48	4.06	4.80	4.60	1.82	4.48	3.82	3.40	4.48	4.20
0.1	6.90	12.9	11.3	9.82	11.9	11.5	5.04	11.4	9.72	7.96	10.6	10.2
0.2	23.4	34.7	31.9	27.8	31.6	31.6	17.6	30.4	27.4	24.0	28.7	28.2
0.3	46.9	60.6	57.9	52.8	57.8	57.4	38.4	55.6	51.6	46.2	53.2	52.6
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.06	4.42	3.42	2.64	4.26	4.04	0.54	2.46	1.98	1.48	2.62	2.52
0.1	3.02	9.46	7.80	6.46	9.42	8.92	1.00	5.68	4.40	2.98	5.96	5.70
0.2	10.6	24.6	21.5	17.4	24.5	23.7	3.60	15.1	12.0	8.10	15.5	15.0
0.3	24.6	45.6	40.6	35.0	44.7	43.9	8.90	28.5	24.8	17.4	29.4	29.1
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	5.10	4.74	4.94	5.48	5.66	5.48						
0.1	7.32	9.00	8.78	8.76	9.66	9.68						
0.2	18.1	21.4	21.2	18.7	22.1	21.9						
0.3	33.4	38.4	37.7	33.8	38.3	37.6						

Table 2b: GMM LM test results for Model 2b: $u_t \sim i.i.d.t_7/\sqrt{7/5}$. Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. The nominal level of the tests is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.68	4.26	3.48	3.46	3.86	3.80	1.44	4.04	3.26	3.34	4.28	4.26
0.1	3.66	8.08	6.82	6.78	8.14	8.02	2.78	7.92	6.52	6.06	8.06	7.94
0.2	9.56	18.4	15.9	16.3	19.6	19.0	7.20	16.5	14.4	13.8	17.7	17.3
0.3	19.9	34.5	31.2	31.7	36.7	36.2	15.5	30.0	26.7	26.4	32.4	31.8
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	0.64	3.22	2.40	2.24	3.62	3.36	0.42	1.72	1.30	1.40	2.54	2.46
0.1	1.30	5.42	4.14	3.38	6.10	5.78	0.32	3.10	2.00	1.64	4.12	3.72
0.2	3.04	12.7	9.92	8.68	14.4	13.7	0.88	6.70	5.14	3.54	8.64	8.12
0.3	7.66	22.3	19.1	18.3	25.6	24.9	1.98	12.3	9.20	7.20	15.4	14.8
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	4.88	4.98	4.94	5.26	5.06	4.96						
0.1	5.90	7.08	6.86	6.40	7.66	7.34						
0.2	11.0	13.5	13.3	13.0	14.9	14.6						
0.3	18.0	21.7	21.1	19.8	23.4	23.0						

Table 2c: GMM LM test results for Model 2c: $u_t \sim i.i.d.t_5/\sqrt{5/3}$. Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. The nominal level of the tests is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.54	4.18	3.06	3.76	4.60	4.40	1.16	3.94	2.98	3.20	4.32	4.08
0.1	2.50	6.80	5.64	6.02	7.36	7.18	2.04	6.22	5.10	4.96	6.62	6.52
0.2	6.06	13.5	11.9	12.4	16.1	15.6	4.48	12.4	10.3	10.8	14.8	14.3
0.3	13.0	23.8	21.4	25.1	29.2	28.8	9.28	21.4	18.3	21.2	27.0	26.5
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	0.54	3.10	2.24	2.84	4.16	3.72	0.34	1.52	1.10	1.04	2.60	2.44
0.1	0.82	4.58	3.30	3.44	5.92	5.54	0.40	2.66	1.84	1.50	3.90	3.66
0.2	1.94	8.82	6.84	7.48	11.6	10.9	0.48	4.92	3.38	2.88	7.00	6.74
0.3	4.22	16.5	13.0	14.0	21.3	20.1	1.14	8.64	6.48	5.16	12.0	11.4
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	4.62	4.70	4.64	4.94	5.26	5.30						
0.1	5.32	6.50	6.24	5.56	6.52	6.52						
0.2	8.24	10.5	10.2	9.86	11.5	11.3						
0.3	13.3	16.9	16.5	16.9	19.6	19.3						

Table 2d: GMM LM test results for Model 2d: $u_t \sim i.i.d.CN(0.05, 0, 3) / \sqrt{1.4}$.

Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. The nominal level of the tests is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.22	4.32	3.24	3.56	4.48	4.18	0.76	3.44	2.50	2.94	4.02	3.82
0.1	2.02	6.14	5.16	6.68	8.22	8.04	1.30	5.70	4.38	5.34	7.84	7.62
0.2	5.04	12.0	10.1	14.7	18.0	17.5	3.94	11.5	9.22	13.6	17.2	17.0
0.3	10.0	21.0	18.7	29.7	33.9	33.7	7.64	20.0	17.2	25.0	30.8	30.1
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	0.58	2.52	1.66	2.32	3.58	3.22	0.30	1.46	1.06	1.20	2.24	2.06
0.1	0.62	4.06	5.02	3.66	6.02	5.62	0.40	2.32	1.56	1.54	4.14	3.90
0.2	1.98	9.04	6.48	8.92	13.9	13.3	0.56	4.78	3.56	3.50	7.78	7.40
0.3	4.42	16.0	12.9	17.3	24.7	23.8	1.56	8.98	6.88	6.60	14.1	13.5
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	4.56	4.56	4.70	4.62	4.38	4.50						
0.1	5.62	6.60	6.40	6.68	7.50	7.50						
0.2	9.38	11.6	11.4	11.7	14.1	14.2						
0.3	14.8	17.6	17.6	19.4	21.9	22.0						

Table 2e: GMM LM test results for Model 2e: $u_t \sim i.i.d.CN(0.1, 0, 3) / \sqrt{1.8}$.

Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. The nominal level of the tests is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.08	4.10	3.32	2.76	3.34	3.38	0.68	3.86	2.64	2.72	4.24	4.04
0.1	1.78	5.34	4.18	4.84	6.32	6.04	1.14	5.24	3.88	5.04	6.76	6.62
0.2	3.50	9.20	7.86	10.1	12.1	11.8	2.58	8.92	6.80	9.34	12.1	11.9
0.3	6.78	15.3	13.2	18.5	22.0	21.7	4.66	13.8	11.5	16.1	20.9	20.2
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	0.48	2.42	1.72	2.30	3.40	3.16	0.28	1.44	1.12	1.32	2.76	2.54
0.1	0.58	3.30	3.30	2.86	4.88	4.58	0.38	1.50	1.16	1.38	3.14	3.02
0.2	1.00	5.38	5.38	5.50	8.58	8.20	0.42	2.70	1.86	2.22	5.44	5.14
0.3	1.74	8.80	8.80	9.26	14.8	14.4	0.52	4.36	3.10	3.28	8.64	8.28
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	4.66	4.54	4.54	4.70	4.66	4.62						
0.1	4.98	5.50	5.50	5.28	6.02	6.00						
0.2	6.02	7.54	7.34	8.24	10.2	9.90						
0.3	8.92	11.2	11.0	13.0	15.2	14.8						

Table 2f: GMM LM test results for Model 2f: the observations from a sample of Model 1a are replaced with probability 0.025 by an outlier $y_t \sim 4N(0, 0.01/0.4)$. Parameters: $\alpha_0 = 0.01$, $\alpha_1 = 0.6$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 6.18$. $p_0 = 0.45$, 5000 Monte Carlo simulations. Nominal size is 5%.

