Finite Sample Properties of Bootstrap GMM *for Nonlinear Conditional Moment Models*

Rachida Ouysse[a]

[a] *School of Economics, The University of New South Wales, Sydney, NSW 2052, Australia*

We investigate the finite sample performance of block bootstrap tests for over-identification (*J test*) for nonlinear conditional models estimated using Generalized Methods of Moment (GMM). Overall, block bootstrap methods with fixed length blocks outperform the stationary bootstrap which uses a random block length. Randomizing the block length decreases the sensitivity of the distribution of the bootstrap *J test* and GMM estimators to the choice of the block size. This sensitivity diminishes with the degree of nonlinearity of the conditional moments. In addition, the accuracy of the block bootstrap approximation diminishes as the dimensionality of the joint test increases, especially in the tails of the distribution of the *J test*.

Keywords: Continuously-updated GMM, iterated GMM, non-overlapping and moving blocks bootstrap, stationary bootstrap, rejection probability, Q-Q-plot, Monte Carlo test.

JEL Subject Classification: 62F40; 62E17; 62G09; 91B84; 91B28; 91B25

1. Introduction

Perhaps the most popular technique for estimating conditional moment models is the generalized method-of-moments (GMM). First introduced by Hansen [1] as an alternative to maximum likelihood for the estimation of models described by conditional moments, GMM estimators have desirable asymptotic properties in several contexts. A leading example occurs in the consumption-based intertemporal asset-pricing model (C-CAPM) where the conditional moments are nonlinear functions of the model parameters; Hansen and Singleton [2].

However, GMM estimators and the associated tests may have poor finite sample properties. The poor performance of the asymptotic approximation may be the result of many reasons, two of which are the focus of this paper.

Firstly, the way the moment conditions are weighted defines two alternative efficient estimators; the iterated GMM (IT-GMM) estimator of Hansen [1] in which the weighting matrix is iterated to convergence, and the continuously updated GMM (CU-GMM) estimator of Hansen et al. [3] in which the weighting matrix is a function of the parameters of the model. Secondly, the number of conditioning variables (*instruments*), used to construct the unconditional moments, determines the degree of over-identification of the model and the distribution of the asymptotic test. This paper pays particular attention to the finite sample bias and variance of CU-GMM and IT-GMM and to the performance of the associated test of over-identifying restrictions (also known as the *J test*) defined by the GMM criterion

Tel.:+61 2 9385 3321; fax:+61 2 9313 6337.
E-mail address: rouysse@unsw.edu.au

function.

Tauchen [4] finds that while increasing the number of instruments reduces the bias, IT-GMM estimators with the smallest number of instruments are the most reliable. In addition to this bias/variance trade-off, the asymptotic chi-square test for over-identifying restrictions tends to reject the model more often than the nominal level. Hansen et al. [3] find that while *J test* is more conservative when the weighting matrix is continuously updated, the CU-GMM fattens the distribution of the GMM estimator.

Alternative hypothesis tests and bias corrected estimators with better finite sample properties are therefore of considerable interest in this context. One approach is to use the bootstrap as a simulation-based alternative to the asymptotic approximation. Studies of the finite sample properties of bootstrap GMM tests are scarce and limited to the case of linear models. See, for example, Hansen [5], Bond and Windmeijer [6], Lagunes [7] and Lahiri [8].

We study the finite sample properties of the bootstrap GMM inference for the class of nonlinear conditional moment models. For dependent data, the distributions of estimators and test statistics generally depend on the joint distribution of the observations; Leger et al. [9] and Lepage and Billard [10]. We consider nonparametric dependent bootstrap methods which use blocking rules to account for the time series dependence in the sample observations. Using the C-CAPM as a laboratory for the Monte Carlo experiments, we compare the finite sample bias of the GMM estimators and the *size distortion* (*i.e.,* the discrepancy between the nominal and the actual probability that a test rejects a true null hypothesis) of the *J test* for three block bootstrap methods. The first is the *non-overlapping blocks* bootstrap (NOB) of Carlstein [11] which uses non-overlapping segments of the data to define the blocks. The second is the *moving blocks* bootstrap (MBB) due to Künsch [12] and Liu and Singh [13], where the blocks are defined using overlapping segments of the data. The bootstrap data are obtained by independent sampling with replacement from among the blocks. The block length in both NOB and MBB is nonrandom and must grow with the sample size in order to achieve consistency. The third is the *stationary bootstrap* (SB) of Politis and Romano [14]. In the SB, the lengths of the resampled blocks are random and have a geometric probability distribution. Politis and Romano [14] show that the bootstrap sample generated by randomizing the block length is stationary. In the context of estimating the bias and the variance of smooth functions of the population mean, Lahiri [8, 15] find that the MBB outperforms the NOB which in turns outperforms the SB.

The paper also explores the usefulness of the block bootstrap as a finite-sample size-reduction for the *J test* when the model parameters are nearly under-identified and the *instruments* are weak.

The findings of the simulation study can be summarized in the following. The bootstrap test based on any of the blocking methods outperforms the asymptotic χ^2-approximation. In terms of approximating the size of the *J test*, the block bootstrap with nonrandom block lengths is more accurate than the stationary bootstrap. The latter is the least sensitive to the choice of the block length. Further, the simulation evidence suggests that the nonlinearity in the moment functions affects the sensitivity of the bootstrap approximation to the blocking rule. Finally, the accuracy of the block bootstrap in approximating the tails of the distribution of *J test* decreases with the dimensionality of the joint tests.

The paper is organized as follows. Section two reviews the GMM estimation and its asymptotic properties. Section three defines the bootstrap methods. Section four describes the calibrated model and the Monte Carlo environment. Section five presents the finite sample properties of the GMM estimators and the asymptotic test for over-identification. Section 6 discusses the statistical properties of the bootstrap tests.

We use the following notation throughout the paper: $\mathbb{E}\left(.|\Im_t\right)$ is the expectation of the argument conditional on some suitable information set $\Im_t$ available at time t, $\mathbb{E}\left(Z_t\right)$ is the unconditional expectation of Z_t, $\mathbb{E}_F(.)$ is the (unconditional) expectation under the probability distribution F, the indicator function $Ind(\mathcal{A})$ takes the value 1 if the statement "$\mathcal{A}$" is correct and 0 otherwise, A^{-1} is the inverse of A, ι_m is a m-vector of ones, I_m is an $m \times m$ identity matrix, $diag(A)$ is the vector consisting of the diagonal elements of A, $tr(A)$ is the sum of the diagonal elements of A, the norm of A is $\|A\| = (tr(A'A))^{1/2}$, $A \otimes B$ is the Kronecker product of A and B, *i.e.,* for $A = [a_{ij}]$, $A \otimes B = [a_{ij}B]$, iid stands for 'independent and identically distributed' and, 'vector' means a column vector.

2. GMM for the class of conditional moment models

Models that are defined in terms of conditional moment restrictions establish that certain parametric functions have zero conditional mean when evaluated at the true parameter values. Let $\{Y_t\}_{t=1}^n$ be an ergodic and stationary time series vector of endogenous and exogenous random variables. The coordinates of Y_t are related by an econometric model that establishes that the true distribution of the data satisfies the *conditional* moment restrictions

$$\mathbb{E}\left[h(Y_{t+1}, \theta_0)|\Im_t\right] = 0, \ t = 1, ..., n-1, \tag{2.1}$$

for a unique value of the k-vector $\theta_0 \in \Theta$, where $\Theta \subset \mathbb{R}^k$ and $h(Y_{t+1}, \theta_0)$ is an $m-$dimensional parametric function. The function h can be understood as the errors measuring the deviation from an equilibrium condition.

Suppose that we can form an $n \times q$ matrix Z with typical row z_t such that all its elements belong to $\Im_t$. The q variables given by the columns of Z are called *instrumental variables*, or simply *instruments*. These instruments are required to be 'predetermined' and not necessarily 'econometrically' exogenous. That is, current and lagged values of Y are valid instruments.

Given the conditional moment restriction (2.1) and the additional assumption that the constituents of $h(Y_{t+1}, \theta_0)$ and the variables in z_t have finite second moments (Hansen and Singleton [2]), a family of (unconditional) population orthogonality conditions

$$\mathbb{E}\left[g(X_t, \theta_0)\right] = 0, \tag{2.2}$$

can be constructed where $X_t \equiv (Y_{t+1}, z_t)$, $g(X_t, \theta_0) = h(Y_{t+1}, \theta_0) \otimes z_t$ and provided that $q \times m$ (henceforth mq) is at least as large as k. The moment restrictions in (2.2) are also known as the *estimating equations*.

The generalized methods-of-moments estimation uses the sample versions of the population orthogonality conditions (2.2) to construct an estimator for θ_0. The GMM estimator $\widehat{\theta}$ is

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} g_n(\mathcal{X}, \theta)' W_n g_n(\mathcal{X}, \theta), \tag{2.3}$$

$$g_n(\mathcal{X}, \theta) = \frac{1}{n} \sum_{t=1}^{n} g(X_t, \theta), \tag{2.4}$$

where W_n is a sequence of symmetric positive-definite weighting matrices which converge to a positive definite matrix W when n goes to infinity, Θ is a compact parameter space, $\Theta \subset \mathbb{R}^k$, and $\mathcal{X} = \{X_1, ..., X_n\}$. Regularity conditions for the consistency of the GMM estimator $\widehat{\theta}$ in (2.3) include: (a) $g_n(\mathcal{X}, \theta)$ converges to $\mathbb{E}\left(g(X_t, \theta)\right)$ uniformly in $\theta \in \Theta$, (b) $\mathbb{E}\left(g(X_t, \theta)\right) \neq 0$ for all $\theta \neq \theta_0$, (c) $\mathbb{E}\left(g(X_t, \theta)\right)$ and $g_n(\mathcal{X}, \theta)$ are continuously differentiable and, $\frac{\partial g_n(\mathcal{X}, \theta)}{\partial \theta}$ converges to $\frac{\partial \mathbb{E}\left(g(X_t, \theta)\right)}{\partial \theta}$. If in addition, (d) $\sqrt{n} g_n(Y, \theta_0)$ converges in distribution to a normal distribution with mean zero and variance $V_g > 0$ and, (e) the $mq \times k$ matrix $G_0 = \frac{\partial \mathbb{E}\left(g(X_t, \theta)\right)}{\partial \theta'}|_{\theta=\theta_0}$ has full rank k, then $\sqrt{n}(\widehat{\theta} - \theta_0)$ converges in distribution to a normal distribution with mean zero and asymptotic variance

$$AsyV(\widehat{\theta}) = \left(G_0' W G_0\right)^{-1} G_0' W \Omega_0 W G_0 \left(G_0' W G_0\right)^{-1}, \tag{2.5}$$

$$\Omega_0 = E\left[g(X_t, \theta_0) g(X_t, \theta_0)'\right]. \tag{2.6}$$

See Davidson and MacKinnon [16] and Greene [17] for a discussion of the asymptotic properties of minimum distance estimators.

