Efficient Bootstrap with Weakly Dependent Processes

(Article begins on next page)

07 May 2024

THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 12/08

Efficient bootstrap with weakly dependent processes

By

Francesco Bravo and Federico Crudu

Efficient bootstrap with weakly dependent processes

Francesco Bravo[*]

University of York

Federico Crudu[†]

University of Groningen

Abstract

The efficient bootstrap methodology is developed for overidentified moment conditions models with weakly dependent observation. The resulting bootstrap procedure is shown to be asymptotically valid and can be used to approximate the distributions of t-statistics, J statistic for overidentifying restrictions, and Wald, Lagrange multiplier and distance statistics for nonlinear hypotheses. The asymptotic validity of the efficient bootstrap based on a computationally less demanding approximate k-step estimator is also shown. The finite sample performance of the proposed bootstrap is assessed using simulations in an intertemporal consumption based asset pricing model.

Key words and Phrases: α-mixing, Consumption CAPM, GEL, GMM, Hypothesis testing

---

[*]Corresponding author. Department of Economics and Related Studies, University of York, York YO10 5DD, United Kingdom. email fb6@york.ac.uk

[†]Department of Economics and Econometrics, University of Groningen, 9747 AE Groningen, The Netherlands. email: FCrudu@rug.nl

1   Introduction

This paper develops Brown and Newey's (2002) efficient bootstrap methodology to possibly overidentified moment conditions models with weakly dependent observations. The efficient bootstrap differs from the traditional one in that it uses as resampling probabilities those obtained by estimating the unknown distribution of the observations subject to the constraint implied by the moment conditions themselves. The resulting estimator is typically more efficient than the empirical distribution function used in the traditional bootstrap as the nonparametric estimator of the distribution of the observation, hence the term efficient, although sometimes in the statistical literature the same bootstrap methodology is called "biased" (Hall and Presnell, 1999). In this paper the estimator we consider is the generalised empirical likelihood estimator of Newey and Smith (2004). This estimator is very general and includes a number of well-known estimators including empirical likelihood (Owen, 1988), exponential tilting (Efron, 1981), and euclidean likelihood (Owen, 1991).

In this paper we make two main contributions: first we generalise the efficient bootstrap to weakly dependent observations. To be specific we prove the asymptotic validity of the efficient bootstrap approximation to the true distribution of the generalised method of moment (GMM) estimator. We also consider testing and prove the asymptotic validity of the bootstrap approximation for t-statistics, Hansen's (1982) J statistic for overidentifying restrictions, and for Wald, Lagrange multiplier and distance statistics for nonlinear hypotheses. This extension is theoretically interesting and empirically relevant in economics and finance where most of the observed time series exhibit some form of temporal dependence and most of the hypotheses of interest are typically composite. Second we provide Monte Carlo evidence about the finite sample performance of the proposed bootstrap and compare it with that of the standard bootstrap. The results of the simulations are encouraging and suggest that the proposed bootstrap has competitive finite sample properties compared to those

of the standard bootstrap.

The results of this paper complement those obtained by Flachaire (2005) and Godfrey and Tremayne (2005) among others. These authors recommend using wild (block) bootstrap in the context of (dynamic) heteroskedastic linear regression models. The wild bootstrap however cannot accommodate potential endogeneity of regressors, and, more generally, it requires a regression type of model. In contrast the method proposed in this paper applies to more general statistical models and can accommodate endogeneity; for example nonlinear instrumental variable estimation is allowed.

It is important to note that the results of this paper are related to those obtained by Allen, Gregory and Shimotsu (2005). They propose to use the same type of efficient bootstrap used in this paper. There are however a number of important differences between their paper and ours. First, our efficient bootstrap uses the estimated probabilities to resample the moment indicators, whereas Allen et al. (2005) use the estimated probabilities only to centre the resampled moment indicator. Thus our bootstrap method is the direct extension of that proposed by Brown and Newey (2002). Second we consider efficient bootstrap GMM inferences for possibly nonlinear statistical hypotheses. Third we consider k-step versions of the efficient bootstrap GMM estimators. Finally we resample using the overlapping blocks scheme (the so-called moving block bootstrap) as opposed to nonoverlapping blocks scheme used by Allen et al. (2005). On the other hand we consider stationary α-mixing observations, instead of the more general possibly nonstationary near epoch dependent specification used by Allen et al. (2005).

The rest of the paper is structured as follows: Section 2 briefly introduces the statistical model and GMM estimation and inference. Section 3 describes the efficient bootstrap and develops the necessary asymptotic theory. Section 4 reports the results of the simulations and some concluding remarks. An appendix contains all the proofs and some details about the artificial data used in the simulations.

2 GMM estimation and inference

Let $\{z_t\}_{t\in\mathbb{Z}}$ denote a sequence of $\mathbb{R}^{d_z}$-valued random vectors defined on some probability space $(\Omega, \mathcal{F}, P)$. Let $\beta \in B \subset \mathbb{R}^k$ denote a parameter vector, and let $g(z_t, \beta)$: $\mathbb{R}^d \times B \to \mathbb{R}^l$ $(l \geq k)$ denote a vector of ($\mathcal{F}\backslash$Borel-measurable for each $\beta \in B$) functions satisfying the possibly overidentified moment conditions

$$E[g(z_t, \beta_0)] = 0, \tag{1}$$

where the expectation is with respect to the unknown distribution F of z_t and β_0 is the unique unknown parameter.

Given an observed sample $\{z_t\}_{t=1}^T$, the two-step (efficient) GMM estimators $\widehat{\beta}$ for β_0 is defined as

$$\widehat{\beta} = \arg\min_{\beta \in B} \widehat{g}(\beta)' \widehat{\Omega}\left(\widetilde{\beta}\right)^{-1} \widehat{g}\left(\widehat{\beta}\right),$$

where $\widehat{g}(\beta) = \sum_{t=1}^T g_t(\beta)/T$, $g_t(\beta) = g(z_t, \beta)$, $\widehat{\Omega}\left(\widetilde{\beta}\right)$ is a consistent estimator of the covariance matrix $\Omega(\beta_0) := \lim_{T\to\infty} V\left[T^{1/2}\widehat{g}(\beta_0)\right]$ and $\widetilde{\beta}$ any preliminary $T^{1/2}$-consistent estimator. Under mild regularity conditions Hansen (1982) shows that

$$T^{1/2}\left(\widehat{\beta} - \beta_0\right) \xrightarrow{d} N\left(0, \Sigma(\beta_0)^{-1}\right),$$

where $\Sigma(\beta_0) := G(\beta_0)' \Omega(\beta_0)^{-1} G(\beta_0)$ and $G(\beta_0) := E[\partial\widehat{g}(\beta_0)/\partial\beta']$. Associated with the two-step GMM estimator $\widehat{\beta}$ there is the so-called J-statistic for overidentifying restriction $J\left(\widehat{\beta}\right)$, where $J(\beta) = T\widehat{g}(\beta)'\widehat{\Omega}(\beta)^{-1}\widehat{g}(\beta)$, which can be used to test the correct specification of (1) since Hansen (1982) shows that

$$J\left(\widehat{\beta}\right) \xrightarrow{d} \chi^2_{l-k}.$$

Let $h(\beta) : B \to \mathbb{R}^p$ $(p \leq k)$ denote a vector of continuously differentiable on B functions, and suppose that we want to test the hypothesis $H_0 : h(\beta_0) = 0$. As in Newey and West (1987a) we can define three GMM analogues of the Wald, Lagrange

4

multiplier and likelihood ratio statistics:

$$W\left(\widehat{\beta}\right) = Th\left(\widehat{\beta}\right)' \left[H\left(\widehat{\beta}\right)\widehat{\Sigma}\left(\widehat{\beta}\right)^{-1}H\left(\widehat{\beta}\right)'\right]^{-1}h\left(\widehat{\beta}\right), \tag{2}$$

$$LM\left(\widetilde{\beta}\right) = T\widehat{g}\left(\widetilde{\beta}\right)'\widehat{\Omega}\left(\widetilde{\beta}\right)^{-1}\widehat{G}\left(\widetilde{\beta}\right)\widehat{\Sigma}\left(\widetilde{\beta}\right)^{-1}\widehat{G}\left(\widetilde{\beta}\right)'\widehat{\Omega}\left(\widetilde{\beta}\right)^{-1}\widehat{g}\left(\widetilde{\beta}\right)$$

and

$$D\left(\widetilde{\beta}, \widehat{\beta}\right) = J_T\left(\widetilde{\beta}\right) - J_T\left(\widehat{\beta}\right),$$

where $H\left(\beta\right) = \partial h\left(\beta\right)/\partial\beta'$, $\widetilde{\beta}$ is the constrained GMM estimator for β_0 defined as

$$\widetilde{\beta} = \arg\min_{\beta\in B}\widehat{g}\left(\beta\right)'\widehat{\Omega}\left(\widetilde{\beta}\right)^{-1}\widehat{g}\left(\beta\right) \text{ subject to } h\left(\beta\right) = 0, \tag{3}$$

$\widehat{G}\left(\beta\right) = \sum_{t=1}^{T}\partial g_t\left(\beta\right)/\partial\beta'T$. Under mild regularity conditions Newey and West (1987a) show that

$$W\left(\widehat{\beta}\right), LM\left(\widetilde{\beta}\right), D\left(\widetilde{\beta}, \widehat{\beta}\right) \overset{d}{\rightarrow} \chi_p^2.$$

GMM is widely used in empirical economics and finance -see the special issue of the *Journal of Business and Economic Statistics* 2002 and especially the monograph of Hall (2005) for a survey of recent applications and development of GMM. There exists however Monte Carlo evidence, see for example the special issue of the *Journal of Business and Economic Statistics* 1996, showing that asymptotic theory might not provide a good approximation to the finite sample behaviour of GMM estimators and associated statistics.

In order to improve the finite sample behaviour of GMM statistics one possibility is to use bootstrap methods. Indeed Hall and Horowitz (1996), Andrews (2002) and more recently Inoue and Shintani (2006) use the block bootstrap to obtain asymptotic refinements to the distributions of Hansen's (1982) J statistic for overidentifying restrictions and symmetrical t statistics. All these authors base the bootstrap estimation on centred sample moment conditions. Centring is not necessary to obtain the asymptotic validity of the bootstrap GMM t-statistic (Hahn, 1996), but it is necessary to obtain asymptotic refinements (Hall and Horowitz, 1996). It is also necessary to obtain asymptotic validity of the bootstrap J statistic (Brown and Newey, 2002).

An alternative approach to centring is to use a different estimator of the unknown distribution of the observations that automatically centres the resampled moment indicators, as originally suggested by Brown and Newey (2002) in the context of for identically and independently observations.

