Implied probability kernel block bootstrap for times series moment condition models

Paulo M.D.C. Parente
Richard J. Smith

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP07/24

Implied Probability Kernel Block Bootstrap for Time Series Moment Condition Models*

Paulo M.D.C. Parente
ISEG- Lisbon School of Economics & Management, Universidade de Lisboa
REM - Research in Economics and Mathematics;
CEMAPRE- Centro de Matemática Aplicada à Previsão e Decisão Económica
pparente@iseg.ulisboa.pt

Richard J. Smith
*c*emmap, U.C.L and I.F.S.
Faculty of Economics, University of Cambridge
Department of Economics, University of Melbourne
National Institute for Economic and Social Research
rjs27@econ.cam.ac.uk

This Draft: April 2024

Abstract

This article generalizes and extends the kernel block bootstrap (KBB) method of Parente and Smith (2018, 2021) to provide a comprehensive treatment of its use for GMM estimation and inference in time-series models formulated in terms of moment conditions. KBB procedures that employ bootstrap distributions with generalised empirical likelihood implied probabilities as probability mass points are also considered. The first-order asymptotic validity of new KBB estimators and test statistics for over-identifying moments, additional moment constraints and parametric restrictions is established. Their empirical distributions may serve as practical alternative approximations to those of GMM estimators and statistics and to other bootstrap distributions in the extant literature. Simulation experiments reveal that critical values arising from the empirical distributions of some KBB test statistics are more accurate than those from standard first-order asymptotic theory.

JEL Classification: C14, C15, C32
Keywords: bootstrap; generalised empirical likelihood; generalised method of moments; heteroskedastic and autocorrelation consistent inference.

*This paper is a complete revision of and provides a substantial generalisation of the results in Parente, P.M.D.C. and R.J. Smith (2018): "Generalised empirical likelihood kernel block bootstrapping," REM Working Paper 055-2018, REM, ISEG - Lisbon School of Economics and Management, Universidade de Lisboa. Paulo Parente is grateful for the financial support of Fundação para a Ciência e a Tecnologia (Project PTDC/IIM-ECO/3097/2014 and Project CEMAPRE/REM - UIDB/05069/2020).

1 Introduction

The recent papers Parente and Smith (2018, 2021) introduce a novel bootstrap method, the kernel block bootstrap (KBB), for the analysis of time-series data. This article generalizes and extends KBB to provide a comprehensive treatment of its use for GMM estimation and inference in time-series models formulated in terms of moment conditions. KBB procedures that employ bootstrap distributions with generalised empirical likelihood implied probabilities as probability mass points are also considered. The paper details new KBB estimators and test statistics whose empirical distributions can serve as alternative approximations to those offered by standard and other bootstrap methods for GMM estimators and test statistics.

The particular focus of the paper is generalized method of moments (GMM) proposed in the seminal paper Hansen (1982) which, because of its wide-spread applicability in many and varied contexts, has become the main workhorse for estimation and inference in the analysis of economic data. As is widely appreciated, however, for the sample sizes usually available in practice, the standard large sample distributions of GMM estimators and statistics typically poorly approximate their respective empirical distributions, a situation worsened in the time-series context due to dependence; see *inter alia* Burnside and Eichenbaum (1996), Christiano and den Haan (1996) and Hansen, Heaton and Yaron (1996). The bootstrap method originally proposed in the landmark paper Efron (1979) offers an alternative approach to ameliorate this problem and has often been found to be successful in this regard. From a practical standpoint the bootstrap, being a resampling method, has the benefit of not requiring the application of complicated formulae. Theoretically, the bootstrap may admit asymptotic refinements if the statistic of interest is asymptotically pivotal and a smooth function of the data.

This article departs from the dominant paradigm of bootstrap resampling of moving blocks; see, e.g., the review paper Kreiss and Paparoditis (2011) and associated discussion and the monographs Shao and Tu (1995) and Lahiri (2003). We introduce resampling schemes based on the KBB method of Parente and Smith (2018, 2021), which, rather than work with the observational sample moment indicators or functions directly, resamples suitable kernel function-based weighted transformations of the sample moment indicators, an idea borrowed from the (generalised) empirical likelihood ((G)EL) literature, see, e.g., Kitamura and Stutzer (1997) and Smith (1997, 2011).[1] Furthermore, we implement KBB by independently resampling from bootstrap distributions which employ either the standard empirical measure or GEL implied probabilities as mass points, the latter thereby potentially exploiting efficient moment estimation, cf. Brown and Newey (1998) and Smith (2011). The KBB method itself is a generalisation

[1] Such transformations in the presence of weakly dependent data induce large sample efficiency for GEL as with randomly sampled data. The sample mean and, moreover, the standard random sample variance of the transformed sample moment indicators respectively provide a consistent estimator for the population mean, Smith (2011, Lemma A.1, p.1217), and a heteroskedastic and autocorrelation (HAC) consistent and automatically positive semidefinite estimator for the variance of the standardized mean of the original sample moment indicators, Smith (2005, Section 2, pp.161-165, and 2011, Lemma A.3, p.1219).

of the tapered block bootstrap (TBB), Paparoditis and Politis (2001), but, in contradistinction, allows kernel functions with unbounded support and includes the incomplete blocks at the beginning and end of the sample. While both KBB and TBB admit familiar kernels with finite support, e.g., rectangular, Bartlett and Tukey-Hanning, KBB also allows non-monotonic truncated kernels in the positive quadrant, e.g., flat-top cosine windows (D'Antona and Ferrero, 2006, p.40), excluded by TBB, cf. Paparoditis and Politis (2001, Assumption 2, p.1107). For a more detailed comparison, see Parente and Smith (2018, section 4.1, pp.6-8).[2]

Bootstrap methods for moment condition models have been developed in numerous contributions; see *inter alia* Hahn (1996) and Brown and Newey (1992, 2002) for randomly sampled data and Hall and Horowitz (1996) and Andrews (2002) for weakly dependent data. Two strands to this literature may be discerned: first, `i.i.d.` resampling and, secondly, GEL implied probability resampling. In the former Hahn (1996), for `i.i.d.` data, proves the consistency of the 2-step (2S) GMM bootstrap distribution for the limiting distribution of standard 2SGMM. Camponovo (2016) investigates asymptotic refinements of an `i.i.d.` bootstrap for quasi-likelihood ratio type tests of nonlinear restrictions which are applicable in a GMM framework. Hall and Horowitz (1996), with weakly dependent data, after centering the bootstrap sample moment indicators at their sample mean, apply the non-overlapping moving blocks bootstrap (MBB) method of Carlstein (1986) to the 2SGMM estimator, a t-test statistic for a single parametric restriction and the Hansen (1982) test statistic for over-identifying moment restrictions. Their resultant bootstrap statistics admit asymptotic refinements after appropriate re-scaling. Andrews (2002) extends Hall and Horowitz (1996) to standard overlapping MBB (Künsch, 1989, and Liu and Singh, 1992), and the k-step bootstrap (Davidson and Mackinnon, 1999). Both papers, however, require a form of m-dependence to achieve higher order refinements, an assumption relaxed in Inoue and Shintani (2006) for the instrumental variable linear model. In the latter literature, Brown and Newey (1992, 2002), for `i.i.d.` data, suggests independent resampling from a bootstrap distribution with GEL implied probabilities replacing the standard empirical measure, hazarding that this bootstrap method might offer theoretical improvements over standard `i.i.d.` resampling. More recently, the (G)EL implied probability bootstrap was extended to the time series context in Allen et al. (2011) and Bravo and Crudu (2012) using the rectangular kernel-weighted observational sample moment indicators and GEL implied probabilities as bootstrap distribution mass point probabilities. These papers differ in a number of respects: first, Allen et al. (2011) studies EL implied probabilities while Bravo and Crudu (2011) uses GEL implied probabilities, Smith (2011, (3.1), p.1205); secondly, Allen et al. (2011) analyses both non-overlapping and overlapping MBB whereas Bravo and Crudu (2012) only studies the latter; thirdly, Allen et al. (2011) investigates first order validity for general GMM estimators whereas Bravo and

[2] In the case of the sample mean, a particular choice of the kernel with unbounded support yields a bootstrap variance estimator that is asymptotically equivalent in mean square to the optimal quadratic spectral estimator of the long run variance, see Parente and Smith (2018, Corollary 3.1, p.6) and fn.4 below; cf. Andrews (1991, Theorem 2, p.829).

Crudu (2011) only considers the efficient 2SGMM estimator, with both articles addressing the first-order asymptotic behaviour of their respective bootstrapped over-identifying moment test statistics and Wald test statistics for parametric restrictions; finally, Bravo and Crudu (2011) proposes bootstrap Lagrange multiplier and distance statistics for parametric restrictions using null hypothesis rather than maintained hypothesis GEL implied probabilities. A recent related paper, La Vecchia et al. (2023), constructs higher order correct confidence regions for the full parameter vector based on the unrestricted GMM or GEL criterion, employing `i.i.d.` resampling of kernel function-based weighted sample moment indicators evaluated at the respective first order efficient GMM or GEL estimator.

The KBB bootstrap approach taken here employs general forms for both the kernel function-based weighted sample moment indicators and the GEL implied probabilities that define the bootstrap distribution probability mass points. We examine both unrestricted and restricted moment condition models, the latter subject to additional moment constraints and parametric restrictions expressed generally in mixed form (Gouriéroux and Monfort, 1989); cf. Newey and McFadden (1994, section 9, pp.2215-2241), Ruud (2000, chapter 22, pp.564-607) and Smith (2011, section 5, pp.1209-1213). The restricted GEL implied probability GMM-KBB estimator is presented whose distribution suitably centred is first order asymptotically valid, consistently estimating and approximating the asymptotic distribution of the corresponding restricted GMM estimator. The GEL implied probability GMM estimator provides the appropriate centring; this GMM estimator minimises the corresponding GMM criterion but, to compute the requisite sample average moment indicator, replaces the standard GMM empirical measure by the bootstrap mass point probabilities as weights. The asymptotic validity result is specialised for unrestricted GMM estimation and both efficient unrestricted and restricted GMM estimation.[3] We explore the impact of using unrestricted and restricted GEL implied probabilities, efficient or otherwise, and the standard empirical measure as bootstrap mass point probabilities. Correspondingly, we describe appropriately centred GEL implied probability GMM-KBB overidentifying moment restrictions test statistics and likelihood ratio- (distance), score-, Lagrange multiplier- and generalised Wald-like test statistics for mixed form additional moment constraints and parametric restrictions. We also define alternative likelihood ratio-like and distance statistics which adapt and generalise the bootstrap statistic proposed in Camponovo (2016) for the dependent data context and inference setting considered here.

The effect of efficient unrestricted and restricted GEL implied probabilities and the standard empirical measure as bootstrap mass point probabilities on the form of the GEL implied probability GMM-KBB test statistics is explored. We establish the first order asymptotic validity of these GEL implied probability GMM-KBB test statistics for their non-bootstrap counterparts. Computationally simpler alternative GEL implied probability GMM-KBB estimators and statistics are presented which avoid

[3]We note that the consistency proof for the EL block bootstrap of Allen et al. (2011) is in error if applied to the inefficient GMM estimator. The bootstrap distribution of the GMM estimator should be centred at efficient 2SGMM. Hence Allen et al. (2011, Theorems 1 and 2, p.114) are invalid in general although these results continue to hold if the GMM weighting matrix is the efficient GMM metric.

the necessity of computing the GEL implied probability GMM-KBB estimator. Although the proofs in the paper are developed for KBB they may be straightforwardly adapted analogously for other block bootstrap methods.

This paper is organized as follows. Section 2 briefly surveys and summarises unrestricted and restricted GMM estimation subject to additional moment constraints and parametric restrictions expressed in mixed form together with associated GMM inference. GEL implied probabilities are introduced in section 3.1 with the GEL implied probability GMM estimator described in section 3.2. GEL implied probability GMM-KBB estimation and inference are discussed in sections 3.3 and 3.4. Section 4 presents simulation evidence on the usefulness of the empirical distributions of GEL implied probability GMM-KBB test statistics as descriptions of those of the corresponding GMM test statistics. Finally section 5 concludes. Appendix A provides preliminary lemmas and their proofs which are required for the proofs given in Appendix B of the results of the main text. Appendix C details limit results for KBB heteroskedastic and autocorrelation consistent variance matrix estimation.

2 GMM Preliminaries

Let z_t, $(t = 1, ..., T)$, denote a sample of T observations on the stationary and strong mixing real valued d_z-dimensional vector process $\{z_t\}_{t=1}^{\infty}$. Also let $\mathrm{E}[\cdot]$ and $\mathrm{var}[\cdot]$ denote expectation and variance taken with respect to the unknown probability measure $\mathcal{P}$ of the process $\{z_t\}_{t=1}^{\infty}$.

Consider the moment indicator $q(z_t, \theta)$, a d_q-vector of known functions of the data observation z_t and the d_θ-vector $\theta \in \Theta$ of unknown parameters, where $\Theta \subset \mathcal{R}^{d_\theta}$ denotes the parameter space. The moment vector $q(z_t, \theta)$ is partitioned as $q(z_t, \theta) = (g(z_t, \beta)', h(z_t, \theta)')'$, where $g(z_t, \beta)$ and $h(z_t, \theta)$ are d_g- and d_h-subvectors, and the parameter vector θ is partitioned $\theta = (\alpha', \beta')'$, where $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ are d_α- and d_β-subvectors of θ, $\Theta = \mathcal{A} \times \mathcal{B}$, $\mathcal{A} \subset \mathcal{R}^{d_\alpha}$, $\mathcal{B} \subset \mathcal{R}^{d_\beta}$, $d_g \geq d_\beta$. We maintain the moment condition

$$\mathrm{E}[g(z_t, \beta)] = 0 \tag{2.1}$$

at the true value β_0 of β throughout the paper whereas the additional moment constraints and parametric restrictions

$$\mathrm{E}[h(z_t, \theta)] = 0, \; r(\theta) = 0, \tag{2.2}$$

satisfied by the true value $\theta_0 = (\alpha_0', \beta_0')'$, constitute a hypothesis of particular interest, cf. Smith (2011, Section 5. pp.1209-1213); both the moment indicator $h(z_t, \cdot)$ and the d_r-vector of parametric constraints $r(\cdot)$ are expressed in mixed form (Gouriéroux and Monfort, 1989) depending on both the additional parameter vector α as well as β.

Let $q_t(\theta) = q(z_t, \theta)$, $(t = 1, ..., T)$, and define the sample mean $\hat{q}(\theta) = \sum_{t=1}^{T} q_t(\theta)/T$ and long-run variance $\Xi(\theta) = \lim_{T\to\infty} \mathrm{var}[T^{1/2}\hat{q}(\theta)]$, similarly $g_t(\beta) = g(z_t, \beta)$ and $h_t(\theta) = h(z_t, \theta)$, $(t = 1, ..., T)$,

$\hat{g}(\beta) = \sum_{t=1}^{T} g_t(\beta)/T$ and $\hat{h}(\theta) = \sum_{t=1}^{T} h_t(\theta)/T$, and long-run variance $\Sigma(\beta) = \lim_{T\to\infty} \text{var}[T^{1/2}\hat{g}(\beta)]$. Also let $\Xi = \Xi(\theta_0)$ and $\Sigma = \Sigma(\beta_0)$.

The GEL implied probability GMM-KBB sampling schemes, the main concern of this paper, make use of GEL implied probabilities, see Section 3.1, defined in terms of the transformed kernel-weighted moment indicator

$$q_{tT}(\theta) = \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=t-T}^{t-1} k(\frac{s}{S_T}) q_t(\theta), \tag{2.3}$$

partitioned as $q_{tT}(\theta) = (g_{tT}(\beta)', h_{tT}(\theta)')'$, $(t = 1, ..., T)$, where S_T is a bandwidth parameter and $k(\cdot)$ a kernel function with $k_j = \int_{-\infty}^{\infty} k(x)^j dx$, $(j = 1, 2)$, and $k = k_1^2/k_2$.

Remark 2.1. A different scaling is employed here to that in Kitamura and Stutzer (1997) and Smith (1997, 2011), namely $(k_2 S_T)^{-1/2}$ rather than $(k_2 S_T)^{-1}$, which permits the standard outer product form to be employed as a consistent estimator of the long-run variance matrix Ξ; see Remark 2.4 below. The standard sample average $\hat{q}(\theta)$ ($\hat{g}(\beta)$, $\hat{h}(\theta)$) and the sample average $\hat{q}_T(\theta) = \sum_{t=1}^{T} q_{tT}(\theta)/T$ ($\hat{g}_T(\beta) = \sum_{t=1}^{T} g_{tT}(\beta)/T$, $\hat{h}_T(\theta) = \sum_{t=1}^{T} h_{tT}(\theta)/T$) are first order asymptotically equivalent after scaling under the assumptions stated below, i.e., both obey UWLs, e.g., $\sup_{\theta\in\theta} \|\hat{q}_T(\theta)/S_T^{1/2} - \text{E}[q_t(\theta)]/k^{1/2}\| \to 0$, prob-$\mathcal{P}$, cf. $\sup_{\theta\in\theta} \|\hat{q}(\theta) - \text{E}[q_t(\theta)]\| \to 0$, prob-$\mathcal{P}$, and CLTs, e.g., $(T/S_T)^{1/2}(\hat{q}_T(\theta) - \text{E}[\hat{q}_T(\theta)]) \to^{d_\mathcal{P}} \mathcal{N}(0, \Xi(\theta)/k)$, cf. $T^{1/2}(\hat{q}(\theta) - \text{E}[q_t(\theta)]) \to^{d_\mathcal{P}} N(0, \Xi(\theta))$; see Smith (2011, Lemmas A.1, p.1217, and A.2, p.1219).

2.1 GMM Estimation

Since the GEL sample average $\hat{q}_T(\theta)$ forms the basis of GEL implied probability GMM-KBB estimators and statistics described below, the following discussion is conducted, without loss of generality, in terms of $\hat{q}_T(\theta)$ rather than $\hat{q}(\theta)$ used in standard GMM analysis.

Let W_{qT} denote a (d_q, d_q) p.s.d. matrix such that $W_{qT} \to W_q$, prob-$\mathcal{P}$, W_q p.d. A restricted GMM estimator for θ_0 is defined by

$$\breve{\theta}_T = \arg\min_{\theta\in\Theta_r} \breve{\mathcal{Q}}_T(\theta) \tag{2.4}$$

where $\Theta_r = \{\theta \in \Theta : r(\theta) = 0\}$ with associated Lagrangean function $\breve{\mathcal{L}}_T(\theta, \mu) = \breve{\mathcal{Q}}_T(\theta)/S_T - k\mu' r(\theta)$, Lagrange multiplier estimator $\breve{\mu}_T$ and GMM criterion

$$\breve{\mathcal{Q}}_T(\theta) = \hat{q}_T(\theta)'(W_{qT})^{-1}\hat{q}_T(\theta). \tag{2.5}$$

Cf. the standard GMM criterion which substitutes $\hat{q}(\theta)$ for $\hat{q}_T(\theta)$; see, e.g., the seminal paper Hansen (1982). Cf. Smith (2005, Section 3, pp.165-166).

The following regularity conditions, cf. Gonçalves and White (2004), are adaptations for the bootstrap moment condition context of Parente and Smith (2021, Assumptions 3.1-3.6, p.382) and Smith (2011, Assumptions 2.1, 2.2, p.1199, and 5.1, 5.3, p.1210).

Let $G = \mathrm{E}[\partial g_t(\beta_0)/\partial\beta']$, $Q = \mathrm{E}[\partial q_t(\theta_0)/\partial\theta']$, $H_\alpha = \mathrm{E}[\partial h_t(\theta_0)/\partial\alpha']$, $R = \partial r(\theta_0)/\partial\theta'$ and $R_\alpha = \partial r(\theta_0)/\partial\alpha'$.

Assumption 2.1. **(a)** $(\Omega, \mathcal{F}, \mathcal{P})$ is a complete probability space; **(b)** the finite d_z-dimensional stochastic process z_t: $\Omega \longmapsto \mathcal{R}^{d_z}$, $(t = 1, 2, ...)$, is stationary and strong mixing with mixing coefficients of size $-3v/(v-1)$ for some $\nu > 1$ and is measurable for all t, $(t = 1, 2, ...)$.

Let $\mathbb{I}(x \geq 0)$ denote the indicator function, i.e., $\mathbb{I}(A) = 1$ if A true and 0 otherwise.

Assumption 2.2. **(a)** $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $\frac{1}{6} < \eta < \frac{1}{2}$; **(b)** $k(\cdot)$: $\mathcal{R} \to [-k_{\max}, k_{\max}]$, $k_{\max} < \infty$, $k(0) \neq 0$, $k_1 \neq 0$, and is continuous at 0 and almost everywhere; **(c)** $\int_{-\infty}^{\infty} \bar{k}(x)dx < \infty$ where $\bar{k}(x) = \mathbb{I}(x \geq 0)\sup_{y \geq x}|k(y)| + \mathbb{I}(x < 0)\sup_{y \leq x}|k(y)|$; **(d)** $|K(\lambda)| \geq 0$ for all $\lambda \in \mathcal{R}$, where $K(\lambda) = (2\pi)^{-1}\int k(x)\exp(-ix\lambda)dx$.

Assumption 2.3. **(a)** q: $\mathcal{R}^{d_z} \times \Theta \longmapsto \mathcal{R}^{d_q}$ is $\mathcal{F}$-measurable for each $\theta \in \Theta$, $\Theta = \mathcal{A} \times \mathcal{B}$, $\mathcal{A}$ and $\mathcal{B}$ compact subsets of $\mathcal{R}^{d_\alpha}$ and $\mathcal{R}^{d_\beta}$ respectively; **(b)** $q_t(\cdot)$: $\Theta \longmapsto \mathcal{R}^{d_q}$ and r: $\Theta \to \mathcal{R}^{d_r}$ are continuous on Θ a.s.-$\mathcal{P}$; **(c)** $\theta_0 = (\alpha_0', \beta_0')' \in \Theta$ is the unique solution to $\mathrm{E}[q_t(\theta)] = 0$ and $r(\theta) = 0$; **(d)** $\mathrm{E}[\sup_{\theta\in\Theta}\|q_t(\theta)\|^\alpha] < \infty$ for some $\alpha > \max(4v, \frac{1}{\eta})$**;** **(e)** $q_t(\theta)$ is global Lipschitz continuous on Θ, i.e., for all $\theta, \theta^0 \in \Theta$, $\|q_t(\theta) - q_t(\theta^0)\| \leq L_t\|\theta - \theta^0\|$ a.s.-$\mathcal{P}$ and $\sup_T \mathrm{E}[\sum_{t=1}^T L_t/T] < \infty$; **(f)** $\Xi(\theta)$ is finite and p.d. for all $\theta \in \Theta$.