Assumption (b) is the global identification which requires that the population moment condition only holds at one parameter value in the entire parameter space Θ. In nonlinear models, it is rarely possible to derive testable conditions for global identification. This assumption is replaced by assumption (e), which is a local identification condition defined in a neighborhood of θ_0. It is also known as the *first order identification* condition, see Dovonon and Renault [18].

The covariance matrix of $g(X_t, \theta_0)$ has the form in (2.6) because the moment functions $\{h(Y_{t+1}, \theta) \otimes z_t\}_{t=1}^{\infty}$ are martingale first differences. This is a direct implication of the conditional moment restriction (2.1). With the asset pricing application in mind, this moment condition is consistent with an economy where investors hold assets for one period. The moment condition can be made more general by considering an economy where assets are held to maturity $s > 1$

$$\mathbb{E}\left[h(Y_{t+s}, \theta_0) | \Im_t\right] = 0, \ t = 1, ..., n. \tag{2.7}$$

The assumption of $s > 1$ does not affect the asymptotic properties of the GMM estimators. However in finite samples, the difficulty in accurately estimating the spectral density matrix (long run variance) of the moment functions is an additional source of poor finite sample performance of the asymptotic approximation. See, for example, Burnside and Eichenbaum [19].

The optimal weight matrix W_0 which minimizes (2.5) is $W_0 = \Omega_0^{-1}$. The covariance matrix of the *efficient* GMM estimator is $AsyV(\widehat{\theta}) = \left(G_0' \Omega_0^{-1} G_0\right)^{-1}$ and is optimal in the class of GMM estimators with this set of moment conditions.

2.1. *The iterative* GMM *estimator*

In practice, the efficient GMM estimator is unfeasible since Ω_0^{-1} is not known. Hansen [1] shows that a consistent estimator of Ω_0 is sufficient for asymptotic efficiency. If $\widetilde{\theta}$ is a consistent estimator for θ_0, then

$$\Omega_n(\widetilde{\theta}) = \frac{1}{n} \sum_{t=1}^{n} g(X_t, \widetilde{\theta}) g(X_t, \widetilde{\theta})', \tag{2.8}$$

is a consistent estimator for Ω_0.

An efficient two-step GMM estimator, denoted $\widehat{\theta}^{(2)}$, is based on a weight matrix

$$W_n(\widehat{\theta}^{(1)}) = \Omega_n\left(\widehat{\theta}^{(1)}\right)^{-1},$$

where $\widehat{\theta}^{(1)}$ is a consistent one-step estimator for θ_0 based on a weighting matrix equal to the identity matrix.

The iterative GMM estimator IT-GMM, denoted $\widehat{\theta}_{it}$, continues from the two-step estimator by reestimating the weighting matrix. For each subsequent step $l = 3, .., L$, the weighting matrix is updated using $W_n(\widehat{\theta}^{(l-1)}) = \Omega_n(\widehat{\theta}^{(l-1)})^{-1}$, where $\widehat{\theta}^{(l-1)}$ is the consistent estimator when $W_n = W_n(\widehat{\theta}^{(l-2)})$. This is repeated until l attains some large value L, we choose $L = 15$, or until convergence defined as $\|W_n(\widehat{\theta}^{(l+1)}) - W_n(\widehat{\theta}^{(l)})\| < 1E-4$.

2.2. *The continuous updating* GMM *estimator*

Instead of taking the weighting matrix as given in each iteration, Hansen et al. [3] propose an estimator in which the weighting matrix is continuously updated. Formally, the CU-GMM estimator, denoted $\widehat{\theta}_{cu}$, is

$$\widehat{\theta}_{cu} = \arg\min_{\theta} g_n(\mathcal{X}, \theta)'\Omega_n(\theta)^{-1} g_n(\mathcal{X}, \theta), \tag{2.9}$$

where $\Omega_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} g(X_t, \theta)g(X_t, \theta)'$.

2.3. *Wald test for over-identifying restrictions*

The first order conditions for minimizing (2.3) is a system of k equations with k unknowns

$$\frac{\partial g_n(\mathcal{X}, \theta)'}{\partial \theta} W_n g_n(\mathcal{X}, \theta) = 0.$$

There are $mq - k$ remaining linearly independent moment conditions that are not set to zero in the estimation. These must be close to zero if the model is correct. In an over-identified model, $mq > k$ and there may not be a parameter value θ that satisfies (2.2). The model and the moment restrictions are therefore testable. The standard test statistic for over-identifying restrictions (also called *J test*) is based on the minimized GMM criterion function,

$$J_n(\widehat{\theta}) = n g_n(\mathcal{X}, \widehat{\theta})'\Omega_n(\widehat{\theta})^{-1} g_n(\mathcal{X}, \widehat{\theta}). \tag{2.10}$$

When the moment conditions are valid, the *J test* has an asymptotic chi-squared distribution with $mq - k$ degrees of freedom. See Hansen [1] for an exposition of the general theory of GMM estimation and testing. It is worth noting that the *J test* statistic is a Wald test for the hypothesis, $\mathbb{E}\left(g(X_t, \theta_0)\right) = 0$. The latter is a joint hypothesis with mq individual moment restrictions. In finite samples, the number of unconditional moments mq affects the accuracy of the chi-squared approximation. The finite sample size (*i.e.,* rejection of a true null) of the *J test* increases uniformly as the dimension of joint tests increases and is drastically larger than the asymptotic size of the test. See Burnside and Eichenbaum [19].

Economic theory often provides information about m, the number of conditional moment restrictions. For example, the Euler equation of a decision theoretic problem. However, the number of instruments q is often arbitrarily determined by the practitioner. This paper investigates whether the bootstrap can provide improved finite sample inference in over-identified models.

3. Bootstrap GMM

3.1. *Preliminary*

Let $X_1, X_2, \cdots$ be a sequence of stationary random variables with unknown joint probability distribution F_0 indexed by an unknown real-valued parameter θ_0. Consider a statistic of interest, $T_n(\theta_0)$, which is possibly a function of F_0 through θ_0,

$$T_n(\theta_0) = T_n(\mathcal{X}, F_0), \tag{3.1}$$

where $\mathcal{X} = \{X_t, t = 1, \cdots, n\}$ is the original sample. The bootstrap uses a nonparametric estimate F_n of F_0 to approximate the distribution of T_n using $T_n^* = T_n(\mathcal{X}, F_n)$. To obtain asymptotic refinements, bootstrap sampling must take into account the dependence of $\{X_t\}$ in the population data generating process.

In an attempt to reproduce different aspects of the dependence structure of the observed data in the bootstrap sample, different block bootstrap methods have been proposed in the literature. We briefly describe three of these methods.

Non-overlapping blocks bootstrap, NOB

The base sample, $\mathcal{X}$, is divided into ϕ blocks of pre-specified length ω such that $\phi\omega = n$. Denote these blocks by $\{B_i, i = 1, 2, \cdots, \phi\}$ where $B_1 = \{X_t, t = 1, \cdots, \omega\}$, $B_2 = \{X_t, t = \omega + 1, \cdots, 2\omega\}$, $B_s = \{X_t, t = \omega(s-1) + 1, \cdots, \omega s\}$, and so forth. The NOB is implemented by randomly sampling with replacement ϕ blocks from $\{B_i, i = 1, 2, \cdots, \phi\}$. The selected blocks are laid end to end to form the bootstrap sample $\mathcal{X}^*$.

Moving blocks bootstrap, MBB

We use the following overlapping block bootstrap procedure, which is originally attributed to Kunsch [12]. Let $I = \{1, 2, \cdots, n-\omega+1\}$ denote the set of observations that can begin a block of ω observations. The construction of the bootstrap sample, $\{B_i, i = 1, \cdots, \phi\}$, begins by random sampling from I with replacement ϕ times. Let $\{Q_i, i = 1, ..., \phi\}$ be the random sample from I. The MBB sample consists of the ϕ blocks which begin with the observations $\{Q_i, i = 1, \cdots, \phi\}$. The first block in the bootstrap sample is $B_1 = \{X_{Q_1+i}, i = 1, \cdots, \omega\}$, the j^{th} block is $B_j = \{X_{Q_j+i}, i = 1, \cdots, \omega\}$ for $j = 1, \cdots, \phi$. These blocks put together form the bootstrap sample, $\mathcal{X}^* = \{B_j, j = 1, \cdots, \phi\}$.

Stationary bootstrap, SB

Unlike the NOB and MBB, the stationary bootstrap of Politis and Romano [14] uses a *random* block length to generate the bootstrap sample. Let $L_1, L_2, \cdots$ be a sequence of iid random variables having a geometric distribution with parameter $p = \omega^{-1} \in (0, 1)$. That is, for $m = 1, 2, \cdots$, the probability of an event, $\{L_i = m\}$, is $(1-p)^{m-1}p$.

Independent of X_i and L_i, let $I_1, I_2, \cdots$ be a sequence of iid variables that have a discrete uniform distribution on $\{1, \cdots, n\}$. The first block B_1 consists of L_1 observations, $B_1 = \{X_{I_1}, \cdots, X_{I_1+L_1-1}\}$. The next sampled block B_2 consists of L_2 observations, $B_2 = \{X_{I_2}, \cdots, X_{I_2+L_2-1}\}$, and so forth. In other words, the first observation X_1^* is sampled randomly from $\mathcal{X}$; $X_1^* = X_{I_1}$. For $j = 2, \cdots, n$, if the

$(j-1)^{th}$ observation in the bootstrap sample is $X^*_{j-1} = X_k$, then

$$X^*_j = \begin{cases} X_{k+1} & \text{with probability } p \\ X_{I_j} & \text{with probability } 1-p \end{cases}$$

Blocking rule

Hall et al. [20] show that the asymptotically optimal blocking rule is $\omega \sim n^\kappa$ where κ minimizes the mean-square error (MSE) of the block bootstrap estimator. They find that $\kappa = \frac{1}{5}$ is optimal for estimating a double sided distribution, such as the $t-$ *statistic*, and $\kappa = \frac{1}{4}$ is optimal for estimating a one-sided distribution, such as the *J test* for over-identification.

The three bootstrap methods have different asymptotic properties in terms of efficiency and MSE. Politis and White [21] show that the asymptotic relative efficiency of MBB relative to SB is always bounded away from zero. They argue that although the MBB is asymptotically more efficient, it is more sensitive to the choice of block size. Using expansions of the bias, the variance and the MSE, Lahiri [15] finds that MBB is to be preferred to NOB and that the random block length leads to MSE larger than those for nonrandom block length.

In this paper, a grid of values of p and ω are used to compare the three bootstrap methods in terms of the mean and median bias, and the size of the *J test.* To link the fixed block length ω of the NOB and the MBB to the random block length of the stationary bootstrap, we choose ω equal to the expected block size of the SB; $\omega = E(L_k)$ and $p = 1/\omega = \phi/n$.