3   Efficient block bootstrap

In this section we introduce a modification of Brown and Newey's (2002) efficient bootstrap that is based on the generalised empirical likelihood (GEL) estimator of Newey and Smith (2004) and can be used with weakly dependent observations. Let $\rho(v)$ denote a function of a scalar v that is concave on its domain, an open interval V containing 0, and let $\rho_j(v) = d^j \rho(v)/dv^j$. Examples of $\rho(v)$ are $\log(1-v)$ (empirical likelihood), $-\exp(v)$ (exponential tilting) and $-(1+v)^2/2$ (euclidean likelihood).

To capture the weakly dependent structure of the observations we consider overlapping blocks of observations; let $m = m(T)$ and $b_i = [z'_i, ..., z'_{i+m-1}]'$ be the i-th block of m consecutive observations for $1 \leq i \leq Q = T - m + 1$. Define now the blockwise moment function

$$\psi(b_i, \beta) := \psi_i(\beta) = \sum_{j=1}^m g(z_{i+j-1}, \beta)/m, \tag{4}$$

and note that if (1) holds then $E[\psi_i(\beta_0)] = 0$. The (blockwise) GEL estimator of the unknown distribution F consistent with (1) is

$$\widehat{F}_{\widehat{\pi}}(z) = \sum_{i=1}^Q \sum_{j=1}^m \widehat{\pi}_i^\rho I\{z_{i+j-1} \leq z\}/m,$$

where

$$\widehat{\pi}_i^\rho = \rho_1\left(\widehat{\lambda}'\psi_i\left(\widehat{\beta}\right)\right) / \sum_{i=1}^Q \rho_1\left(\widehat{\lambda}'\psi_i\left(\widehat{\beta}\right)\right) \tag{5}$$

are the so-called GEL implied probabilities, $\widehat{\lambda} = \arg\max_{\lambda \in \widehat{V}_Q} \sum_{i=1}^Q \rho\left(\lambda'\psi_i\left(\widehat{\beta}\right)\right)/Q$, $\widehat{V}_Q := \{\lambda : \lambda'\psi_i(\beta) \in V, i = 1, ..., Q\}$ and $\widehat{\beta}$ is any efficient estimator for β_0, such as

the efficient GMM or any asymptotically equivalent GEL estimators defined as $\widehat{\beta} = \arg\min_{\beta \in B} \sum_{i=1}^{Q} \rho\left(\widehat{\lambda}' \psi_i(\beta)\right)/Q$. Note that the computation of $\widehat{\lambda}$ is straightforward because of the global concavity of $\rho(\cdot)$, and that

$$\widehat{\lambda} = \Omega(\beta_0)^{-1} \widehat{\psi}\left(\widehat{\beta}\right) + o_p(1), \tag{6}$$

where $\widehat{\psi}(\cdot) = \sum_{i=1}^{Q} \psi_i(\cdot)/Q$ (Bravo, 2009).

The efficient block bootstrap (EBB henceforth) uses the GEL implied probabilities $\widehat{\pi}_i^\rho$ to resample the blocks b_i to obtain n blocks b_i^* with $n = \lfloor T/m \rfloor$ where $\lfloor \cdot \rfloor$ is the integer part function, that is each bootstrap block b_i^* is drawn independently with replacement with probability $\Pr\left(b_j^* = b_i\right) = \widehat{\pi}_i^\rho$ $i = 1, ..., Q$, $j = 1, ..., n$. Given b_i^* $(i = 1, ..., n)$ we can construct EBB analogues of the blockwise moment indicators (4), that is $\psi_i^*(\beta)$ $(i = 1, ..., n)$. These moment indicators satisfy the sample moment condition because $E^*[\psi_i^*(\beta)] = 0$ when $\beta = \widehat{\beta}$ by construction (by (6) and Lemma 4 in the Appendix), where E^* denote the expectation relative to the EBB bootstrap distribution conditional on the original sample.

The EBB two-step GMM estimator $\widehat{\beta}^*$ is defined as

$$\widehat{\beta}^* = \arg\min_{\beta \in B} \widehat{\psi}^*(\beta)' \widehat{\Omega}^*\left(\widetilde{\beta}^*\right)^{-1} \widehat{\psi}^*(\beta), \tag{7}$$

where $\widetilde{\beta}^*$ is a preliminary consistent EBB GMM estimator, such as an EBB one-step GMM estimator. The latter can be defined in an analogous way as

$$\widehat{\beta}^* = \arg\min_{\beta \in B} \widehat{\psi}^*(\beta)' \widehat{W} \widehat{\psi}^*(\beta),$$

where $\widehat{W}$ is a possibly random positive semidefinite matrix. Furthermore we can define the EBB t-statistic and J statistic for overidentifying restrictions as

$$t^* = T^{1/2}\left(\widehat{\beta}_j^* - \widehat{\beta}_j\right)/\left(\widehat{\Sigma}^*\left(\widehat{\beta}^*\right)_{jj}^{-1}\right)^{1/2} \quad j = 1, ..., k, \tag{8}$$

$$J^*\left(\widehat{\beta}^*\right) = T\widehat{\psi}^*\left(\widehat{\beta}^*\right)' \widehat{\Omega}^*\left(\widehat{\beta}^*\right)^{-1} \widehat{\psi}^*\left(\widehat{\beta}^*\right),$$

7

where, with a slight abuse of notation, $mn = T$. Thus an EBB t-test and J test reject, respectively, if $|t| \geq \widehat{q}_\alpha^t$ and $J \geq \widehat{q}_\alpha^J$, where $\widehat{q}_\alpha^t$ and $\widehat{q}_\alpha^J$ are the $1 - \alpha$ percentile of the distributions of t^* and $J^*\left(\widehat{\beta}^*\right)$ obtained by computing (8) B times.

To define EBB analogues of the three GMM based statistics (2) that can be used to test $H_0 : h\left(\beta_0\right) = 0$ let

$$\widetilde{\pi}_i^\rho = \rho_1\left(\widetilde{\lambda}'\psi_i\left(\widetilde{\beta}\right)\right) / \sum_{i=1}^Q \rho_1\left(\widetilde{\lambda}'\psi_i\left(\widetilde{\beta}\right)\right) \tag{9}$$

denote the restricted implied probabilities, where $\widetilde{\lambda} = \arg\max_{\lambda \in \widehat{V}_Q} \sum_{i=1}^Q \rho\left(\lambda'\psi_i\left(\widetilde{\beta}\right)\right) / Q$ a and $\widetilde{\beta}$ is any two-step constrained estimator for β_0, such as the constrained GMM estimator defined in (3)or any asymptotically equivalent blockwise GEL estimators defined as $\widetilde{\beta} = \arg\min_{\beta \in B} \sum_{i=1}^Q \rho\left(\widetilde{\lambda}'\psi_i\left(\beta\right)\right) / Q$ subject to $h\left(\beta\right) = 0$. As with the unconstrained EBB two-step GMM estimator we use $\widetilde{\pi}_i^\rho$ to obtain moment functions $\psi_i^*\left(\beta\right)$ $(i = 1, ..., n)$ that by construction $E^*\left[\psi_i^*\left(\beta\right)\right] = 0$ when $\beta = \widetilde{\beta}$.

The EBB constrained two step GMM estimator is

$$\widetilde{\beta}^* = \arg\min_{\beta \in B} \widehat{\psi}^*\left(\beta\right)' \widehat{\Omega}^*\left(\overline{\beta}^*\right)^{-1} \widehat{\psi}^*\left(\widehat{\beta}\right),$$

where $\overline{\beta}^*$ is a preliminary consistent EBB constrained GMM estimator. The EBB analogues of (2) are

$$W^*\left(\widehat{\beta}^*\right) = T\left[h\left(\widehat{\beta}^*\right) - h\left(\widehat{\beta}\right)\right]'\left[H\left(\widehat{\beta}^*\right)\widehat{\Sigma}^*\left(\widehat{\beta}^*\right)^{-1}H\left(\widehat{\beta}^*\right)'\right]^{-1}\left[h\left(\widehat{\beta}^*\right) - h\left(\widehat{\beta}\right)\right],$$

$$LM^*\left(\widetilde{\beta}\right) = T\widehat{\psi}^*\left(\widetilde{\beta}^*\right)'\widehat{\Omega}^*\left(\widetilde{\beta}^*\right)^{-1}\widehat{G}^*\left(\widetilde{\beta}^*\right)\widehat{\Sigma}^*\left(\widetilde{\beta}^*\right)^{-1}\widehat{G}^*\left(\widetilde{\beta}^*\right)'\widehat{\Omega}^*\left(\widetilde{\beta}^*\right)^{-1}\widehat{\psi}^*\left(\widetilde{\beta}^*\right)$$

and $$\tag{10}$$

$$D^*\left(\widetilde{\beta}, \widehat{\beta}\right) = J^*\left(\widetilde{\beta}^*\right) - J^*\left(\widehat{\beta}^*\right).$$

Thus an EBB Wald, Lagrange multiplier and distance tests, say S, reject if $S \geq \widehat{q}_\alpha^s$ where $\widehat{q}_\alpha^s$ is the $1 - \alpha$ percentile of the distribution obtained by computing B times any of the three statistics (10).

Like any other resampling method EBB can be computationally very demanding when applied to nonlinear moment conditions models. One way to reduce the computational cost is to follow Davidson and MacKinnon's (1999) suggestion and use an approximate k-step $(k = 1, 2, ...)$ EBB two-step GMM estimator alternative to (the fully optimised) $\overline{\beta}^*$, that is

$$\overline{\beta}^{*(j)} = \overline{\beta}^{*(j-1)} - \widehat{\Sigma}^* \left(\overline{\beta}^{*(j-1)}\right)^{-1} \widehat{G}^* \left(\overline{\beta}^{*(j-1)}\right)' \widehat{\Omega}^* \left(\overline{\beta}^{*(j-1)}\right)^{-1} \widehat{\psi}^* \left(\overline{\beta}^{*(j-1)}\right) \ \ 1 \leq j \leq k$$

where $\overline{\beta}^{*(0)} = \overline{\beta}$ and $\overline{\beta}$ can be either the unconstrained or constrained two-step GMM estimator.

Asymptotic theory

The following assumptions are standard in the GMM/GEL literature on nonlinear (differentiable) moment condition models with stationary weakly dependent observations - see for example Wooldridge (1994), Hall (2005), and Politis and Romano (1992), and Goncalves and White (2004) for a bootstrap analogue.