Assumption 2.4. **(a)** $\theta_0 \in \mathrm{int}(\Theta)$; **(b)** $q_t(\theta)$: $\Theta \longmapsto \mathcal{R}^{d_q}$ is continuously differentiable on a neighborhood $\mathcal{N}$ of θ_0, $(t = 1, 2, ...)$, and $E[\sup_{\theta\in\mathcal{N}}\|\partial q_t(\theta)/\partial\theta'\|^{\alpha/(\alpha/(1+\varepsilon)-1)}] < \infty$ for some $\varepsilon > 0$; **(c)** $r(\cdot)$: $\Theta \to \mathcal{R}^{d_r}$ is continuously differentiable in a neighbourhood $\mathcal{N}$ of θ_0 and $\sup_{\theta\in\mathcal{N}}\|\partial r(\theta)/\partial\theta'\| < \infty$; **(d)** $\mathrm{rank}(G) = d_\beta$, $\mathrm{rank}(R) = d_r$ and $\mathrm{rank}((H_\alpha', R_\alpha')') = d_\alpha$.

Remark 2.2. These assumptions are generally stronger than required for the consistency and asymptotic normality of the restricted GMM estimator $\breve{\theta}_T$ (2.4) but are imposed here to provide for a unified treatment of GMM and GMM-KBB; e.g, the non-bootstrap results allow strong mixing to be replaced by ergodicity, see Hansen (1982, Assumption 2.1, p.1032) and also Hall (2005, Assumption 3.8, p.66). Andrews (1991, Lemma 1, p.824) and Smith (2011, Assumption 2.1, p.1199) require the weaker $\sum_{j=1}^{\infty} j^2\alpha(j)^{(v-1)/v} < \infty$ implied by the mixing condition Assumption 2.1(b); see Andrews (1991, Comment, p.824). Assumption 2.4(b) is slightly stronger than $\mathrm{E}[\sup_{\theta\in\mathcal{N}}\|\partial q_t(\theta)/\partial\theta'\|^{\alpha/(\alpha-1)}] < \infty$, Smith (2011, Assumption 2.5(b), p.1200), and is required for heteroskedastic autocorrelation consistent (HAC) bootstrap estimation of Ξ; see Lemma C.3 of Appendix C. See Smith (2011, pp.1199-1201, p.1210) for further discussion of Assumptions 2.1-2.4.

For W_q p.d., $M_{W_q} = (Q'(W_q)^{-1}Q + R'R)^{-1}$ p.d. by Assumption 2.4(d). Let $H_{W_q} = K_{W_q}Q'(W_q)^{-1}$, $J_{W_q} = (RM_{W_q}R')^{-1}RM_{W_q}Q'(W_q)^{-1}$ and $K_{W_q} = M_{W_q} - M_{W_q}R'(RM_{W_q}R')^{-1}RM_{W_q}$. Also let $P_{W_q} = (W_q)^{-1} - (W_q)^{-1}QH_{W_q}$.

Proposition 2.1. (Consistency and Limiting Distribution of $\breve{\theta}_T$.) Let p.s.d. $W_{qT} \to W_q$, prob-$\mathcal{P}$,

W_q p.d. If Assumptions 2.1, 2.2 and 2.3(a)-(d)(f) are satisfied, then **(a)** $\breve{\theta}_T \to \theta_0$, $\breve{\mu}_T \to 0$, prob-$\mathcal{P}$. If, in addition, Assumption 2.4 holds, then **(b)**

$$T^{1/2}(\breve{\theta}_T - \theta_0) \to^{d_\mathcal{P}} \mathcal{N}(0, H_{W_q}\Xi H_{W_q}'), T^{1/2}\breve{\mu}_T \to^{d_\mathcal{P}} \mathcal{N}(0, J_{W_q}\Xi J_{W_q}').$$

Remark 2.3. The first order asymptotic representations for $\breve{\theta}_T$ and $\breve{\mu}_T$ are $T^{1/2}(\breve{\theta}_T - \theta_0)/k^{1/2} + H_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0, T^{1/2}\breve{\mu}_T/k^{1/2} - J_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$, respectively; cf. Smith (2001, (B.15) and (B.17), p.1229).

Remark 2.4. An efficient restricted GMM estimator of θ_0 replaces W_{qT} in (2.5) by a p.s.d. variance estimator Ξ_T for Ξ such that $\Xi_T \to \Xi$, prob-$\mathcal{P}$. The efficient restricted GMM estimator $\tilde{\theta}_T = \arg\min_{\theta \in \Theta_r} \tilde{\mathcal{Q}}_T(\theta)$ with Lagrangean function $\tilde{\mathcal{L}}_T(\theta, \mu) = \tilde{\mathcal{Q}}_T(\theta)/S_T - k\mu' r(\theta)$, Lagrange multiplier estimator $\tilde{\mu}_T$ and efficient GMM criterion $\tilde{\mathcal{Q}}_T(\theta) = \hat{q}_T(\theta)'(\Xi_T)^{-1}\hat{q}_T(\theta)$. Hence, cf. Proposition 2.1, $\tilde{\theta}_T \to \theta_0$, $\tilde{\mu}_T \to 0$, prob-$\mathcal{P}$, and $T^{1/2}(\tilde{\theta}_T - \theta_0) \to^{d_\mathcal{P}} N(0, K_\Xi)$, $T^{1/2}\tilde{\mu}_T \to^{d_\mathcal{P}} N(0, J_\Xi \Xi J_\Xi')$ are asymptotically independent. Cf. Smith (2011, Theorem 5.1, p.1210). Numerous HAC estimators Ξ_T have been proposed under various assumptions; see *inter alia* Newey and West (1987), Gallant (1987), Andrews (1991) and Ng and Perron (1996). Smith (2005, 2011) suggest an approach using the transformed moment indicator vectors $q_{tT}(\theta)$ (2.3), $(t = 1, ..., T)$. In particular, under the assumptions stated above, given an initial $T^{1/2}$-consistent estimator $\breve{\theta}_T$ for θ_0, the standard outer product and centred variance estimators $\sum_{t=1}^T q_{tT}(\breve{\theta}_T)q_{tT}(\breve{\theta}_T)'/T$, $\sum_{t=1}^T (q_{tT}(\breve{\theta}_T) - \hat{q}_T(\breve{\theta}_T))(q_{tT}(\breve{\theta}_T) - \hat{q}_T(\breve{\theta}_T))'/T \to \Xi$, prob-$\mathcal{P}$, from Assumption 2.2(a), since $(T/S_T)^{1/2}\hat{q}_T(\breve{\theta}_T) = O_p(1)$; see Smith (2005, Theorem 2.1, p.165) and Smith (2011, Lemma A.7, p.1223).[4] Other HAC consistent estimators of Ξ obtain if the empirical measure T^{-1} is replaced by implied probabilities π_{tT}, $(t = 1, ..., T)$, satisfying Assumption 3.2, see Sections 3.1 and 3.2. The scaling constants k_1 and k_2 may also be replaced by their respective sample counterparts $\hat{k}_j = \sum_{s=1-T}^{T-1} k\left(\frac{s}{S_T}\right)^j / S_T$, $(j = 1, 2)$.

Remark 2.5. In the absence of α, i.e., (2.2) $\mathrm{E}[h(z_t, \beta_0)] = 0$, $r(\beta_0) = 0$, M_{W_q} is replaced by $(Q'(W_q)^{-1}Q)^{-1}$ where now $Q = \mathrm{E}[\partial q_t(\beta_0)/\partial\beta']$. Assumption 2.4(d) $\mathrm{rank}((H_\alpha', R_\alpha')') = d_\alpha$ is no longer required.

[4] Smith (2011, Section 2.5, pp.1201-1202) establishes the first order asymptotic equivalence between $\Xi_T(\theta_0) = \sum_{t=1}^T q_{tT}(\theta_0)q_{tT}(\theta_0)'/T$ and HAC consistent estimators of Ξ based on the induced p.s.d. kernel $k^*(\cdot) = \int k(b - \cdot)k(b)db/k_2$, cf. Andrews (1991, p.822). Moreover, Parente and Smith (2018, Corollary 3.1, p.7) demonstrates the higher order optimality of $\Xi_T(\theta_0)$ with the kernel

$$k(x) = \left(\frac{5\pi}{8}\right)^{1/2} \frac{1}{x} J_1\left(\frac{6\pi x}{5}\right) \text{ if } x \neq 0 \text{ and } (\frac{5\pi}{8})^{1/2}\frac{3\pi}{5} \text{ if } x = 0,$$

which induces the optimal quadratic spectral or Bartlett-Priestley-Epanechnikov kernel

$$k_{\mathrm{QS}}^*(y) = \frac{3}{(6\pi y/5)^2}(\frac{\sin 6\pi y/5}{6\pi y/5} - \cos 6\pi y/5).$$

The Bessel function $J_\nu(\cdot)$, see Gradshteyn and Ryzhik (1980, 8.402, p.951), is given by

$$J_\nu(z) = \frac{z^\nu}{2^\nu}\sum_{k=0}^\infty (-1)^k \frac{z^{2k}}{2^{2k}k!\Gamma(\nu + k + 1)}.$$

Remark 2.6. Assumptions 2.3, 2.4 and Proposition 2.1 are easily adapted for unrestricted GMM estimation of β_0. Given W_{gT}, a (d_g, d_g) p.s.d. matrix such that $W_{gT} \to W_g$, prob-$\mathcal{P}$, W_g p.d., an unrestricted GMM estimator $\check{\beta}_T = \arg\min_{\beta \in \mathcal{B}} \check{\mathcal{Q}}_T(\beta)/k$, where the unrestricted GMM criterion $\check{\mathcal{Q}}_T(\beta) = \hat{g}_T(\beta)'(W_{gT})^{-1}\hat{g}_T(\beta)$. The matrices above Proposition 2.1 relevant for unrestricted GMM estimation of β_0 are $K_{W_g} = M_{W_g} = (G'(W_g)^{-1}G)^{-1}$, $H_{W_g} = K_{W_g}G'(W_g)^{-1}$, $J_{W_g} = 0$ and $P_{W_g} = (W_g)^{-1} - (W_g)^{-1}H_{W_g}$. Thus, $\check{\beta}_T \to \beta_0$, prob-$\mathcal{P}$, $T^{1/2}(\check{\beta}_T - \beta_0) \to^{d_{\mathcal{P}}} \mathcal{N}(0, H_{W_g}\Sigma H'_{Wg})$, cf. Proposition 2.1(a)(b), and $T^{1/2}(\check{\beta}_T - \beta_0)/k^{1/2} + H_{W_g}(T/S_T)^{1/2}\hat{g}_T(\theta_0) \to 0$, prob-$\mathcal{P}$, cf. Remark 2.3. Cf. Hansen (1982, Theorems 2.1, p.1035, and 3.1, p.1042) and Hall (2005, Theorems 3.1, p.68, and 3.2, p. 71). Substituting p.s.d. $\Sigma_T \to \Sigma$, prob-$\mathcal{P}$, for W_{gT}, the efficient unrestricted GMM estimator $\hat{\beta}_T \to \beta_0$, prob-$\mathcal{P}$, $T^{1/2}(\hat{\beta}_T - \beta_0) \to^{d_{\mathcal{P}}} \mathcal{N}(0, (G'\Sigma^{-1}G)^{-1})$ and $T^{1/2}(\hat{\beta}_T - \beta_0)/k^{1/2} + H_\Sigma(T/S_T)^{1/2}\hat{g}_T(\theta_0) \to 0$, prob-$\mathcal{P}$, where $\hat{\beta}_T = \arg\min_{\beta \in \mathcal{B}} \hat{\mathcal{Q}}_T(\beta)/k$ with efficient unrestricted GMM criterion $\hat{\mathcal{Q}}_T(\beta) = \hat{g}_T(\beta)'(\Sigma_T)^{-1}\hat{g}_T(\beta)$. Smith (2005, Theorems 3.1, p.165, and 3.2, p.166) demonstrates that two-step, CUE and iterated GMM estimators based on $\hat{\mathcal{Q}}_T(\beta)$ are asymptotically equivalent to the optimal GMM estimator $\hat{\beta}_T$.

2.2 GMM Inference

2.2.1 Overidentification Tests

To test the validity of the maintained over-identifying moment conditions (2.1), Hansen (1982) proposed the $\mathcal{J}$-statistic, *viz.*

$$\mathcal{J}_T = (T/S_T)\hat{g}_T(\hat{\beta}_T)'(\Sigma_T)^{-1}\hat{g}_T(\hat{\beta}_T),$$

where $\Sigma_T \to \Sigma$, prob-$\mathcal{P}$, e.g., $\Sigma_T = \sum_{t=1}^T g_{tT}(\check{\beta}_T)g_{tT}(\check{\beta}_T)'/T$, cf. Remarks 2.4 and 2.6.

Proposition 2.2. Under Assumptions 2.1, 2.2, 2.3(a)-(d)(f) and 2.4, if $\Sigma_T \to \Sigma$, prob-$\mathcal{P}$, and $m > p$, $\mathcal{J}_T \to^{d_{\mathcal{P}}} \chi^2(d_g - d_\beta)$.

Cf. Hansen (1982, Lemma 4.2, pp.1049-1050).

2.2.2 Specification Tests

Smith (2011, Section 5, pp.1209-1213) proposes a number of GEL classical-like statistics to test the additional moment conditions and parametric restrictions (2.2). Correspondingly, we consider a non-negative likelihood ratio-like statistic,

$$\mathcal{LR}_T/k = (T/S_T)(\hat{q}_T(\tilde{\theta}_T) - \Xi_T S_g(\Sigma_T)^{-1}\hat{g}_T(\hat{\beta}_T))'(\Xi_T)^{-1}(\hat{q}_T(\tilde{\theta}_T) - \Xi_T S_g(\Sigma_T)^{-1}\hat{g}_T(\hat{\beta}_T)), \qquad (2.6)$$

where $S_g = (I_{d_g}, 0)'$ is a (d_q, d_g) selection matrix, i.e., $S'_g\hat{q}_T(\theta) = \hat{g}_T(\beta)$, cf. Smith (2011, (5.6), p.1212); and the more common, but potentially negative valued, difference statistic as the normalised difference of GMM criteria

$$\mathcal{D}_T/k = (T/S_T)(\hat{q}_T(\tilde{\theta}_T)'(\Xi_T))^{-1}\hat{q}_T(\tilde{\theta}_T) - \hat{g}_T(\hat{\beta}_T)'(\Sigma_T)^{-1}\hat{g}_T(\hat{\beta}_T)), \qquad (2.7)$$

cf. Newey (1985) and Smith (2011, (5.6), p.1212), where Σ_T is the (d_g, d_g) top left diagonal block of p.s.d. Ξ_T, $\Xi_T \to \Xi$, prob-$\mathcal{P}$; a simplified score-like statistic

$$\mathcal{S}_T/k = (T/S_T)\hat{q}_T(\tilde{\theta}_T)'((\Xi_T))^{-1} - S_g P_{\Sigma T} S_g')\hat{q}_T(\tilde{\theta}_T), \tag{2.8}$$

where $P_{\Sigma T} \to P_\Sigma$, prob-$\mathcal{P}$, cf. Remark 2.6, cf. Smith (2011, p.1213); a Lagrange multiplier-like statistic, cf. the GEL LM-type statistic $\mathcal{LM}_a$, Smith (2011, (5.4), p.1211),

$$\mathcal{LM}_T/k = T(\hat{q}_T(\tilde{\theta}_T)'(\Xi_T)^{-1}/S_T^{1/2}, \tilde{\mu}_T'/k^{1/2})S_{h,\mu}(S_{h,\mu}'\Psi_{\Xi T}S_{h,\mu})^{-1}S_{h,\mu}'(\hat{q}_T(\tilde{\theta}_T)'(\Xi_T)^{-1}/S_T^{1/2}, \tilde{\mu}_T'/k^{1/2})', \tag{2.9}$$

where the $(d_q + d_r, d_h + d_r)$ selection matrix $S_{h,\mu} = \mathtt{diag}(S_h, I_{d_r})$, $S_h = (I_{d_h}, 0)'$ is a (d_q, d_h) selection matrix, i.e., $S_h'\hat{q}_T(\cdot) = \hat{h}_T(\cdot)$; a generalised Wald-type statistic

$$\mathcal{GW}_T/k = T(\breve{q}_T(\hat{\theta}_T)'/S_T^{1/2}, r(\hat{\theta}_T)'/k^{1/2})\Psi_{\Xi T}(\breve{q}_T(\hat{\theta}_T)'S_T^{1/2}, r(\hat{\theta}_T)'/k^{1/2})', \tag{2.10}$$

where $\breve{q}_T(\cdot) = (I_{d_q} - \Xi_T S_g(\Sigma_T)^{-1}S_g')\hat{q}_T(\cdot)$, cf. Smith (2011, (5.5), p.1211), and the unrestricted estimator $\hat{\theta}_T = (\tilde{\alpha}_T', \hat{\beta}_T')'$ defined to circumvent the lack of identification of α under (2.1). Note that the moment vector $\breve{q}_T(\cdot)$ forming $\mathcal{GW}_T$ may be expressed as $(0', (\hat{h}_T(\cdot) - (\Xi_T)_{hg}(\Sigma_T)^{-1}\hat{g}_T(\cdot))')'$. An alternative form for the averaged moment indicator $\breve{q}_T(\cdot)$ is the implied probability weighted average $\hat{q}_T^\pi(\cdot) = \sum_{t=1}^T \pi_{tT} q_{tT}(\cdot)$ below (3.4), where the implied probabilities π_{tT}, $(t = 1, ..., T)$, satisfy Assumption 3.2; see Sections 3.1 and 3.2.

The statistics (2.6), (2.8), (2.9) and (2.10) require consistent estimators Ξ_T of Ξ and $\Psi_{\Xi T}$ of

$$\Psi_\Xi = \begin{pmatrix} P_\Xi & -J_\Xi' \\ -J_\Xi & J_\Xi \Xi J_\Xi' \end{pmatrix},$$

where P_Ξ, K_Ξ, J_Ξ and M_Ξ are defined above Proposition 2.1, together with consistent estimators of Q and R; note $J_\Xi \Xi J_\Xi' = (RM_\Xi R')^{-1} - I_{d_r}$. E.g., with $T^{1/2}$-consistent restricted $\breve{\theta}_T$ and unrestricted $\check{\theta}_T = (\check{\alpha}_T', \check{\beta}_T')'$ estimators of θ_0, $\Xi_T = \sum_{t=1}^T q_{tT}(\breve{\theta}_T)q_{tT}(\breve{\theta}_T)'/T$ or $\sum_{t=1}^T q_{tT}(\check{\theta}_T)q_{tT}(\check{\theta}_T)'/T$, cf. Remark 2.4, $Q_T = \sum_{t=1}^T Q_{tT}(\breve{\theta}_T)/T$ or $\sum_{t=1}^T Q_{tT}(\check{\theta}_T)/T$, where $Q_{tT}(\theta) = \partial q_{tT}(\theta)/\partial\theta'$, $(t = 1, ..., T)$, cf. Smith (2011, p.1201), and $R_T = R(\breve{\theta}_T)$ or $R(\check{\theta}_T)$, where $R(\theta) = \partial r(\theta)/\partial\theta'$.

Proposition 2.3. Let p.s.d. $\Xi_T \to \Xi$ and p.s.d. $\Psi_{\Xi T} \to \Psi_\Xi$, prob-$\mathcal{P}$. If Assumptions 2.1, 2.2, 2.3(a)-(d)(f) and 2.4 are satisfied, then

$$\mathcal{T}_T \to^{d_\mathcal{P}} \chi^2(d_q - d_g + d_r - d_\alpha),$$

where $\mathcal{T} = \mathcal{LR}$, $\mathcal{D}$, $\mathcal{S}$, $\mathcal{LM}$ or $\mathcal{GW}$. Moreover, $\mathcal{LR}_T$, $\mathcal{D}_T$, $\mathcal{S}_T$, $\mathcal{LM}_T$ and $\mathcal{GW}_T$ are asymptotically equivalent.

Cf. Smith (2011, Theorem 5.2, p.1212).

Remark 2.7. In the absence of α, i.e., $\mathrm{E}[h(z_t, \beta_0)] = 0$, $r(\beta_0) = 0$, $\mathcal{GW}_T$ is now based solely on the unrestricted GMM estimator $\hat{\beta}_T$. The moment function $\breve{h}_T(\hat{\beta}_T) = \hat{h}_T(\hat{\beta}_T) - (\Xi_T)_{hg}(\Sigma_T)^{-1}\hat{g}_T(\hat{\beta}_T)$ is a GMM-efficient estimator of $\mathrm{E}[h(z_t, \beta_0)]$ under the maintained hypothesis (2.1) and is asymptotically equivalent to $\sum_{t=1}^{T} \hat{\pi}_{tT} h_{tT}(\hat{\beta}_T)$, see Smith (2011, (3.2), p.1205), where $\hat{\pi}_{tT}$, $(t = 1, ..., T)$, are the efficient unrestricted GEL implied probabilities defined in (3.3) below; cf. Ruud (2000, (22.23), p.575). Cf. Remarks 2.5 and 2.6.

Remark 2.8. See *inter alia* Newey and West (1987) for GMM tests of parametric restrictions $r(\beta_0) = 0$ maintaining (2.1).[5] Newey (1985), Eichenbaum et al. (1988) and Ruud (2000) detail GMM tests of additional moment restrictions $\mathrm{E}[h(z_t, \beta)] = 0$, cf. Remark 2.5; also see Smith (1997, Section II.3, pp.513-514) and Smith (2011, Section 5, pp.1211-1213) for GEL-based tests.[6]

3 GEL Implied Probability GMM-KBB

Let $\mathcal{P}^*_\omega$ denote the bootstrap probability measure conditional on the observational data $\{z_t\}_{t=1}^T$ with E^* and var^* the corresponding conditional expectation and variance respectively.