3.2. *Bootstrap test for over-identification*

Bootstrapping GMM is not standard; in over-identified models the population moment conditions (2.1) do not hold exactly in the bootstrap sample. The bootstrap sample $\mathcal{X}^*$ does not satisfy the same moment conditions as the population distribution. The recentering proposed by Hall and Horowitz [22] replaces the bootstrap moment functions, $g(X^*_t, \theta)$, with

$$g^*(X^*_t, \theta) = g(X^*_t, \theta) - g_n(\mathcal{X}, \widehat{\theta}), \ t = 1, \cdots, n. \tag{3.2}$$

This recentering makes the *J test* asymptotically pivotal and preserves the higher order refinements of the bootstrap. In this paper, we adopt this recentering of the moment functions to compute the bootstrap IT-GMM and CU-GMM estimators and the three bootstrap tests for over-identification.

Let $g^*_n(\mathcal{X}^*, \theta) = \frac{1}{n}\sum_{t=1}^n g^*(X^*_t, \theta)$, the bootstrap GMM estimator, $\widehat{\theta}^*$, solves

$$\widehat{\theta}^* = \arg\min_{\theta\in\Theta} g^*_n(\mathcal{X}^*, \theta)' W^*_n g^*_n(\mathcal{X}^*, \theta). \tag{3.3}$$

The bootstrap CU-GMM estimator, $\widehat{\theta}^*_{cu}$, solves (3.3) for $W^*_n = \Omega^*_n(\theta)^{-1}$, where

$$\Omega^*_n(\theta) = \frac{1}{n}\sum_{t=1}^n g^*(X^*_t, \theta) g^*(X^*_t, \theta)'. \tag{3.4}$$

The bootstrap IT-GMM estimator, $\widehat{\theta}^*_{it}$, solves (3.3) for $W^*_n = \Omega^*_n(\widetilde{\theta}^{l-1})^{-1}$, where $\widetilde{\theta}^{(l-1)}$ is a consistent estimator that solves (3.3) for $W_n = \Omega^*_n(\widetilde{\theta}^{(l-2)})^{-1}$. This is iterated until convergence is achieved or a maximum number of iterations, $L = 15$, is reached.

Hall and Horowitz [22] propose new formulae to correct the *J test* for the differences between the asymptotic covariances of the original sample and the bootstrap variances. These corrections preserve the higher order expansions that produce the higher order refinements. Andrews [23] generalizes the correction factors to the case of MBB bootstrap. The distribution of the *J test* statistic is approximated using the following statistic,

$$J_n^*(\theta) = n g_n^*(\mathcal{X}^*, \theta)' W_n^{*\frac{1}{2}} V_n^+ W_n^{*\frac{1}{2}} g_n^*(\mathcal{X}^*, \theta),$$

where V_n^+ denotes the Moore-Penrose generalized inverse of the correction matrix V_n,

$$V_n = \widehat{M}_n W_n^{\frac{1}{2}} \widetilde{W}_n W_n^{\frac{1}{2}} \widehat{M}_n, \tag{3.5}$$

$$\widehat{M}_n = I - \widehat{\Omega}_n^{-\frac{1}{2}} \widehat{G}_n [\widehat{G}_n' \widehat{\Omega}_n^{-1} \widehat{G}_n]^{-1} \widehat{G}_n' \widehat{\Omega}_n^{-\frac{1}{2}}, \tag{3.6}$$

where $\widehat{\Omega}_n = \Omega_n(\widehat{\theta})$, $\widehat{G}_n = \frac{\partial g_n(\mathcal{X},\theta)}{\partial \theta}|_{\theta=\widehat{\theta}}$, $\widehat{\theta}$ is the GMM estimator calculated using the original sample $\mathcal{X}$, and

$$\widetilde{W}_n = \frac{1}{n} \sum_{i=0}^{\phi-1} \sum_{j=1}^{\omega} \sum_{s=1}^{\omega} g^*(X_{i\omega+j}, \widehat{\theta}) g^*(X_{i\omega+s}, \widehat{\theta})' \tag{3.7}$$

for the non-overlapping blocks, and

$$\widetilde{W}_n = \phi n^{-1} (n - \omega + 1)^{-1} \sum_{i=0}^{n-\omega} \sum_{j=1}^{\omega} \sum_{s=1}^{\omega} g^*(X_{i+j}, \widehat{\theta}) g^*(X_{i+s}, \widehat{\theta})' \tag{3.8}$$

for the overlapping blocks.

For the stationary bootstrap, $V_n = I_{mq}$. In the case of martingale first difference moment functions, $\widetilde{W}_n = \Omega_n^*(\widehat{\theta})$ in (3.7) and the correction factor is equal to the identity matrix, $V_n = I_{mq}$. See the proof in Appendix 1.

4. The Monte Carlo environment

4.1. *The calibrated model*

Consider a representative consumer with intertemporally separable constant risk aversion preferences, and an economy with m assets traded in complete and frictionless markets. The Euler equation in the consumption capital asset pricing model is as given in equation (2.1) with

$$h(Y_{t+1}, \theta) = \beta (c_{t+1})^{-\gamma} \otimes R_{t+1} - \iota_m, \tag{4.1}$$

where $Y_t = (c_t, R_t')$, $\theta = (\gamma, \beta)'$, c_t is the growth rate of consumption, R_t is the m dimensional vector of gross stock returns, γ is the risk aversion parameter (> 0), β is the impatience parameter, and ι_m is an $m \times 1$ vector of ones. The one period gross return from holding one unit of stock j is defined as:

$$R_{j,t+1} = \frac{P_{j,t+1} + D_{j,t+1}}{P_{j,t}}, j = 1, \cdots, m,$$

where $D_{j,t+1}$ is the dividend yield on stock j from period t to $t+1$.

We follow Kocherlakota [24] and consider three assets; $m = 3$. The Risk free R^f which pays one unit of consumption, the market portfolio MP which pays C_t units of consumption in period t, and the stock market SM with dividend payoffs D_t in period t.

To simulate series of consumption growth and stock returns that satisfy the moment conditions (2.1) and (4.1), we use the quadrature approximation originally introduced by Tauchen and Hussey [25]. See also Tauchen [4, 26], Kocherlakota [24] and Wright [27].

Let $\mathbf{y}_t = (log(c_t), log(d_t))'$, where $d_t = D_t/D_{t-1}$ is the dividend growth, be a vector of jointly stationary first order Markov processes. The quadrature approximation fits an N state Markov chain to $\mathbf{y_t}$ so as to approximate the first order vector autoregression (VAR)

$$\mathbf{y}_t = \mu + \mathbf{\Phi}\mathbf{y}_{t-1} + \epsilon_t, \tag{4.2}$$

where $\mathbf{\epsilon_t}$ are iid with $\mathbb{E}(\epsilon_\mathbf{t}) = 0$ and $V(\epsilon_t) = V_\epsilon$, $\mu = \mathbb{E}(\mathbf{y}_t)$ and $\mathbf{\Phi} = \{\Phi_{ij}\}, i, j = 1, 2$, is the VAR matrix of coefficients. For computational considerations, we choose $N = 8$. Tauchen and Hussey [25] find that an $8-$point rule generally yields accuracy close to four digits in terms of mean-squared error. Appendix 2 provides a detailed description of the quadrature procedure, and a description of how the series of assets returns R^f, MP and SM are generated using the conditional moments (4.1), and the series c_t and d_t.

[Table 1 about here.]

4.2. *Parameter and preference setting*

Table 1 lists the combinations of the parameters γ, β, $\mathbf{\Phi}$, V_ϵ, and μ for the models we consider in the Monte Carlo simulations. We are particularly interested in three characteristics of the data generating process of $\mathbf{y}_t$. First, the serial correlation in the consumption series, Φ_{11}, and in the dividends growth series, Φ_{22}. Second, the interaction between $log(c_t)$ and $log(d_t)$ in the VAR determined by the parameters Φ_{12} and Φ_{21}. This is also known as the *Granger-Sims* effect allowing for feedback from $log(d_{t-1})$ to $log(c_t)$ through Φ_{12}, and from $log(c_{t-1})$ to $log(d_t)$ through Φ_{21}. Third, the degree of nonlinearity in the moment function measured by the magnitude of the coefficient of relative risk aversion γ.

We consider two sets of values for θ. The first is $\theta_1 = (\,1.30\; 0.97\,)'$, used in Tauchen [4], and the second is $\theta_2 = (\,13.7\; 1.139\,)'$ from Kocherlakota [24].

The first two models, $M1$ and $M2$, are used as benchmarks where the conventional asymptotic theory is expected to work well. $M2$ is the full rank model described in Wright [27]. It satisfies the first order identification assumption of full rank matrix G_0. Models $M3$ and $M5$, represent VAR specification as estimated by Kocherlakota [24] when the VAR is fitted to U.S. annual data. These are weakly identified models where the matrix G_0 has rank 2 but is close to being rank deficient, see Wright [27]. Models $M4$, $M6$ and $M7$ are introduced for comparative purposes as cases of unrealistically strong serial correlation while maintaining the *Granger-Sims* feedback between the VAR series.

To study the effect of the degree of over-identification $mq - k$, we consider instruments with small and large number of variables. In addition to a column of ones, the case of $Z_t = Z2_t$ consists of the series of lagged consumption growth c_{t-1} and the case of $Z_t = Z1_t$ consists of 4 variables, c_{t-1} and lagged returns $R_{t-1} = (r^f_{t-1}, r^{MP}_{t-1}, r^{SM}_{t-1})'$. In practice the market portfolio is not observable, therefore we omit r^{MP}_{t-1} in $Z3_t$. The case of $Z4_t$ represents a situation where the number of instruments is increased by taking additional lags of the same variable. We consider two sample sizes $n = 120$ and $n = 500$.

[Table 2 about here.]

4.3. *Monte Carlo algorithm*

This paper compares the rejection probability of the asymptotic and block bootstrap *J test* under the null data generating process (DGP). Let μ_0 denote the true DGP, μ^* the bootstrap DGP generated by F_n, and τ an asymptotically pivotal test statistic. The rejection probability of the bootstrap test at nominal level α under μ_0 is $RP = P_{\mu_0}(\tau < Q(\alpha, \mu^*))$, where $Q(\alpha, \mu^*)$ is the α quantile of τ under μ^*. Using M Monte Carlo replications, an estimate $\widehat{RP}$ of this probability can be obtained,

$$\widehat{RP} = \frac{1}{M}\sum_{s=1}^{M} Ind(\tau_s \geq Q^*(\alpha, \mu^*)), \tag{4.3}$$

where $Q^*(\alpha, \mu^*)$ is the α quantile of τ^*, the bootstrap statistic, under μ^*. For each replication, the quantity (4.3) requires the calculation of B bootstrap samples. This implies, $M(B+1)$ test statistic calculated for each Monte Carlo experiment. This turns out to be computationally prohibitive because of the nonlinearity of the moment conditions and the computational demand of the CU-GMM algorithm. Davidson and MacKinnon [28, 29] propose very fast methods to obtain approximations to RP with the computational effort of only $2M + 1$ visits to (2.3) and (3.3). See, for example, Lamarche [30], Richard [31], and Ahlgren and Antell [32] for discussions of the performance of fast bootstrap inference and applications.

For a given DGP in Table 1 and a given value of the random block probability p, the (fast bootstrap) algorithm below is implemented for the NOB, the MBB and the SB bootstrap under both the iterated and the continuously updated GMM. We consider 18 values of p; $p \in \{0.05, 0.10, \cdots, 1\}$ and $\omega = 1/p$. We describe the fast bootstrap algorithm as follows.