A1 $\{z_t\}_{t \in \mathbb{Z}}$ is a strictly stationary strong mixing sequence of size $-\alpha/(\alpha - 2)$ where $\alpha > 2$,

A2 (i) The parameter space B is compact, (ii) $\beta_0 \in B$ is the unique solution to $E\left[g_t\left(\beta_0\right)\right] = 0$, (iii) $g_t\left(\beta\right)$ is continuous $a.s.$ at each $\beta \in B$, (iv) (a) $E\left[\sup_{\beta \in B} \|g_t\left(\beta\right)\|\right] < \infty$, (b) $E\left[\|g_t\left(\beta_0\right)\|^{2\alpha+\delta}\right] < \infty$ for some $\delta > 0$ (v) $\Omega\left(\beta_0\right) := \lim_{T \to \infty} V\left[n^{1/2}\widehat{g}\left(\beta_0\right)\right]$ is positive definite,

A3 (i) $\beta_0 \in int\left(B\right)$, (ii) $g_t\left(\beta\right)$ is continuously differentiable $a.s.$ in a convex neighbourhood $\mathcal{N}$ of β_0 $\forall t$ and $\forall \beta \in \mathcal{N}$ (iii) $E\sup_{\beta \in \mathcal{N}} \|\partial g_t\left(\beta\right)/\partial\beta'\|^2 < \infty$ (iv) $rank\left[G\left(\beta_0\right)\right] = k$ where $G\left(\beta_0\right) = E\left[\partial g_t\left(\beta_0\right)/\partial\beta'\right]$,

A4 $\rho\left(\cdot\right)$ is twice continuously differentiable in an open neighbourhood of 0, and $\rho_k\left(0\right) = -1$ for $k = 1, 2$.

The following theorem establishes the asymptotic validity of the EBB two-step GMM estimator $\widehat{\beta}^*$ and of the J statistic for overidentifying restrictions $J^*\left(\widehat{\beta}^*\right)$.

Theorem 1 *Suppose that A1-A4 hold. If $m = o\left(T^{1/2}\right)$ then*

$$\sup_{x \in \mathbb{R}^k} \left| P^* \left[T^{1/2} \left(\widehat{\beta}^* - \widehat{\beta} \right) \leq x \right] - P \left[T^{1/2} \left(\widehat{\beta} - \beta_0 \right) \leq x \right] \right| \xrightarrow{p} 0,$$

$$\sup_{x \in \mathbb{R}^+} \left| P^* \left[J^* \left(\widehat{\beta}^* \right) \leq x \right] - P \left[J \left(\widehat{\beta} \right) \leq x \right] \right| \xrightarrow{p} 0.$$

The following theorem establishes the asymptotic validity of the EBB Wald, Lagrange multiplier and distance statistics $W^*\left(\widehat{\beta}^*\right)$, $LM^*\left(\widetilde{\beta}\right)$ and $D^*\left(\widetilde{\beta},\widehat{\beta}\right)$.

Theorem 2 *Suppose that A1-A4 hold. If $rank\left[H\left(\beta_0\right)\right] = p$ and $m = o\left(T^{1/2}\right)$ then*

$$\sup_{x \in \mathbb{R}^+} \left| P^* \left[W^* \left(\widehat{\beta}^* \right) \leq x \right] - P \left[W \left(\widehat{\beta} \right) \leq x \right] \right| \xrightarrow{p} 0,$$

$$\sup_{x \in \mathbb{R}^+} \left| P^* \left[LM^* \left(\widetilde{\beta}^* \right) \leq x \right] - P \left[LM \left(\widetilde{\beta} \right) \leq x \right] \right| \xrightarrow{p} 0$$

and

$$\sup_{x \in \mathbb{R}^+} \left| P^* \left[D^* \left(\widetilde{\beta}^*, \widehat{\beta}^* \right) \leq x \right] - P \left[D \left(\widetilde{\beta}, \widehat{\beta} \right) \leq x \right] \right| \xrightarrow{p} 0.$$

Finally the following theorem shows that the k-step $(k = 1, 2, ..)$ EBB two-step GMM estimator $\overline{\beta}^{*(k)}$ achieves the same asymptotic accuracy as that of the fully optimised one $\overline{\beta}^*$.

Theorem 3 *Suppose that A1-A4 hold. If $m = o\left(T^{1/2}\right)$ then*

$$\sup_{x \in \mathbb{R}^k} \left| P^* \left[T^{1/2} \left(\overline{\beta}^{*(k)} - \overline{\beta}^* \right) \leq x \right] - P \left[T^{1/2} \left(\overline{\beta}^* - \overline{\beta} \right) \leq x \right] \right| \xrightarrow{p} 0.$$

4 Monte Carlo evidence

In this section we use simulations to evaluate the finite sample properties of the EBB and compare them with those obtained by the standard block bootstrap (BB henceforth) and by standard asymptotic approximations. We focus on the t and J

statistics partly because they are routinely used in empirical work and partly because of their well documented finite sample over-rejections problems.

We consider an intertemporal consumption based asset pricing model used by Tauchen (1986), Kocherlakota (1990) and Wright (2003) among others. Consumption and dividend growth are assumed to follow a first order vector autoregression

$$
\begin{bmatrix} \log\left(c_t/c_{t-1}\right) \\ \log\left(d_t/d_{t-1}\right) \end{bmatrix} = \begin{bmatrix} \mu_{c0} \\ \mu_{d0} \end{bmatrix} + \Phi_0 \begin{bmatrix} \log\left(c_{t-1}/c_{t-2}\right) \\ \log\left(d_{t-1}/d_{t-2}\right) \end{bmatrix} + \begin{bmatrix} \varepsilon_{ct} \\ \varepsilon_{dt} \end{bmatrix}, \tag{11}
$$

where c_t is consumption, d_t is dividend, Φ_0 is a 2×2 matrix of constants and $[\varepsilon_{ct}, \varepsilon_{dt}]' \sim N\left(0, \Sigma_0\right)$. Returns are generated so as to satisfy the stochastic Euler equation

$$
E\left[\beta_{10}\left(c_t/c_{t-1}\right)^{-\beta_{20}} r_t - 1_s | \mathcal{I}_{t-1}\right] = 0 \tag{12}
$$

where $\beta_0 = [\beta_{10}, \beta_{20}]'$ is the unknown parameters vector, r_t is an s-dimensional vector of returns, 1_s is an s-dimensional vector of ones and $\mathcal{I}_{t-1}$ is the information set at time $t-1$. To generate consumption and returns time series consistent with both (11) and (12) we use the same method proposed by Tauchen (1986) and Tauchen and Hussey (1991). This method fits a 16 state Markov chain to $[\log\left(c_t/c_{t-1}\right), \log\left(d_t/d_{t-1}\right)]'$ to approximate (11) and then uses numerical methods to approximate the expectation in (12). The resulting (discretised) system of equations is then used to obtain the prices p_t (and hence the returns r_t) of stocks and risk-free bonds in each time period (see the Appendix for some details).

We consider two returns: one based on a stock, say r_t^s, and one risk free, say r_t^f. Estimation of β_0 is based on

$$
\widehat{g}\left(\beta_0\right) = \sum_{t=1}^{T} f_t\left(\beta_0\right) r_t \otimes z_{t-1}/T, \tag{13}
$$

where $f_t\left(\beta_0\right) = \beta_{10}\left(c_t/c_{t-1}\right)^{-\beta_{20}}$, $r_t = \left[r_t^s, r_t^f\right]'$, $\otimes$ is the Kronecker product, and $z_t = [1, r_t', c_t/c_{t-1}]'$ is a vector of so-called instruments. Thus (13) consists of 8 estimating equations for 2 unknown parameters, that is the degree of overidentification

11

is 6. To compute the covariance $\Omega(\beta_0)$ in the two-step GMM we use the Newey-West estimator (Newey and West, 1987b) , whereas we use a blockwise bootstrap covariance estimator either centred as in Politis and Romano (1992) or with the implied probabilities $\widehat{\pi}_i^\rho$ in the bootstrap two-step GMM. These estimators are asymptotically equivalent for $m = o\left(T^{1/2}\right)$ and have the same optimal block length parameter $m^* = \left\lfloor \gamma T^{1/3} \right\rfloor$ for any choice of finite $\gamma > 0$. We consider two $\widehat{\pi}_i^\rho$, namely

$$
\begin{aligned}
\widehat{\pi}_i^{EL} &= Q^{-1}\left(1 - \widehat{\lambda}'\psi_i\left(\widehat{\beta}\right)\right)^{-1}, \\
\widehat{\pi}_i^{EU} &= 1 - \widehat{\psi}\left(\widehat{\beta}\right)'\widehat{\Omega}\left(\widehat{\beta}\right)^{-1}\psi_i\left(\widehat{\beta}\right) / \left[Q\left(1 - J_\psi\left(\widehat{\beta}\right)\right)\right],
\end{aligned}
$$

which correspond to empirical likelihood (EL) and euclidean likelihood (EU), respectively. Note that the latter does not require to numerically find $\widehat{\lambda}$ because in this case $\widehat{\lambda} = \widehat{\Omega}\left(\widehat{\beta}\right)^{-1}\widehat{\psi}\left(\widehat{\beta}\right)$ exactly.

In the simulations we consider two parameterisations of (11) and (12) namely

Case 1. $\beta_0 = [0.97, 1.36]'$, $\mu_0 = [0.018, 0.013]'$,

$$
\Phi_0 = \begin{bmatrix} -0.5 & 0.00 \\ 0.00 & -0.5 \end{bmatrix}, \ \Sigma_0 = \begin{bmatrix} 0.01 & 0.005 \\ 0.005 & 0.01 \end{bmatrix},
$$

and

Case 2. $\beta_0 = [0.97, 0.36]'$, $\mu_0 = [0.02, 0.03]'$,

$$
\Phi_0 = \begin{bmatrix} -0.1 & 0.05 \\ 0.20 & 0.12 \end{bmatrix}, \ \Sigma_0 = \begin{bmatrix} 0.01 & 0.02 \\ 0.02 & 0.05 \end{bmatrix},
$$

which are in the same spirit of those used by Tauchen (1986) and Kocherlakota (1990), respectively. The sample sizes are $T = 100$ and $T = 400$, and the block length parameter m is chosen using Newey and West's (1994) method. The number of bootstrap repetitions is 500 and the number of Monte Carlo replications is 5000.