	$T = 1000$						$T = 500$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	1.94	5.34	4.08	3.94	4.70	4.58	1.46	4.38	3.30	3.16	3.78	3.76
0.1	2.96	8.02	6.98	7.36	8.86	8.84	2.20	6.78	5.62	6.16	8.38	7.88
0.2	7.60	14.3	12.8	18.8	22.6	22.1	6.70	14.5	12.7	15.8	20.4	19.9
0.3	15.4	25.9	23.5	37.9	42.7	42.4	13.6	25.1	22.8	32.1	38.0	37.5
	$T = 250$						$T = 125$					
	GMM			RGMM			GMM			RGMM		
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp	sup	ave	exp	sup	ave	exp
0.0	0.84	3.70	2.64	2.42	3.74	3.58	0.92	2.54	2.02	1.62	2.98	2.82
0.1	1.76	7.02	5.50	4.72	7.60	7.20	0.92	4.32	3.48	2.70	5.48	5.38
0.2	5.04	14.5	12.2	11.8	14.3	16.0	1.10	9.72	7.80	6.18	12.1	11.7
0.3	11.1	25.6	22.1	22.7	30.5	29.6	4.16	16.7	13.5	12.1	21.1	20.8
	$T = 125$ Bootstrap											
	GMM			RGMM								
$\frac{\alpha_2 - \alpha_1}{\sqrt{1000/T}}$	sup	ave	exp	sup	ave	exp						
0.0	5.10	5.06	5.08	4.58	4.58	4.50						
0.1	7.28	8.86	8.70	7.60	8.72	8.84						
0.2	13.3	15.9	15.6	14.2	16.9	16.6						
0.3	24.0	27.3	27.1	25.1	28.5	28.1						

Table 3a: GMM LM test results for Model 3a: $u_{i,t} \sim i.i.d.N(0,1)$. Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	4.42	4.90	5.08	4.14	5.00	5.22	0.0	4.04	4.94	5.14	4.32	5.00	5.14
0.1	11.3	14.3	14.6	11.5	14.4	14.7	0.1	12.1	15.2	14.9	12.6	15.2	14.8
0.2	30.1	38.2	36.7	31.9	39.3	38.3	0.2	37.3	43.8	42.9	37.3	43.8	42.9
0.3	57.3	65.9	64.6	59.7	67.6	67.0	0.3	69.5	75.8	75.5	69.0	75.0	74.5
$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	3.28	5.22	4.94	3.18	4.98	4.94	0.0	4.72	4.50	4.54	5.10	4.90	5.02
0.1	7.36	12.2	11.4	8.16	12.9	12.5	0.1	10.4	11.4	11.1	10.7	12.3	11.7
0.2	21.1	29.7	28.9	23.4	32.8	31.7	0.2	25.9	29.2	28.1	27.6	31.4	29.9
0.3	39.7	51.1	49.8	44.6	55.9	55.0	0.3	47.1	51.0	49.5	49.4	54.4	52.7

Table 3b: GMM LM test results for Model 3a: $u_{i,t} \sim i.i.d.t_7/\sqrt{7/5}$. Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	3.44	4.82	4.84	3.84	4.62	4.40	0.0	4.08	4.96	4.94	3.90	4.60	4.54
0.1	6.94	9.54	9.40	8.42	11.5	11.1	0.1	8.48	10.4	10.3	10.1	12.6	12.3
0.2	18.3	24.8	23.6	23.2	29.5	29.0	0.2	22.4	27.7	26.7	27.6	27.6	31.8
0.3	35.9	44.6	43.6	44.2	55.1	50.7	0.3	44.2	51.5	50.6	52.9	52.9	58.8
$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	2.68	4.20	4.40	2.88	4.32	4.44	0.0	5.28	5.42	5.14	4.76	5.14	4.86
0.1	5.50	8.80	8.74	6.12	9.78	9.66	0.1	9.84	10.3	10.3	9.08	9.88	9.62
0.2	13.2	20.0	19.0	16.7	24.9	23.8	0.2	19.9	21.4	21.0	21.8	24.5	24.0
0.3	25.6	35.4	34.0	32.5	43.5	42.6	0.3	32.6	36.5	34.9	39.7	43.3	42.6