[Figure 1 about here.]

Step 1 For $s = 1, ..., M$, where $M = 10,000$,

a) We generate a random sample of size n of consumption and dividends growth. We use the quadrature approximation to generate returns series that satisfy (2.1) and use the sample data, $\mathcal{X}$, to compute the GMM estimates $\widehat{\theta}^{(s)}$ and the *J test* value $\widehat{J}^{(s)}$.

b) We generate one single bootstrap sample $\mathcal{X}^*$ and compute the bootstrap GMM estimates $\widehat{\theta}^{*(s)}$ and the bootstrap *J test* , $\widehat{J}^{*(s)}$.

Step 2 At the end of Step 1, we have a sequence of sample statistics $\{\widehat{\theta}^{(s)}, \widehat{J}^{(s)}\}$, and bootstrap statistics $\{\widehat{\theta}^{*(s)}, \widehat{J}^{*(s)}\}$ for $s = 1, \cdots, M$.

Let $\widehat{Q}^*(\alpha)$ be the α quantile of the bootstrap *J test*, $\widehat{J}^*$. Using the approximation proposed by Davidson and MacKinnon [29], we estimate the rejection probability, RP, and the error in rejection probability, ERP,

$$\widehat{RP}_A \equiv \frac{1}{M}\sum_{s=1}^{M} Ind\left(\widehat{J}^{(s)} \geq \widehat{Q}^{*}(\alpha)\right), \tag{4.4}$$

$$\widehat{ERP} \equiv \widehat{RP}_A - \alpha. \tag{4.5}$$

5. Finite sample properties of GMM estimators and *J test*

5.1. *Statistical properties of* $\widehat{J}$

Figure 1 is a plot of the kernel density estimates for $\widehat{J}$ along with the density of the asymptotic chi-squared test (χ^2_{df}, $df = mq - k$). Table 3 reports sample summary statistics for central location (mean, median and mode), standard deviation and sample skewness for $\widehat{J}$ under a selection of DGPs from Table 1.

[Table 3 about here.]

The results support the existing evidence that the asymptotic approximation performs very poorly in finite samples. In Table 3, regardless of the weighting matrix used to construct the GMM criterion in (2.3), the sample mean, median and mode of $\widehat{J}$ are significantly higher than those of the χ^2 distribution. In addition, the distribution of $\widehat{J}$ is more dispersed and generally less skewed than the asymptotic approximation. The kernel density estimates in Figure 1 provide further evidence for the departure of the finite sample distribution of *J test* from its asymptotic approximation. The asymptotic χ^2 test suffers severe size distortion leading to spurious rejection of the C-CAPM model.

A number of interesting distributional patterns emerges from our study of the effect of the DGP properties on the distribution of the *J test.* We find that at least four factors impact the quality of the asymptotic approximation.

First, consider Figure 1a where the distribution of CU-GMM test, $\widehat{J}_{cu}$, is generated using the DGP in experiments $M1$ and $M6$. In these experiments, the transmission of the dynamics between the variables in the VAR is enabled by non-zero VAR parameters Φ_{12} and Φ_{21}.

In the remainder of the paper, we use the terms *Granger-Sims* and *transmission channels* to indicate this interaction between lc_t and d_t through their lagged values. In the DGP of model $M1$, there is no correlation between log consumption and log dividends growth while in $M6$ there is positive feedback from $log(c_t)$ to $log(d_t)$ ($\Phi_{21} = 0.414$). The distribution of $\widehat{J}_{cu}$ moves further away from the χ^2 as more interaction is introduced between the VAR variables. This result is similar to that of Tauchen [4] who finds that introducing positive feedback from consumption to dividend growth tends to raise the bias of the iterated GMM estimator.

Second, to illustrate the effect of the degree of nonlinearity, γ, in the moment function (4.1), Figure 1b plots the kernel density estimates of $\widehat{J}_{cu}$ under DGP $M3$ (small $\gamma_0 = 1.30$) and DGP $M5$ (large $\gamma = 13.7$). Our findings suggest that increased nonlinearity causes the distribution of $\widehat{J}$ to shift further away from the χ^2 approximation. This is consistent with Tauchen [4] finding that larger values of γ tend to move the bias in the IT-GMM of γ downward.

[Table 4 about here.]

Third, the weighting matrix W_n in (2.3) impacts the distribution of the over-identification test. Figure 1c plots the kernel density estimates for the IT-GMM statistic, $\widehat{J}_{it}$, and the CU-GMM statistic, $\widehat{J}_{cu}$, and Table 3 reports their summary statistics. The iterated GMM test, $\widehat{J}_{it}$, has significantly larger mean, median and standard deviation than $\widehat{J}_{cu}$. The distribution of the continuously updated GMM is however more skewed with heavy tails.

Finally, Figure 1d shows that increasing the serial correlation in the log of consumption growth series (from $-.161$ in $M3$ to $-.677$ in $M4$) shifts the body of the kernel density estimate of $\widehat{J}$ to the right with higher median and more skewness. In Figure 1a-1d, as the body of the kernel density shifts towards the right, the density curve becomes more centered and less skewed.

5.2. *Statistical properties of the GMM estimators*

Our analysis hereafter focuses on the continuously updated GMM estimator. Evidence from Monte Carlo simulations (not reported here but available upon request) suggests that the main conclusions are qualitatively valid for both the CU-GMM and IT-GMM estimators.

Table 4 presents the summary statistics for the distribution of $\widehat{\theta}_{cu}$ for an economy calibrated with preference parameters $\gamma_0 = 13.7$ and $\beta_0 = 1.139$. Panel A highlights the feedback effect discussed earlier. The GMM estimator under the DGP of model $M2$ has the least mean and median bias for γ. Although the serial correlation in $M5$ is weaker, the specification introduces strong feedback from lc_t to d_t. The estimates are biased upward suggesting similar conclusions to those of Tauchen [4] who argues that "positive association between dividend and consumption growth tends to counteract the downward bias and may produce upward bias."

Panel B highlights the effect of an increase in the sample size from 120 to 500. While the mean and median bias decrease marginally with the number of observations, there is a significant drop in the standard errors and the overall mean squared error of the estimates of the model parameters.

[Table 5 about here.]

It is worth noting that in addition to its poor approximation in finite samples, the asymptotic theory developed in Hansen [1] is unreliable in large samples under *weak identification* or *weak instruments*, which may be manifested in the Monte Carlo experiments.

First, models $M3$ and $M5$ in Table 1 are nearly rank deficient (Wright [27]). In a model with nonlinear moment conditions, global identification (assumption (b)) can hold without the rank condition (assumption (e)) being satisfied. Dovonon and Renault [18] derive the asymptotic properties of the *J test* when the parameter value is globally identified but the matrix G_0 is rank deficient. They find that the *J test* is asymptotically distributed as half and half mixture of χ_{mq-k} and χ_{mq-k-1}. Hansen [1]'s asymptotic χ^2_{mq-k} assumes first order identification and leads to significant over-rejection rate.

Earlier, we noted the small improvement in the bias and standard errors when the sample size increases. This may be attributed to the slow rate of consistency, $O_P(n^{-1/4})$ instead of $O_P(n^{-1/2})$, reported by Dovonon and Renault [18].

Second, the instruments used to construct the unconditional moments (2.2), may be weak. This means the moment restriction $\mathbb{E}\left[g(X_t,\theta)\right]$ is uniformly close to zero over the parameter space and does not permit to identify θ. In their analysis of instruments' weakness, Antoine and Renault [33] note that the moment restrictions vary quite substantially with the constant term as instrument while remains fairly small when Z_t consisted of lagged consumption growth and asset returns.

6. Block bootstrap finite sample inference

The simulation evidence suggests that the estimates of the bootstrap rejection probability $\widehat{RP}_A$ depend on the number of over-identifying restrictions (degrees of freedom), the VAR structure (serial correlation and *Granger-Sims* feedback), and the degree of nonlinearity of the moment functions (parameter γ). This is also true for the sample mean, median, mode, standard deviation and skewness of the kernel density estimate of the distribution of $\widehat{J}^*$.

[Figure 2 about here.]

6.1. *Bootstrap* GMM *inference and the number of instruments*

Table 5 reports the sample measures of location, dispersion and skewness for the block bootstrap $\widehat{J}^*$. Consider first the case where the null DGP is calibrated using $\gamma_0 = 1.30$. Panel A reports sample summary statistics when $Z1$ is used to construct the unconditional moment restrictions (2.2). The number of over-identifying restrictions in this case is $df = 10$. Panel C presents the results for the case of $Z = Z2$ with smaller number of instruments and $df = 4$. Additional moment restrictions that result from increasing the number of instruments improves the bootstrap approximation for the mean, the median and the mode of the distribution of $\widehat{J}$. However, increasing the number of instruments almost always results in higher standard errors.

[Figure 3 about here.]

In panels B and D of Table 5, the moment functions are highly nonlinear as in the case of $\gamma_0 = 13.7$. Additional moment restrictions in the form of additional instruments worsens the bootstrap approximation. In fact, the bootstrap approximation performs better when $Z = Z2$. In this case, the measures of central location for $\widehat{J}^*$ are the closest to those of $\widehat{J}$.

[Figure 4 about here.]

We find similar results when analyzing the estimated bootstrap rejection probability $\widehat{RP}_A$. Figure 2 plots the estimates of $\widehat{RP}_A$ as a function of the nominal level α and the block length probability p. The DGP in Figure 2 is model $M2$ with weakly nonlinear moment functions ($\gamma_0 = 1.139$). The severe over-rejection of the asymptotic approximation is evident. In Figure 2a, the three bootstrap methods tend to under-reject for the case of $df = 1$. The stationary bootstrap shows the largest discrepancies between the actual rejection probability and the nominal level α (the 45°-line).

In Figure 2c, more instruments are included in the estimation in the form of four additional lags of the consumption growth variable. The approximation based on the block bootstrap methods with nonrandom block lengths (NOB and MBB) is very accurate with very low departures from the 45°-line especially for low nominal levels. The stationary bootstrap performs very poorly with rejection probabilities in the order of 0.28 to 0.30 for the 5% nominal level.

Increasing the degrees of freedom to $df = 7$ results in high sensitivity of the estimates of $\widehat{RP}_A$ to p for the stationary bootstrap. In figures 2c-2d, the sample size is increased to $n = 500$. The stationary bootstrap approximation of the rejection probability improves and its variability with P diminishes.

We further demonstrate these findings by a plot of the rejection probability, under the null model $M2$, as a function of p for nominal levels 5% and 10%. The evidence in Figure 3 further supports the good performance of the NOB and MBB approximation regardless of the length of the blocks.

[Figure 5 about here.]

[Table 6 about here.]

6.2. *Sensitivity of the bootstrap inference to the block length*

Figure 4 plots estimates of the bootstrap error in rejection probability, $\widehat{ERP}$, as a function of the random blocks probability p. The null DGP is model $M2$ with low degree of nonlinearity ($\gamma = 1.30$). We make the following observations.