The results of the simulations are presented using the graphical methods proposed by Davidson and MacKinnon (1998). To save space we report only the results for Case 2 and $T = 100$. The results for Case 1 and $T = 400$ are very similar and
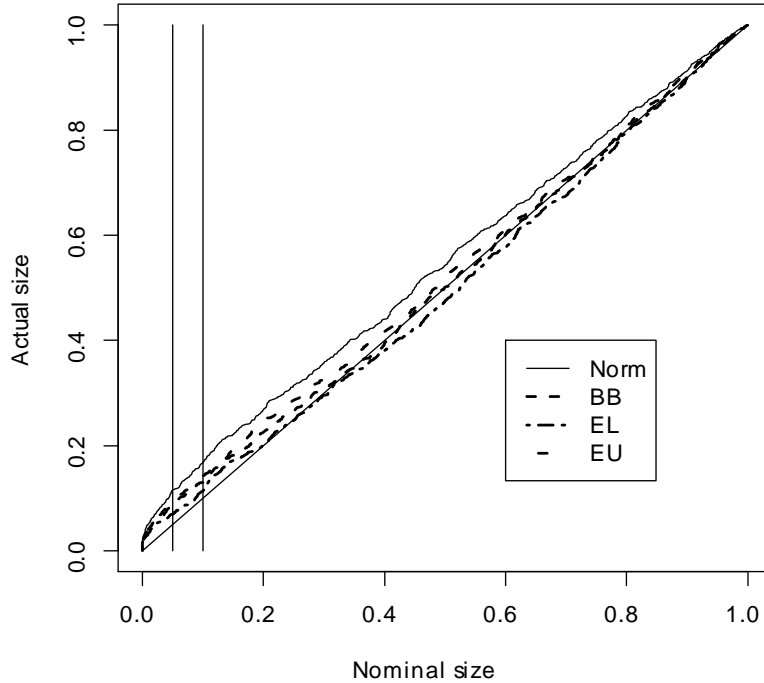
Figure 1: P-value plots of the t-statistics for $H_0 : \beta_1 = \beta_{10}$. The two vertical lines correspond to the 0.05 and 0.10 nominal size.

available upon request. Figures 1-3 show the p-value plots of the two t and J statistics. These plots show the empirical distribution function $\widehat{F}(x_i)$ of the p-values of the simulated statistics against the set of points x_i $(i = 1, ..., l)$ in the $(0, 1)$ interval with $l = 1000$. The closer is the plot to the 45-degree line the more accurate is the corresponding approximation. In the plots the solid lines correspond to the asymptotic approximation ("Norm" or "χ_6^2" in the legend), the dashed lines to the block bootstrap approximation ("BB" in the legend), the two-dash lines to the empirical likelihood based efficient bootstrap approximation ("EL" in the legend) and the dot-dash lines to the euclidean likelihood efficient bootstrap approximation ("EU" in the legend).
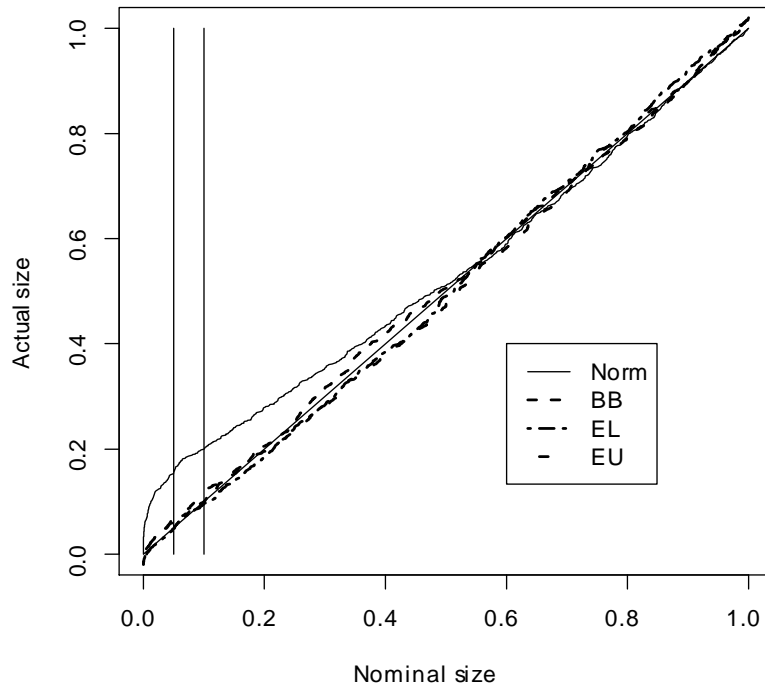
13

Figure 2: P-value plots of the t-statistics for $H_0 : \beta_2 = \beta_{20}$. The two vertical lines correspond to the 0.05 and 0.10 nominal size.
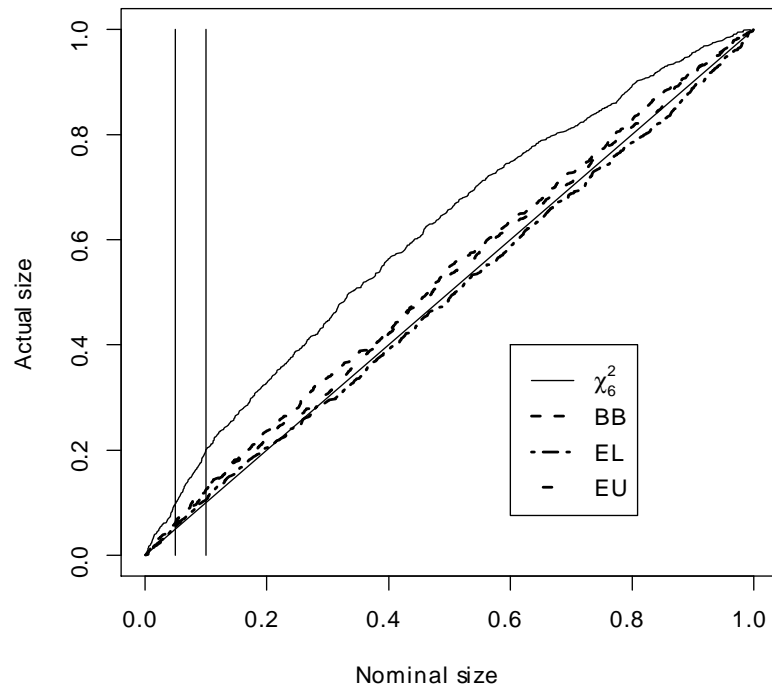
Figure 3: P-value plots of the J-statistics. The two vertical lines correspond to the 0.05 and 0.10 nominal size.

Figures 1-3 show that the t and J statistics based on the asymptotic approximation over-reject, whereas those based on the three block bootstrap approximations perform significantly better, particularly those based on EL. For example the actual size of the J-statistic at the 0.05 nominal level is 0.113 whereas the size of the three bootstrapped J-statistics are 0.052 for that based on EL, 0.059 for that based on EU and 0.061 for that based on BB. Figures 4-6 report the p-value discrepancy plots, which show the discrepancy of $\widehat{F}(x_i) - x_i$ against x_i. The figures also feature the 0.05 critical value of the Kolmogorov-Smirnov (KS)-type statistic

$$\max_i \left| \widehat{F}(x_i) - x_i \right|, \tag{14}$$

which is used to assess whether the discrepancies can be explained by experimental randomness. Figures 4-6 show that the discrepancies for the three bootstrap procedures are not significant. Indeed the p-values of (14) are typically above 0.50. The only exception is for the J statistic based on BB whose p-value is around 0.13, which indicates that in this case the BB approximation is less satisfactory. Figures 4-6 also confirm that among the three different block bootstraps the two based on EBB have smaller p-value discrepancies (with those based on EL having the smallest discrepancies), implying an overall better finite sample approximation to the unknown distributions of both the t and J statistics. As a further indication of the better quality of approximations obtained using EBB we have computed the probabilities of BB leading to size distortions that could have been avoided using both EL-EBB and EU-EBB. Table 1 reports these probabilities for the conventional 0.05 and 0.10 nominal sizes and both the t and J statistics.

Table 1

	EL		EU	
nominal size	0.05	0.10	0.05	0.10
t_1	0.24	0.17	0.14	0.11
t_2	0.12	0.09	0.09	0.05
J	0.19	0.14	0.12	0.09

t_j is the BB t-statistic for $H_0{:}\beta_j{=}\beta_{j0}$ $(j{-}1{,}2)$ and J is the BB J-statistic.

Table 1 shows that if we use for example EL-EBB instead of BB for testing $H_0 : \beta_1 = \beta_{10}$ we are 24% less likely to have a size distortion at the 0.05 nominal size. Likewise if we use EU-EBB for the J-statistic we are 9% less likely to have a size distortion at the 0.10 nominal size.

Before we consider the finite sample power of the t and J statistics, it should be noted that although the various block bootstrap procedures improve considerably their finite sample behaviour, some small size distortions are still present, particularly for the J statistic with BB. However this fact seems to be typical of overidentified moment conditions models and is consistent for example with the findings of Hall and Horowitz (1996).