Table 3c: GMM LM test results for Model 3a: $u_{i,t} \sim i.i.d.t_5/\sqrt{5/3}$. Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	3.40	4.94	4.74	3.70	4.42	4.76	0.0	4.26	5.22	5.00	4.36	5.00	5.20
0.1	6.52	9.44	9.24	8.16	11.0	10.7	0.1	7.78	9.48	9.54	9.22	11.3	11.3
0.2	15.4	20.8	20.3	20.6	25.8	25.2	0.2	18.0	22.3	21.3	24.5	29.1	28.1
0.3	30.2	38.0	37.3	39.1	47.0	45.9	0.3	33.9	40.7	39.6	46.3	53.3	52.4
$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	3.42	4.92	4.78	3.44	5.16	5.16	0.0	5.52	5.50	5.64	5.28	5.64	5.28
0.1	5.70	8.24	7.96	6.56	9.82	9.88	0.1	8.24	8.56	8.08	8.94	9.82	9.46
0.2	11.8	17.6	17.4	15.9	22.4	21.6	0.2	15.6	17.8	17.3	18.7	21.3	20.1
0.3	21.5	30.5	29.6	29.1	38.2	37.0	0.3	26.6	29.3	28.6	32.6	37.3	35.4

Table 3d: GMM LM test results for Model 3a: $u_{i,t} \sim i.i.d.CN(0.05, 0, 3) / \sqrt{1.4}$.

Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$.

$p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	4.08	5.34	5.34	4.36	5.46	5.50	0.0	4.38	5.08	5.26	4.12	5.12	5.06
0.1	6.96	9.14	9.30	9.32	11.7	11.4	0.1	6.82	8.38	8.34	10.3	12.2	12.2
0.2	15.1	20.0	19.6	24.2	29.7	29.0	0.2	15.9	20.7	20.1	28.1	33.5	32.2
0.3	29.3	36.8	35.7	46.1	53.4	52.7	0.3	30.7	37.6	36.7	54.1	60.6	59.7
$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	3.30	5.32	5.06	3.32	5.20	5.22	0.0	4.54	4.64	4.82	4.72	5.22	5.06
0.1	5.56	8.82	8.62	7.24	10.4	10.4	0.1	7.86	8.32	8.16	10.3	11.7	11.2
0.2	12.5	18.1	18.0	17.5	24.2	23.7	0.2	16.7	18.9	17.9	23.3	26.3	25.2
0.3	22.9	32.8	31.7	32.5	42.7	41.6	0.3	29.4	32.5	31.3	40.4	45.7	43.6

Table 3e: GMM LM test results for Model 3a: $u_{i,t} \sim i.i.d.CN(0.1, 0, 3)/\sqrt{1.8}$.

Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$.

$p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	3.92	5.22	5.22	4.44	5.34	5.58	0.0	4.30	5.04	5.18	4.26	5.08	5.12
0.1	5.92	7.70	7.66	8.26	10.6	10.3	0.1	6.02	8.00	7.50	8.48	10.6	10.2
0.2	11.6	15.7	15.2	18.9	23.8	22.9	0.2	12.7	15.9	15.6	22.2	26.3	25.5
0.3	20.8	26.7	26.0	35.7	43.0	42.1	0.3	22.3	28.0	26.8	42.6	48.5	48.1
$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	3.28	5.52	5.40	3.68	5.58	5.26	0.0	4.46	4.36	4.30	4.72	5.18	5.18
0.1	4.56	7.72	7.70	6.22	9.10	8.82	0.1	6.46	6.50	6.60	8.70	9.86	9.36
0.2	9.48	14.2	13.7	13.9	19.8	19.4	0.2	12.0	13.0	12.6	18.4	20.8	19.9
0.3	16.2	22.7	22.3	25.4	34.8	34.0	0.3	19.9	22.2	21.8	32.3	36.3	34.5

Table 3f: GMM LM test results for Model 3f: the observations of components $i = 1, \dots, 5$ from a sample of model 1a are replaced with probability 0.05 by an outlier $W_{i,t} \sim (3 \cdot N(0,1))^2$. Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
$\beta_2 - \beta_1$	sup	ave	exp	sup	ave	exp	$(\beta_2 - \beta_1) \cdot \sqrt{2}$	sup	ave	exp	sup	ave	exp
0.0	3.76	5.04	5.18	3.76	5.12	5.02	0.0	4.06	5.10	4.82	4.28	5.26	5.14
0.1	9.38	12.2	11.9	10.1	13.6	13.1	0.1	9.03	11.9	11.4	11.2	13.7	13.3
0.2	22.8	29.9	28.8	27.5	34.4	33.5	0.2	24.9	31.0	29.9	31.1	36.6	36.0
0.3	44.3	53.4	52.6	51.5	59.7	58.5	0.3	49.1	56.2	55.3	59.7	65.4	64.4