Firstly, the choice of p has very little effect on the over-rejection rate of the stationary bootstrap. The latter performs very poorly even with large sample size. The estimates of $\widehat{ERP}$ are more sensitive to the choice of the number of blocks for the NOB and MBB. There is a downward but "volatile" behavior of $\widehat{ERP}$ in response to the increase in p.

Second, the degree of nonlinearity in the moment functions affects the sensitivity of the bootstrap test to the block length. Figure 5 is a quantile-quantile plot (Q-Q plot) aimed at comparing the probability distribution of the approximating bootstrap test $\widehat{J}^*_{cu}$ and the actual sampling distribution of $\widehat{J}_{cu}$.

The Q-Q plot is a plot of the sample order statistics of $\widehat{J}^*_{cu}$, with varying block length $\omega = 1/p$, against the sample order statistics of $\widehat{J}_{cu}$. If the two sampling distributions have the same shape, the Q-Q plot should be approximately a straight line. Any substantial deviations from a straight line indicate that the two distributions are basically not the same and therefore the approximation is very poor.

Overall, the quality of the approximation of the block bootstrap is satisfactory. The Q-Q plots are concentrated around a straight line except for the higher order quantiles. This suggests that there is a discrepancy between the sampling distribution of the $\widehat{J}_{cu}$ and the bootstrap approximation $\widehat{J}^*$ at the tails. The plots suggest that the distribution of the bootstrap J *test* is skewed with long and heavy tails.

The DGP in Panels (a)-(c) corresponds to the highly nonlinear case with $\gamma_0 = 13.7$. In these plots, it is almost impossible to distinguish the quantile lines for the different choices of p as they all lie on the top of each other except at the tails of the distribution. However, in Panels (d)-(f), where $\gamma_0 = 1.30$, we observe high sensitivity of the distribution of $\widehat{J}^*_{cu}$ to the choice of p for both NOB and MBB tests. The quantile plots are almost parallel to each other and are further away from the straight lines in the tails of the distribution.

The Q-Q plots for the SB go to corroborates Politis and White [21] theoretical finding that the stationary bootstrap is less sensitive to the choice of the randomizing probability p.

We observe similar patterns in the bootstrap distribution of the CU-GMM estimator for γ. Table 6 reports the summary statistics for the distribution of CU-GMM $\widehat{\gamma}$ and the bootstrap CU-GMM estimators for $\widehat{\gamma}^*$ for different choices number of blocks ϕ. The mode and the skewness of the distribution of the block bootstrap GMM is more sensitive to the block size.

7. Concluding Remarks

This paper examines the small sample properties of the dependent bootstrap Generalized Method of Moment (GMM) estimators and over-identification test (J *test*) for a class of nonlinear conditional moment model.

We compare the finite sample bias-correction of the GMM estimators and size distortion of *J test* based on the *non-overlapping block* bootstrap (NOB), the *moving block* bootstrap (MBB), and the *stationary* bootstrap (SB).

First, our findings corroborate the well known results that the small sample size of the *J test* severely exceeds its asymptotic size. The over-rejection rate worsens as the dimensionality of the joint tests increases with the number of instruments. All the three bootstrap methods outperform the asymptotic approximation and provide significant reduction in the size of the test.

Second, in spite of the differences in their resampling mechanisms, the NOB and MBB have the same order of magnitude for the over-rejection rate of the resulting *J test*. The two methods outperform the stationary bootstrap which in turns has the least sensitivity to the choice of the block randomizing probability.

Third, we provide new evidence that the degree of nonlinearity of the conditional moment function affects the distribution of the bootstrap GMM estimators and *J test*. For highly nonlinear moment functions, the bootstrap *J test* is less sensitive to the choice of the block size. When the moment function is close to linear, the choice of the block size significantly affects the accuracy of the bootstrap approximation, especially in the tails of the distribution of the *J test*.

Finally, we add to the understanding of the usefulness of the block bootstrap in GMM inference under near under-identification. We find evidence of fat tails and skewness in the distribution of the *J test* that corroborate the asymptotic chi-square mixture of Dovonon and Renault [18]. The bootstrap provides improvements over the standard asymptotic theory with very slow rates of convergence when the instruments are week.

Acknowledgments

I would like to thank an anonymous referee for valuable comments and suggestions that led to substantial improvement of the paper.

References

[1] L. P. Hansen, Large Sample Properties of Generalized Method of Moments Estimators, Econometrica 50 (4) (1982) 1029–1054.
[2] L. Hansen, K. Singleton, Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, Econometrica 50 (5) (1982) 1269–1286.
[3] L. Hansen, J. Heaton, A. Yaron, Finite Sample Properties of Some Alternative GMM Estimators, Journal of Business and Economic Statistics 14 (3) (1996) 262–280.
[4] G. Tauchen, Statistical Properties of Generalized Method of Moments Estimators of Structural Parameters Obtained from Financial Market Data, Journal of Business and Economic Statistics 4 (4) (1986) 397–416.
[5] E. B. Hansen, Non-Parmetric Dependent Data Bootstrap for Conditional Moment Models, university of Wisconsin, www.ssc.wisc.edu/~bhansen, 1999.
[6] S. Bond, F. Windmeijer, Reliable Inference for GMM Estimators? Finite Sample Properties of Alternative Test Procedures in Linear Panel Data Models, Econometric Reviews 24 (1) (2005) 1–37.
[7] A. F. Lagunes, Finite Sample Evidence of IV Estimators under Weak Instruments, Journal of Applied Econometrics 22 (2007) 677–694.
[8] S. N. Lahiri, Resampling Methods for Dependent Data, Springer Series in Statistics Springer-Velag New York, Inc. (NY) 2003.
[9] C. Leger, D. N. Politis, J. P. Romano, Bootstrap Technology and Applications, Technometrics 34 (1992) 378–398.
[10] R. Lepage, L. Billard, Exploring the Limits of Bootstrap, Wiley and Sons, New York, 1992.
[11] E. Carlstein, The use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series, Annals of Statistics 14 (1986) 1171–1179.
[12] H. R. Kunsch, The Jacknife and the Bootstrap for General Stationary Observations, Annals of Statistics 17 (1989) 1217–1241.
[13] R. Y. Liu, K. Singh, Moving Blocks Jacknife and Bootstrap Capture Weak Dependence, In Exploring the Limits of Bootstrap eds. R. Lepage and L. Billard Wiley and Sons, New York (1992) 225–248.

[14] D. N. Politis, J. P. Romano, The Stationary Bootstrap, Journal of the American Statistical Association 89 (428) (1994) 1003–1313.
[15] S. N. Lahiri, Theoretical Comparison of Block Bootstrap Methods, Annals of Statistics 27 (1999) 386–404.
[16] R. Davidson, J. MacKinnon, Econometric Theory and Methods, New York: Oxford University Press.
[17] W. H. Greene, Econometric Analysis, Sixth Edition, Pearson Prentice Hall New Jersey.
[18] P. Dovonon, E. Renault, GMM overidentification test with first order underidentification, Unpublished Manuscript www4.ncsu.edu/ dpellet/tec/renault.pdf.
[19] C. Burnside, M. Eichenbaum, Small Sample Properties of Generalized Method of Moments Based Wald Tests, NBER Technical Working Paper 155.
[20] P. Hall, L. J. Horowitz, Y. Jing, Bing, On Blocking Rules for the Bootstrap with Dependent Data, Biometrika 82 (1995) 561–574.
[21] D. N. Politis, H. White, Automatic Block-Length Selection for the Dependent Bootstrap, Econometric Reviews 23 (1) (2004) 53–70.
[22] P. Hall, L. J. Horowitz, Bootstrap Critical Values for Tests Based on Generalized Method of Moments Estimators, Econometrica 64 (1996) 891–916.
[23] D. W. K. Andrews, Higher-order improvements of a Computationally Cttractive k-step Bootstrap for Extremum Estimators, Econometrica 70 (2002) 119–162.
[24] N. Kocherlakota, On Tests of Representative Consumer Asset Pricing Models, Journal of Monetary Economics 26 (1990) 285–304.
[25] G. Tauchen, R. Hussey, Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models, Econometrica 59 (2) (1991) 371–396.
[26] G. Tauchen, Finite State Space Markov-Chain Approximations to Univariate and Vector Autoregressions, Economics Letters 20 (1986) 177–181.
[27] J. Wright, H., Detecting Lack of Identification in GMM, Econometric Theory 19 (2003) 322–330.
[28] R. Davidson, J. G. MacKinnon, Fast Double Bootstrap Tests of Nonnested Linear Regression Models, Econometric Reviews 21 (2002) 417–427.
[29] R. Davidson, J. G. MacKinnon, Improving the Reliability of Bootstrap Tests with the Fast Double Bootstrap, Computational Statistics and Data Analysis 51 (7) (2007) 3259–3281.
[30] J. F. Lamarche, The Numerical Performance of Fast Bootstrap Procedures, Computational Economics 23 (4) (2004) 379–389.
[31] P. Richard, Modified fast double sieve bootstraps for ADF tests, Computational Statistics and Data Analysis In Press (2009) doi:10.1016/j.csda.2009.07.008.
[32] N. Ahlgren, J. Antell, Bootstrap and Fast Double Bootstrap Tests of Cointegration Rank with Financial Time Series, Computational Statistics and Data Analysis 52 (10) (2008) 4754–4767.
[33] B. Antoine, E. Renault, Efficient GMM with Nearly-Weak Instruments, The Econometrics Journal forthcoming.

Appendix 1

Proof: Let $\{g_t^*\}$, where $g_t^* = g^*(X_t, \widehat{\theta})$, be a martingale first difference sequence satisfying

$$\mathbb{E}\,(g_t^* g_s^{*'}) = 0, \text{for } t \neq s.$$

This implies that the matrix $\widetilde{W}_n$ for the non-overlapping blocks reduces to

$$\begin{aligned}\widetilde{W}_n &= \frac{1}{n}\sum_{i=0}^{\phi-1}\sum_{j=1}^{\omega} g^*_{i\omega+j} g^{*\prime}_{i\omega+j} \\ &= \frac{1}{n}\sum_{t=1}^{n} g_t^* g_t^{*\prime} = \Omega_n^*(\widehat{\theta}).\end{aligned}$$

Plugging $\widetilde{W}_n$ in V_n and using $W_n = \Omega^*(\widehat{\theta})^{-1}$ leads to

$$V_n = \widehat{M}_n \widehat{M}_n. \tag{7.1}$$

Let $\widehat{M_n} = I_{mq} - \widehat{P}_n$ where $\widehat{P}_n = \widehat{\Omega}_n^{-\frac{1}{2}} \widehat{G}_n [\widehat{G}'_n \widehat{\Omega}_n^{-1} \widehat{G}_n]^{-1} \widehat{G}'_n \widehat{\Omega}_n^{-\frac{1}{2}}$. $\widehat{P}_n$ is an idempotent matrix, that is $\widehat{P}_n = \widehat{P}'_n$ and $\widehat{P}_n = \widehat{P}_n \widehat{P}_n$. Therefore,

$$\begin{aligned} V_n &= \widehat{M_n} \widehat{M_n} \\ &= I_{mq} - \widehat{P}_n + \widehat{P}_n \widehat{P}_n = I_{mq} \end{aligned}$$

□

Appendix 2: Quadrature approximation

The Euler equation for the calibrated economy can be written in terms of prices and dividends,

$$E_\theta \left[\beta c_{t+1}^{-\gamma} (1 + v_{t+1}) d_{t+1} \mid \Im_t \right] = v_t \tag{7.2}$$

where we denote by $d_t = \frac{D_t}{D_{t-1}}$ the dividend growth, $v_t = \frac{P_t}{D_t}$ the price-dividend ratio and $c_t = \frac{C_{t+1}}{C_t}$ the consumption growth.