Figures 7-9 show the size-power curves, which plot the power of a test statistic against its true size. The figures show that in terms of power EL based EBB (EL-EBB henceforth) uniformly dominates the other procedures for a given true size; for example in Figure 7 the t-statistic based on EL-EBB is on average about 18% more powerful than the one based on the normal approximation, whereas in Figure 10 the J statistic based on EL-EBB is on average around 32% more powerful than the one based on BB. These plots also show that neither of the other two block bootstrap approximations dominate the one based on the asymptotic distribution. However the statistics based on EU-EBB are more powerful than those based on BB approximation (from around 5% (on average) in Figure 7 to around 12% (on average) in Figure 9).
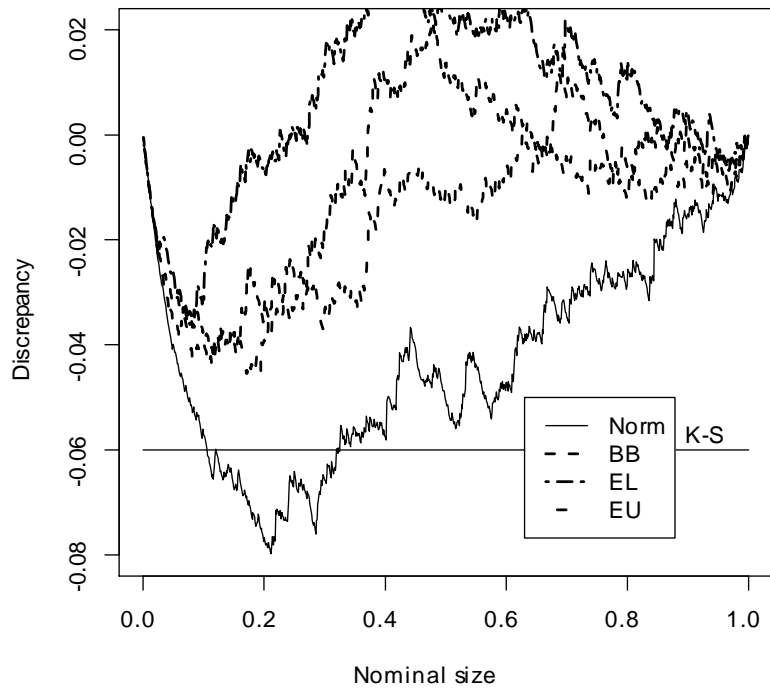
Figure 4: P-value discrepancy plots of the t statistic for $H_0 : \beta_1 = \beta_{10}$. The horizontal line corresponds to the 0.05 critical value of the K-S statistic (14).
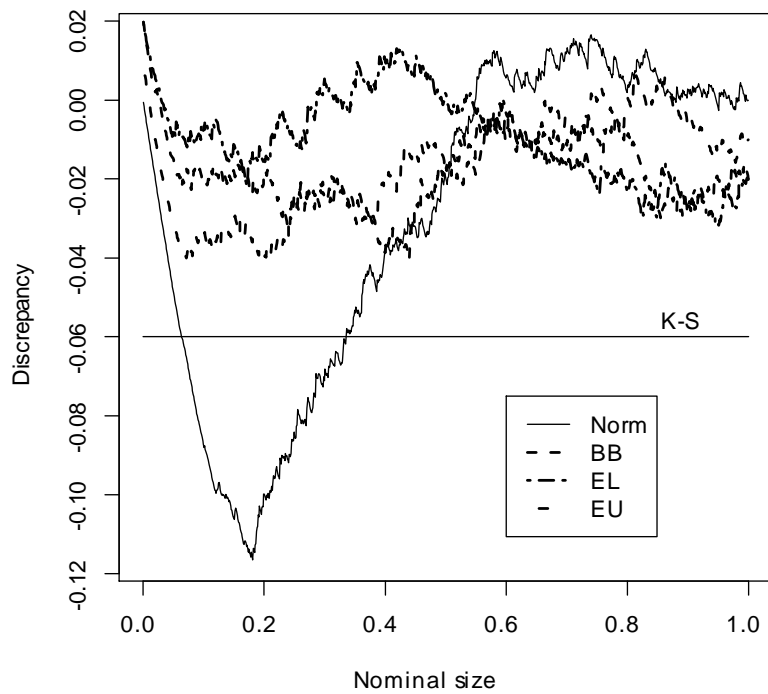
Figure 5: P-value discrepancy plots of the t statistic for $H_0 : \beta_2 = \beta_{20}$. The horizontal line corresponds to the 0.05 critical value of the K-S statistic (14).
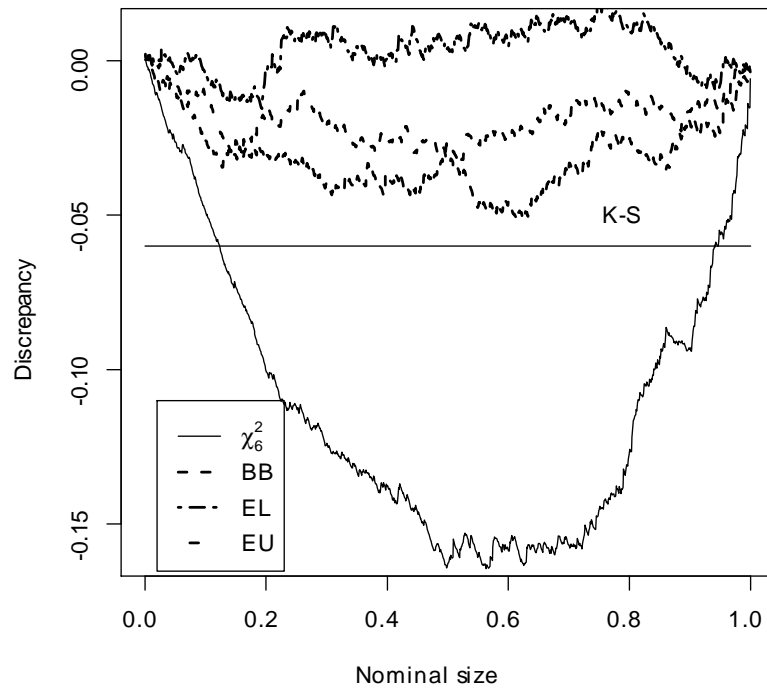
Figure 6: P-value discrepancy plots of the J statistic. The horizontal line corresponds to the 0.05 critical value of the K-S statistic (14).
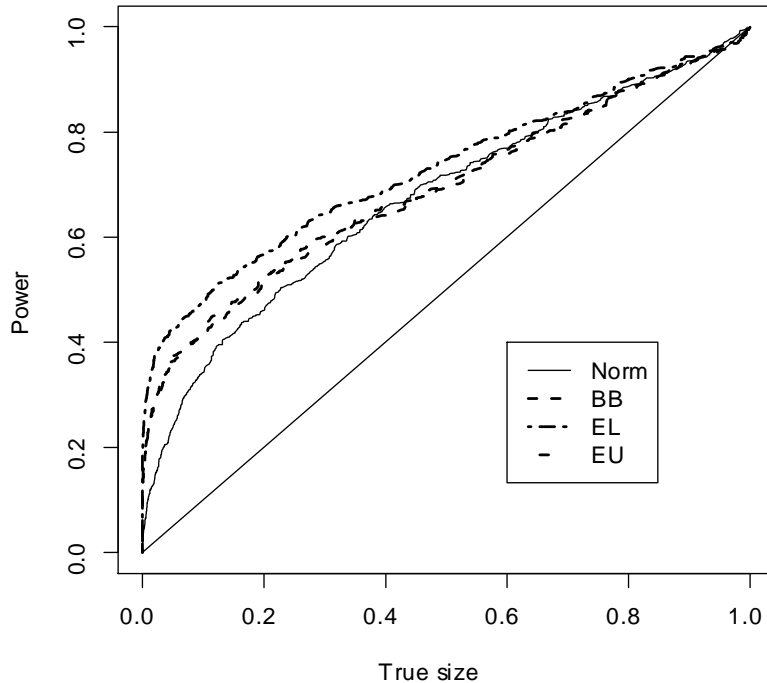
Figure 7: Size-power curves of the t statistic for $H_0 : \beta_1 = \beta_{10}$.

This suggests that, in general, the efficient bootstrap method of this paper has a clear advantage over the standard bootstrap in terms of power.

We now consider the one-step version of both BB and EBB based GMM. These estimators are computationally very attractive because they are simply the first (bootstrap) iteration from the original GMM estimator. Overall the finite sample properties of the resulting bootstrapped t-statistics are very similar. Therefore we only report the results for the t-statistic for $H_0 : \beta_2 = \beta_{20}$ and the J-statistic. Figures 10-11 show the differences between the p-value discrepancy plots of the fully optimised with those based on the one-step version of the bootstrap, with a negative value indicating a larger discrepancy for the one-step estimator. It is clear that BB has the largest discrepancy difference, however all of the differences are statistically
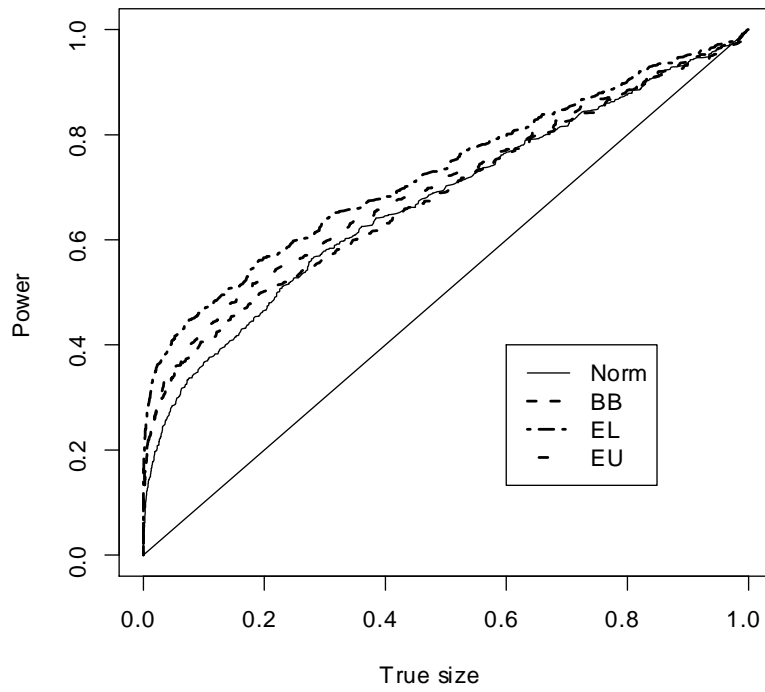
21

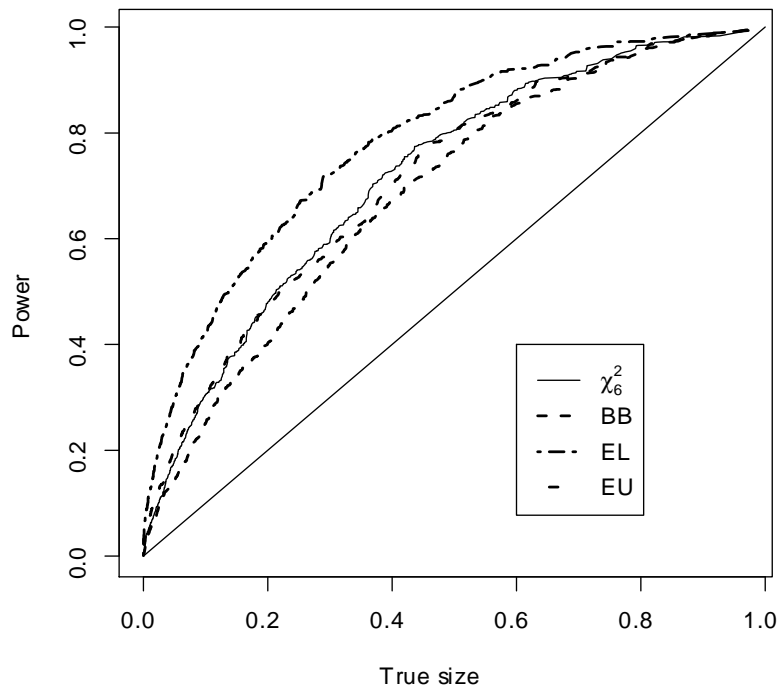Figure 8: Size-power curves of the t statistic for $H_0 : \beta_2 = \beta_{20}$.

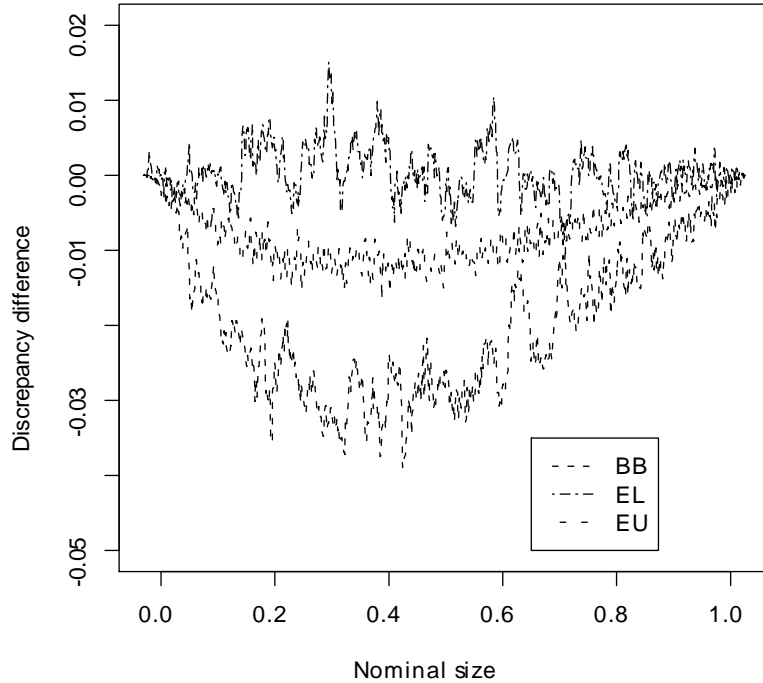Figure 9: Size-power curves of the J statistic.