Let $m_T = [T/S_T]$ denote the integer part of T/S_T. The indices t_s^* and the consequent bootstrap sample $q^{\pi}_{t_s^* T}(\theta)$, $(s = 1, ..., m_T)$, denote m_T independent draws with replacement from the index set $\mathcal{T}_T = \{1, ..., T\}$ and the bootstrap sample space $\{q_{tT}(\theta)\}_{t=1}^T$ with bootstrap sampling probabilities $\mathcal{P}^*_\omega(q^{\pi}_{t_s^* T}(\theta) = q_{tT}(\theta)) = \pi_{tT}$, $(t = 1, ..., T)$. The probabilities π_{tT} can depend on the data and satisfy $0 \leq \pi_{tT} \leq 1$ w.p.a.1, $\sum_{t=1}^T \pi_{tT} = 1$ and $\max_{1 \leq t \leq T} |T\pi_{tT} - 1| \to 0$, prob-$\mathcal{P}$; see Assumption 3.2 below. The identically sampled KBB method of Parente and Smith (2018, 2021) sets the bootstrap sampling probabilities as the standard empirical GMM measures, i.e., $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$.

3.1 GEL Implied Probabilities

Of particular interest here are the bootstrap sampling probabilities π_{tT}, $(t = 1, ..., T)$, generated from GEL criteria. The relevant GEL criteria corresponding to (2.1) and (2.2) are $\hat{\mathcal{P}}^{\rho}_{gT}(\beta, \lambda_g) = \sum_{t=1}^T \rho(\lambda_g' g_{tT}(\beta)/k^{1/2})/T$ and $\hat{\mathcal{P}}^{\rho}_{qT}(\theta, \lambda_q) = \sum_{t=1}^T \rho(\lambda_q' q_{tT}(\theta)/k^{1/2})/T$ respectively, where $\rho(\cdot)$ denotes a concave function normalised so that $\rho(0) = 0$ with derivatives $\rho_j(\cdot) = \partial^j \rho(\cdot)/\partial v^j$, $(j = 0, 1, 2, ...)$; we set, without loss of generality, $\rho_j(0) = -1$, $(j = 1, 2)$. Given unrestricted, $\breve{\beta}_T$, see Remark 2.6, and restricted

[5] $\mathcal{LR}_T$ (2.6) and $\mathcal{D}_T$ (2.7) set $S_g = I_{d_g}$ and Ξ_T as Σ_T with $\hat{q}_T(\tilde{\theta}_T)$ replaced by $\hat{g}_T(\tilde{\beta}_T)$ where $\tilde{\beta}_T$ is the restricted estimator for β_0. $\mathcal{S}_T$ (2.8) is similarly re-defined as $\mathcal{S}_T/k = (T/S_T)\hat{g}_T(\tilde{\beta}_T)'(\Sigma_T)^{-1}G_T M_{\Sigma T} G_T'(\Sigma_T)^{-1}\hat{g}_T(\tilde{\beta}_T)$ with $G_T \to G$, $M_{\Sigma T} \to M_\Sigma$, prob-$\mathcal{P}$. $\mathcal{LM}_T$ (2.9) sets $S_{h,\mu} = \mathtt{diag}(0, I_{d_r})$ and, hence, $\mathcal{LM}_T = T\tilde{\mu}_T' R_T M_{\Sigma T} R_T' \tilde{\mu}_T$ with $\tilde{\mu}_T$ the Lagrange multiplier estimator associated with the parametric restrictions $r(\beta_0) = 0$; note $G_T'(\Sigma_T)^{-1}(T/S_T)^{/12}\hat{g}_T(\tilde{\beta}_T) - R_T' T^{1/2}\tilde{\mu}_T \to 0$, prob-$\mathcal{P}$, from the first order conditions. $\mathcal{GW}_T$ (2.10) sets $\breve{q}_T(\hat{\theta}_T)$ as $\breve{g}_T(\hat{\beta}_T)$, i.e., 0, and is, thus, rendered as $\mathcal{GW}_T = Tr(\hat{\theta}_T)'(R_T M_{\Sigma T} R_T')^{-1} r(\hat{\theta}_T)$.

[6] $\mathcal{LR}_T$ (2.6) and $\mathcal{D}_T$ (2.7) replace $\hat{q}_T(\tilde{\theta}_T)$ by $\hat{q}_T(\tilde{\beta}_T)$ where $\tilde{\beta}_T$ is the restricted estimator for β_0. $\mathcal{S}_T$ (2.8) is similarly re-defined. $\mathcal{LM}_T$ (2.9) sets $S_{h,\mu} = \mathtt{diag}(S_h, 0)$ and $R = 0$ eliminating $\tilde{\mu}_T$, i.e., thus, $\mathcal{LM}_T/k = (T/S_T)\hat{q}_T(\tilde{\beta}_T)'(\Xi_T)^{-1}S_h(S_h' P_{\Xi T} S_h)^{-1}S_h'(\Xi_T)^{-1}\hat{q}_T(\tilde{\beta}_T)$. $\mathcal{GW}_T$ (2.10) sets $\breve{q}_T(\hat{\theta}_T)$ as $\breve{q}_T(\hat{\beta}_T)$ and $r(\hat{\theta}_T) = 0$, i.e., $\mathcal{GW}_T/k = (T/S_T)\breve{h}_T(\hat{\beta}_T)' S_h' P_{\Xi T} S_h \breve{h}_T(\hat{\beta}_T)$.

$\breve{\theta}_T$ (2.4) GMM estimators, the corresponding GEL estimators of λ_o, $(o = g, q)$, are

$$\breve{\lambda}_{gT} = \arg\sup_{\lambda_g \in \Lambda_{gT}} \hat{\mathcal{P}}^{\rho}_{gT}(\breve{\beta}_T, \lambda_g), \breve{\lambda}_{qT} = \arg\sup_{\lambda_q \in \Lambda_{qT}} \hat{\mathcal{P}}^{\rho}_{qT}(\breve{\theta}_T, \lambda_q), \tag{3.1}$$

where the parameter spaces Λ_{oT}, $(o = g, q)$, are defined in Assumption 3.1(b) below.

Assumption 3.1. **(a)** $\rho(\cdot)$ is twice continuously differentiable and concave on its domain, an open interval $\mathcal{V}$ containing 0, $\rho_1(0) = \rho_2(0) = -1$; **(b)** $\lambda_o \in \Lambda_{oT}$ where $\Lambda_{oT} = \{\lambda_o : \|\lambda_o\| \leq D(T/S_T)^{-\zeta}\}$, $(o = g, q)$, for some $D > 0$ with $\frac{1}{2\alpha\eta} < \zeta < \frac{1}{2}$.

Cf. Smith (2011, Assumption 2.4, p.1200).

Proposition 3.1. If Assumptions 2.1, 2.2, 2.3(a)–(d)(f) and 3.1 hold, **(a)** $\breve{\lambda}_{qT} \to 0$, prob-$\mathcal{P}$, and, **(b)** if in addition Assumption 2.4 holds, $(T/S_T)^{1/2}\breve{\lambda}_{qT} \to^{d_{\mathcal{P}}} \mathcal{N}(0, \Xi^{-1}(I_{d_q} - QH_{W_q})\Xi(I_{d_q} - QH_{W_q})'\Xi^{-1})$.

Remark 3.1. The first order asymptotic representation for $\breve{\lambda}_{qT}$ is $(T/S_T)^{1/2}\breve{\lambda}_{qT}/k^{1/2} + \Xi^{-1}W_q P_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$; see Proposition B.1. With the efficient restricted GMM estimator $\tilde{\theta}_T$, see Remark 2.4, substituted in the requisite GEL criterion, i.e., $\hat{P}^{\rho}_{qT}(\tilde{\theta}_T, \lambda_q)$, the resultant GEL estimator $(T/S_T)^{1/2}\tilde{\lambda}_{qT} \to^{d_{\mathcal{P}}} \mathcal{N}(0, P_{\Xi})$ which is asymptotically distributed independently of $T^{1/2}(\tilde{\theta}_T - \theta_0)$, prob-$\mathcal{P}$; cf. Smith (2011, Theorem 5.1, p.1210). A similar result to Proposition 3.1 obtains with the unrestricted GEL criterion $\hat{\mathcal{P}}^{\rho}_{gT}(\breve{\beta}_T, \lambda_g)$, *viz.*, $\breve{\lambda}_{gT} \to 0$, prob-$\mathcal{P}$, $(T/S_T)^{1/2}\breve{\lambda}_{gT} \to^{d_{\mathcal{P}}} \mathcal{N}(0, \Sigma^{-1}(I_{d_g} - GH_{W_g})\Sigma(I_{d_g} - GH_{W_g})'\Sigma^{-1})$ with $(T/S_T)^{1/2}\breve{\lambda}_{gT}/k^{1/2} + \Sigma^{-1}W_g P_{W_g}(T/S_T)^{1/2}\hat{g}_T(\beta_0) \to 0$, prob-$\mathcal{P}$, see Remark B.1. When the efficient unrestricted GMM estimator $\hat{\beta}_T$ is substituted for $\breve{\beta}_T$, $(T/S_T)^{1/2}\hat{\lambda}_{gT} \to \mathcal{N}(0, P_{\Sigma})$ and is asymptotically distributed independently of $T^{1/2}(\hat{\beta}_T - \beta_0)$, prob-$\mathcal{P}$; cf. Smith (2011, Theorem 2.3, p.1201).

We define unrestricted (2.1) and restricted (2.2) GEL implied probabilities

$$\breve{\pi}_{tT} = \frac{\rho_1(\breve{\lambda}'_{gT} g_{tT}(\breve{\beta}_T)/k^{1/2})}{\sum_{s=1}^{T} \rho_1(\breve{\lambda}'_{gT} g_{sT}(\breve{\beta}_T)/k^{1/2})}, \breve{\pi}_{tT} = \frac{\rho_1(\breve{\lambda}'_{qT} q_{tT}(\breve{\theta}_T)/k^{1/2})}{\sum_{s=1}^{T} \rho_1(\breve{\lambda}'_{qT} q_{sT}(\breve{\theta}_T)/k^{1/2})}, \tag{3.2}$$

$(t = 1, ..., T)$, respectively; cf. Smith (2011, (3.1), p.1205).

Remark 3.2. Particular examples are : EL $\breve{\pi}_{tT} = 1/T(1 + \breve{\lambda}'_{gT} g_{tT}(\breve{\beta}_T)/k^{1/2})$, $\breve{\pi}_{tT} = 1/T(1 + \breve{\lambda}'_{qT} q_{tT}(\breve{\theta}_T)/k^{1/2})$, cf. Owen (1988); ET: $\breve{\pi}_{tT} = \exp(\breve{\lambda}'_{gT} g_{tT}(\breve{\beta}_T)/k^{1/2})/\sum_{s=1}^{T}\exp(\breve{\lambda}'_{gT} g_{sT}(\breve{\beta}_T)/k^{1/2})$, $\breve{\pi}_{tT} = \exp(\breve{\lambda}'_{qT} q_{tT}(\breve{\theta}_T)/k^{1/2})/\sum_{s=1}^{T}\exp(\breve{\lambda}'_{qT} q_{sT}(\breve{\theta}_T)/k^{1/2})$, cf. Kitamura and Stutzer (1997); CUE $\breve{\pi}_{tT} = (1 + \breve{\lambda}'_{gT} g_{tT}(\breve{\beta}_T)/k^{1/2})/\sum_{s=1}^{T}(1 + \breve{\lambda}'_{gT} g_{sT}(\breve{\beta}_T)/k^{1/2})$, $\breve{\pi}_{tT} = (1 + \breve{\lambda}'_{qT} q_{tT}(\breve{\theta}_T)/k^{1/2})/\sum_{s=1}^{T}(1 + \breve{\lambda}'_{qT} q_{sT}(\breve{\theta}_T)/k^{1/2})$, cf. Back and Brown (1993). See also Brown and Newey (1992, 2002) and Newey and Smith (2004) for the general `i.i.d.` case. Members of the Cressie-Read (1984) family of discrepancies also have a dual counterpart in the GEL class; see Newey and Smith (2004, Theorem 2.2, p.224). The ratios $\breve{\pi}_{tT}$ and $\breve{\pi}_{tT}$, $(t = 1, ..., T)$, sum to unity, are bounded between zero and unity w.p.a.1, prob-$\mathcal{P}$, and satisfy Assumption 3.2 below; see Lemma B.1.[7] GEL implied probabilities induce empirical measure

[7] Implied probabilities π_{tT}, $(t = 1, ..., T)$, may fail to be non-negative in finite samples. Antoine et al. (2007, (2.8), (2.9), p.466) provide a remedy without affecting the analysis by defining appropriate shrinkage estimators. E.g., replace $\breve{\pi}_{tT}$ by $(\breve{\pi}_{tT} + T^{-1}\breve{\varepsilon}_T)/(1 + \breve{\varepsilon}_T)$, $(t = 1, ..., T)$, where $\breve{\varepsilon}_T = -T\min[\min_{1\leq t\leq T}\breve{\pi}_{tT}, 0]$.

counterparts to the expectation operator in (2.1) and (2.2) ensuring that the moment conditions are satisfied in the sample, i.e., $\sum_{t=1}^{T}\breve{\pi}_{tT}g_{tT}(\breve{\beta}_T)=0$, $\sum_{t=1}^{T}\breve{\pi}_{tT}q_{tT}(\breve{\theta}_T)=0$.

Of particular interest are the efficient unrestricted (2.1) and restricted (2.2) GEL implied probabilities

$$\hat{\pi}_{tT}=\frac{\rho_1(\hat{\lambda}_{gT}'g_{tT}(\hat{\beta}_T)/k^{1/2})}{\sum_{s=1}^{T}\rho_1(\hat{\lambda}_{gT}'g_{sT}(\hat{\beta}_T)/k^{1/2})},\tilde{\pi}_{tT}=\frac{\rho_1(\tilde{\lambda}_{qT}'q_{tT}(\tilde{\theta}_T)/k^{1/2})}{\sum_{s=1}^{T}\rho_1(\tilde{\lambda}_{qT}'q_{sT}(\tilde{\theta}_T)/k^{1/2})}, \tag{3.3}$$

$(t=1,...,T)$, where the efficient GMM estimators $\hat{\beta}_T$ and $\tilde{\theta}_T$ are substituted in the respective GEL criteria above; see (3.1).

Remark 3.3. Moment estimators with the efficient implied probabilities (3.3) substituting for the standard empirical measure are optimal; see, respectively, Brown and Newey (1998, Corollary 1, p.458) and Smith (2011, p.1206) for the `i.i.d.` and weakly dependent contexts.

3.2 GEL Implied Probability GMM Estimation

The restricted GEL implied probability (π-GEL) GMM estimator $\breve{\theta}_T^{\pi}$ and associated GMM criterion play important roles in the analysis elucidated below. In particular, the restricted π-GEL GMM estimator provides the appropriate centring for both inefficient and efficient restricted GMM bootstrap estimators under the KBB sampling schemes considered in this paper. This π-GEL GMM estimator minimises a corresponding GMM criterion which, rather than using the standard sample average moment indicator $\hat{q}_T(\cdot)$, replaces the standard GMM empirical measure T^{-1} by the bootstrap mass point probabilities π_{tT}, $(t=1,...,T)$, as weights. The restricted π-GEL GMM estimator is asymptotically first order equivalent to and, thus, its empirical distribution provides an alternative, to that of the corresponding GMM estimator.

Redefine the GMM criterion $\breve{\mathcal{Q}}_T(\theta)$ (2.5) as the π-GEL GMM criterion

$$\breve{\mathcal{Q}}_T^{\pi}(\theta)=\hat{q}_T^{\pi}(\theta)'(W_{qT})^{-1}\hat{q}_T^{\pi}(\theta) \tag{3.4}$$

with the GEL implied probability-weighted sample moment $\hat{q}_T^{\pi}(\theta)=\sum_{t=1}^{T}\pi_{tT}q_{tT}(\theta)$ replacing the empirical sample moment $\hat{q}_T(\theta)$. The restricted π-GEL GMM estimator $\breve{\theta}_T^{\pi}$ is then defined by

$$\breve{\theta}_T^{\pi}=\arg\min_{\theta\in\Theta_r}\breve{\mathcal{Q}}_T^{\pi}(\theta), \tag{3.5}$$

with the corresponding Lagrange multiplier estimator $\breve{\mu}_T^{\pi}$ of μ associated with the parametric constraint $r(\theta)=0$ in the π-GEL GMM Lagrangean $\breve{\mathcal{L}}_T^{\pi}(\theta)=\breve{\mathcal{Q}}_T^{\pi}(\theta)/S_T-2\mu'r(\theta)/k$.

Assumption 3.2. **(a)** $0\leq\pi_{tT}\leq 1$ w.p.a.1, $\sum_{t=1}^{T}\pi_{tT}=1$; **(b)** $\max_{1\leq t\leq T}|T\pi_{tT}-1|\rightarrow 0$, prob-$\mathcal{P}$.

Proposition 3.2. Let $W_{qT}\rightarrow W_q$, prob-$\mathcal{P}$, W_q p.d. and W_{qT} p.s.d. Then, under Assumptions 2.1, 2.2(b)-(d), 2.3(a)-(d) and 3.2, if $S_T\rightarrow\infty$ and $S_T=o(T^{1/2})$, **(a)** $\breve{\theta}_T^{\pi}-\theta_0\rightarrow 0$, $\breve{\mu}_T^{\pi}\rightarrow 0$, prob-$\mathcal{P}$.

If, in addition, Assumptions 2.2(a) and 2.4 hold and if $(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) = O_p(1)$, then **(b)** $T^{1/2}(\breve{\theta}_T^{\pi} - \theta_0)/k^{1/2} + H_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) \to 0$, prob-$\mathcal{P}$, and $T^{1/2}\breve{\mu}_T^{\pi}/k^{1/2} - J_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) \to 0$, prob-$\mathcal{P}$.

Cf. Remark 2.3.

Corollary 3.1. Let $\pi_{tT} = \breve{\pi}_{tT}$ (3.2), $(t = 1, ..., T)$. Also let $W_{qT} \to W_q$, prob-$\mathcal{P}$, W_{qT} p.s.d. and W_q p.d. Then, under Assumptions 2.1, 2.2, 2.3(a)-(d)(f), 2.4 and 3.1, $T^{1/2}(\breve{\theta}_T^{\breve{\pi}} - \breve{\theta}_T) \to 0$ and $T^{1/2}\breve{\mu}_T^{\breve{\pi}} \to 0$, prob-$\mathcal{P}$.

Remark 3.4. As noted in Remark 3.2 the restricted GEL implied probabilities $\breve{\pi}_{tT}$, $(t = 1, ..., T)$, satisfy Assumption 3.2. Corollary 3.1 follows directly from Proposition 3.2 since $(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\theta_0) - QH_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$; see Corollary B.1 with $\sigma = 1$. Consequently, the first order asymptotic properties of the $\breve{\pi}$-GEL GMM estimator $\breve{\theta}_T^{\breve{\pi}}$ (3.5), being equivalent to the restricted GMM estimator $\breve{\theta}_T$, are invariant to choice of limiting weighting matrix W_q used in the π-GEL GMM criterion $\breve{\mathcal{Q}}_T^{\pi}(\theta)$ (3.7). With the efficient restricted GEL implied probabilities $\pi_{tT} = \tilde{\pi}_{tT}$ (3.3), $(t = 1, ..., T)$, $(T/S_T)^{1/2}\hat{q}_T^{\tilde{\pi}}(\theta_0) - QH_{\Xi}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$. Thus, from Proposition 3.2, since $H_{W_q}QH_{\Xi} = H_{\Xi}$, Section B.2(b)(c), and $J_{W_q}QH_{\Xi} = 0$, Section B.2(d), the restricted $\tilde{\pi}$-GEL GMM estimator $\breve{\theta}_T^{\tilde{\pi}}$ is first order asymptotically equivalent to the efficient two-step restricted GMM estimator $\tilde{\theta}_T$, i.e., $T^{1/2}(\breve{\theta}_T^{\tilde{\pi}} - \tilde{\theta}_T) \to 0$, and $T^{1/2}\breve{\mu}_T^{\tilde{\pi}} \to 0$, prob-$\mathcal{P}$, a result again asymptotically invariant to the choice of weighting matrix W_q.

Remark 3.5. The unrestricted π-GEL GMM estimator $\breve{\beta}_T^{\pi} = \arg\min_{\beta \in \mathcal{B}} \breve{\mathcal{Q}}_T^{\pi}(\beta)$ where the π-GEL GMM criterion $\breve{\mathcal{Q}}_T^{\pi}(\beta) = \hat{g}_T^{\pi}(\beta)'(W_{gT})^{-1}\hat{g}_T^{\pi}(\beta)$, W_{gT} p.s.d., with the π-weighted sample average $\hat{g}_T^{\pi}(\beta) = \sum_{t=1}^{T}\pi_{tT}g_{tT}(\beta)$ replacing $\hat{g}_T(\beta)$ in the GMM criterion $\breve{\mathcal{Q}}_T(\beta)$, see Remark 2.6. If $W_{gT} \to W_g$, prob-$\mathcal{P}$, W_g p.d., $T^{1/2}(\breve{\beta}_T^{\pi} - \beta_0)/k^{1/2} + H_{W_g}(T/S_T)^{1/2}\hat{g}_T^{\pi}(\beta_0) \to 0$, prob-$\mathcal{P}$, cf. Proposition 3.2. Moreover, if $\pi_{tT} = \breve{\pi}_{tT}$ (3.2), $(t = 1, ..., T)$, since $(T/S_T)^{1/2}\hat{g}_T^{\breve{\pi}}(\beta_0) - GH_{W_g}(T/S_T)^{1/2}\hat{g}_T(\beta_0) \to 0$, prob-$\mathcal{P}$, by Remark B.6(b), $T^{1/2}(\breve{\beta}_T^{\breve{\pi}} - \breve{\beta}_T) \to 0$, prob-$\mathcal{P}$, cf. Corollary 3.1, invariant to the choice of metric W_g, cf. Remark 3.4. Similarly, the $\hat{\pi}$-GEL GMM estimator $\breve{\beta}_T^{\hat{\pi}}$ is first order asymptotically equivalent to the efficient two-step GMM estimator $\hat{\beta}_T$, i.e., $T^{1/2}(\breve{\beta}_T^{\hat{\pi}} - \hat{\beta}_T) \to 0$, prob-$\mathcal{P}$. With efficient restricted implied probabilities $\tilde{\pi}_{tT}$, $(t = 1, ..., T)$, (3.3), since $H_{\Sigma}(T/S_T)^{1/2}\hat{g}_T^{\tilde{\pi}}(\beta_0) - S_gH_{\Xi}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$, by Lemma B.2(b), $T^{1/2}(\hat{\beta}_T^{\tilde{\pi}} - \tilde{\beta}_T) \to 0$, prob-$\mathcal{P}$, i.e., the $\tilde{\pi}$-GEL GMM estimator $\hat{\beta}_T^{\tilde{\pi}}$ with efficient unrestricted metric $(\Sigma_T)^{-1}$ is first order asymptotically equivalent to the efficient restricted two-step GMM estimator $\tilde{\beta}_T$.