$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp	$\frac{\beta_2 - \beta_1}{\sqrt{2}}$	sup	ave	exp	sup	ave	exp
0.0	3.40	5.34	5.22	3.46	5.10	5.54	0.0	5.04	5.08	5.22	4.82	4.70	4.72
0.1	6.46	10.4	9.96	7.50	11.4	11.1	0.1	9.26	9.92	9.86	10.4	12.0	11.5
0.2	17.2	24.7	23.9	19.3	27.7	26.8	0.2	20.6	23.6	22.7	24.7	27.6	26.5
0.3	32.3	44.0	43.0	38.1	49.0	48.8	0.3	37.2	40.8	40.0	44.1	48.5	46.9

Table 3g: GMM LM test results for Model 3g: the observations of components $i = 1, \dots, 5$ from a sample of model 1a are replaced with probability η by an outlier $W_{i,t}$ such that $W_{i,t} \sim (3 \cdot N(0, 1))^2$ for $t = 1, \dots, 0.5 \cdot T$. Parameter: $\beta_1 = 1$. Break date: $t_0 = 0.5 \cdot T$. Tuning constant for RGMM: $c = 4.18$. $p_0 = 0.25$, 5000 Monte Carlo simulations. The nominal level of the tests is $\alpha_0 = 5\%$.

$T = 100$	GMM			RGMM			$T = 200$	GMM			RGMM		
η	sup	ave	exp	sup	ave	exp	η	sup	ave	exp	sup	ave	exp
0.0	57.3	65.9	64.6	59.7	67.6	67.0	0.0	69.5	75.8	75.5	69.0	75.0	74.5
0.025	41.4	49.2	48.3	44.7	53.3	52.4	0.025	39.5	46.9	45.7	47.4	54.1	53.0
0.05	27.8	35.4	34.8	31.8	39.2	38.3	0.05	22.5	27.8	27.3	29.2	35.7	34.6
0.075	18.4	24.9	23.9	20.1	26.7	25.7	0.075	13.5	17.2	16.9	16.5	20.3	19.8

$T = 50$	GMM			RGMM			$T = 50$ Bootstrap	GMM			RGMM		
η	sup	ave	exp	sup	ave	exp	η	sup	ave	exp	sup	ave	exp
0.0	39.7	51.1	49.8	44.6	55.9	55.0	0.0	47.1	51.0	49.5	49.4	54.4	52.7
0.025	30.3	42.1	40.9	34.4	45.6	44.4	0.025	39.1	42.6	41.4	41.3	45.9	44.3
0.05	24.4	34.4	33.5	26.6	36.8	36.0	0.05	30.6	33.8	33.2	33.0	36.7	35.2
0.075	19.4	28.5	27.3	20.5	29.3	28.3	0.075	25.3	28.0	27.3	25.1	28.7	27.9