Figure 10: P-value discrepancy difference plots of the t-statistic for $H_0 : \beta_2 = \beta_{20}$.

insignificant, as indicated by p-values above 0.60 of the K-S statistics for the equality of their distributions. Thus Figures 10-11 indicate that in terms of accuracy the one-step bootstrap approximation is a valid alternative to that based on the fully optimised bootstrap.

Figures 12-13 report the size-power difference curves between the fully optimised and the one-step version of the three block bootstrap procedures. For the t-statistic the differences are rather small, particularly for that based on EL-EBB. For the J-statistic however there is a clear loss in power in the case of BB, which is on average about 15% and 11.5% less powerful than EBB-EL and EBB-EU, respectively.

The Monte Carlo results of this section suggest that EBB and in particular EL-
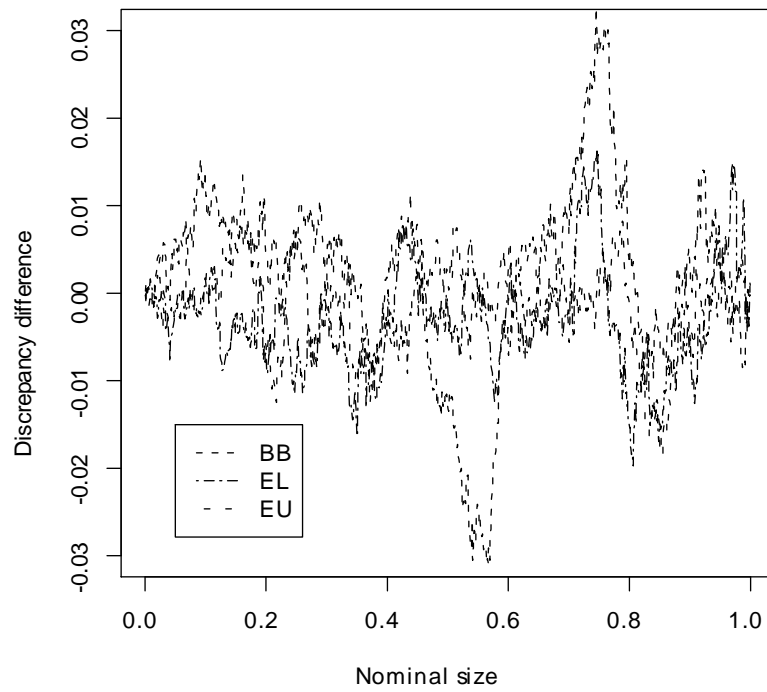
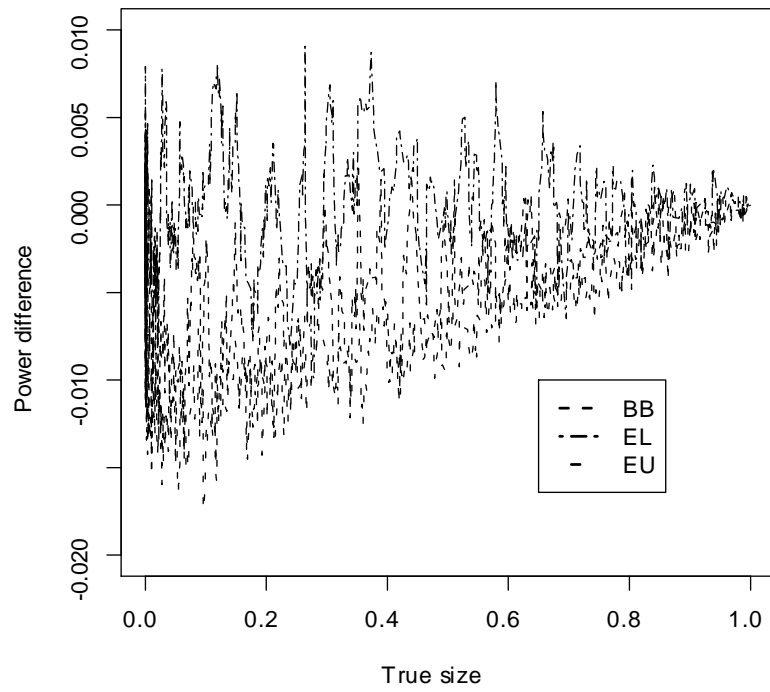Figure 11: P- value discrepancy difference plots of the J statistic.

Figure 12: Size-power difference plots of the t statistic for $H_0 : \beta_2 = \beta_{20}$.
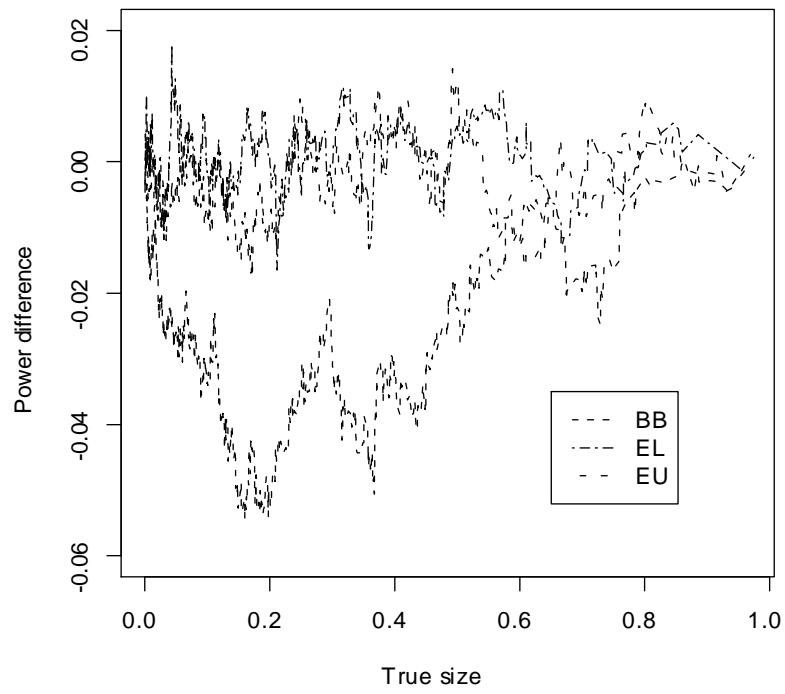
Figure 13: Size-power difference plots of the J statistic.

EBB is a valid alternative to BB. Compared to the latter, EBB requires in general an additional maximisation. On the other hand the overall quality of the resulting approximation seems to be superior to that based on BB. The graphical analysis indicates that EBB provides a slightly more accurate finite sample approximation to the unknown distributions of both t and J statistics than that obtained by BB. However the real advantage of using EBB comes when considering the power properties of the resulting statistics. The graphical analysis indeed indicates that statistics based on EL-EBB not only outperform those based on BB, but, perhaps more remarkably, also those based on standard asymptotic approximations. The results of the simulations also suggest that a one-step version of the EBB is an accurate and computationally convenient alternative to its fully optimised analogue. This should be particularly convenient when estimation is numerically difficult and/or very time consuming.

Acknowledgements

Appendix

Throughout the Appendix we use the following abbreviations: BCLT, CMT, ULLN stand for bootstrap central limit theorem as in Goncalves and White (2004), continuous mapping theorem, and uniform law of large numbers as in Wooldridge (1994). "CS", "M" and "T" stand for Cauchy-Schwarz, Markov and Triangle inequalities; "$\overset{p^*-p}{\rightarrow}$", "$\overset{d^*-p}{\rightarrow}$" denote, respectively, convergence in bootstrap probability and in bootstrap distribution in probability, "$O_{p^*-p}(\cdot)$" and "$o_{p^*-p}(\cdot)$" are the bootstrap stochastic order of magnitude in probability. Finally "$\overset{a^*-p}{\rightarrow}$" denotes asymptotically equivalent bootstrap random vectors, i.e. $X^* \overset{a^*-p}{=} Y^* \Rightarrow X^* = Y^* + o_{p^*-p}(1)$, when X^* and

Y^* are $O_{p^*-p}(1)$.

Preliminary lemmas

Lemma 4 *Suppose A1-A4 hold. Then for $m = o(T)$*

$$\max_i \left| \widehat{\pi}_i - \left(1 + \widehat{\lambda}' \psi_i \left(\widehat{\beta} \right) \right) / Q \right| \xrightarrow{p} 0.$$

Proof. By Bravo (2009) $\max_i \sup_{\beta \in \mathcal{B}} \left| \widehat{\lambda}' \psi_i(\beta) \right| = o_p(1)$ and the result follows by a mean value expansion of $\widehat{\pi}_i$, results of Fitzenberger (1997), ULLN and simple algebra. ∎

Lemma 5 *Suppose A1-A4 hold. Then for $m = o(T)$, $l = 0, 1$ and $j = 1, ..., k$*

$$E^* \left[\sup_{\beta \in B} \left\| \partial^l \psi^*(\beta) / \partial \beta_j^l \right\| \right] < \infty \quad \text{in probability.}$$

Proof. By Lemma 4, results of Fitzenberger (1997), and ULLN

$$E^* \left[\sup_{\beta \in B} \left\| \partial^l \psi^*(\beta) / \partial \beta_j^l \right\| \right] = \sum_i \widehat{\pi}_i \sup_{\beta \in B} \left\| \partial^l \psi_i(\beta) / \partial \beta_j^l \right\|$$

$$\leq \sum_t \left(\sup_{\beta \in B} \left\| \partial^l g_t(\beta) / \partial \beta_j^l \right\| \right) / T + O_p(m/T) = O_p(1).$$

∎

Lemma 6 *Suppose A1-A4 hold. Then for $m = o(T)$, $l = 0, 1$ and $j = 1, ..., k$*

$$\sup_{\beta \in B} \left\| \partial^l \widehat{\psi}^*(\beta) / \partial \beta_j^l - E^* \left[\partial^l \psi_i^*(\beta) / \partial \beta_j^l \right] \right\| \xrightarrow{p^*-p} 0.$$

Proof. Let $\partial^l \overline{\psi}^*(\beta) / \partial \beta_j^l = \partial^l \widehat{\psi}^*(\beta) / \partial \beta_j^l - E^* \left[\partial^l \psi_i^*(\beta) / \partial \beta_j^l \right]$, and

$$\partial^l \psi_\delta^*(\beta, \beta') / \partial \beta_j^l = \sup_{\beta \in B} \sup_{\beta' \in \mathcal{N}(\beta, \delta)} \sum \left\| \partial^l \psi_i^*(\beta) / \partial \beta_j^l - \partial^l \psi_i^*(\beta') / \partial \beta_j^l \right\| / Q.$$