Remark 3.6. The discussion of Remark 3.4 may easily be adapted for the unrestricted GEL implied probabilities $\breve{\pi}_{tT}$, $(t = 1, ..., T)$. The first order asymptotic properties of the $\breve{\pi}$-GEL GMM estimator $\tilde{\theta}_T^{\breve{\pi}}$ from the π-GEL criterion $\breve{\mathcal{Q}}_T^{\pi}(\theta)$ (3.4) with efficient restricted metric $(\Xi_T)^{-1}$ follow from the first order conditions $H_{\Xi}(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\breve{\theta}_T^{\breve{\pi}}) \to 0$, prob-$\mathcal{P}$. Now, by Remark B.6(b), $(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\theta_0) - ((T/S_T)^{1/2}\hat{q}_T(\theta_0) - \Xi S_g\Sigma^{-1}W_gP_{W_g}S_g'(T/S_T)^{1/2}\hat{q}_T(\theta_0)) \to 0$, prob-$\mathcal{P}$, yielding $T^{1/2}(\tilde{\theta}_T^{\breve{\pi}} - \tilde{\theta}_T) - K_{\Xi}S_g(M_{\Sigma})^{-1}T^{1/2}(\hat{\beta}_T - \breve{\beta}_T) \to 0$, prob-$\mathcal{P}$, cf. Remark 2.6. Thus, with efficient unrestricted probabilities $\hat{\pi}_{tT}$, $(t = 1, ..., T)$,

$T^{1/2}(\tilde{\theta}_T^{\breve{\pi}} - \tilde{\theta}_T) \to 0$, prob-$\mathcal{P}$, i.e., $\tilde{\theta}_T^{\breve{\pi}} - \tilde{\theta}_T$ is first order equivalent to the efficient restricted estimator $\tilde{\theta}_T$.

3.3 GEL Implied Probability GMM-KBB Estimation

We adopt the notation of Gonçalves and White (2004). For a bootstrap statistic W_T^* we write $W_T^* \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, if for any $\varepsilon > 0$ and any $\delta > 0$, $\lim_{T\to\infty} \mathcal{P}(\mathcal{P}_\omega^*(|W_T^*| > \varepsilon) > \delta) = 0$.

Let $W_{qm_T}^{\pi*} - W_{qT} \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, $W_{qm_T}^{\pi*}$ p.s.d. The restricted π-GEL GMM-KBB estimator $\breve{\theta}_{m_T}^{\pi*}$ is defined by

$$\breve{\theta}_{m_T}^{\pi*} = \underset{\theta\in\Theta_r}{\arg\min}\, \breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta), \tag{3.6}$$

with π-GEL GMM-KBB Lagrangean $\breve{\mathcal{L}}_{m_T}^{\pi*}(\theta) = \breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta)/S_T - 2\mu' r(\theta)/k$, estimator $\breve{\mu}_{m_T}^{\pi*}$ of μ, the Lagrange multiplier associated with the parametric constraint $r(\theta) = 0$, and π-GEL GMM-KBB criterion

$$\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta) = \hat{q}_{m_T}^{\pi*}(\theta)'(W_{qm_T}^{\pi*})^{-1}\hat{q}_{m_T}^{\pi*}(\theta). \tag{3.7}$$

where $\hat{q}_{m_T}^{\pi*}(\theta) = \sum_{s=1}^{m_T} q_{t_s^* T}(\theta)$.

Theorem 3.1 states the consistency of the π-GEL GMM-KBB bootstrap estimator $\breve{\theta}_{m_T}^{\pi*}$.

Theorem 3.1. (π-GEL GMM-KBB Estimator Consistency.) Let $W_{qm_T}^{\pi*} - W_{qT} \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, $W_{qm_T}^{\pi*}$ p.s.d., $W_{qT} - W_q \to 0$, prob-$\mathcal{P}$, W_{qT} p.s.d. and W_q p.d. Under Assumptions 2.1, 2.2(b)-(d), 2.3(a)-(e) and 3.2, if $S_T \to \infty$, $S_T = o(T^{1/2})$ and $T^{1/\alpha}/m_T \to 0$, then $\breve{\theta}_{m_T}^{\pi*} - \breve{\theta}_T^{\pi} \to 0$ and $\breve{\mu}_{m_T}^{\pi*} - \breve{\mu}_T^{\pi} \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$.

Theorem 3.2 demonstrates the uniform convergence of the distribution of the appropriately centred π-GEL GMM-KBB estimator $\breve{\theta}_{m_T}^{\pi*}$ (3.6) to that of the centred restricted GMM estimator $\breve{\theta}_T$ (2.4).

Theorem 3.2. (π-GEL GMM-KBB Estimator Distribution Consistency.) Let Assumptions 2.1-2.4 and 3.2 hold. Then, if $W_{qm_T}^{\pi*} - W_{qT} \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, $W_{qm_T}^{\pi*}$ p.s.d., $W_{qT} - W_q \to 0$, prob-$\mathcal{P}$, W_{qT} p.s.d. and W_q p.d.,

$$\sup_{x\in\mathcal{R}^{d_\theta}} |\mathcal{P}_\omega^*(T^{1/2}(\breve{\theta}_{m_T}^{\pi*} - \breve{\theta}_T^{\pi}) \le x) - \mathcal{P}(T^{1/2}(\breve{\theta}_T - \theta_0) \le x)| \to 0, \text{ prob-}\mathcal{P},$$

$$\sup_{x\in\mathcal{R}^{d_r}} |\mathcal{P}_\omega^*(T^{1/2}(\breve{\mu}_{m_T}^{\pi*} - \breve{\mu}_T^{\pi}) \le x) - \mathcal{P}(T^{1/2}\breve{\mu}_T \le x)| \to 0, \text{ prob-}\mathcal{P}.$$

Note the centring of the π-GEL GMM-KBB distributions for $\breve{\theta}_{m_T}^{\pi*}$ and $\breve{\mu}_{m_T}^{\pi*}$ at the π-GEL GMM estimators $\breve{\theta}_T^{\pi}$ and $\breve{\mu}_T^{\pi}$ respectively.

Remark 3.7. If $\pi_{tT} = \breve{\pi}_{tT}$ (3.2), $(t = 1, ..., T)$, the $\breve{\pi}$-GEL GMM-KBB estimator $\breve{\theta}_T^{\breve{\pi}}$ may be replaced by the restricted GMM estimator $\breve{\theta}_T$ and $\breve{\mu}_T^{\breve{\pi}}$ omitted in the statement of Theorem 3.2; see Corollary 3.1. The first result is not specific to $\breve{\pi}$-GEL GMM-KBB and also holds for other block bootstrap methods; e.g., Theorems 3.1 and 3.2 apply to standard KBB, cf. Parente and Smith (2018, 2021), which uses the standard GMM bootstrap sampling probabilities $\pi_{tT}=T^{-1}$, $(t=1,...,T)$, thus, also entailing $\breve{\theta}_T^{\pi}$ and $\breve{\mu}_T^{\pi}$ being replaced by $\breve{\theta}_T$ and $\breve{\mu}_T$ respectively. Moreover, with the efficient GEL implied probabilities $\pi_{tT}=\tilde{\pi}_{tT}$ (3.3), $(t=1,...,T)$, see Remark 3.4, possibly counterintuitively, the $\tilde{\pi}$-GEL GMM-KBB estimator $\breve{\theta}_{m_T}^{\tilde{\pi}*}$ is centered at the efficient GMM estimator $\tilde{\theta}_T$ in Theorem 3.2, rather than the inefficient $\breve{\theta}_T$, whether or not an efficient metric $\Xi_{m_T}^{\pi*}$, $\Xi_{m_T}^{\pi*}-\Xi_T\rightarrow 0$, prob-$\mathcal{P}_{\omega}^*$, prob-$\mathcal{P}$, is employed in the $\tilde{\pi}$-GEL GMM-KBB criterion $\breve{\mathcal{Q}}_{m_T}^{\tilde{\pi}*}(\theta)$.

Remark 3.8. Lemma C.3 of Appendix C establishes that, under Assumptions 2.1-2.4 and 3.2, if $(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0)=O_p(1)$, then $\Xi_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})-\Xi_T\rightarrow 0$, prob-$\mathcal{P}_{\omega}^*$, prob-$\mathcal{P}$, $\Xi_T-\Xi\rightarrow 0$, prob-$\mathcal{P}$, where $\Xi_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})=\sum_{s=1}^{m_T}q_{t_s^*T}^{\pi}(\breve{\theta}_{m_T}^{\pi*})q_{t_s^*T}^{\pi}(\breve{\theta}_{m_T}^{\pi*})'/m_T$ or $\Xi_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})=\sum_{s=1}^{m_T}(q_{t_s^*T}^{\pi}(\breve{\theta}_{m_T}^{\pi*})-\hat{q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*}))(q_{t_s^*T}^{\pi}(\breve{\theta}_{m_T}^{\pi*})-\hat{q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*}))'/m_T$.

Remark 3.9. Theorems 3.1 and 3.2 are easily adapted for unrestricted GMM estimation with suitable modifications of Assumptions 2.1-2.4. With bootstrap sample $g_{t_s^*T}^{\pi}(\beta)$, $(s=1,...,m_T)$, m_T independent draws with replacement from the sample space $\{g_{tT}(\theta)\}_{t=1}^T$ with bootstrap sampling probabilities $\mathcal{P}_{\omega}^*(g_{t_s^*T}^{\pi}(\beta)=g_{tT}(\beta))=\pi_{tT}$, $(t=1,...,T)$, define the unrestricted π-GEL GMM-KBB estimator $\breve{\beta}_{m_T}^{\pi*}=\arg\min_{\beta\in\mathcal{B}}\breve{\mathcal{Q}}_{m_T}^{\pi*}(\beta)$ where $\breve{\mathcal{Q}}_{m_T}^{\pi*}(\beta)=\hat{g}_{m_T}^{\pi*}(\beta)'(W_{gm_T}^{\pi*})^{-1}\hat{g}_{m_T}^{\pi*}(\beta)$, $W_{gm_T}^{\pi*}$ p.s.d. and $\hat{g}_{m_T}^{\pi*}(\beta)=\sum_{s=1}^{m_T}g_{t_s^*T}(\beta)/m_T$. If $W_{gm_T}^{\pi*}-W_{gT}\rightarrow 0$, prob-$\mathcal{P}_{\omega}^*$, prob-$\mathcal{P}$, $W_{gT}-W_g\rightarrow 0$, prob-$\mathcal{P}$, W_g p.d., then $\breve{\beta}_{m_T}^{\pi*}-\breve{\beta}_T^{\pi}\rightarrow 0$, prob-$\mathcal{P}_{\omega}^*$, prob-$\mathcal{P}$, where $\breve{\beta}_T^{\pi}$ is defined in Remark 3.5, and $\sup_{x\in\mathcal{R}^{d_\beta}}|\mathcal{P}_{\omega}^*(T^{1/2}(\breve{\beta}_{m_T}^{\pi*}-\breve{\beta}_T^{\pi})\leq x)-\mathcal{P}(T^{1/2}(\breve{\beta}_T-\beta_0)\leq x)|\rightarrow 0$, prob-$\mathcal{P}$. Similarly to Remark 3.5 above, employing the unrestricted GEL implied probabilities $\pi_{tT}=\check{\pi}_{tT}$ (3.2), $(t=1,...,T)$, the unrestricted $\check{\pi}$-GEL GMM estimator $\breve{\beta}_T^{\check{\pi}}$ is replaced by $\breve{\beta}_T$, a result that also applies with standard GMM bootstrap sampling probabilities $\pi_{tT}=T^{-1}$, $(t=1,...,T)$. With efficient unrestricted GEL implied probabilities $\pi_{tT}=\hat{\pi}_{tT}$ (3.3), $(t=1,...,T)$, see Remark 3.5, $\breve{\beta}_T^{\hat{\pi}}$ is replaced by the efficient unrestricted GMM estimator $\hat{\beta}_T$ contradicting Allen et al. (2011, Theorems 1 and 2, p.114) for their EL MBB method, both estimators coinciding only if $W_g=\Sigma$.

Alternative less computationally intensive bootstraps may be constructed given the π-GEL GMM estimator $\breve{\theta}_T^{\pi}$ thereby avoiding the use of the π-GEL GMM-KBB estimator $\breve{\theta}_{m_T}^{\pi*}$. Cf. the fast subsampling procedure of Hong and Scaillet (2006); see also Camponovo et al. (2012).

Corollary 3.2. Suppose $W_{qT}-W_q\rightarrow 0$, $H_{W_qT}-H_{W_q}\rightarrow 0$ and $J_{W_qT}-J_{W_q}\rightarrow 0$, prob-$\mathcal{P}$, W_{qT} p.s.d. and W_q p.d. Then, under the conditions of Theorem 3.2, if $W_{qm_T}^{\pi*}-W_{qT}\rightarrow 0$, $H_{W_qm_T}^{\pi*}-H_{W_qT}\rightarrow 0$ and $J_{W_qm_T}^{\pi*}-J_{W_qT}\rightarrow 0$, prob-$\mathcal{P}_{\omega}^*$, prob-$\mathcal{P}$,

$$\sup_{x\in\mathcal{R}^{d_\theta}}|\mathcal{P}_{\omega}^*(-H_{W_qm_T}^{\pi*}(T/S_T)^{1/2}\hat{q}_{m_T}^{\pi*}(\breve{\theta}_T^{\pi})\leq x)-\mathcal{P}(T^{1/2}(\breve{\theta}_T-\theta_0)/k^{1/2}\leq x)|\rightarrow 0,\ \text{prob-}\mathcal{P},$$

$$\sup_{x\in\mathcal{R}^{d_r}}|\mathcal{P}_{\omega}^*(J_{W_qm_T}^{\pi*}(T/S_T)^{1/2}(\hat{q}_{m_T}^{\pi*}(\breve{\theta}_T^{\pi})-\hat{q}_T^{\pi}(\breve{\theta}_T^{\pi}))\leq x)-\mathcal{P}(T^{1/2}\breve{\mu}_T/k^{1/2}\leq x)|\rightarrow 0,\ \text{prob-}\mathcal{P}.$$

Suitable estimators $H_{W_qm_T}^{\pi*}$ and $J_{W_qm_T}^{\pi*}$ may be constructed straightforwardly from, e.g., $\hat{Q}_{m_T}^{\pi*}(\breve{\theta}_T)$

or $\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T)$, where $\hat{Q}^{\pi*}_{m_T}(\cdot) = \sum_{s=1}^{m_T} \partial q^{\pi}_{t^*_s T}(\cdot)/\partial\theta'/m_T$, $R(\breve{\theta}_T)$ or $R(\breve{\theta}^{\pi}_T)$ and $W^{\pi*}_{qm_T}$.

Remark 3.10. Cf. the influence function corresponding to the π-GEL GMM criterion $\breve{\mathcal{Q}}^{\pi}_T(\theta)$ (3.4), *viz.*,

$$IF^{\pi}_{W_qT}(z_t, \hat{F}_T) = -H^{\pi}_{W_qT}(T/S_T)^{1/2}(\hat{q}^{\pi}_{tT}(\breve{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)), (t = 1, ..., T),$$

where $H^{\pi}_{W_qT}$ denotes an estimator of H_{W_q} using $\hat{Q}^{\pi}_T(\breve{\theta}^{\pi}_T)$, $R(\breve{\theta}^{\pi}_T)$ and W_{qT}. Note that the term $\hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)$ is omitted in Corollary 3.2 since, from the π-GEL GMM first order condition for $\breve{\theta}^{\pi}_T$, $H^{\pi}_{W_qT}(T/S_T)^{1/2}\hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T) \to 0$, prob-$\mathcal{P}$. Cf. Gonçalves and White (2004, Corollary 2.1, p.203); also see Davidson and MacKinnon (1999).

Remark 3.11. Computation may further be simplified by exploiting a suitable adaptation of the GEL-KBB Local UWL Lemma A.4 in Appendix A, i.e., $\sup_{\theta\in\mathcal{N}} \|\hat{Q}^{\pi*}_{m_T}(\theta) - \hat{Q}^{\pi}_T(\theta)/S_T^{1/2}\| \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and replacing, e.g., $\hat{Q}^{\pi*}_{m_T}(\cdot)$ by $\hat{Q}^{\pi}_T(\cdot)$ evaluated at $\breve{\theta}_T$ or $\breve{\theta}^{\pi}_T$ and $W^{\pi*}_{qm_T}$ by W_{qT} in the definition of $H^{\pi*}_{W_qm_T}$ and $J^{\pi*}_{W_qm_T}$ to form $H^{\pi}_{W_qT}$ and $J^{\pi}_{W_qT}$,

$$\sup_{x\in\mathcal{R}^{d_\theta}} |\mathcal{P}^*_\omega(-H^{\pi}_{W_qT}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T) \le x) - \mathcal{P}(T^{1/2}(\breve{\theta}_T - \theta_0)/k^{1/2} \le x)| \to 0,\ \text{prob-}\mathcal{P},$$

$$\sup_{x\in\mathcal{R}^{d_r}} |\mathcal{P}^*_\omega(J^{\pi}_{W_qT}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)) \le x) - \mathcal{P}(T^{1/2}\breve{\mu}_T/k^{1/2} \le x)| \to 0,\ \text{prob-}\mathcal{P}.$$

A similar argument based on the UWLs $\sup_{\theta\in\Theta} \|\hat{Q}^{\pi}_T(\theta) - \hat{Q}_T(\theta)\|/S_T^{1/2}\| \to 0$, prob-$\mathcal{P}$, $\sup_{\theta\in\Theta} \|\hat{Q}_T(\theta)/S_T^{1/2} - Q(\theta)/k^{1/2}\| \to 0$, prob-$\mathcal{P}$, cf. Lemma A.1, with H_{W_qT} and J_{W_qT} constructed using the estimator $\hat{Q}_T(\theta)$ evaluated at $\breve{\theta}_T$ or $\breve{\theta}^{\pi}_T$, yields

$$\sup_{x\in\mathcal{R}^{d_\theta}} |\mathcal{P}^*_\omega(-H_{W_qT}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T) \le x) - \mathcal{P}(T^{1/2}(\breve{\theta}_T - \theta_0)/k^{1/2} \le x)| \to 0,\ \text{prob-}\mathcal{P},$$

$$\sup_{x\in\mathcal{R}^{d_r}} |\mathcal{P}^*_\omega(J_{W_qT}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)) \le x) - \mathcal{P}(T^{1/2}\breve{\mu}_T/k^{1/2} \le x)| \to 0,\ \text{prob-}\mathcal{P}.$$

Strengthening Assumption 2.4(d) to include Q_α (and, thus, Q) f.c.r. permits other confidence regions for the full vector θ_0 to be formulated; conservative confidence regions for a smooth function of θ_0 may be obtained by the "projection" method, see, e.g., Dufour (2003, Section 6.2, 791-792). Let $\chi^2_{d_\theta}(\cdot)$ denote a chi-square c.d.f. with d_θ degrees of freedom. Also let $Q_T \to Q$, Q_α f.c.r., and $\Xi_T \to \Xi$, prob-$\mathcal{P}$. To test $\theta = \theta_0$, the score statistic $\mathcal{S}_T(\theta_0) = (T/S_T)\hat{q}_T(\theta_0)'(\Xi_T)^{-1}Q_T(Q'_T(\Xi_T)^{-1}Q_T)^{-1}Q'_T(\Xi_T)^{-1}\hat{q}_T(\theta_0)$ converges in distribution to $\chi^2_{d_\theta}(\cdot)/k$, prob-$\mathcal{P}$, and is, thus, asymptotically pivotal.[8]

Corollary 3.3. Under the conditions of Theorem 3.2, if $Q_T \to Q$, Q_α f.c.r., and $\Xi_T \to \Xi$, prob-$\mathcal{P}$,

$$\sup_{x\in\mathcal{R}_+} |\mathcal{P}^*_\omega((T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T))(\Xi_T)^{-1}Q_T(Q'_T(\Xi_T)^{-1}Q_T)^{-1}Q'_T(\Xi_T)^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)) \le x) - \mathcal{P}(\mathcal{S}_T(\theta_0) \le x)| \to 0,\ \text{prob-}\mathcal{P}.$$

[8] With Q f.c.r., $M_W = (Q'W^{-1}Q)^{-1}$ and $H_W = M_W Q'W^{-1}$. To test $\theta = \theta_0$, the appropriate form of S_T (2.8) sets $S_g = I_{d_\theta}$, $P_\Xi = \Xi^{-1} - \Xi^{-1}QM_\Xi Q'\Xi^{-1}$ and $\hat{q}_T(\theta_0)$ substitutes for $\hat{q}_T(\tilde{\theta}_T)$.