References

- Altonji, J. and L. Segal (1996), Small sample bias in GMM estimation of covariance structures, *Journal of Business and Economic Statistics*, 14, 353-366.
- Andrews, D.W.K. (1993), Tests for parameter instability and structural change with unknown change point, *Econometrica*, 61, 821-856.
- Andrews, D.W.K. (2003), Tests for parameter instability and structural change with unknown change point: a corrigendum, *Econometrica*, 71, 395-397.
- Andrews, D. W. K., and R. C. Fair (1988), Inference in nonlinear econometric models with structural change, *Review of Economic Studies*, 55, 615-640.
- Andrews, D.W.K. and W. Ploberger, (1994), Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 62, 1383-1414.
- Bednarski, T. (1993), Fréchet differentiability of statistical functionals and implications to Robust Statistics, in: Morgenthaler, S., E. Ronchetti, and W. Stahel (eds.): *New directions in statistical data analysis and robustness*, Birkhäuser, Basel, 26-34.
- Burnside, C. and M. Eichenbaum (1996), Small sample properties of Generalized Method of Moments Wald tests, *Journal of Business and Economic Statistics*, 14, 294-308.
- Clarke, B. R. (1986), Nonsmooth analysis and Fréchet differentiability of M-functionals, *Probability Theory and Related Fields*, 73, 197-209.
- Diebold, F. X. and C. Chen (1996), Testing structure stability with endogenous breakpoint: a size comparison of analytic and bootstrap procedures, *Journal of Econometrics*, 70, 221-241.
- Engle, R. F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987-1008.
- Fiteni, I., (2002), Robust estimation of structural break points, *Econometric Theory*, 18, 349-386.
- Genton, M. G. and E. Ronchetti, (2003), Robust Indirect Inference, *Journal of the American Statistical Association*, 98, 67-76.
- Ghysels, E., Guay, A., and A. Hall, (1997), Predictive tests for structural change with unknown breakpoint, *Journal of Econometrics*, 82, 209-233.
- Ghysels, E., and A. Hall (1990), A test for structural stability of Euler conditions parameters estimated via the Generalized Method of Moments estimator, *International Economic Review*, 31, 355-364.
- Hall, A. R., and A. Inoue (2003), The large sample behaviour of the generalized method of moments estimator in misspecified models, *Journal of Econometrics*, 114, 361-394.
- Hall, P., and J. L. Horowitz (1996), Bootstrap critical values for tests based on Generalized-Method-of-Moment estimators, *Econometrica*, 64, 891-916.
- Hampel, F. R. (1974), The influence curve and its role in robust estimation, *Journal of the American Statistical Association*, 69, 383-393.
- Hampel F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (1986), *Robust statistics: the approach based on influence functions*. Wiley, New York.

- Hansen, L. P. (1982), Large sample properties of Generalized Method of Moments estimators, *Econometrica*, 50, 1029-1054.
- Hansen, B., (1997) Asymptotic P values for structural change tests, *Journal of Business and Economic Statistics*, Vol 15, No. 1, 60-67.
- Hansen, L. P., J. Heaton and A. Yaron (1996), Finite sample properties of some alternative GMM estimators, *Journal of Business and Economic Statistics*, 14, 262-280.
- Heritier, S., and E. Ronchetti (1994), Robust bounded-influence tests in general parametric models, *Journal of the American Statistical Association*, 89, 897-904.
- Huber, P., (1981), *Robust statistics*, Wiley, New York.
- Jensen, J. L., (1995), *Saddlepoint Approximations*, Clarendon Press, Oxford.
- Krishnakumar, J. and E. Ronchetti, (1997), Robust estimators for simultaneous equations models, *Journal of Econometrics*, 78, 295-314.
- Koenker, R. W., and G. Bassett (1978), Regression quantiles, *Econometrica*, 46, 33-50.
- Künsch, H., (1984), Infinitesimal robustness for autoregressive processes, *Annals of Statistics*, 12, 843-863.
- Mancini, L., E. Ronchetti and F. Trojani (2003), Optimal conditionally unbiased bounded-influence inference in dynamic location and scale models, Working paper, University of Lugano, Switzerland.
- Peracchi, F. (1990), Robust M-estimators, *Econometric Reviews*, 9, 1-30.
- Peracchi, F. (1991), Robust M-Tests, *Econometric Theory*, 7, 69-84.
- Pollard, D. (1984), *Convergence of stochastic processes*. New York, Springer Verlag.
- Ronchetti, E. and F. Trojani, (2001), Robust inference with GMM estimators, *Journal of Econometrics*, 101, 37-69.
- Sakata, S. and H. White, 1998. High Breakdown Point Conditional Dispersion Estimation With Application to S&P 500 Daily Returns Volatility, *Econometrica* 66, 529-567.
- von Mises, R. (1947), On the asymptotic distribution of differentiable statistical functions, *Annals of Mathematical Statistics*, 18, 309-348.