By Lemma 5 $E^* \left[\partial^l \psi_\delta^*(\beta, \beta') / \partial \beta_j^l \right] < \infty$ in probability and thus $E^* \left[\partial^l \psi_\delta^*(\beta, \beta') / \partial \beta_j^l \right] \xrightarrow{p} 0$ as $\delta \to 0$. Note that

$$\sup_{\beta \in B} \sup_{\beta' \in \mathcal{N}(\beta, \delta)} \left\| \partial^l \overline{\psi}^*(\beta) / \partial \beta_j^l - \partial^l \overline{\psi}^*(\beta') / \partial \beta_j^l \right\| \leq \partial^l \psi_\delta^*(\beta, \beta') / \partial \beta_j^l + E^* \left[\partial^l \psi_\delta^*(\beta, \beta') / \partial \beta_j^l \right]$$

and thus by M

$$\lim P^* \left(\sup_{\beta \in B} \sup_{\beta' \in \mathcal{N}(\beta, \delta)} \left\| \partial^l \overline{\psi}^* (\beta) / \partial \beta_j^l - \partial^l \overline{\psi}^* (\beta') / \partial \beta_j^l \right\| > \varepsilon \right) \leq$$

$$\lim 2 E^* \left[\partial^l \psi_\delta^* (\beta, \beta') / \partial \beta_j^l \right] / \varepsilon \overset{p}{\to} 0,$$

which implies that $\partial^l \overline{\psi}^* (\beta) / \partial \beta_j^l$ is stochastically equicontinuous in probability. Note that by Lemma 4 and M applied twice

$$\left\| \partial^l \widehat{\psi}^* (\beta) / \partial \beta_j^l - E^* \left[\partial^l \psi_i^* (\beta) / \partial \beta_j^l \right] \right\| \overset{p^*-p}{\to} 0,$$

and thus the conclusion follows. ∎

Lemma 7 *Suppose A1-A4 hold, and that $T^{1/2} \left(\overline{\beta}^* - \beta_0 \right) = O_{p^*-p} (1)$ (or $T^{1/2} \left(\overline{\beta} - \beta_0 \right) = O_p (1)$). Then for $m = o \left(T^{1/2} \right)$*

$$\left\| \widehat{\Omega}^* \left(\widehat{\beta}^* \right)^{-1} - \Omega (\beta_0)^{-1} \right\| \overset{p^*-p}{\to} 0.$$

Proof. By mean value expansion, CS, results of Fitzenberger (1997) and ULLN

$$\left\| \widehat{\Omega}^* \left(\widehat{\beta}^* \right) - \widehat{\Omega}^* (\beta_0) \right\| \leq 2 \left(\sum \sup_{\beta \in \mathcal{N}} \| g_t (\beta) \|^2 / T \right)^{1/2} \left(\sum \sup_{\beta \in \mathcal{N}} \| \partial g_i^* (\beta) / \partial \beta' \|^2 / T \right)^{1/2}$$

$$\left(m / T^{1/2} \right) T^{1/2} \left\| \widehat{\beta}^* - \beta_0 \right\| + O_{p^*} \left(m^2 / T \right) = o_{p^*-p} (1),$$

and the result follow by T and CMT since $\left\| \widehat{\Omega}^* (\beta_0) - \Omega (\beta_0) \right\| = o_{p^*-p} (1)$ (Politis and Romano, 1992). ∎

Proof of the main theorems

Proof of Theorem 1. We first show the consistency of $\widehat{\beta}^*$. This follows by the standard arguments based on the uniqueness of β_0, implied by $E \left[g (\beta)' \Omega (\beta_0)^{-1} g (\beta) \right] > 0$ for all $\beta \neq \beta_0$, and

$$\sup_{\beta \in B} \left| \widehat{\psi}^* (\beta)' \widehat{\Omega}^* \left(\overline{\beta}^* \right)^{-1} \widehat{\psi}^* (\beta) - E^* \left[\psi_i^* (\beta) \right]' \widehat{\Omega}^* \left(\overline{\beta}^* \right)^{-1} E^* \left[\psi_i^* (\beta) \right] \right| = o_{p^*-p} (1)$$

for any $T^{1/2}\left(\overline{\beta}^* - \widehat{\beta}\right) = O_{p^*-p}(1)$ implied by Lemmas 6, 7 and CMT. The asymptotic normality of $T^{1/2}\left(\widehat{\beta}^* - \widehat{\beta}\right)$ follows by mean value expansion about $\widehat{\beta}$ of the EBB FOCs: $0 = \partial\left(\widehat{\psi}^*\left(\widehat{\beta}^*\right)'\widehat{\Omega}^*\left(\overline{\beta}^*\right)^{-1}\widehat{\psi}^*\left(\widehat{\beta}^*\right)\right)/\partial\beta$, noting that by Lemmas 6, 7 and CMT $\left\|\left[\partial\widehat{\psi}^*\left(\widehat{\beta}^*\right)/\partial\beta'\right]'\widehat{\Omega}^*\left(\overline{\beta}^*\right)^{-1} - G\left(\beta_0\right)'\Omega\left(\beta_0\right)^{-1}\right\| = o_{p^*-p}(1)$ and that $T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}\right) \overset{d^*-p}{\to} N\left(0,\Omega\left(\beta_0\right)\right)$. The latter follows by $T^{1/2}\left(\widehat{\psi}^*\left(\beta_0\right) - E^*\left[\psi_i^*\left(\beta_0\right)\right]\right) \overset{d^*-p}{\to} N\left(0,\Omega\left(\beta_0\right)\right)$ (BCLT), Lemma 7 combined with a mean value expansion

$$T^{1/2}\left\|\widehat{\psi}^*\left(\widehat{\beta}\right) - \widehat{\psi}^*\left(\beta_0\right) - \widehat{\psi}\left(\widehat{\beta}\right) + \widehat{\psi}\left(\beta_0\right)\right\| \leq$$
$$\sup_{\beta\in\mathcal{N}}\left\|\partial\widehat{\psi}^*\left(\beta\right)/\partial\beta' - E^*\left[\partial\psi^*\left(\beta\right)/\partial\beta'\right]\right\|T^{1/2}\left\|\widehat{\beta} - \beta_0\right\| = o_{p^*-p}(1),$$

and $T^{1/2}E^*\left[\psi_i^*\left(\beta_0\right)\right] \overset{a}{=} T^{1/2}\left(\widehat{\psi}\left(\beta_0\right) - \widehat{\psi}\left(\widehat{\beta}\right)\right)$ (Lemma 4). Thus by CMT

$$T^{1/2}\left(\widehat{\beta}^* - \widehat{\beta}\right) \overset{d^*-p}{\to} N\left(0,\Sigma\left(\beta_0\right)^{-1}\right),$$

and the first conclusion follows. By mean value expansion about $\widehat{\beta}$ of $T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}^*\right)$, the asymptotic normality of $T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}\right)$ and standard arguments imply that $J^*\left(\widehat{\beta}^*\right) \overset{d^*-p}{\to} \chi^2_{l-k}$ and the second conclusion follows. ∎

Proof of Theorem 2. To prove the first result note that by mean value expansion about $\widehat{\beta}$, the results of Theorem 1 and CMT $T^{1/2}\left[h\left(\widehat{\beta}^*\right) - h\left(\widehat{\beta}\right)\right] \overset{d^*-p}{\to} N\left(0,H\left(\beta_0\right)\Sigma\left(\beta_0\right)^{-1}H\left(\beta_0\right)'\right)$. By Lemma 7 and CMT

$$\left\|H\left(\widehat{\beta}^*\right)\Sigma\left(\widehat{\beta}^*\right)^{-1}H\left(\widehat{\beta}^*\right)' - H\left(\beta_0\right)\Sigma\left(\beta_0\right)^{-1}H\left(\beta_0\right)'\right\| = o_{p^*-p}(1).$$

Thus by standard arguments $W^*\left(\widehat{\beta}^*\right) \overset{d^*-p}{\to} \chi^2_p$. To prove the second result we fist note that the consistency of $\widetilde{\beta}^*$ follows as in the Theorem 1 (using the modified compact parameter space $B\cap h\left(\beta\right) = 0$). Then by a standard Lagrangian argument, a mean value expansion about $\widetilde{\beta}$, Lemmas 6 and 7 and CMT

$$T^{1/2}\left(\widetilde{\beta}^* - \widetilde{\beta}\right) \overset{a^*-p}{=} -\Sigma\left(\beta_0\right)^{-1}\left(I - \Pi\left(\beta_0\right)\Sigma\left(\beta_0\right)^{-1}\right)G\left(\beta_0\right)'\Omega\left(\beta_0\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widetilde{\beta}\right),$$

where $\Pi(\beta_0) = H(\beta_0)'\left[H(\beta_0)\Sigma(\beta_0)^{-1}H(\beta_0)'\right]^{-1}H(\beta_0)$. Thus by a further mean value expansion of the constrained EBB GMM FOCs about $\widetilde{\beta}$ we obtain

$$\widehat{G}^*\left(\widetilde{\beta}^*\right)'\widehat{\Omega}\left(\widetilde{\beta}^*\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widetilde{\beta}^*\right) \stackrel{a^*=^p}{=} \Psi(\beta_0)\Sigma(\beta_0)^{-1}G(\beta_0)'\Omega(\beta_0)^{-1}T^{1/2}\widehat{\psi}^*\left(\widetilde{\beta}\right),$$

so that using similar arguments as those used in the proof of Theorem 1

$$\widehat{G}^*\left(\widetilde{\beta}^*\right)'\widehat{\Omega}\left(\widetilde{\beta}^*\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widetilde{\beta}\right) \stackrel{d^*=^p}{\rightarrow} N\left(0, \Pi(\beta_0)\right),$$

and therefore by standard arguments $LM^*\left(\widehat{\beta}^*\right) \stackrel{d^*=^p}{\rightarrow} \chi_p^2$. Finally to prove the third result note that by a mean value expansion of $\widehat{\psi}^*\left(\widetilde{\beta}\right)$ about $\widehat{\beta}^*$, some algebra, Lemma 6 and CMT

$$D^*\left(\widetilde{\beta}^*, \widehat{\beta}^*\right) \stackrel{a^*=^p}{=} T\left(\widetilde{\beta}^* - \widehat{\beta}^*\right)'G(\beta_0)'\Omega(\beta_0)^{-1}G(\beta_0)\left(\widetilde{\beta}^* - \widehat{\beta}^*\right) +$$
$$2T^{1/2}\left(\widetilde{\beta}^* - \widehat{\beta}^*\right)'\widehat{G}\left(\widehat{\beta}^*\right)'\widehat{\Omega}\left(\widetilde{\beta}^*\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widetilde{\beta}^*\right).$$