Remark 3.12. Let x^*_α denote the α-level critical value from the empirical distribution of the π-GEL GMM-KBB statistic of Corollary 3.3. Inversion of the non-rejection region $\{\mathcal{S}_T(\theta)) \leq x^*_\alpha\}$ gives a nominal α-level confidence region for θ_0. The Corollary 3.3 statistic potentially provides the basis for asymptotic refinements but exploration of this topic is beyond the scope of the paper; cf. LaVecchia et al. (2023).

3.4 GEL Implied Probability GMM-KBB Inference

The following subsections detail π-GEL GMM-KBB versions of the test statistics for the over-identifying moment conditions (2.1) and the parametric restrictions and additional moment conditions (2.2) outlined in subsections 2.2.1 and 2.2.2. These forms of statistic exploit the GEL first order conditions which impose the population moment conditions on sample moments through re-weighting sample moment indicators with the GEL implied probabilities rather than the standard empirical GMM weights in the construction of sample averages.

3.4.1 Overidentification Tests

Let $\Sigma^{\pi *}_{m_T} - \Sigma_T \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$; cf. Remark 3.8. Consider the π-GEL GMM-KBB over-identifying moments test statistic

$$\mathcal{J}^{\pi *}_{m_T}/k = (T/S_T)(\hat{g}^{\pi *}_{m_T}(\hat{\beta}^{\pi *}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi *}_{m_T})^{-1}(\hat{g}^{\pi *}_{m_T}(\hat{\beta}^{\pi *}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)) \tag{3.8}$$

where $\hat{\beta}^{\pi}_T$ and $\hat{\beta}^{\pi *}_{m_T}$ are the efficient unrestricted π-GEL GMM and π-GEL GMM-KBB estimators respectively, see Remarks 3.5 and 3.9.

The following Theorem states the asymptotic validity of the π-GEL GMM-KBB statistic $\mathcal{J}^{\pi *}_{m_T}$ for overidentifying moment restrictions.

Theorem 3.3. With Assumptions 2.1-2.4 and 3.2 appropriately restated, if $\Sigma^{\pi *}_{m_T} - \Sigma_T \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Sigma^{\pi *}_{m_T}$ p.s.d., $\Sigma_T - \Sigma \to 0$, prob-$\mathcal{P}$, Σ_T p.s.d.,

$$\sup_{x \in \mathcal{R}_+} |\mathcal{P}^*_\omega(\mathcal{J}^{\pi *}_{m_T} \leq x) - \mathcal{P}(\mathcal{J}_T \leq x)| \to 0, \text{ prob-}\mathcal{P}.$$

Remark 3.13. With efficient unrestricted implied probabilities $\pi_{tT} = \hat{\pi}_{tT}$ (3.3) or standard GMM weights $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$, the efficient unrestricted estimator $\hat{\beta}_T$ may replace the π-GEL GMM estimator $\hat{\beta}^{\pi}_T$; cf. Remark 3.5. Moreover, with the former weights, the centring term $\hat{g}^{\hat{\pi}}_T(\hat{\beta}_T)$ in $\mathcal{J}^{\pi *}_{m_T}$ (3.8) may also be omitted since, from the first order conditions, $(T/S_T)^{1/2}\hat{g}^{\hat{\pi}}_T(\hat{\beta}_T) \to 0$, prob-$\mathcal{P}$, cf. Proof of Theorem 3.2. Thus, re-defining $\mathcal{J}^{\hat{\pi} *}_{m_T} = (T/S_T)\hat{g}^{\hat{\pi} *}_{m_T}(\hat{\beta}^{\hat{\pi} *}_{m_T})'(\Sigma^{\hat{\pi} *}_{m_T})^{-1}\hat{g}^{\hat{\pi} *}_{m_T}(\hat{\beta}^{\hat{\pi} *}_{m_T})$, $\sup_{x \in \mathcal{R}_+} |\mathcal{P}^\star(\mathcal{J}^{\hat{\pi} *}_{m_T} \leq x) - \mathcal{P}(\mathcal{J}_T/k \leq x)| \to 0$, prob-$\mathcal{P}$. As noted in Remark 3.5, with inefficient unrestricted implied probabilities $\pi_{tT} = \check{\pi}_{tT}$ (3.2), $(t = 1, ..., T)$, although the efficient metric $(\Sigma_T)^{-1}$ is employed the resultant $\check{\pi}$-GEL GMM estimator $\hat{\beta}^{\pi}_T$ is equivalent to the unrestricted but inefficient estimator $\check{\beta}_T$.

Remark 3.14. Hall and Horowitz (1996, p.898) proposes a non-overlapping GMM-MBB overidentifying moment restrictions test statistic with standard GMM empirical weights $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$, cf. Remark 3.9, similar to $\mathcal{J}^{\pi*}_{m_T}$ (3.8), an adjustment to which achieves higher order validity, see Hall and Horowitz (1996, Theorems 2 and 3, p.902). Bravo and Crudu (2011, (8), p.3447) details an overlapping $\hat{\pi}$-GEL GMM-MBB overidentifying moment restrictions test statistic with the efficient unrestricted implied probabilities $\pi_{tT} = \hat{\pi}_{tT}$ (3.3), $\hat{\beta}_T$ replacing $\hat{\beta}^{\pi}_T$ in and dropping $(T/S_T)^{1/2}\hat{g}^{\hat{\pi}}_T(\hat{\beta}_T)$ from $\mathcal{J}^{\pi*}_{m_T}$ (3.8), see Bravo and Crudu (2011, Theorem 1, p.3448); cf. Remark 3.13. Allen et al. (2011, pp.112-114) implements a similar uncentred statistic with both non-overlapping and overlapping MBB, implied probabilities $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$, and EL, $\pi_{tT} = 1/T(1 + \hat{\lambda}'_{gT} g_{tT}(\hat{\beta}_T))$, $(t = 1, ..., T)$, and flat kernel $k(x) = 1$, $|x| \leq 1$, and 0, $|x| > 1$, defining the kernel weighted moment indicators $g_{tT}(\cdot)$, $(t = 1, ..., T)$.

Similarly to Corollary 3.2, the use of the unrestricted π-GEL GMM-KBB estimator $\hat{\beta}^{\pi*}_{m_T}$ may be circumvented by the appropriate substitution of the efficient unrestricted π-GEL GMM estimator $\hat{\beta}^{\pi}_T$ since $(T/S_T)^{1/2}\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \Sigma P_{\Sigma}(T/S_T)^{1/2}\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi}_T) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, cf. Proof of Lemma B.3, and $H_{\Sigma}(T/S_T)^{1/2}\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T) \to 0$, prob-$\mathcal{P}$; *viz.*

$$\mathcal{J}^{\pi*}_{m_T}(\hat{\beta}^{\pi}_T)/k = (T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi}_T) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))' P^{\pi*}_{\Sigma T}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi}_T) - \hat{g}^{\pi}_T(\hat{\beta}_T))$$

Corollary 3.4. Under the Assumptions of Theorem 3.3, if $P^{\pi*}_{\Sigma T} - P_{\Sigma T} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and $P_{\Sigma T} - P_{\Sigma} \to 0$, prob-$\mathcal{P}$,

$$\sup_{x \in \mathcal{R}_+} |\mathcal{P}^*_\omega(\mathcal{J}^{\pi*}_{m_T}(\hat{\beta}^{\pi}_T) \leq x) - \mathcal{P}(\mathcal{J}_T \leq x)| \to 0, \text{ prob-}\mathcal{P}.$$

Remark 3.15. Further computational simplifications may be made; cf. Remark 3.11. Similarly to Remark 3.13, the efficient unrestricted estimator $\hat{\beta}_T$ may substitute for the π-GEL GMM estimator $\hat{\beta}^{\pi}_T$ with the efficient unrestricted implied probabilities $\pi_{tT} = \hat{\pi}_{tT}$ (3.3) or standard GMM weights $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$. Also $\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)$ can be omitted if $\pi_{tT} = \hat{\pi}_{tT}$, $(t = 1, ..., T)$. The estimator $P^{\pi*}_{\Sigma T}$ may be replaced by $P^{\pi}_{\Sigma T}$ or $P_{\Sigma T}$, where $P^{\pi}_{\Sigma T}$, $P_{\Sigma T} \to P_{\Sigma}$, prob-$\mathcal{P}$.

3.4.2 Specification Tests

Consider the following π-GEL GMM-KBB statistics for testing the additional moment restrictions $\mathrm{E}[h(z_t, \theta_0)] = 0$ and parametric constraints $r(\theta_0) = 0$ (2.2): a likelihood ratio-like statistic

$$\begin{aligned}\mathcal{LR}^{\pi*}_{m_T}/k &= (T/S_T)((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi^{\pi*}_{m_T} S_g(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)))' \\ &\quad \times (\Xi^{\pi*}_{m_T})^{-1}((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi^{\pi*}_{m_T} S_g(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)),\end{aligned} \tag{3.9}$$

cf. (2.6); the π-GEL GMM-KBB distance statistic

$$\begin{aligned}\mathcal{D}^{\pi*}_{m_T}/k &= (T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))'(\Xi^{\pi*}_{m_T})^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) \\ &\quad -(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)),\end{aligned} \tag{3.10}$$

where $\Sigma^{\pi*}_{m_T}$ is the (d_g, d_g) top left diagonal block of p.s.d. $\Xi^{\pi*}_{m_T}$, cf. below (2.6); a simplified score-like statistic

$$\mathcal{S}^{\pi*}_{m_T}/k = (T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))'((\Xi^{\pi*}_{m_T})^{-1} - S_g P^{\pi*}_{\Sigma m_T} S'_g)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T), \tag{3.11}$$

cf. (2.8); a Lagrange multiplier-like statistic

$$\begin{aligned}\mathcal{LM}^{\pi*}_{m_T}/k &= T\begin{pmatrix} (\Xi^{\pi*}_{m_T})^{-1}(\breve{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \breve{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))/S_T^{1/2} \\ (\tilde{\mu}^{\pi*}_{m_T} - \tilde{\mu}^{\pi}_T)/k^{1/2}\end{pmatrix}' S_{h,\mu}(S'_{h,\mu}\Psi^{\pi*}_{m_T}S_{h,\mu})^{-1}S'_{h,\mu} \\ &\times \begin{pmatrix} (\Xi^{\pi*}_{m_T})^{-1}(\breve{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \breve{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))/S_T^{1/2} \\ (\tilde{\mu}^{\pi*}_{m_T} - \tilde{\mu}^{\pi}_T)/k^{1/2}\end{pmatrix},\end{aligned} \tag{3.12}$$

cf. (2.9); a generalized Wald-like statistic

$$\mathcal{GW}^{\pi*}_{m_T}/k = T\begin{pmatrix} (\breve{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - \breve{q}^{\pi}_T(\hat{\theta}^{\pi}_T))/S_T^{1/2} \\ (r(\hat{\theta}^{\pi*}_{m_T}) - r(\hat{\theta}^{\pi}_T))/k^{1/2}\end{pmatrix}' \Psi^{\pi*}_{m_T}\begin{pmatrix} (\breve{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - \breve{q}^{\pi}_T(\hat{\theta}^{\pi}_T))/S_T^{1/2} \\ (r(\hat{\theta}^{\pi*}_{m_T}) - r(\hat{\theta}^{\pi}_T))/k^{1/2}\end{pmatrix}, \tag{3.13}$$

cf. (2.10), where $\breve{q}^{\pi*}_{m_T}(\cdot) = (I_{d_q} - \Xi^{\pi*}_{m_T}S_g(\Sigma^{\pi*}_{m_T})^{-1}S'_g)\hat{q}^{\pi*}_{m_T}(\cdot)$ and $\breve{q}^{\pi}_T(\cdot) = (I_{d_q} - \Xi^{\pi}_T S_g(\Sigma^{\pi}_T)^{-1}S'_g)\hat{q}^{\pi}_T(\cdot)$. The definition of $\mathcal{GW}^{\pi*}_{m_T}$ requires π-GEL GMM and π-GEL GMM-KBB estimators of the unidentified parameter α_0 under the maintained moment constraint $\mathrm{E}[g(z_t, \beta_0]$ (2.1); to avoid this difficulty, we set $\hat{\alpha}^{\pi}_T = \tilde{\alpha}^{\pi}_T$ and $\hat{\alpha}^{\pi*}_{m_T} = \tilde{\alpha}^{\pi*}_{m_T}$. Note that $\breve{q}^{\pi*}_{m_T}(\cdot) = (0', (\hat{h}^{\pi*}_{m_T}(\cdot) - (\Xi^{\pi*}_{m_T})_{hg}(\Sigma^{\pi*}_{m_T})^{-1}\hat{g}^{\pi*}_{m_T}(\cdot))')'$ and $\breve{q}^{\pi}_T(\cdot) = (0', (\hat{h}^{\pi}_T(\cdot) - (\Xi^{\pi}_T)_{hg}(\Sigma^{\pi}_T)^{-1}\hat{g}^{\pi}_T(\cdot))')'$. Cf. (2.10).

Theorem 3.4. Under Assumptions 2.1-2.4 and 3.2, if $\Xi^{\pi*}_{m_T} - \Xi_T \to 0$, $\Xi^{\pi*}_{m_T}$ p.s.d., $\Psi^{\pi*}_{\Xi m_T} - \Psi_{\Xi T} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Xi_T \to \Xi$, $\Psi_{\Xi T} \to \Psi_\Xi$, prob-$\mathcal{P}$,

$$\sup_{x\in\mathcal{R}_+}|\mathcal{P}^*_\omega(\mathcal{T}^{\pi*}_{m_T} \le x) - \mathcal{P}(\mathcal{T}_T \le x)| \to 0, \text{ prob-}\mathcal{P},$$

where $\mathcal{T} = \mathcal{LR}$, $\mathcal{D}$, $\mathcal{S}$, $\mathcal{LM}$ or $\mathcal{GW}$. Moreover, $\mathcal{LR}^{\pi*}_{m_T}$, $\mathcal{D}^{\pi*}_{m_T}$, $\mathcal{S}^{\pi*}_{m_T}$, $\mathcal{LM}^{\pi*}_{m_T}$ and $\mathcal{GW}^{\pi*}_{m_T}$ are asymptotically equivalent.

Remark 3.16. As above, with efficient implied probabilities $\pi_{tT} = \tilde{\pi}_{tT}$ (3.3) or standard GMM weights $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$, the efficient restricted estimator $\tilde{\theta}_T$ may substitute for the π-GEL GMM estimator $\tilde{\theta}^{\pi}_T$. Moreover, with the former weights, the efficient restricted estimator $\tilde{\beta}_T$ $(\tilde{\theta}_T)$ can replace $\hat{\beta}^{\pi}_T$ $(\hat{\theta}^{\pi}_T)$ in $\mathcal{T}^{\tilde{\pi}*}_{m_T}$, $\mathcal{T} = \mathcal{LR}$ and $\mathcal{D}$ $(\mathcal{GW})$, see Remark 3.5, and the centring term $\hat{q}^{\tilde{\pi}}_T(\tilde{\theta}^{\tilde{\pi}}_T)$ in $\mathcal{T}^{\tilde{\pi}*}_{m_T}$, $\mathcal{T} = \mathcal{LR}$, $\mathcal{D}$, $\mathcal{S}$, $\mathcal{LM}$ and $\mathcal{GW}$, may also be omitted since, from the first order conditions, $(T/S_T)^{1/2}\hat{q}^{\tilde{\pi}}_T(\tilde{\theta}^{\tilde{\pi}}_T) \to 0$, prob-$\mathcal{P}$, cf. Proof of Theorem 3.2. Similarly, with efficient unrestricted implied probabilities $\pi_{tT} = \hat{\pi}_{tT}$ (3.3) or standard GMM weights $\pi_{tT} = T^{-1}$, $\hat{\beta}_T$ may substitute for $\hat{\beta}^{\pi}_T$ and $\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)$ omitted in $\mathcal{T}^{\hat{\pi}*}_{m_T}$, $\mathcal{T} = \mathcal{LR}$, $\mathcal{D}$ and $\mathcal{GW}$. Additionally, since $H_\Xi \Xi S_g P_\Sigma = 0$, with $\pi_{tT} = \hat{\pi}_{tT}$, $(t = 1, ..., T)$, $\hat{\theta}_T$ may replace $\hat{\theta}^{\hat{\pi}}_T$ in $\mathcal{GW}^{\hat{\pi}*}_{m_T}$; see Remark 3.5.

Remark 3.17. Hall and Horowitz (1996, p.898), Bravo and Crudu (2011, (8), p.3447) and Allen et al. (2011, pp.112-114) give t-statistics $T^{1/2}(\hat{\beta}^{\pi*}_{m_T} - \hat{\beta}_T)_j/(((G^{\pi*}_{m_T})'(\Sigma^{\pi*}_{m_T})^{-1}(G^{\pi*}_{m_T}))^{jj})^{1/2}$ under their respective bootstrap designs described in Remark 3.14 for the parametric hypothesis $\beta_{0j} = 0$; cf. $\mathcal{GW}^{\pi*}_{m_T}$ (3.13) with no additional moment restrictions $\mathrm{E}[h(z_t,\beta_0)]=0$, noting $M_\Sigma=(G'\Sigma^{-1}G)^{-1}$, $J_\Sigma G=(RM_\Sigma R')^{-1}$ and, thus, $J_\Sigma\Sigma J_\Sigma=(RM_\Sigma R')^{-1}$, cf. Section B.2(d). Uncentred distance, Lagrange multiplier and Wald statistics for tests of $r(\beta_0)=0$ are proposed by Bravo and Crudu (2011, (10), p.3447). Their distance statistic is a version of $\mathcal{D}^{\pi*}_{m_T}$ (3.10) with $\hat{g}^{\tilde{\pi}*}_{m_T}(\tilde{\beta}^{\tilde{\pi}*}_{m_T})$ replacing $\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})$ with efficient restricted implied probabilities $\pi_{tT}=\tilde{\pi}_{tT}$ (3.3), $(t=1,...,T)$, under $r(\beta_0)=0$, $\Sigma^{\pi*}_{m_T}$ substituting for $\Xi^{\pi*}_{m_T}$, and $\hat{g}^{\hat{\pi}*}_{m_T}(\hat{\beta}^{\hat{\pi}*}_{m_T})$ replacing $\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})$ with efficient unrestricted implied probabilities $\pi_{tT}=\hat{\pi}_{tT}$ (3.3), omitting $\hat{q}^{\tilde{\pi}}_T(\tilde{\theta}^{\tilde{\pi}}_T)$ and $\hat{g}^{\hat{\pi}}_T(\hat{\beta}^{\hat{\pi}}_T)$. Their Lagrange multiplier statistic is $\mathcal{S}^{\pi*}_{m_T}$ (3.11) with $\hat{g}^{\tilde{\pi}*}_{m_T}(\tilde{\beta}^{\tilde{\pi}*}_{m_T})$ replacing $\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})$, $S_g=I_{d_g}$, $\Sigma^{\pi*}_{m_T}$ substituting for $\Xi^{\pi*}_{m_T}$ and dropping $\hat{q}^{\tilde{\pi}}_T(\tilde{\theta}^{\tilde{\pi}}_T)$. The Bravo and Crudu (2011, (10), p.3447) Wald statistic is $\mathcal{GW}^{\pi*}_{m_T}$ (3.13) with efficient unrestricted implied probabilities $\pi_{tT}=\hat{\pi}_{tT}$ (3.3), $(t=1,...,T)$, $M_\Sigma=(G'\Sigma^{-1}G)^{-1}$ and $J_\Sigma\Sigma J_\Sigma=(RM_\Sigma R')^{-1}$; note $(T/S_T)^{1/2}\breve{g}^{\hat{\pi}*}_{m_T}(\hat{\beta}^{\hat{\pi}*}_{m_T})\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $(T/S_T)^{1/2}\breve{g}^{\hat{\pi}}_T(\hat{\beta}^{\hat{\pi}}_T)\to 0$, prob-$\mathcal{P}$, replacing $(T/S_T)^{1/2}\breve{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T})$ and $(T/S_T)^{1/2}\breve{q}^{\pi}_T(\hat{\theta}^{\pi}_T)$ respectively, from the first order conditions. Cf. fn. 5.