Our economy is similar to the one described by Kocherlakota [24] with three assets: the Risk free R_f which pays one unit of consumption, the market portfolio MP which pays C_t in period t and the stock market SM with dividend payoffs D_t in period t. Equation (7.2) implies a conditional moment restriction for each asset $j \in \{MP, SM, R_f\}$.

The only driving random processes in the model are c_t and d_t, and so, conditional on the past, P_t, or equivalently, v_t is a deterministic function of c_t and d_t implicitly given by the Euler equation (7.2).

The deterministic function that gives v_t as a solution to (7.2) cannot be found in closed analytic form. The quadrature approximation uses a finite state Markov process to approximate the bivariate vector autoregression, and enables the approximate solution to be obtained by matrix inversion.

The approximation involves fitting a 8 state Markov chain to log consumption growth and log dividend growth calibrated so as to approximate the first order VAR in (4.2).

Let $\widetilde{c}(l)$ and $\widetilde{d}(l)$, $l = 1, ..., 16$, denote the abscissa for the $8 - point$ quadrature rule. Each combination of abscissa $\{(k, k'), k = 1, .., 8; k' = 1, ..., 8\}$ defines a state s_j, for $j = 1, ..., \widetilde{N}$, where $\widetilde{N} = 8^2$. Let $\widetilde{c}_{s_j}$ and $\widetilde{d}_{s_j}$ denote the values of c and d in state s_j and let $\widetilde{Y}_{s_j} = (\widetilde{c}_{s_j}, \widetilde{d}_{s_j})$.

The transition matrix $\mathbf{\Pi}$ for the Markov process defined by $\mathbf{\Pi}_{k,j} = P\left(Y_{t+1} = \widetilde{y}_{s_j} | Y_t = \widetilde{y}_{s_k}\right)$ where,

$$\mathbf{\Pi}_{k,j} = \frac{p(\widetilde{y}_{s_j} | \widetilde{y}_{s_k})}{S(\widetilde{y}_{s_k}) p(\widetilde{y}_{s_j})} w_j \tag{7.3}$$

where $S(x) = \sum_{l=1}^{\widetilde{N}} \frac{p(\widetilde{y}_{s_l}|x)}{p(\widetilde{y}_{s_l})} w_l$. The Gaussian rule defines $p(x|\widetilde{y}_{s_k})$ and $p(x)$ as density functions for the bi-variate normals $N(\widetilde{y}_{s_k}, \Sigma)$ and $N(\mu, \Sigma_Y)$ respectively, where Σ_Y solves $\Sigma_Y = \mathbf{\Phi} \Sigma_Y \mathbf{\Phi}' + \Sigma$. The weights w_j are computed using a Hermite Gauss rule.

The solution to the integral in the Euler equation (7.2) is characterized by the solution to the system of $\widetilde{N}$ linear equations given by the discrete approximation in

$$\beta \sum_{j=1}^{\widetilde{N}} \Pi_{k,j} (\widetilde{c}_{s_j})^{-\gamma} (1 + \widetilde{v}_{s_j}) \widetilde{d}_{s_j} = \widetilde{v}_{s_k} \ k = 1, .., \widetilde{N} \tag{7.4}$$

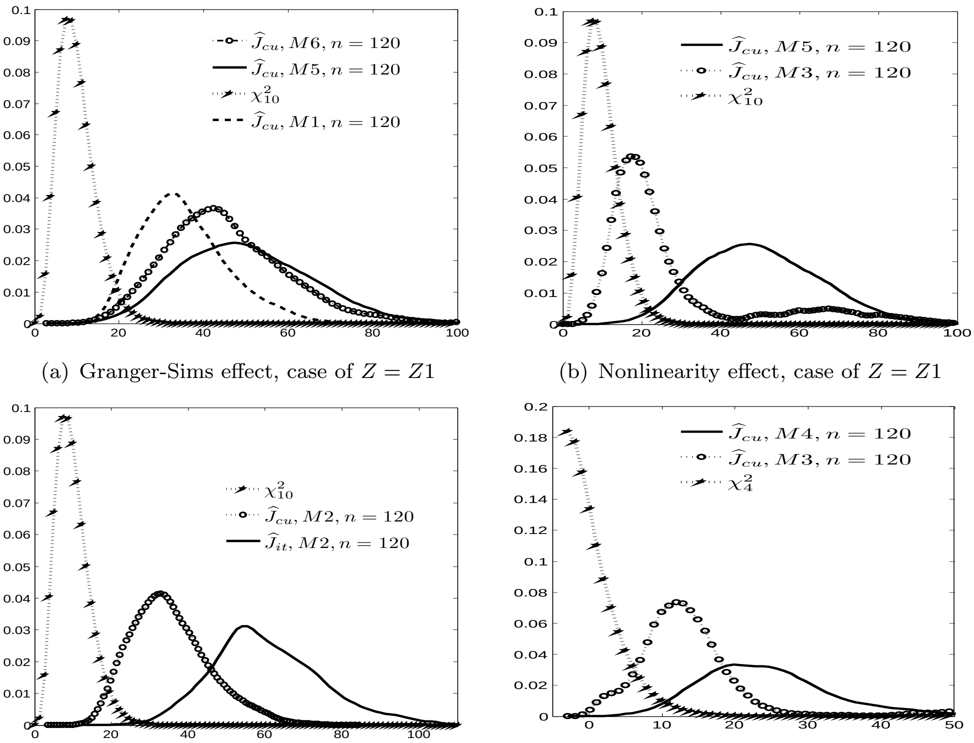
The solution exists if all the eigenvalues of the $\widetilde{N} \times \widetilde{N}$ matrix $\mathcal{S}$ defined by the elements $\mathcal{S}_{k,j} = \beta \Pi_{k,j} (\widetilde{c}_{s_j})^{-\gamma} \widetilde{d}_{s_j}$ lie within the unit circle. The solution is characterized by,

$$\widetilde{v} = (\mathbf{I}_{\widetilde{\mathbf{N}}} - \mathcal{S})^{-\mathbf{1}} \mathcal{S} \iota_{\widetilde{\mathbf{N}}} \tag{7.5}$$

where $\iota_{\widetilde{N}}$ is a $\widetilde{N} \times 1$ column vector of ones. For the market portfolio, the dividend ratio is equal to $\widetilde{d}_{MP,s_j} = \widetilde{c}_{s_j}$ and therefore $\mathcal{S}_{k,j} = \beta \Pi_{k,j} (\widetilde{c}_{s_j})^{1-\gamma}$. From the series of the equilibrium price-dividend ratios $\widetilde{v}_i, i \in \{F, MP, SM\}$, the corresponding returns are computed as:

$$\begin{aligned}
\widetilde{r}_{MP,s_k s_j} &= \widetilde{c}_{s_j} \frac{1 + \widetilde{v}_{MP,s_j}}{\widetilde{v}_{MP,s_k}} \\
\widetilde{r}_{SM,s_k s_j} &= \widetilde{d}_{s_j} \frac{1 + \widetilde{v}_{SM,s_j}}{\widetilde{v}_{SM,s_k}} \\
\widetilde{r}_{F,s_k} &= 1 / \left(\sum_{j=1}^{\widetilde{T}} \beta \Pi_{k,j} (\widetilde{c}_{s_j})^{-\gamma} \right)
\end{aligned}$$

See, for example, Tauchen [26], Tauchen and Hussey [25] and Kocherlakota [24] for further details about these derivations.



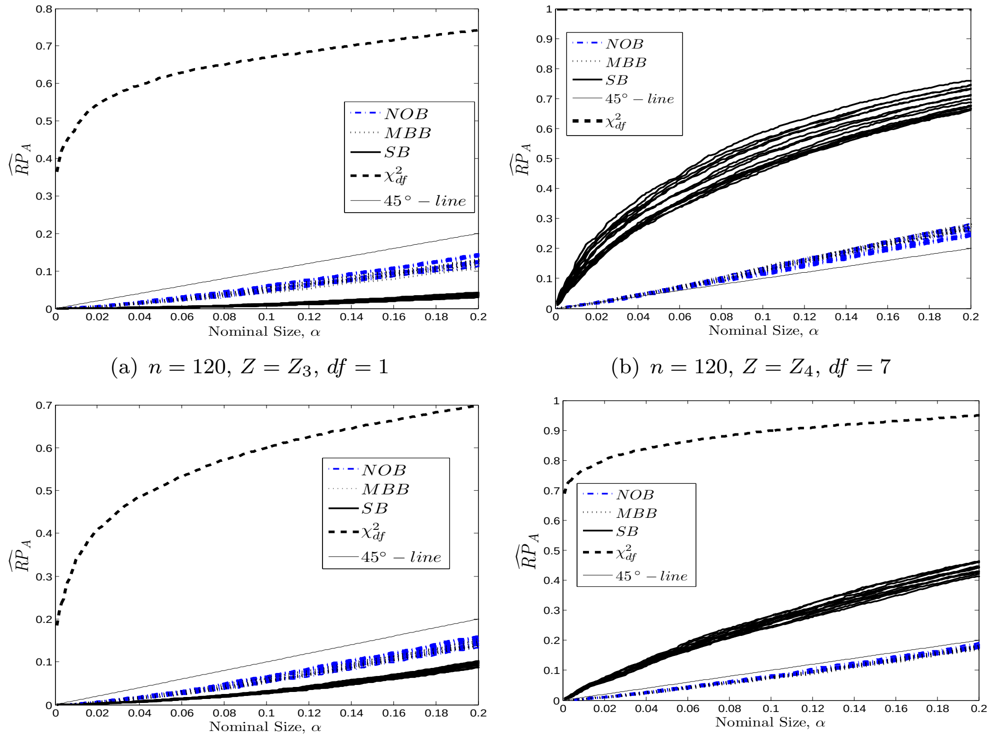
(a) Granger-Sims effect, case of $Z = Z1$

(b) Nonlinearity effect, case of $Z = Z1$

(c) Choice of weight matrix W_n, case of $Z = Z1$

(d) Serial correlation Effect, case of $Z = Z2$

Figure 1. Kernel density estimate for the sampling distribution of the *J test* statistic, $\widehat{J}$, where $\widehat{J}_{it}$ is for the IT-GMM and, $\widehat{J}_{cu}$ is for the CU-GMM. The sample size is $n = 120$, the number of Monte Carlo simulations is $M = 10,000$. Z_1 and Z_2 are the instruments used in the estimation. The plots also show the asymptotic χ^2_{df} probability density, where df is the number of degrees of freedom.