By the EBB FOCs $0 = \widehat{G}\left(\widehat{\beta}^*\right)'\widehat{\Omega}\left(\widetilde{\beta}^*\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}^*\right)$ the second term on the right hand side is $o_{p^*-p}(1)$. Some algebra shows that

$$T^{1/2}\left(\widetilde{\beta}^* - \widehat{\beta}^*\right) \stackrel{d^*=^p}{\rightarrow} \Sigma(\beta_0)^{-1}\Pi(\beta_0)\Sigma(\beta_0)^{-1}G(\beta_0)'\Omega(\beta_0)^{-1}T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}^*\right),$$

from which by the same arguments of Theorem 1

$$T^{1/2}\left(\widetilde{\beta}^* - \widehat{\beta}^*\right) \stackrel{d^*=^p}{\rightarrow} N\left(0, \Sigma(\beta_0)^{-1}\Pi(\beta_0)\Sigma(\beta_0)^{-1}\right),$$

and therefore by standard arguments $D^*\left(\widetilde{\beta}^*, \widehat{\beta}^*\right) \stackrel{d^*=^p}{\rightarrow} \chi_p^2$. $\blacksquare$

Proof of Theorem 3. We first show the consistency of the one-step estimator $\widehat{\beta}^{*(1)}$. By the consistency of $\widehat{\beta}^{*(0)} = \widehat{\beta}$, Lemmas 6, 7 and CMT we have that $\left\|\widehat{\Sigma}^*\left(\widehat{\beta}\right)^{-1} - \Sigma(\beta_0)^{-1}\right\| = o_{p^*-p}(1)$ and $\left\|\widehat{G}^*\left(\widehat{\beta}\right)'\widehat{\Omega}^*\left(\widehat{\beta}\right)^{-1} - G(\beta_0)'\Omega(\beta_0)^{-1}\right\| = o_{p^*-p}(1)$. The same arguments of Theorem 1 applied to

$$\widehat{\Sigma}^*\left(\widehat{\beta}\right)^{-1}\widehat{G}^*\left(\widehat{\beta}\right)'\widehat{\Omega}^*\left(\widehat{\beta}\right)^{-1}T^{1/2}\widehat{\psi}^*\left(\widehat{\beta}\right)$$

and the definition of $\widehat{\beta}^{*(1)}$ can be used to show that

$$T^{1/2}\left(\widehat{\beta}^{*(1)} - \widehat{\beta}\right) \stackrel{d^*=^p}{\rightarrow} N\left(0, \Sigma(\beta_0)^{-1}\right),$$

32

hence the conclusion. For any other k-step estimator $\widehat{\beta}^{*(k)}$ $(k \geq 2)$ the result follows by the same arguments applied recursively using the fact that $T^{1/2}\left(\widehat{\beta}^{*j} - \widehat{\beta}^{*(j-1)}\right) = O_{p^* - p}(1)$ $(j = 1, 2, ..., k-1)$. ∎

Data generating process

The method and design of the data generation process is the same as that proposed by Tauchen (1986) and Tauchen and Hussey (1991). The basic idea is to approximate a continuous process through a finite-state Markov chain that mimics closely the underlying process. The distribution of the resulting Markov chain can then be used to approximate the integral operator that arises in a number of stochastic optimisation problems, such as, for example, those arising in dynamic assets pricing. More specifically let $x_{it} = d_{it}/d_{it-1}$, $w_t = c_t/c_{t-1}$ and let $v_{it} = p_{it}/d_{it}$ denote the price dividend ratio for the i-th asset $(i = 1, ..., s)$. Note that

$$r_{it} = (p_{it} + d_{it})/p_{it-1},$$

so that (12) can be written as

$$\beta_1 E\left[w_t^{-\beta_2}(1 + v_{it}) x_{it} | \mathcal{I}_{t-1}\right] = v_{it-1} \quad (i = 1, ..., s). \tag{15}$$

Under the assumption that x_t and w_t are a (jointly stationary) first order Markov process with conditional cumulative probability distribution

$$F\left(x^1, w^1 | x, w\right) = \Pr\left(x_t \leq x^1, w_t \leq w^1 | x_{t-1} = x, w_{t-1} = w\right)$$

(with "1" denoting one period ahead), the values x, w when the event $\{w_{t-1} = w, x_{t-1} = x\}$ occurs characterise completely the state of the system (15) so the equilibrium v_{it} will be a function $v_i(x, w)$ of x and w for $i = 1, ..., s$. These s functions are the solutions to the following set of asset pricing equations (integral equations)

$$\beta_1 \int \left(w^1\right)^{-\beta_2}\left(1 + v_i\left(x^1, w^1\right) x_i^1 dF\left(x^1, w^1 | x, w\right)\right) = v_i(x, w), \tag{16}$$

33

which under certain regularity conditions (see, for example, Lucas (1978)) admit a unique positive solution for $v_i(x, w)$. Let $n = 1, 2, .., N$ denote the states of nature, $x(n)$ and $w(n)$ denote the values of x and w in the state n, and let

$$\pi\left(n, n^1\right) = \Pr\left(x_t = x\left(n^1\right), w_t = w\left(n^1\right) | x_{t-1} = x(n), w_{t-1} = w(n)\right),$$

denote the transition probabilities for $[x_t', w_t]'$. Then (16) can be written as

$$\beta_1 \sum_{n^1=1}^{N} \pi\left(n, n^1\right)\left(w\left(n^1\right)\right)^{-\beta_2}\left(1 + v_i\left(n^1\right)\right) x_i\left(n^1\right) = v_i(n). \qquad (17)$$

Tauchen (1986) and Tauchen and Hussey (1991) propose to use numerical methods to compute $\pi(n, n')$, from which the equilibrium price dividend ratio $v_i =: v_i(n)$ $(n = 1, ..., N)$ (solution of (17)) is simply

$$v_i = (I_N - P)^{-1} P 1_N$$

where $P =: P_{n,n^1} = \beta_1 \pi\left(n, n^1\right)\left(w\left(n^1\right)\right)^{-\beta_2} x_i\left(n^1\right)$ $(n, n^1 = 1, ..., N)$. Then the return for the ith asset r_i^s can be computed simply as

$$r_i^s\left(n, n^1\right) = x_i\left(n^1\right)\left(1 + v_i\left(n^1\right)\right) / v_i(n),$$

whereas the return for the risk free asset r^f is

$$r^f\left(n, n^1\right) = \left(\beta_1 \sum_{n^1=1}^{N} \pi\left(n, n^1\right)\left(w\left(n^1\right)\right)^{-\beta_2}\right)^{-1}$$

- see Kocherlakota (1990) for further details.

References

Allen, J., Gregory, A. and Shimotsu, K. (2005), 'Empirical likelihood block bootstrap'. Mimeo.

Andrews, D. W. K. (2002), 'Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators', *Econometrica* **70**, 119–162.

Bravo, F. (2009), 'Blockwise generalised empirical likelihood inference for nonlinear dynamic moment conditions models', *Econometrics Journal* **12**, 208–231.

Brown, B. W. and Newey, W. K. (2002), 'Generalized method of moments, efficient bootstrapping, and improved inference', *Journal of Business and Economic Statistics* **20**, 507–517.

Davidson, R. and MacKinnon, D. (1999), 'Bootstrap testing in nonlinear models', *International Economic Review* **40**, 487–508.

Davidson, R. and MacKinnon, J. (1998), 'Graphical methods for investigating the size and power of hypothesis tests', *The Manchester School* **66**, 1–26.

Efron, B. (1981), 'Nonparametric standard errors and confidence intervals (with discussion)', *Canadian Journal of Statistics,* **9**, 139–172.

Fitzenberger, B. (1997), 'The moving blocks bootstrap and robust inference for linear least squares and quantile regressions', *Journal of Econometrics* **82**, 235–287.

Flachaire, E. (2005), 'Bootstrapping heteroskedastic regression models: Wild bootstrap vs. pairs bootstrap', *Computational Statistics and Data Analysis* **49**, 361–376.

Godfrey, L. and Tremayne, A. (2005), 'The wild bootstrap and heteroskedasticity-robust tests for serial correlation in dynamic models', *Computational Statistics and Data Analysis* **49**, 377–395.

Goncalves, S. and White, H. (2004), 'Maximum likelihood and the bootstrap for nonlinear dynamic models', *Journal of Econometrics* **119**, 199–219.

Hahn, J. (1996), 'A note on bootstrapping generalized method of moments estimators', *Econometric Theory* **12**, 187–197.

Hall, A. R. (2005), *Generalized Method of Moments*, Oxford University Press.

Hall, P. and Horowitz, J. L. (1996), 'Bootstrap critical values for tests based on generalized-method of moment estimators', *Econometrica* **64**, 891–916.

Hall, P. and Presnell, B. (1999), 'Intentionally biased bootstrap methods', *Journal of the Royal Statistical Society B* **61**, 143–158.

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

Inoue, A. and Shintani, M. (2006), 'Bootstrapping GMM estimators for time series', *Journal of Econometrics* **133**, 531–555.

Kocherlakota, N. (1990), 'On tests of raprasentative consumer asset pricing models', *Journal of Monetary Economics* **26**, 285–304.

Lucas, R. (1978), 'Asset prices in an exchange economy', *Econometrica* **46**, 1429–1445.

Newey, W. K. and Smith, R. J. (2004), 'Higher order properties of GMM and generalized empirical likelihood estimators', *Econometrica* **72**, 219–256.

Newey, W. K. and West, K. D. (1987*a*), 'Hypothesis testing with efficient method of moments testing', *International Economic Review* **28**, 777–787.

Newey, W. K. and West, K. D. (1987*b*), 'A simple positive semi-definite hateroskedasticity and autocorrelation consistent covariance matrix', *Econometrica* **55**, 703–708.

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

Owen, A. (1988), 'Empirical likelihood ratio confidence intervals for a single functional', *Biometrika* **36**, 237–249.

Owen, A. (1991), 'Empirical likelihood for linear models', *Annals of Statistics* **19**, 1725–1747.

Politis, D. N. and Romano, J. P. (1992), 'A general resampling scheme for triangular arrays of α - mixing random variables with application to the problem of spectral density estimation', *Annals of Statistics* **20**, 1985–2007.

Tauchen, G. (1986), 'Statistical properties of generalized method-of-moments estimators of structural parameters obtained from financial market data', *Journal of Business and Economic Statistics* **4**, 387–416.

Tauchen, G. and Hussey, R. (1991), 'Quadrature based methods for otaining approximate solutions to nonlinear asset pricing models', *Econometrica* **59**, 371–396.

Wooldridge, J. (1994), Estimation and inference for dependent processes, *in* R. Engle and D. McFadden, eds, 'Handbook of Econometrics IV', Amsterdam: North Holland, pp. 2639–2738.

Wright, J. (2003), 'Detecting lack of identification in GMM', *Econometric Theory* **19**, 322–330.