We also define alternative π-GEL GMM-KBB likelihood ratio-like and distance statistics which adapt and generalise the bootstrap statistic proposed in Camponovo (2016) for the dependent data context and inference setting considered here. Let $\Xi_T-\Xi\to 0$, prob-$\mathcal{P}$, Ξ_T p.s.d. Rewrite the parametric constraint in (2.2) as $r(\theta)=r(\alpha,\hat{\beta}^{\pi}_T)$. Correspondingly, an alternative restricted π-GEL GMM estimator $\dot{\theta}^{\pi}_T$ is then defined by $\dot{\theta}^{\pi}_T=\arg\min_{\theta\in\Theta_{\dot{r}}}\tilde{\mathcal{Q}}^{\pi}_T(\theta)$, where $\Theta_{\dot{r}}=\{\theta\in\Theta:r(\theta)=r(\alpha,\hat{\beta}^{\pi}_T)\}$, with associated π-GEL GMM Lagrangean $\dot{\mathcal{L}}^{\pi}_T(\theta)=\tilde{\mathcal{Q}}^{\pi}_T(\theta)/S_T-2\mu'(r(\theta)-r(\alpha,\hat{\beta}^{\pi}_T))/k$, Lagrange multiplier estimator $\dot{\mu}^{\pi}_T$, and π-GEL GMM criterion $\tilde{\mathcal{Q}}^{\pi}_T(\theta)=\hat{q}^{\pi}_T(\theta)'(\Xi_T)^{-1}\hat{q}^{\pi}_T(\theta)$, cf. Remark 2.4. Define $\ddot{\theta}^{\pi}_T=(\dot{\alpha}^{\pi\prime}_T,\hat{\beta}^{\pi\prime}_T)'$. Also let $\Xi^{\pi*}_{m_T}-\Xi_T\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Xi^{\pi*}_{m_T}$ p.s.d. Then, cf. Camponovo (2016, (3.6), p. 37), with parametric constraint now $r(\theta)=r(\ddot{\theta}^{\pi}_T)$, the corresponding restricted π-GEL GMM-KBB estimator $\mathring{\theta}^{\pi*}_{m_T}=\arg\min_{\theta\in\Theta_{\mathring{r}}}\tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta)$, where $\Theta_{\mathring{r}}=\{\theta\in\Theta:r(\theta)=r(\ddot{\theta}^{\pi}_T)\}$, with π-GEL GMM-KBB Lagrangean $\tilde{\mathcal{L}}^{\pi*}_{m_T}(\theta)=\tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta)/S_T-2\mu'(r(\theta)-r(\ddot{\theta}^{\pi}_T)/k$, Lagrange multiplier estimator $\mathring{\mu}^{\pi*}_{m_T}$, and π-GEL GMM-KBB criterion $\tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta)=\hat{q}^{\pi*}_{m_T}(\theta)'(\Xi^{\pi*}_{qm_T})^{-1}\hat{q}^{\pi*}_{m_T}(\theta)$. Thus, cf. Proof of Proposition 3.2, $T^{1/2}(\mathring{\theta}^{\pi}_T-\ddot{\theta}^{\pi}_T)/k+H_\Xi(T/S_T)^{1/2}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T)\to 0$, prob-$\mathcal{P}$. The requisite π-GEL GMM-KBB likelihood ratio-like and distance statistics are respectively

$$\begin{aligned}\overline{\mathcal{LR}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T},\ddot{\theta}^{\pi}_T)/k &= (T/S_T)((\hat{q}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T})-\Xi^{\pi*}_{m_T}P^{\pi*}_{\Xi m_T}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))-\Xi^{\pi*}_{m_T}S_g(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)))'\\ &\quad\times(\Xi^{\pi*}_{m_T})^{-1}((\hat{q}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T})-\Xi^{\pi*}_{m_T}P^{\pi*}_{\Xi m_T}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))-\Xi^{\pi*}_{m_T}S_g(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)),\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}\overline{\mathcal{D}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T},\ddot{\theta}^{\pi}_T)/k &= (T/S_T)(\hat{q}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T})-\Xi^{\pi*}_{m_T}P^{\pi*}_{\Xi m_T}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))'(\Xi^{\pi*}_{m_T}))^{-1}(\hat{q}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T})-\Xi^{\pi*}_{m_T}P^{\pi*}_{\Xi m_T}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))\\ &\quad-(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)).\end{aligned} \tag{3.15}$$

Theorem 3.5. Under Assumptions 2.1-2.4 and 3.2, if $\Xi^{\pi*}_{m_T} - \Xi_T \to 0$, $\Xi^{\pi*}_{m_T}$ p.s.d., $P^{\pi*}_{\Xi m_T} - P_{\Xi T} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Xi_T \to \Xi$, $P_{\Xi T} - P_\Xi \to 0$, prob-$\mathcal{P}$,

$$\sup_{x\in\mathcal{R}_+} |\mathcal{P}^*_\omega(\overline{\mathcal{T}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T}, \ddot{\theta}^\pi_T) \le x) - \mathcal{P}(\mathcal{T}_T \le x)| \to 0, \text{ prob-}\mathcal{P},$$

where $\overline{\mathcal{T}} = \overline{\mathcal{LR}}$ or $\overline{\mathcal{D}}$. Moreover, $\overline{\mathcal{T}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T}, \ddot{\theta}^\pi_T)$ and $\mathcal{T}^{\pi*}_{m_T}$ of Theorem 3.4 are asymptotically equivalent.

Remark 3.18. The complexity of the π-GEL GMM-KBB likelihood ratio-like and distance statistics, $\overline{\mathcal{LR}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T}, \ddot{\theta}^\pi_T)$ and $\overline{\mathcal{D}}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T}, \ddot{\theta}^\pi_T)$, is reduced if only the parametric restrictions $r(\theta_0) = 0$ are subject to test and the additional moment constraints $\mathrm{E}[h(z_t, \theta_0)] = 0$ are maintained, cf. (2.2). In this case the parametric constraint becomes $r(\theta) = r(\hat{\theta}^\pi_T)$ as now $\ddot{\theta}^\pi_T = \hat{\theta}^\pi_T$, where $\hat{\theta}^\pi_T$ denotes the efficient unrestricted π-GEL GMM estimator of θ_0 with moment condition $\mathrm{E}[q(z_t, \theta_0)] = 0$. It may be straightforwardly demonstrated that $(T/S_T)^{1/2}\hat{q}^\pi_T(\hat{\theta}^\pi_T) - \Xi P_\Xi (T/S_T)^{1/2}\hat{q}^\pi_T(\hat{\theta}^\pi_T) \to 0$, prob-$\mathcal{P}$, since $H_\Xi (T/S_T)^{1/2}\hat{q}^\pi_T(\hat{\theta}^\pi_T) \to 0$, prob-$\mathcal{P}$, noting $M_\Xi = (Q'\Xi^{-1}Q)^{-1}$, $H_\Xi = M_\Xi Q'\Xi^{-1}$ and $H_\Xi(T/S_T)^{1/2}\hat{q}^\pi_T(\hat{\theta}^\pi_T) \to 0$, prob-$\mathcal{P}$, enabling the deletion of the matrix $\Xi^{\pi*}_{m_T} P^{\pi*}_{\Xi m_T}$ in both statistics.

Computationally less burdensome π-GEL GMM-KBB statistics may be defined in terms of the π-GEL GMM estimators $\tilde{\theta}^\pi_T$ or $\hat{\theta}^\pi_T$. Cf. Corollaries 3.2 and 3.4. Namely,

$$\begin{aligned}
\mathcal{LR}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T, \tilde{\beta}^\pi_T)/k &= (T/S_T)((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T)) - \Xi^{\pi*}_{m_T} S_g (\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^\pi_T) - \hat{g}^\pi_T(\hat{\beta}^\pi_T)))' \\
&\quad \times P^{\pi*}_{\Xi m_T}((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T)) - \Xi^{\pi*}_{m_T} S_g P^{\pi*}_{\Sigma m_T}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^\pi_T) - \hat{g}^\pi_T(\hat{\beta}^\pi_T));
\end{aligned}$$

$$\begin{aligned}
\mathcal{D}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T, \tilde{\beta}^\pi_T)/k &= (T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T))'(\Xi^{\pi*}_{m_T}))^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T)) \\
&\quad -(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^\pi_T) - \hat{g}^\pi_T(\hat{\beta}^\pi_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^\pi_T) - \hat{g}^\pi_T(\hat{\beta}^\pi_T));
\end{aligned}$$

$$\mathcal{S}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T)/k = (T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T))'(P^{\pi*}_{\Xi m_T} - S_g P^{\pi*}_{\Sigma m_T} S'_g)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T);$$

$$\begin{aligned}
\mathcal{LM}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T, \tilde{\mu}^\pi_T)/k &= T\begin{pmatrix} (\Xi^{\pi*}_{m_T})^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T))/S^{1/2}_T \\ J^{\pi*}_{\Xi m_T}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T)/S^{1/2}_T - \tilde{\mu}^\pi_T/k^{1/2} \end{pmatrix}' S_{h,\mu}(S'_{h,\mu}\Psi^{\pi*}_{m_T}S_{h,\mu})^{-1}S'_{h,\mu} \\
&\quad \times \begin{pmatrix} (\Xi^{\pi*}_{m_T})^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T) - \hat{q}^\pi_T(\tilde{\theta}^\pi_T))/S^{1/2}_T \\ J^{\pi*}_{\Xi m_T}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^\pi_T)/S^{1/2}_T - \tilde{\mu}^\pi_T/k^{1/2} \end{pmatrix};
\end{aligned}$$

$$\mathcal{GW}^{\pi*}_{m_T}(\hat{\theta}^\pi_T)/k = T\begin{pmatrix} (\breve{q}^{\pi*}_{m_T}(\hat{\theta}^\pi_T) - \breve{q}^\pi_T(\hat{\theta}^\pi_T))/S^{1/2}_T \\ (r(\hat{\theta}^\pi_T) - r(\hat{\theta}^\pi_T))/k^{1/2} \end{pmatrix}' \Psi^{\pi*}_{m_T} \begin{pmatrix} (\breve{q}^{\pi*}_{m_T}(\hat{\theta}^\pi_T) - \breve{q}^\pi_T(\hat{\theta}^\pi_T))/S^{1/2}_T \\ (r(\hat{\theta}^\pi_T) - r(\hat{\theta}^\pi_T))/k^{1/2} \end{pmatrix};$$

cf. (3.9), (3.10), (3.11), (3.12) and (3.13).

Corollary 3.5. Under the Assumptions of Theorem 3.4, if $P^{\pi*}_{\Xi m_T} - P_{\Xi T} \to 0$, $P^{\pi*}_{\Sigma T} - P_{\Sigma T} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and $P_{\Xi T} - P_\Xi \to 0$, $P_{\Sigma T} - P_\Sigma \to 0$, prob-$\mathcal{P}$,

$$\sup_{x\in\mathcal{R}_+} |\mathcal{P}^*_\omega(\mathcal{T}^{\pi*}_{m_T}(\cdot) \le x) - \mathcal{P}(\mathcal{T}_T \le x)| \to 0, \text{ prob-}\mathcal{P},$$

where $\mathcal{T} = \mathcal{LR}$, $\mathcal{D}$, $\mathcal{S}$, $\mathcal{LM}$ or $\mathcal{GW}$. Moreover, $\mathcal{LR}^{\pi *}_{m_T}(\tilde{\theta}^{\pi}_T, \hat{\beta}^{\pi}_T)$, $\mathcal{D}^{\pi *}_{m_T}(\tilde{\theta}^{\pi}_T, \hat{\beta}^{\pi}_T)$, $\mathcal{S}^{\pi *}_{m_T}(\tilde{\theta}^{\pi}_T)$, $\mathcal{LM}^{\pi *}_{m_T}(\tilde{\theta}^{\pi}_T, \tilde{\mu}^{\pi}_T)$ and $\mathcal{GW}^{\pi *}_{m_T}(\hat{\theta}^{\pi}_T)$ are asymptotically equivalent.

Remark 3.14. Further computational simplifications may be made similar to those in Remark 3.11.

4 Simulation Evidence

This section compares the performance of inference methods based on first order asymptotic theory, cf. Section 2.2, with those that rely on bootstrap methods, in particular, the π-GEL GMM-KBB statistics of Section 3.4. We consider over-identifying moments $\mathcal{J}$-statistics, see Sections 2.2.1 and 3.4.1, and test statistics for parametric restrictions only, see Sections 2.2.2 and 3.4.2. Empirical rejection rates for the various statistics are presented and are compared with their corresponding nominal sizes. Where required, we employ a prewhitened version of the Newey and West (1987) HAC estimator of the long run variance of the moment indicator throughout which, in our simulation design, behaved better than the standard Newey-West (1987) HAC estimator; cf. Andrews and Monahan (1992).[9] We consider the standard moving blocks bootstrap (MBB) alongside KBB methods based on different kernel-weighted moment indicators together with implied probability versions of MBB and KBB. All experiments are based on a simulation design similar to that of Inoue and Shintani (2006).

4.1 Design

We examine an instrumental variable model with intercept and regressor x_t, *viz.*

$$y_t = \beta_1 + \beta_2 x_t + u_t,\ (t = -49, ..., T),$$

with instrument vector $z_t = (1, x_t, x_{t-1}, x_{t-2})'$. We set $\beta_1 = 0$ and $\beta_2 = 0$ and consider two distinct data generating processes for x_t and u_t.

Model 1 Homoskedasticity. This model corresponds to the design considered in the simulation study in Inoue and Shintani (2006), the regressor x_t and regression error u_t are each generated by independent AR(1) processes with common autoregressive parameter ρ, i.e.,

$$u_t = \rho u_{t-1} + \varepsilon_{1t},\ u_{-49} = \varepsilon_{1,-49}\left(1-\rho^2\right)^{-1/2}$$

and

$$x_t = \rho x_{t-1} + \varepsilon_{2t},\ x_{-49} = \varepsilon_{2,-49}\left(1-\rho^2\right)^{-1/2},$$

[9] Empirical rejection rates of the various tests using an estimator of the moment indicator long-run variance based on the quadratic spectral kernel were also computed; cf. fn.4. Results from these experiments are available from the authors upon request.

where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \sim N(0, I_2)$, $(t = -49, ..., T)$.

Model 2 Heteroskedasticity. Here u_t is now generated by an AR(1)-GARCH(1,1) process, while, as in Model 1, x_t is generated by an AR(1) process independent of u_t, i.e.,

$$u_t = \rho u_{t-1} + \sigma_t \varepsilon_{1t},\ u_{-49} = \varepsilon_{1,-49}\left[\left(1-\rho^2\right)\right]^{-1/2},$$

where

$$\sigma_t^2 = 0.1 + 0.3\varepsilon_{1t-1}^2\sigma_{t-1}^2 + 0.6\sigma_{t-1}^2, \sigma_{-49}^2 = 1,$$

and

$$x_t = \rho x_{t-1} + \varepsilon_{2t},\ x_{-49} = \varepsilon_{2,-49}\left(1-\rho^2\right)^{-1/2}.$$

Again $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \sim N(0, I_2)$, $(t = -49, ..., T)$.

Remark 4.1. The Model 2 design choice is suggested by that in Allen et al. (2011, Section 5.1.2, p.115). Those authors consider a GARCH(1,1) specification for the error term u_t but without dependence. Dependence here in the conditional heteroskedasticity specification for σ_t^2 arises from the introduction of the additional AR(1) component to give an AR(1)-GARCH(1,1) specification for u_t.

Both models employ a sample $\{(y_t, x_t')'\}_{t=1}^T$ with autocorrelation parameter values $\rho \in \{0.5, 0.9\}$ and sample sizes $T \in \{64, 128\}$. Note that, given our choice of instrument vector z_t, two data points are lost.

Each experiment employs 499 bootstrap replications with 5000 random samples.

4.2 Bootstrap Methods

As in Section 3 above $m_T = [T/S_T]$ denotes the integer part of T/S_T. Here the indices t_s^* and the consequent bootstrap sample $g_{t_s^*T}^{\pi}(\beta)$, $(s = 1, ..., m_T)$, denote m_T independent draws with replacement from the index set $\mathcal{T}_T = \{1, ..., T\}$ and the bootstrap sample space $\{g_{tT}(\beta)\}_{t=1}^T$ with implied probability bootstrap sampling probabilities π_{tT}, $(t = 1, ..., T)$, i.e., $\mathcal{P}_\omega^*(g_{t_s^*T}^{\pi}(\beta) = g_{tT}(\beta)) = \pi_{tT}$, $(t = 1, ..., T)$, where the transformed kernel-weighted moment indicator $g_{tT}(\beta) = \sum_{s=t-T}^{t-1} k(\frac{s}{S_T}) g_t(\beta)/(k_2 S_T)^{1/2}$, $(t = 1, ..., T)$. We consider implied probability bootstrap sampling schemes with the standard GMM empirical measure, i.e., $\pi_{tT} = T^{-1}$, and unrestricted and restricted ET implied probabilities, i.e., $\hat{\pi}_{tT} = \exp(\hat{\lambda}_{gT}' g_{tT}(\hat{\beta}_T))/\sum_{s=1}^T \exp(\hat{\lambda}_{gT}' g_{sT}(\hat{\beta}_T))$ and $\tilde{\pi}_{tT} = \exp(\tilde{\lambda}_{gT}' g_{tT}(\tilde{\beta}_T))/\sum_{s=1}^T \exp(\tilde{\lambda}_{gT}' g_{sT}(\tilde{\beta}_T))$ respectively, $(t = 1, ..., T)$, where $\hat{\beta}_T$ and $\tilde{\beta}_T$ denote the efficient unrestricted and restricted GMM estimators of β_0; cf. Remarks 2.4 and 2.6. The ET implied probability bootstrap is adopted to avoid the well-known convex-hull problem associated with EL estimation.[10]

[10] Tables 7 and 8 provide an indication of the likely empirical failure of the convex hull condition necessary for the application of ET empirical probabilities. The convex hull condition is considered to hold if $\| \sum_{t=1}^T \hat{\pi}_{tT} g_{tT}(\hat{\beta}_T) \| < 10^{-6}$. Overall, these tables indicate that it is a relatively rare occurrence across all KBB and MBB methods for both values of ρ and both sample sizes.

4.3 Test Statistics

Recall the GEL sample average $\hat{g}_T(\beta) = \sum_{t=1}^{T} g_{tT}(\beta)/T$, the π-weighted sample average $\hat{g}^{\pi}_T(\beta) = \sum_{t=1}^{T} \pi_{tT} g_{tT}(\beta)$ and the π-GEL GMM-KBB sample average $\hat{g}^{\pi*}_{m_T}(\beta) = \sum_{s=1}^{m_T} g^{\pi}_{t^*_s T}(\beta)/m_T$. Define the centred long-run variance estimator $\Sigma^{\pi*}_{m_T} = \sum_{s=1}^{m_T}(g^{\pi}_{t^*_s T}(\breve{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))(g^{\pi}_{t^*_s T}(\breve{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'/m_T$, cf. Remark 3.7, where the first step standard GMM-KBB estimator $\breve{\beta}^{\pi*}_{m_T} = \arg\min_{\beta\in\mathcal{B}} \hat{g}^{\pi*}_{m_T}(\beta)'\hat{g}^{\pi*}_{m_T}(\beta)$, setting $\pi_{tT} = T^{-1}$ The efficient centred standard GMM-KBB estimator $\hat{\beta}^{*}_{m_T}$ is then given by

$$\hat{\beta}^{\pi*}_{m_T} = \arg\min_{\beta\in\mathcal{B}} \hat{g}^{\pi*}_{m_T}(\beta)'(\Sigma^{\pi*}_{m_T})^{-1}\hat{g}^{\pi*}_{m_T}(\beta),$$

cf. (3.7).

4.3.1 Overidentification Tests

Recall the π-GEL GMM-KBB over-identifying moments test statistic (3.8)

$$\mathcal{J}^{\pi*}_{m_T} = (T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))$$

where $\hat{\beta}^{\pi}_T$ and $\hat{\beta}^{\pi*}_{m_T}$ are the efficient unrestricted π-GEL GMM and π-GEL GMM-KBB estimators respectively.

With both standard GMM empirical measure, i.e., $\pi_{tT} = T^{-1}$, and efficient unrestricted ET implied probability, i.e., $\hat{\pi}_{tT} = \exp(\hat{\lambda}'_{gT} g_{tT}(\hat{\beta}_T))/\sum_{s=1}^{T}\exp(\hat{\lambda}'_{gT} g_{sT}(\hat{\beta}_T))$, cf. (3.3), $(t = 1, ..., T)$, bootstrap sampling schemes, the efficient unrestricted GMM estimator $\hat{\beta}_T$ is substituted for the π-GEL GMM estimator $\hat{\beta}^{\pi}_T$; cf. Inoue and Shintani (2006). Moreover, with the efficient unrestricted ET implied probabilities, $\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)$ is omitted. See Remark 3.14.

4.3.2 Specification Tests

The parametric restriction $\beta_2 = 0$ is of interest here. We examine forms of π-GEL GMM-KBB distance, generalised Wald and Camponovo-like (2016) distance statistics adapted for parametric restrictions $r(\beta_0) = 0$, cf. (3.10), (3.13) and (3.15), in the absence of α and additional moment constraints $\mathrm{E}[h(z_t, \beta_0)] = 0$, cf. (2.2). Namely

$$\begin{aligned}\mathcal{D}^{\pi*}_{m_T}/k &= (T/S_T)(\hat{g}^{\pi*}_{m_T}(\tilde{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\tilde{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\tilde{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\tilde{\beta}^{\pi}_T)) \\ &\quad -(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T));\end{aligned}$$

$$\mathcal{GW}^{\pi*}_{m_T} = T(r(\hat{\beta}^{\pi*}_{m_T}) - r(\hat{\beta}^{\pi}_T))'(R^{\pi*}_{m_T}(G^{\pi*\prime}_{m_T}(\Sigma^{\pi*}_{m_T})^{-1}G^{\pi*}_{m_T})^{-1}R^{\pi*\prime}_{m_T})^{-1}(r(\hat{\beta}^{\pi*}_{m_T}) - r(\hat{\beta}^{\pi}_T))$$

or, equivalently, to test $\beta_2 = 0$, the t-statistic $T^{1/2}(\hat{\beta}^{\pi*}_{m_T} - \hat{\beta}^{\pi}_T)_2/(G^{\pi*\prime}_{m_T}(\Sigma^{\pi*}_{m_T})^{-1}G^{\pi*}_{m_T})^{22})^{1/2}$, see Remark 3.16;

$$\begin{aligned}\overline{\mathcal{LR}}^{\pi*}_{m_T}(\dot{\beta}^{\pi*}_{m_T}, \hat{\beta}^{\pi}_T)/k &= (T/S_T)(\hat{g}^{\pi*}_{m_T}(\dot{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\dot{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T)) \\ &\quad -(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T));\end{aligned}$$

Here $\tilde{\beta}_T^{\pi}$ and $\tilde{\beta}_{m_T}^{\pi*}$ are the efficient restricted π-GEL GMM and π-GEL GMM-KBB estimators respectively. Likewise $\hat{\beta}_T^{\pi}$ and $\hat{\beta}_{m_T}^{\pi*}$ are the efficient unrestricted π-GEL GMM and π-GEL GMM-KBB estimators respectively and $\dot{\beta}_{m_T}^{\pi*}$ the restricted π-GEL GMM-KBB estimator subject to the constraint $r(\beta) = r(\hat{\beta}_T^{\pi})$, see Remark 3.17.

With the standard GMM empirical measures, $\pi_{tT} = T^{-1}$, $(t = 1, ..., T)$, the efficient unrestricted GMM estimator $\hat{\beta}_T$ is substituted for the π-GEL GMM estimators $\hat{\beta}_T^{\pi}$; see Remark 3.5 and Proposition 3.2 cf. Remark 2.5. With efficient unrestricted ET implied probabilities, $\hat{\pi}_{tT} = \exp(\hat{\lambda}_{gT}' g_{tT}(\hat{\beta}_T)) / \sum_{s=1}^{T} \exp(\hat{\lambda}_{gT}' g_{sT}(\hat{\beta}_T))$, $(t = 1, ..., T)$, $\hat{g}_T^{\pi}(\hat{\beta}_T^{\pi})$ is omitted. Similarly, with efficient restricted ET implied probabilities $\tilde{\pi}_{tT} = \exp(\tilde{\lambda}_{gT}' g_{tT}(\tilde{\beta}_T)) / \sum_{s=1}^{T} \exp(\tilde{\lambda}_{gT}' g_{sT}(\tilde{\beta}_T))$, $(t = 1, ..., T)$, $\hat{g}_T^{\pi}(\tilde{\beta}_T^{\pi})$, see Remark 3.15, is omitted. The efficient restricted GMM estimator $\tilde{\beta}_T$ is substituted for the π-GEL GMM estimator $\tilde{\beta}_T^{\pi}$ in all bootstrap sampling schemes; cf. Remark 2.4.