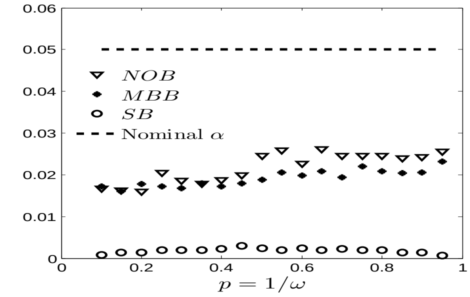
(a) $n = 120,\ Z = Z_3,\ df = 1$

(b) $n = 120,\ Z = Z_4,\ df = 7$

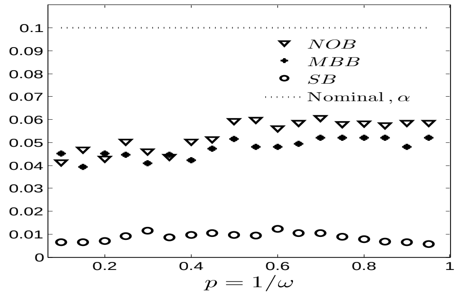
(c) $n = 500,\ Z = Z_3,\ df = 1$

(d) $n = 500,\ Z = Z_4,\ df = 7$

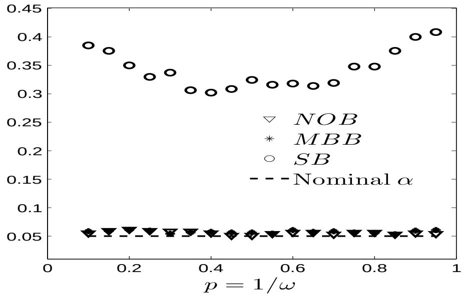
Figure 2. Fast Bootstrap estimate of the rejection probability approximation, $\widehat{RP}_A$, for the bootstrap CU-GMM *J test* under the null model $M2$. Each bootstrap method is represented by a different line-style, and each line represents a specific choice of the block length $\omega = 1/p$, where $p = 0.05, 0.10, 0.15, \cdots, 1$. The rejection probability for the asymptotic *J test* is represented by the χ^2_{df} line, where *df* is the number of degrees of freedom.



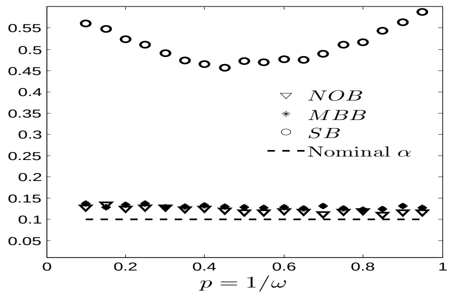
(a) $n = 500,\ Z_3$



(b) $n = 500,\ Z_3$



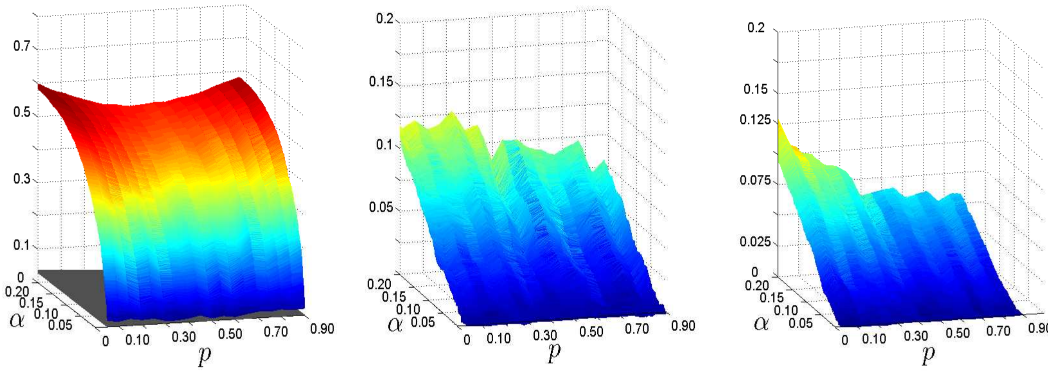
(c) $n = 500,\ Z_4$



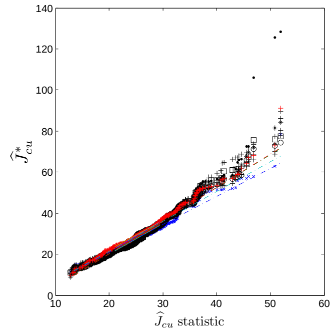
(d) $n = 500,\ Z_4$

Figure 3. Bootstrap approximation for the error in rejection probability, $\widehat{ERP}(\alpha)$ for $\alpha = 0.05$, Fig 3a and Fig 3c, and $\alpha = 0.10$, Fig 3b and Fig 3d. The null DGP is $M2$ with $\gamma_0 = 1.30$ and $\beta = 0.97$. The sample size is $n = 500$, $Z3$ and $Z4$ are defined in Table 2.

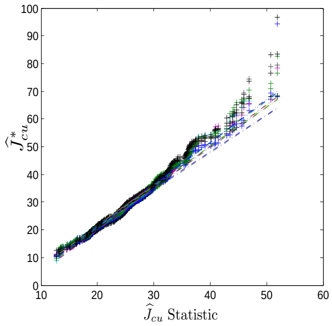
Figure 4. Estimates of the error in the rejection probability, $\widehat{ERP}$ for the bootstrap GMM test, $\widehat{J}^*_{cu}$ as a function of the block size, ω, where $\omega = \frac{1}{p}$ and nominal size α. The null model is model M_2, instruments are $Z = Z4$ and sample size $n = 500$.



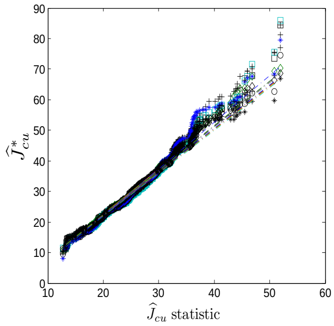
(a) SB (b) MBB (c) NOB



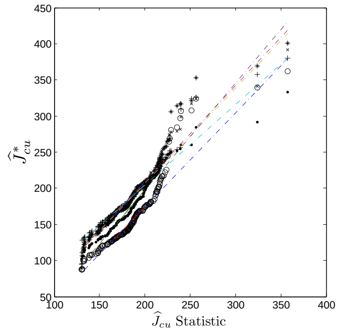
(d) NOB, $\gamma = 13.7$, $\beta = 1.139$



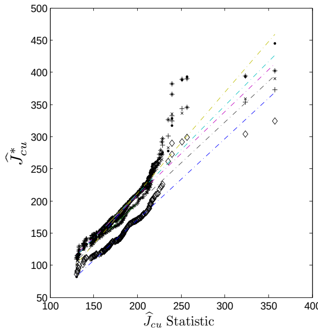
(e) MB, $\gamma = 13.7$, $\beta = 1.139$



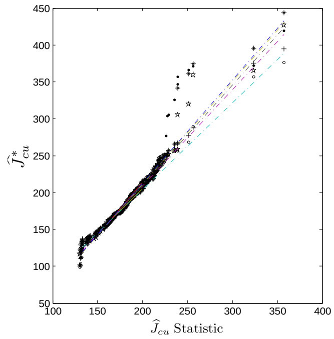
(f) SB, $\gamma = 13.7$, $\beta = 1.139$



(g) NOB, $\gamma = 1.30$, $\beta = 0.97$



(h) MB, $\gamma = 1.30$, $\beta = 0.97$



(i) SB, $\gamma = 1.30$, $\beta = 0.97$

Figure 5. QQ-plot of the quantiles of the distribution of the bootstrap $\widehat{J}^*$. The $x - axis$ represents the quantiles of $\widehat{J}$ and the $y - axis$ represents the quantiles for the bootstrap GMM J test. Each plot shows the quantiles for different choices of the blocking rule. The plots are for experiment $M6$ of *Table 1* and instruments' set $Z1$ of *Table 2*. The sample size for this experiment is $n = 500$

Table 1. Parameter and preference settings for the Monte Carlo experiments. The DGP is calibrated using the risk aversion parameter γ and the discount factor β. The consumption and dividends growth series are calibrated using a VAR with mean μ, matrix coefficient Φ and error covariance matrix V_ϵ.

Model	γ	β	Φ	V_ϵ	μ
M1	13.7	1.139	$\begin{pmatrix} -.50 & 0 \\ 0 & -.50 \end{pmatrix}$	$\begin{pmatrix} .01 & .00 \\ .00 & .01 \end{pmatrix}$	$\begin{pmatrix} .00 \\ .00 \end{pmatrix}$
M2	1.30	.97	$\begin{pmatrix} -.50 & 0 \\ 0 & -.50 \end{pmatrix}$	$\begin{pmatrix} .01 & .00 \\ .00 & .01 \end{pmatrix}$	$\begin{pmatrix} .00 \\ .00 \end{pmatrix}$
M3	1.30	.97	$\begin{pmatrix} -.161 & .017 \\ .414 & .117 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$
M4	1.30	.97	$\begin{pmatrix} -.677 & .017 \\ .414 & .117 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$
M5	13.7	1.139	$\begin{pmatrix} -.161 & .017 \\ .414 & .117 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$
M6	13.7	1.139	$\begin{pmatrix} -.677 & .017 \\ .414 & .117 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$
M7	13.7	1.139	$\begin{pmatrix} -.85 & .414 \\ .03 & -.50 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$
M8	1.30	.97	$\begin{pmatrix} -.85 & .414 \\ .03 & -.50 \end{pmatrix}$	$\begin{pmatrix} .00120 & .00177 \\ .00177 & .01400 \end{pmatrix}$	$\begin{pmatrix} 0.021 \\ 0.004 \end{pmatrix}$

Table 2. Variables used in the GMM estimation of the Calibrated economy

	Returns R_t	Instruments Z_t	df
$Z1_t$	$r^f_{t-1}, r^{MP}_{t-1}, r^{SM}_{t-1}$	$1, c_{t-1}, r^f_{t-1}, r^{MP}_{t-1}, r^{SM}_{t-1}$	10
$Z2_t$	$r^f_{t-1}, r^{MP}_{t-1}, r^{SM}_{t-1}$	$1, c_{t-1}$	4
$Z3_t$	r^{SM}_{t-1}	$1, c_{t-1}$	1
$Z4_t$	r^{SM}_{t-1}	$1, c_{t-1}, c_{t-2}, c_{t-3}, c_{t-4}$	7

Table 3. Sample mean, sample median, sample mode, standard deviation (std) and sample skewness of the distribution of *J test* for the iterated and the continuously updated GMM. The number of Monte Carlo replications is $M = 10,000$, df is the number of degrees of freedom of the asymptotic χ^2 test and the sample size is $n = 120$. The null DGP of experiments $M3$, $M4$ and $M6$ is described in Table 1. For the asymptotic test, the mean, median, standard deviation, mode and skewness correspond to their theoretical values for a χ^2_{df}.