4.4 Computational Issues

The choice of the bandwidth/block size S_T is important. Gonçalves and White (2004) notes the equivalence between the MBB variance estimator of the mean and the HAC variance matrix estimator using the Bartlett kernel. Consequently, Gonçalves and White (2004) bases the choice of MBB block size on the optimal automatic bandwidth for the latter estimator; see Andrews (1991, Section 5, p.830-832). Although this equivalence result is only valid for the mean, Gonçalves and White (2004) also adopts this choice for the quasi-maximum likelihood estimator. We also follow this approach.[11]

Here, in contradistinction to Parente and Smith (2021), we use a prewhitened version of the Newey and West (1987) HAC estimator of Σ to obtain efficient unrestricted and restricted GMM estimators $\hat{\beta}_T$ and $\tilde{\beta}_T$; see Andrews and Monahan (1992). Cf. Remarks 2.4 and 2.6. A vector (V) AR(1) specification is applied to the elements $z_t \breve{u}_t$ where $\breve{u}_t$, $(t = 1, ..., T)$, are the VAR(1) residuals obtained from first step GMM estimation. We then use a version of the automatic bandwidths described in Andrews (1991, Section 6, pp.832-837) but we employ the non-parametric estimation approach due to Politis and Romano (1995) applied to the elements of the VAR(1) residuals avoiding specification and estimation of parametric univariate ARMA models as in Andrews (1991); cf. Parente and Smith (2021, Section 4.3, pp. 387-388). Given the bandwidth/block size estimator $\hat{S}_T$ the prewhitened Newey and West (1987) HAC estimator of Σ is the recoloured Newey and West (1987) HAC estimator based on the VAR(1) residuals; see Andrews and Monahan (1992, (2.4), p.995).

Additionally, since the computed automatic bandwidth $\hat{S}_T$ might induce values of bootstrap sample size $m_T = [T/S_T]$ greater than T or equal to 1, we substituted the censored version $\min\left\{\max\left\{\hat{S}_T, 1\right\}, T/10\right\}$. An additional advantage of using the censored bandwidth estimator is an apparent improvement in the convergence properties of the ET estimation algorithm.

[11] Smith (2011, Lemma A.3, p.1219) describes an equivalence between the KBB variance estimator and the corresponding HAC estimator based on the induced kernel function $k^*(\cdot)$. Also see Smith (2005, Lemma 2.1, p.164).

On a few rare occasions π-GEL GMM-KBB Σ estimates are poorly conditioned. These bootstrap samples are discarded and replaced by additional samples so that only those with well-conditioned π-GEL GMM-KBB estimates of Σ are retained.[12]

4.5 Notation

The subscripts TR, BT, QS and PP indicate use of, respectively, the truncated kernel, the Bartlett kernel, the kernel that induces the quadratic-spectral kernel, see Smith (2011, Example 2.3, p.1204), and the kernel version of the Paparoditis and Politis (2001) optimal taper to define the transformed kernel-weighted moment indicator, cf. (2.3).[13] ASYMP denotes results for the standard GMM over-identifying moments test computed using the efficient unrestricted GMM estimator $\hat{\beta}_T$ and the prewhitened Newey and West (1987) HAC covariance matrix estimator for Σ described above. Results KBB are obtained with the KBB method. MBB refers to the MBB method. Additionally, the superscripts π and π_r indicate results with the ET implied probability bootstrap based on, respectively, unrestricted ($\mathrm{E}[g(z_t,\beta)]=0$) and restricted ($\mathrm{E}[g(z_t,\beta)]=0,\ r(\beta)=0$) ET implied probabilities. Finally, the superscript C denotes the Camponovo-like π-GEL GMM-KBB distance statistic $\overline{\mathcal{LR}}^{\pi*}_{m_T}(\dot{\beta}^{\pi*}_{m_T},\hat{\beta}^{\pi}_T)$.

4.6 Results

Tables 1-6 present percentage empirical rejection rates for Models 1 and 2 for the π-GEL GMM-KBB over-identifying moments test statistic $\mathcal{J}^{\pi*}_{m_T}$, generalised Wald $\mathcal{GW}^{\pi*}_{m_T}$ t-statistic $T^{1/2}(\hat{\beta}^{\pi*}_{m_T}-\hat{\beta}^{\pi}_T)_2/(G^{\pi*\prime}_{m_T}(\Sigma^{\pi*}_{m_T})^{-1}G^{\pi*}_{m_T})^{22})^{1/2}$ and the π-GEL GMM-KBB distance and Camponovo-like (2016) distance specification test statistics based on nominal and bootstrap critical values computed at the 0.01, 0.05 and 0.10 levels. Results for the best performing tests are indicated by bold face text.

Tables 1 and 2 about here

Examining Tables 1 and 2, the standard over-identifying moments test ASYMP using critical values based on asymptotic theory performs quite well even if $\rho=0.9$. Similar results were also noted in the simulation studies undertaken in Hall and Horowitz (1996) and Inoue and Shintani (2006). All KBB and MBB over-identifying moment tests under-reject in Models 1 and 2 for both sample sizes $T=64$ and $T=128$. As expected, the performance of all tests improves with increased sample size $T=128$. The π-GEL GMM-KBB tests employing the standard empirical GMM measures perform somewhat better than the corresponding MBB test for both models when $\rho=0.5$ if $T=64$, but only marginally so if

[12] Ill-conditioned $\Sigma^*_{m_t}$ estimates tend to occur when the ET implied probabilities take large values for a few bootstrap sample observations and appear to be problematic only for those estimates using restricted ET implied probabilities for the smaller sample size $T=64$ and for the larger value $\rho=0.9$; see Tables 9 and 10. Correspondingly, Tables 11 and 12 indicate that considerably more bootstrap replications are required in these cases.

[13] The Paparoditis and Politis (2001) taper kernel is defined as $k(x)=\mathrm{I}(-0.5\le x<-0.07)(x+0.5)/0.43+\mathrm{I}(-0.07\le x\le 0.07)+\mathrm{I}(0.07<x\le 0.5)(0.5-x)/0.43$, where $\mathrm{I}(\cdot)$ denotes the indicator function.

$T = 128$. When $\rho = 0.9$, the MBB^{π} test performs better than most π-GEL GMM-KBB tests for both models and sample sizes. In general, π-GEL GMM-KBB tests computed with ET implied probabilities achieve rejection rates closer to nominal levels than other methods apart from ASYMP, in particular, for sample size $T = 128$ and $\rho = 0.5$. Overall, $\text{KBB}^{\pi}_{\text{PP}}$ achieves the best performance with $\text{KBB}^{\pi}_{\text{TR}}$, $\text{KBB}^{\pi}_{\text{QS}}$ and MBB^{π} displaying reasonable size properties for both values of ρ and both sample sizes.

Tables 3 and 4 about here

Tables 3 and 4 show that the specification ASYMP t-tests based on first order asymptotic theory are severely over-sized for all designs and are outperformed by all π-GEL GMM-KBB t-tests. Of the π-GEL GMM-KBB t-tests all KBB_{TR} tests display the worst performance across all designs. Indeed, for either design and for $\rho = 0.9$, all π-GEL GMM-KBB and MBB statistics display poor size properties for $T = 64$ and only the KBB_{PP} test possesses reasonable size properties for $T = 128$. For Model 1 and $\rho = 0.5$, for the smaller sample size $T = 64$ KBB_{PP} and MBB^{π} seem satisfactorily sized whereas for $T = 128$ $\text{KBB}^{\pi}_{\text{BT}}$, MBB^{π} and MBB^{π_r} tests perform reasonably but KBB_{PP} is now undersized. For Model 2 and both sample sizes, $\text{KBB}^{\pi}_{\text{BT}}$ and all MBB tests, especially the MBB^{π_r} test, seem competitive for $\rho = 0.5$ with $\text{KBB}^{\pi}_{\text{QS}}$ satisfactory for the larger sample size $T = 128$. In general there appears to be no obvious ranking of π-GEL GMM-KBB and MBB t-tests in terms of their use of the standard GMM measure, unrestricted or restricted ET implied probabilities. To summarise, for the smaller sample size $T = 64$ and $\rho = 0.5$ MBB^{π} appears the most satisfactory whereas for $T = 128$ $\text{KBB}^{\pi}_{\text{BT}}$, $\text{KBB}^{\pi}_{\text{QS}}$, MBB and MBB^{π_r} t-tests are competitive. No π-GEL GMM-KBB t-test can be recommended for the higher value $\rho = 0.9$ with sample size $T = 64$ although the KBB_{PP} test is satisfactory for the larger sample size $T = 128$. Overall MBB t-test statistics appear most reliable for both sample sizes and the lower value $\rho = 0.5$.

Tables 5 and 6 about here

Similarly to the t-statistics above the specification ASYMP distance tests based on first order asymptotic theory are severely over-sized for all designs and are outperformed by all π-GEL GMM-KBB distance tests; see Tables 5 and 6. Overall, none of the π-GEL GMM-KBB distance tests uniformly dominates the others. For Model 1, all π-GEL GMM-KBB distance tests display poor size properties for $\rho = 0.9$ and both sample sizes with the exception of the (somewhat under-sized) KBB_{BT} and $\text{KBB}^{\pi}_{\text{QS}}$ tests whereas for Model 2 $\text{MBB}^{\pi_r,\text{C}}$ provides a reasonably sized test. For $\rho = 0.5$ and sample size $T = 128$ many π-GEL GMM-KBB, e.g., KBB_{TR}, $\text{KBB}^{\pi}_{\text{TR}}$, $\text{KBB}^{\pi,\text{C}}_{\text{BT}}$, $\text{KBB}^{\pi_r,\text{C}}_{\text{BT}}$, $\text{KBB}^{\pi,\text{C}}_{\text{QS}}$ and $\text{KBB}^{\pi_r,\text{C}}_{\text{QS}}$, and all Camponovo-like (2016) MBB distance specification tests display reasonable size properties for both designs. For $\rho = 0.5$, both sample sizes and across both designs, the Camponovo-like (2016) distance test $\text{KBB}^{\pi_r,\text{C}}_{\text{BT}}$ possesses adequate size properties. For $T = 64$ for Model 1, all MBB tests display poor size properties but MBB^{C}, $\text{MBB}^{\pi,\text{C}}$ and MBB^{π_r} are reasonably sized for $\rho = 0.5$ and sample size $T = 128$. For Model 2 for both sample sizes, the size properties of MBB^{C}, $\text{MBB}^{\pi,\text{C}}$ and MBB^{π_r} tests are adequate for $\rho = 0.5$. In summary and overall, KBB_{BT} and $\text{KBB}^{\pi}_{\text{QS}}$ tests, albeit under-sized, offer reasonably robust test procedures. For larger sample sizes with moderate dependence, a number of KBB and MBB specification tests offer empirical rejection rates close to the nominal ones. However, for high values of ρ, larger sample sizes are required for reliable procedures.

4.7 Summary

Overall, the results from these simulation experiments are rather mixed. The standard overidentifying moments test ASYMP using critical values based on asymptotic theory performs quite well. Of π-GEL GMM-KBB KBB and MBB tests, $\text{KBB}^{\pi}_{\text{PP}}$, and to a lesser degree, $\text{KBB}^{\pi}_{\text{TR}}$, $\text{KBB}^{\pi}_{\text{QS}}$ and MBB^{π}, tests display reasonable size properties for both values of ρ and both sample sizes. In particular, for moderate dependence and larger sample size, both π-GEL GMM-KBB and MBB statistics with unrestricted ET probabilities provide reasonably sized test procedures. The performance of both standard ASYMP t- and distance specification tests is unsatisfactory. Generally, all π-GEL GMM-KBB and MBB specification tests are unsatisfactory for both sample sizes and with strong dependence. For both sample sizes, moderate dependence and across both designs, MBB specification tests with unrestricted ET probabilities offer reasonable test procedures. With a larger sample and moderate dependence, $\text{KBB}^{\pi}_{\text{BT}}$, $\text{KBB}^{\pi}_{\text{QS}}$, MBB and MBB^{π_r} t-tests and KBB_{TR}, $\text{KBB}^{\pi}_{\text{TR}}$, $\text{KBB}^{\pi,\text{C}}_{\text{BT}}$, $\text{KBB}^{\pi_r,\text{C}}_{\text{BT}}$, $\text{KBB}^{\pi,\text{C}}_{\text{QS}}$ and $\text{KBB}^{\pi_r,\text{C}}_{\text{QS}}$ together with all Camponovo-like (2016) MBB distance specification tests are competitive indicating that, in some cases, unrestricted and restricted ET probabilities can improve the size characteristics of specification tests.

5 Conclusion

This article generalizes and extends the kernel block bootstrap (KBB) method introduced in Parente and Smith (2018, 2021) to time-series models formulated in terms of moment conditions. We provide a comprehensive treatment of the use of KBB for GMM estimation and inference for this context. The paper details new KBB estimators and test statistics whose empirical distributions can serve as alternative approximations to those offered by standard and other bootstrap methods for GMM estimators and test statistics for overidentifying moment conditions, parametric restrictions in mixed form and additional moment restrictions. We consider KBB methods that use the standard GMM empirical measure and unrestricted and restricted GEL implied probabilities. The paper establishes the first-order validity of the various methods generalizing Bravo and Crudu (2011) and (correcting) Allen et al. (2011). A set of simulation experiments reveals that a number of the proposed tests perform well in practice in circumstances with moderate dependence and for larger sample sizes.

Appendix: Proofs

Throughout the Appendix, C and Δ will denote generic positive constants that may be different in different uses, and C, M, and T the Chebyshev, Markov, and triangle inequalities respectively.

Appendix A: Preliminary GEL-KBB Lemmas

In the following $X_t(\theta)$ substitutes, where appropriate, for $q_t(\theta)$, $(t = 1, 2, ...)$, in Assumptions 2.1, 2.3 and 2.4.

Let $X_{tT}(\theta) = \sum_{s=t-T}^{t-1} k(s/S_T)X_t(\theta)/(k_2 S_T)^{1/2}$, $(t = 1, ..., T)$, $\bar{X}(\theta) = \sum_{t=1}^{T} X_t(\theta)/T$ and $\bar{X}_T(\theta) = \sum_{t=1}^{T} X_{tT}(\theta)/T$. Also let $\bar{X}_T^{\pi}(\theta) = \sum_{t=1}^{T} \pi_{tT} X_{tT}(\theta)$ and $\bar{X}_{m_T}^{\pi *}(\theta) = \sum_{s=1}^{m_T} X_{t_s^* T}^{\pi}(\theta)/m_T$ where the indices t_s^* and the consequent bootstrap sample $X_{t_s^* T}^{\pi}(\beta)$, $(s = 1, ..., m_T)$, denote m_T independent draws with replacement from the index set $\mathcal{T}_T = \{1, ..., T\}$ and the bootstrap sample space $\{X_{tT}(\theta)\}_{t=1}^{T}$ with sampling probabilities $\mathcal{P}_{\omega}^*(X_{t_s^* T}(\theta) = X_{tT}(\theta)) = \pi_{tT}$, $(t = 1, ..., T)$, with $m_T = [T/S_T]$ the integer part of T/S_T.

Lemma A.1. (UWL.) If $X_t(\theta)$, $(t = 1, 2, ...)$, satisfies Assumptions 2.1, 2.2 and 2.3 (a)(b)(d), then

$$\sup_{\theta \in \Theta} \|\bar{X}_T(\theta)/S_T^{1/2} - \bar{X}(\theta)/k^{1/2}\| \to 0, \text{ prob-}\mathcal{P}.$$

Proof. The hypotheses of the UWLs Smith (2011, Lemma A.1, p.1217) and Newey and McFadden (1994, Lemma 2.4, p.2129) for stationary and mixing (and, thus, ergodic) processes are satisfied under Assumptions 2.1, 2.2 and 2.3 (a)(b)(d). Hence, noting $\sup_{\theta \in \Theta} \|\bar{X}(\theta) - \mathrm{E}[X_t(\theta)]\| \to 0$, prob-$\mathcal{P}$, $\sup_{\theta \in \Theta} \|\bar{X}_T(\theta)/S_T^{1/2} - k^{1/2}\mathrm{E}[X_t(\theta)]\| \to 0$, prob-$\mathcal{P}$. Thus, the result follows by T and $k = O(1)$.■

Lemma A.2. (GEL-KBB Pointwise WLLN.) Suppose Assumptions 2.1, 2.2 and 2.3(a) are satisfied by $X_t(\theta)$, $(t = 1, 2, ...)$, and π_{tT}, $(t = 1, ..., T)$, satisfy Assumption 3.2. Then, if $T^{1/\alpha}/m_T \to 0$ and $\mathrm{E}[\sup_{\theta \in \Theta} \|X_t(\theta)\|^{\alpha}] < \infty$ for some $\alpha > v$, for each $\theta \in \Theta$,

$$\textbf{(a)}\ \|\bar{X}_{m_T}^{\pi *}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} \to 0,\ \textbf{(b)}\ \|\bar{X}_{m_T}^{\pi *}(\theta) - \bar{X}_T^{\pi}(\theta)\|/S_T^{1/2} \to 0, \text{ prob-}\mathcal{P}_{\omega}^*, \text{ prob-}\mathcal{P}.$$

Proof. It is only necessary to prove **(b)** since

$$\begin{aligned}
\sup_{\theta \in \Theta} \|\bar{X}_T(\theta) - \bar{X}_T^{\pi}(\theta)\|/S_T^{1/2} &= \sup_{\theta \in \Theta} \|\sum_{t=1}^{T} (1 - T\pi_{tT})(X_{tT}(\theta)/S_T^{1/2})/T\| \\
&\le \max_{1 \le t \le T} |1 - T\pi_{tT}| \sum_{t=1}^{T} \sup_{\theta \in \Theta} \|X_{tT}(\theta)/S_T^{1/2}\|/T \\
&= o_p(1)O_p(1) = o_p(1)
\end{aligned}$$

as $\max_{1 \le t \le T} |T\pi_{tT} - 1| = o_p(1)$ by Assumption 3.2(b), $\sum_{t=1}^{T} \sup_{\theta \in \Theta} \|X_{tT}(\theta)/S_T^{1/2}\|/T \le O(1) \sum_{t=1}^{T} \sup_{\theta \in \Theta} \|X_t(\theta)\|/T = O_p(1)$, Smith (2011, eq. (A.5), p.1218) and $\mathrm{E}[\sup_{\theta \in \Theta} \|X_t(\theta)\|] < \infty$ by hypothesis.

The argument θ is now suppressed for brevity throughout the remainder of the proof. First, cf. Gonçalves and White (2004, Proof of Lemma A.5, p.215),

$$\bar{X}^{\pi*}_{m_T} - \bar{X}^{\pi}_T = (\bar{X}^{\pi*}_{m_T} - \mathrm{E}^*[\bar{X}^{\pi*}_{m_T}]) - (\bar{X}^{\pi}_T - \mathrm{E}^*[\bar{X}^{\pi*}_{m_T}]).$$

Since $\mathrm{E}^*[\bar{X}^{\pi*}_{m_T}] = \bar{X}^{\pi}_T$, the second term $\bar{X}^{\pi}_T - \mathrm{E}^*[\bar{X}^{\pi*}_{m_T}]$ is zero. Hence, the result follows if, for any $\delta > 0$ and $\xi > 0$ and large enough T, $\mathcal{P}(\mathcal{P}^*_\omega((k_2/S_T)^{1/2}\|\bar{X}^{\pi*}_{m_T} - \mathrm{E}^*[\bar{X}^{\pi*}_{m_T}]\| > \delta) > \xi) < \xi$.