	CU-GMM, $\widehat{J}_{cu}$				IT-GMM, $\widehat{J}_{it}$					
	$df = 10$		$df = 4$		$df = 10$		$df = 4$		Asymptotic test	
	M6	M3	M6	M4	M6	M4	M6	M4	χ^2_{10}	χ^2_4
mean	41	27	18	42	91	64	42	181	10	4
median	40	20	17	40	90	61	40	180	9.34	3.36
std	12	19	6	13	20	17	15	25	4.47	2.82
mode	13	5	6	11	39	23	8	105	8	2
skewness	0.73	1.77	1.09	1.00	0.53	0.91	0.51	0.29	0.89	1.41

Table 4. Measures of Bias and performance for GMM estimators of γ and β. Case of Continuous updating estimator, $n = 120$

	A: Effect of VAR specification Instruments, Z_1, $df = 10$ true $\gamma_0 = 13.7$, $\beta_0 = 1.139$						B: Effect of sample size Instruments, Z_2, $df = 4$ true $\gamma_0 = 13.7$, $\beta_0 = 1.139$			
	n = 120						**Model: M6**			
	M7		M1		M5		$n = 120$		$n = 500$	
	γ	β	γ	β	γ	β	γ	β	γ	β
mean	21.114	1.188	14.964	1.262	15.990	1.202	15.990	1.2026	15.470	1.198
median	21.776	1.184	14.779	1.264	15.713	1.200	15.713	1.2004	15.437	1.199
std	2.922	0.053	1.225	0.099	1.689	0.036	1.689	0.0366	0.574	0.015
RMSE	7.969	0.072	1.760	0.158	2.846	0.073	3.660	0.0636	1.801	0.059

Table 5. Sample measures of location, dispersion and skewness for the distribution of the bootstrap CU-GMM test $\widehat{J}^*$. The DGPs under the null models $M3$ and $M4$ are described in Table 1 and are calibrated using the preference parameters θ_0. The sample size is $n = 120$, the instruments $Z1$ and $Z2$ are described in Table 2, and df is the number of degrees of freedom of the χ^2 approximation. The number of nonrandom blocks for NOB and MBB is $\phi = n/\omega$, where the blocks length ω is $\omega = 1/p$ and p is the stationary bootstrap probability, $p = 0.05, 0.10, \cdots, 1$. For comparison, the last row in each Panel reports the summary statistics for the distribution of CU-GMM test, $\widehat{J}_{cu}$.

Panel A: Null model is $M3$, $Z = Z1$, $df = 10$, $\theta_0 = (1.30, 0.97)'$

		Non-overlapping blocks: NOB					Moving blocks: MBB					Stationary bootstrap: SB				
$p = 1/\omega$	$\phi = n/\omega$	mean	median	mode	std	skewness	mean	median	mode	std	skewness	mean	median	mode	std	skewness
0.05	6	40	28	5	28	1.26	42	31	3	28	1.23	48	37	6	30	1.06
0.15	18	43	30	2	31	1.53	45	33	8	30	1.15	48	37	4	29	1.03
0.25	30	45	32	5	32	2.21	46	34	5	30	1.24	48	36	7	29	0.98
0.35	42	46	35	4	30	1.20	46	34	3	30	1.16	47	36	5	29	1.02
0.45	54	46	35	5	30	1.18	47	35	8	30	1.10	48	36	7	29	1.08
0.55	66	47	35	3	29	1.12	47	36	4	29	1.03	48	37	7	29	0.95
0.65	78	47	36	7	30	1.06	47	36	5	29	1.06	49	38	6	30	1.01
0.75	90	47	37	4	29	1.12	47	36	3	29	0.98	48	36	7	29	0.97
0.85	102	48	36	4	30	1.12	48	36	6	30	1.02	47	37	8	29	0.99
1.00	120	48	37	7	30	1.01	48	36	2	30	0.99	48	37	6	29	0.98
$\widehat{\mathbf{J}}_{\mathbf{cu}}$		**27**	**20**	**5**	**19**	**1.77**										

Panel B: Null model is $M6$, $Z = Z1$, $df = 10$, $\theta_0 = (13.7, 1.139)'$

		Non-overlapping blocks: NOB					Moving blocks: MBB					Stationary bootstrap: SB				
$p = 1/\omega$	$\phi = n/\omega$	mean	median	mode	std	skewness	mean	median	mode	std	skewness	mean	median	mode	std	skewness
0.05	6	53	49	14	22	1.53	55	52	18	20	1.12	59	55	19	21	1.14
0.15	18	57	51	15	27	2.45	57	52	15	23	1.53	59	56	18	21	1.14
0.25	30	56	50	11	25	2.22	58	52	15	27	2.82	59	56	14	21	1.20
0.35	42	57	52	16	24	1.68	57	53	14	24	1.83	59	56	17	21	1.08
0.45	54	59	54	14	24	1.56	59	54	14	23	1.43	59	55	14	21	1.04
0.55	66	57	53	16	24	2.15	57	53	17	23	1.31	59	56	16	21	1.07
0.65	78	58	54	17	22	1.14	57	54	11	21	1.28	59	55	19	22	1.30
0.75	90	58	54	13	23	1.46	56	52	14	20	1.26	59	56	17	21	1.08
0.85	102	58	54	12	20	1.07	59	56	13	22	1.25	58	56	16	20	0.88
1.00	120	59	56	12	21	1.17	59	56	16	20	1.09	60	55	13	22	1.17
$\widehat{\mathbf{J}}_{\mathbf{cu}}$		**41**	**40**	**13**	**12**	**0.73**										

Panel C: Null model is $M3$, $Z = Z2$, $df = 4$, $\theta_0 = (1.30, 0.97)'$

		Non-overlapping blocks: NOB					Moving blocks: MBB					Stationary bootstrap: SB				
$p = 1/\omega$	$\phi = n/\omega$	mean	median	mode	std	skewness	mean	median	mode	std	skewness	mean	median	mode	std	skewness
0.05	6	47	44	9	18	1.35	49	46	15	19	1.11	50	48	9	18	0.73
0.15	18	46	44	10	18	1.14	49	44	8	20	1.44	50	47	14	19	1.05
0.25	30	48	44	7	20	1.24	49	44	15	21	1.18	51	48	11	19	0.92
0.35	42	48	44	13	21	1.19	48	44	10	19	0.97	50	45	13	20	1.52
0.45	54	48	46	13	20	1.31	50	47	11	22	1.44	49	47	11	19	0.81
0.55	66	49	45	9	21	1.29	48	43	8	21	1.04	50	47	10	18	0.81
0.65	78	50	46	12	22	1.13	50	46	13	21	1.33	50	46	17	20	1.31
0.75	90	48	44	5	21	1.27	47	43	12	20	1.12	50	47	16	20	1.18
0.85	102	50	46	11	21	1.05	50	45	10	19	1.20	49	46	13	18	0.92
1.00	120	50	47	11	20	1.05	51	47	12	20	1.17	50	47	12	19	0.99

Panel D: Null model is $M6$, $Z = Z2$, $df = 4$, $\theta_0 = (13.7, 1.139)'$

		Non-overlapping blocks: NOB					Moving blocks: MBB					Stationary bootstrap: SB				
$p = 1/\omega$	$\phi = n/\omega$	mean	median	mode	std	skewness	mean	median	mode	std	skewness	mean	median	mode	std	skewness
0.05	6	22	20	6	9	1.43	23	21	6	10	1.47	24	23	5	10	1.20
0.15	18	23	21	6	10	1.48	23	21	6	11	1.73	24	22	5	10	1.21
0.25	30	23	21	5	11	2.19	24	21	6	11	1.56	24	22	6	10	1.02
0.35	42	23	21	4	11	2.11	23	21	4	11	1.68	24	22	6	10	1.30
0.45	54	23	22	5	10	1.50	24	22	4	10	1.34	24	22	4	10	1.24
0.55	66	24	21	3	11	1.45	23	21	2	10	1.25	24	23	3	11	1.39
0.65	78	24	22	5	11	1.51	23	21	5	11	1.66	24	22	5	11	1.57
0.75	90	24	22	6	10	1.46	23	21	5	10	1.17	24	22	4	10	1.12
0.85	102	24	22	5	10	1.23	24	22	6	10	1.41	24	22	5	11	1.60
1.00	120	24	23	5	10	1.17	25	23	5	11	1.22	24	23	4	10	1.18
$\widehat{\mathbf{J}}_{\mathbf{cu}}$		**18**	**17**	**6**	**6**	**1.09**										

Table 6. Sample measures of location, dispersion and skewness for the distribution of the bootstrap CU-GMM test $\widehat{J}^*$. The DGPs under the null models $M3$ and $M6$ are described in Table 1 and are calibrated using the preference parameters θ_0. The sample size is $n = 500$, the instruments $Z1$ are described in Table 2, and the number of degrees of freedom of the χ^2 approximation is $df = 10$.

				Panel A: $M6$, $\gamma_0 = 13.7$					Panel B: $M3$, $\gamma_0 = 1.139$				
				mean	median	std	mode	skewness	mean	median	std	mode	skewness
			$\widehat{\gamma}$	21.75	21.79	1.87	17.24	-0.13	8.51	8.54	1.16	1.30	-1.10
	p	ϕ		mean	median	std	mode	skewness	mean	median	std	mode	skewness
NOB	0.05	25		22.03	22.59	2.04	16.74	-0.24	8.39	8.57	2.11	-8.71	-3.46
	0.35	175		22.02	22.81	2.04	16.02	-0.47	8.42	8.52	2.29	-7.67	-2.15
	0.65	325		22.04	22.81	2.07	16.44	-0.62	8.08	8.47	2.66	-8.71	-2.44
	0.95	475		22.18	22.86	2.17	16.67	-0.48	7.91	8.32	3.13	1.30	-1.73
	1.00	500		21.97	22.62	2.06	16.20	-0.45	8.35	8.46	1.93	1.30	-1.34
MBB	0.05	25		21.96	22.51	2.07	15.89	-0.42	8.30	8.59	2.33	1.30	-3.36
	0.35	175		22.05	22.83	2.06	15.73	-0.53	8.41	8.59	2.51	-7.19	-0.55
	0.65	325		22.19	22.87	2.07	16.42	-0.51	8.20	8.53	2.77	1.30	-1.90
	0.95	475		22.17	22.96	2.16	16.56	-0.45	8.09	8.43	2.75	1.30	-2.18
	1.00	500		22	23	1.98	17	-0.49	8.22	8.53	2.65	-7.59	-2.07
SB	0.05	25		21.96	22.51	2.07	15.89	-0.42	8.30	8.59	2.33	1.30	-3.36
	0.35	175		22.05	22.83	2.06	15.73	-0.53	8.41	8.59	2.51	-7.19	-0.55
	0.65	325		22.19	22.87	2.07	16.42	-0.51	8.20	8.53	2.77	1.30	-1.90
	0.95	475		22.17	22.96	2.16	16.56	-0.45	8.09	8.43	2.75	1.30	-2.18
	1.00	500		21.95	22.53	1.98	16.79	-0.49	8.22	8.53	2.65	-7.59	-2.07