Write $\mathcal{K}_{tT} = X_{tT}/S_T^{1/2}$, $(t = 1, ..., T)$, $\mathcal{K}^{\pi}_{t^*_s T} = X^{\pi}_{t^*_s T}/S_T^{1/2}$, $(s = 1, ..., m_T)$, and $\bar{\mathcal{K}}_T = \sum_{t=1}^T \mathcal{K}_{tT}/T$, $\bar{\mathcal{K}}^{\pi}_T = \sum_{t=1}^T \pi_{tT}\mathcal{K}_{tT}$. Without loss of generality, set $\mathrm{E}^*[\bar{X}^{\pi*}_{m_T}] = 0$. Hence, $\bar{\mathcal{K}}^{\pi}_T = 0$ and $\bar{\mathcal{K}}_T = 0$ since $\bar{\mathcal{K}}^{\pi}_T = \sum_{t=1}^T \pi_{tT}\mathcal{K}_{tT} = (1 + o_p(1))\bar{\mathcal{K}}_T$ by Assumption 3.2(b). First, note that

$$\begin{aligned}\mathrm{E}^*[\|\mathcal{K}^{\pi}_{t^*_s T}\|] &= \sum_{t=1}^T \pi_{tT}\|\mathcal{K}_{tT}\| = (1 + o_p(1))\frac{1}{T}\sum_{t=1}^T \|\frac{1}{S_T}\sum_{s=t-T}^{t-1} k(\frac{s}{S_T})X_{t-s}\|/(k_2)^{1/2} \\ &\le O_p(1)\frac{1}{T}\sum_{t=1}^T \|X_t\| = O_p(1),\end{aligned}$$

uniformly, $(s = 1, ..., m_T)$, by WLLN, by $\max_{1\le t\le T} |T\pi_{tT} - 1| = o_p(1)$ of Assumption 3.2(b) and $\mathrm{E}[\sup_{\theta\in\Theta} \|X_t(\theta)\|^\alpha] < \infty$, $\alpha > v$ from Assumption 2.3(d). Also, for any $\delta > 0$,

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^T \|\mathcal{K}_{tT}\| - \frac{1}{T}\sum_{t=1}^T \|\mathcal{K}_{tT}\|\mathbb{I}(\|\mathcal{K}_{tT}\| &< m_T\delta) = \frac{1}{T}\sum_{t=1}^T \|\mathcal{K}_{tT}\|\mathbb{I}(\|\mathcal{K}_{tT}\| \ge m_T\delta) \\ &\le \frac{1}{T}\sum_{t=1}^T \|\mathcal{K}_{tT}\| \max_t \mathbb{I}(\|\mathcal{K}_{tT}\| \ge m_T\delta).\end{aligned}$$

Now, by M,

$$\max_t \|\mathcal{K}_{tT}\| = O(1)\max_t \|X_t\| = O_p(T^{1/\alpha});$$

cf. Newey and Smith (2004, Proof of Lemma A1, p.239). Hence, since, by hypothesis, $T^{1/\alpha}/m_T = o(1)$, $\max_t \mathbb{I}(\|\mathcal{K}_{tT}\| \ge m_T\delta) = o_p(1)$ and $\sum_{t=1}^T \|\mathcal{K}_{tT}\|/T = O_p(1)$,

$$\mathrm{E}^*[\|\mathcal{K}^{\pi}_{t^*_s T}\|\mathbb{I}(\|\mathcal{K}^{\pi}_{t^*_s T}\| \ge m_T\delta)] = (1 + o_p(1))\frac{1}{T}\sum_{t=1}^T \|\mathcal{K}_{tT}\|\mathbb{I}(\|\mathcal{K}_{tT}\| \ge m_T\delta) = o_p(1). \tag{A.1}$$

The remaining part of the proof is similar to that for Khinchine's WLLN given in Rao (1973, pp.112-114). For each s define the pair of random variables

$$V^{\pi}_{t^*_s T} = \mathcal{K}^{\pi}_{t^*_s T}\mathbb{I}(\|\mathcal{K}^{\pi}_{t^*_s T}\| < m_T\delta), W^{\pi}_{t^*_s T} = \mathcal{K}^{\pi}_{t^*_s T}\mathbb{I}(\|\mathcal{K}^{\pi}_{t^*_s T}\| \ge m_T\delta),$$

so that $\mathcal{K}^{\pi}_{t^*_s T} = V^{\pi}_{t^*_s T} + W^{\pi}_{t^*_s T}$, $(s = 1, ..., m_T)$. Write $\bar{V}^{\pi*}_{m_T} = \sum_{s=1}^{m_T} V^{\pi}_{t^*_s T}/m_T$, $\bar{W}^{\pi*}_{m_T} = \sum_{s=1}^{m_T} W^{\pi}_{t^*_s T}/m_T$ and $\bar{\mathcal{K}}^{\pi*}_{m_T} = \sum_{s=1}^{m_T} \mathcal{K}^{\pi}_{t^*_s T}/m_T$. Now

$$\mathrm{var}^*[\|\bar{V}^{\pi*}_{m_T}\|] \le \mathrm{E}^*[\|V^{\pi}_{t^*_s T}\|^2]/m_T \le \delta\mathrm{E}^*[\|V^{\pi}_{t^*_s T}\|]. \tag{A.2}$$

Thus, from eq. (A.2), using C,

$$\begin{aligned}\mathcal{P}^*_\omega(\|\bar{V}^{\pi*}_{m_T} - \mathrm{E}^*[V^{\pi}_{t^*_s T}]\| &\ge \varepsilon) \le \frac{\mathrm{var}^*[\|\bar{V}^{\pi*}_{m_T}\|]}{\varepsilon^2} \\ &\le \frac{\delta\mathrm{E}^*[\|V^{\pi}_{t^*_s T}\|]}{\varepsilon^2}.\end{aligned}$$

Also, $\|\bar{\mathcal{K}}_T - \mathrm{E}^*[V^{\pi}_{t_s^*T}]\| = o_p(1)$, i.e., for any $\varepsilon > 0$, T large enough, $\|\bar{\mathcal{K}}_T - \mathrm{E}^*[V^{\pi}_{t_s^*T}]\| \leq \varepsilon$, since by T, noting $\mathrm{E}^*[V^{\pi}_{t_s^*T}] = \sum_{t=1}^{T} \pi_{tT}\mathcal{K}_{tT}\mathbb{I}(|\mathcal{K}_{tT}| < m_T\delta)/T$,

$$\begin{aligned}\|\bar{\mathcal{K}}_T - \mathrm{E}^*[V^{\pi}_{t_s^*T}]\| &= \|\frac{1}{T}\sum_{t=1}^{T} \mathcal{K}_{tT} - (1+o_p(1))\frac{1}{T}\sum_{t=1}^{T} \mathcal{K}_{tT}\mathbb{I}(|\mathcal{K}_{tT}| < m_T\delta)\| \\ &\leq \frac{1}{T}\sum_{t=1}^{T} \|\mathcal{K}_{tT}\|\mathbb{I}(\|\mathcal{K}_{tT}\| \geq m_T\delta) + o_p(1) = o_p(1)\end{aligned}$$

from eq. (A.1). Hence, for T large enough,

$$\mathcal{P}^*_\omega(\|\bar{V}^{\pi*}_{m_T} - \bar{\mathcal{K}}_T\| \geq 2\varepsilon) \leq \frac{\delta \mathrm{E}^*[\|V^{\pi}_{t_s^*T}\|]}{\varepsilon^2}. \tag{A.3}$$

By M,

$$\begin{aligned}\mathcal{P}^*_\omega(W^{\pi}_{t_s^*T} &\neq 0) = \mathcal{P}^*_\omega(\|\mathcal{K}^{\pi}_{t_s^*T}\| \geq m_T\delta) \\ &\leq \frac{1}{m_T\delta}\mathrm{E}^*[\|\mathcal{K}^{\pi}_{t_s^*T}\|\mathbb{I}(\|\mathcal{K}^{\pi}_{t_s^*T}\| \geq m_T\delta)] \leq \frac{\delta}{m_T}.\end{aligned}$$

To see this, $\mathrm{E}^*[\|\mathcal{K}^{\pi}_{t_s^*T}\|\mathbb{I}(\|\mathcal{K}^{\pi}_{t_s^*T}\| \geq m_T\delta)] = o_p(1)$ from eq. (A.1). Thus, for T large enough, $\mathrm{E}^*[\|\mathcal{K}^{\pi}_{t_s^*T}\|\mathbb{I}(\|\mathcal{K}^{\pi}_{t_s^*T}\| \geq m_T\delta)] \leq \delta^2$ w.p.a.1. Write $\bar{W}^{\pi*}_{m_T} = \sum_{s=1}^{m_T} W^{\pi}_{t_s^*T}/m_T$. Thus, from eq. (A.3),

$$\mathcal{P}^*_\omega(\bar{W}^{\pi*}_{m_T} \neq 0) \leq \sum_{s=1}^{m_T} \mathcal{P}^*_\omega(W^{\pi}_{t_s^*T} \neq 0) \leq \delta. \tag{A.4}$$

Therefore,

$$\begin{aligned}\mathcal{P}^*_\omega(\|\bar{\mathcal{K}}^{\pi*}_{m_T} - \bar{\mathcal{K}}_T\| &\geq 4\varepsilon) \leq \mathcal{P}^*_\omega(\|\bar{V}^{\pi*}_{m_T} - \bar{\mathcal{K}}_T\| + \|\bar{W}^{\pi*}_{m_T}\| \geq 4\varepsilon) \\ &\leq \mathcal{P}^*_\omega(\|\bar{V}^{\pi*}_{m_T} - \bar{\mathcal{K}}_T\| \geq 2\varepsilon) + \mathcal{P}^*_\omega(\|\bar{W}^{\pi*}_{m_T}\| \geq 2\varepsilon) \\ &\leq \frac{\delta E^*[\|\bar{V}^{\pi}_{t_s^*T}\|]}{\varepsilon^2} + \mathcal{P}^*_\omega(\|\bar{W}^{*}_{m_T}\| \neq 0) \leq \frac{\delta E^*[\|V^{\pi}_{t_s^*T}\|]}{\varepsilon^2} + \delta.\end{aligned}$$

where the first inequality follows from T, the third from eq. (A.3) and the final inequality from eq. (A.4). Since δ may be chosen arbitrarily small enough and $\mathrm{E}^*[\|V^{\pi}_{t_s^*T}\|] \leq \mathrm{E}^*[\|\mathcal{K}^{\pi}_{t_s^*T}\|] = O_p(1)$, the result follows by M noting $\bar{\mathcal{K}}_T = 0$ by hypothesis.■

Lemma A.3. (GEL-KBB Global UWL.) Suppose Assumptions 2.1, 2.2 and 2.3(a)(b)(e) and 3.2 are satisfied. Then, if $T^{1/\alpha}/m_T \to 0$ and $E[\sup_{\theta\in\Theta}\|X_t(\theta)\|^\alpha] < \infty$ for some $\alpha > v$, for $S_T \to \infty$ and $S_T = o(T^{1/2})$,

$$\textbf{(a)}\ \sup_{\theta\in\Theta}\|\bar{X}^{\pi*}_{m_T}(\theta)/S_T^{1/2} - \bar{X}(\theta)/k^{1/2}\| \to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P},$$

$$\textbf{(b)}\ \sup_{\theta\in\Theta}\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi}_T(\theta)\|/S_T^{1/2} \to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P}.$$

Proof. Similarly to Proof of Lemma A.2, from Lemma A.1, the first result **(a)** follows from **(b)** since $\sup_{\theta\in\Theta}\|\bar{X}_T(\theta) - \bar{X}^{\pi}_T(\theta)\|/S_T^{1/2} = o_p(1)$. Result **(b)** is proven if

$$\sup_{\theta\in\Theta}\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} \to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P}.$$

The following preliminary results are useful in the later analysis. By global Lipschitz continuity of $X_t(\cdot)$, Assumption 2.3(e), and by T, for T large enough,

$$\begin{aligned} \|\bar{X}_T(\theta)) - \bar{X}_T(\theta^0)\|/S_T^{1/2} &\leq \frac{1}{T}\sum\nolimits_{t=1}^{T}\frac{1}{S_T}\sum\nolimits_{s=t=T}^{y-1}|k\left(\frac{s}{S_T}\right)|\|X_{t-s}(\theta) - X_{t-s}(\theta^0)\|/(k_2)^{1/2} \quad \text{(A.5)} \\ &= \frac{1}{T}\sum\nolimits_{t=1}^{T}\|X_t(\theta) - X_t(\theta^0)\||\frac{1}{S_T}\sum\nolimits_{s=1-t}^{T-t}k\left(\frac{s}{S_T}\right)|/(k_2)^{1/2} \\ &\leq C\|\theta-\theta^0\|\frac{1}{T}\sum\nolimits_{t=1}^{T}L_t \end{aligned}$$

since, for some $0 < C < \infty$, $|\sum_{s=1-t}^{T-t}k(s/S_T)/S_T| \leq O(1) < C$ uniformly t for large enough T, see Smith (2011, eq. (A.5), p.1218). Hence, by M, from Assumption 2.3(e),

$$\begin{aligned} \mathcal{P}^*_\omega(\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi*}_{m_T}(\theta^0)\|/S_T^{1/2} &> \varepsilon) \leq \frac{1}{\varepsilon}\mathrm{E}^*[\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi*}_{m_T}(\theta^0)\|/S_T^{1/2}] \quad \text{(A.6)} \\ &= \frac{1}{\varepsilon}\|\bar{X}^{\pi}_T(\theta) - \bar{X}^{\pi}_T(\theta^0)\|/S_T^{1/2} \\ &\leq \frac{C}{\varepsilon}(1+o_p(1))\|\theta-\theta^0\|\frac{1}{T}\sum\nolimits_{t=1}^{T}L_t. \end{aligned}$$

The remaining part of the proof is identical to Gonçalves and White (2000, Proof of Lemma A.2, pp.30-31) and is given here for completeness; cf. Hall and Horowitz (1996, Proof of Lemma 8, p.913). Given $\varepsilon > 0$, let $\{\eta(\theta_i, \varepsilon), (i = 1, ..., I)\}$ denote a finite subcover of Θ where $\eta(\theta_i, \varepsilon) = \{\theta \in \Theta\text{: } \|\theta - \theta_i\| < \varepsilon\}$, $(i = 1, ..., I)$. Now

$$\sup_{\theta\in\Theta}\|\bar{X}^{\pi*}_{m_T}(\theta)) - \bar{X}_T(\theta)\|/S_T^{1/2} = \max_{i=1,\dots,I}\sup_{\theta\in\eta(\theta_i,\varepsilon)}\|\bar{X}^{\pi*}_{m_T}(\theta)) - \bar{X}_T(\theta)\|/S_T^{1/2}.$$

The argument $\omega \in \Omega$ is omitted for brevity as in Gonçalves and White (2000). It then follows that, for any $\delta > 0$ (and any fixed ω),

$$\mathcal{P}^*_\omega(\sup_{\theta\in\Theta}\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} > \delta) \leq \sum\nolimits_{i=1}^{I}\mathcal{P}^*_\omega(\sup_{\theta\in\eta(\theta_i,\epsilon)}\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} > \delta).$$

For any $\theta \in \eta(\theta_i, \varepsilon)$, by T,

$$\begin{aligned} \|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} &\leq \|\bar{X}^{\pi*}_{m_T}(\theta_i) - \bar{X}_T(\theta_i)\|/S_T^{1/2} + \|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi*}_{m_T}(\theta_i)\|/S_T^{1/2} \\ &\quad +\|\bar{X}_T(\theta) - \bar{X}_T(\theta_i)\|/S_T^{1/2}. \end{aligned}$$

Hence, for any $\delta > 0$ and $\xi > 0$,

$$\begin{aligned} \mathcal{P}(\mathcal{P}^*_\omega(\sup_{\theta\in\eta(\theta_i,\varepsilon)}\|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}_T(\theta)\|/S_T^{1/2} > \delta) > \xi) &\leq \mathcal{P}(\mathcal{P}^*_\omega(\|\bar{X}^{\pi*}_{m_T}(\theta_i) - \bar{X}_T(\theta_i)\|/S_T^{1/2} > \frac{\delta}{3}) > \frac{\xi}{3}) \\ &+\mathcal{P}(\mathcal{P}^*_\omega(\sup_{\theta\in\eta(\theta_i,\varepsilon)}\|\bar{X}^{\pi*}_{m_T}(\theta)) - \bar{X}^{\pi*}_{m_T}(\theta_i)\|/S_T^{1/2} > \frac{\delta}{3}) > \frac{\xi}{3}) \\ &+\mathcal{P}(\sup_{\theta\in\eta(\theta_i,\varepsilon)}\|\bar{X}_T(\theta)) - \bar{X}_T(\theta_i)\|/S_T^{1/2} > \frac{\delta}{3}). \quad \text{(A.7)} \end{aligned}$$

By Lemma A.2

$$\mathcal{P}(\mathcal{P}^*_\omega(\|\bar{X}^{\pi*}_{m_T}(\theta_i) - \bar{X}_T(\theta_i)\|/S_T^{1/2} > \frac{\delta}{3}) > \frac{\xi}{3}) < \frac{\xi}{3}$$

for large enough T. Also, by M (for fixed ω) and Assumption 2.3(e), noting $L_t \geq 0$, $(t = 1, ..., T)$, from eq. (A.6),

$$\begin{aligned}
\mathcal{P}^*_\omega(\sup_{\theta\in\eta(\theta_i,\varepsilon)} \|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi*}_{m_T}(\theta_i)\|/S_T^{1/2} &> \frac{\delta}{3}) \leq \frac{3C^*\varepsilon}{\delta}(1 + o_p(1))\frac{1}{T}\sum\nolimits_{t=1}^{T} L_{tT} \\
&\leq \frac{3C^*\varepsilon}{\delta}(1 + o_p(1))\frac{1}{T}\sum\nolimits_{t=1}^{T} L_t \frac{1}{S_T}\sum\nolimits_{s=t=T}^{t-1} |k\left(\frac{s}{S_T}\right)|/(k_2)^{1/2} \\
&\leq \frac{3C^*\varepsilon}{\delta}(1 + o_p(1))O(1)\frac{1}{T}\sum\nolimits_{t=1}^{T} L_t.
\end{aligned}$$

As a consequence, for any $\delta > 0$ and $\xi > 0$, for T sufficiently large,

$$\begin{aligned}
\mathcal{P}(\mathcal{P}^*_\omega(\sup_{\theta\in\eta(\theta_i,\varepsilon)} \|\bar{X}^{\pi*}_{m_T}(\theta)) - \bar{X}^{\pi*}_{m_T}(\theta_i)\|/S_T^{1/2} &> \frac{\delta}{3}) > \frac{\xi}{3}) \leq \mathcal{P}(\frac{3C^*\varepsilon}{\delta}\frac{1}{T}\sum\nolimits_{t=1}^{T} L_t > \frac{\xi}{3}) \\
&= \mathcal{P}(\frac{1}{T}\sum\nolimits_{t=1}^{T} L_t > \frac{\xi\delta}{9C^*\varepsilon}) \\
&\leq \frac{9C^*\varepsilon}{\xi\delta}\mathrm{E}[\frac{1}{T}\sum\nolimits_{t=1}^{T} L_t] \\
&\leq \frac{9C^*\varepsilon\Delta}{\xi\delta} < \frac{\xi}{3}
\end{aligned}$$

for the choice $\varepsilon < \xi^2\delta/27C^*\Delta$, where, since, by hypothesis $\mathrm{E}[\sum_{t=1}^T L_t/T] = O(1)$, the second and third inequalities follow respectively from M and Δ a sufficiently large but finite constant such that $\sup_T \mathrm{E}[\sum_{t=1}^T L_t/T] < \Delta$. Similarly, from eq. (A.5), for any $\delta > 0$ and $\xi > 0$, by Assumption 2.3(e), $P(\sup_{\theta\in\eta(\theta_i,\varepsilon)} \|\bar{X}_T(\theta) - \bar{X}_T(\theta_i)\|/S_T^{1/2} > \delta/3) \leq P(C\varepsilon\sum_{t=1}^T L_t/T > \delta/3) \leq 3C\varepsilon\Delta/\delta < \xi/3$ for T sufficiently large for the choice $\varepsilon < \xi\delta/9C\Delta$.

Therefore, from eq. (A.7), the conclusion of the Lemma follows if

$$\varepsilon = \frac{\xi\delta}{9\Delta}\max\left(\frac{1}{C}, \frac{\xi}{3C^*}\right). \blacksquare$$

Let $\mathcal{N}$ denote a compact neighbourhood of θ_0.

Lemma A.4. (GEL-KBB Local UWL.) Suppose Assumptions 2.1, 2.2 and 2.3(a)(b)(e) and 3.2 are satisfied. Then, if $T^{1/\alpha}/m_T \to 0$ and $E[\sup_{\theta\in\Theta}\|X_t(\theta)\|^\alpha] < \infty$ for some $\alpha > v$, for $S_T \to \infty$ and $S_T = o(T^{1/2})$,

(a) $\sup_{\theta\in\mathcal{N}} \|\bar{X}^{\pi*}_{m_T}(\theta)/S_T^{1/2} - \bar{X}(\theta)/k^{1/2}\| \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$,

(b) $\sup_{\theta\in\mathcal{N}} \|\bar{X}^{\pi*}_{m_T}(\theta) - \bar{X}^{\pi}_T(\theta)\|/S_T^{1/2} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$.

Proof. The Proof of Lemma A.3 may be simply adapted by replacing Θ by $\mathcal{N}$.$\blacksquare$

Without loss of generality, set $\mathrm{E}[X_t(\theta_0)] = 0$.

Lemma A.5. (GEL-KBB CLT.) Let Assumptions 2.1,2.2(b)-(d), 2.3(a)(b)(d)(f) and 3.2 hold. Then, if $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $\frac{1}{6} < \eta < \frac{1}{2}$,

$$\sup_{x\in\mathcal{R}} |\mathcal{P}^*_\omega((T/S_T)^{1/2}(\bar{X}^{\pi*}_{m_T}(\theta_0) - \bar{X}^{\pi}_T(\theta_0)) \le x) - \mathcal{P}(T^{1/2}\bar{X}_T(\theta_0) \le x)| \to 0, \text{ prob-}\mathcal{P}.$$

Proof. The result is proven in Steps 1-5 below; cf. Politis and Romano (1992, Proof of Theorem 2, pp. 1994-5). To simplify exposition, let $m_T = T/S_T$ be integer, $d_q = 1$ and suppress the argument θ_0.

Step 1. $\bar{X}_T \to 0$, prob-$\mathcal{P}$. Follows by Smith (2011, Lemma A.1, p.1217) and $\mathrm{E}[X_t] = 0$.

Step 2. $\mathcal{P}(\Xi^{-1/2}T^{1/2}\bar{X}_T \le x/(k_2)^{1/2}) \to \Phi(x)$, where $\Phi(\cdot)$ is the standard normal distribution function. Follows by Smith (2011, Lemma A.2, p.1219).

Step 3. $\sup_x |\mathcal{P}(\Xi^{-1/2}T^{1/2}\bar{X}_T \le x/(k_2)^{1/2}) - \Phi(x)| \to 0$. Follows by Pólya's Theorem, Serfling (1980, Theorem 1.5.3, p.18), from Step 2 and the continuity of $\Phi(\cdot)$.

Step 4. $\mathrm{var}^*[m_T^{1/2}\bar{X}^{\pi*}_{m_T}] \to \Xi$, prob-$\mathcal{P}$. Note $\mathrm{E}^*[\bar{X}^{\pi*}_{m_T}] = \bar{X}^{\pi}_T$. Thus,

$$\begin{aligned}
\mathrm{var}^*[m_T^{1/2}\bar{X}^{\pi*}_{m_T}] &= \mathrm{var}^*[X^{\pi}_{t^*T}] \\
&= \sum\nolimits_{t=1}^T \pi_{tT}(X_{tT} - \bar{X}^{\pi}_T)^2 \\
&= (1 + o_p(1))(\frac{1}{T}\sum\nolimits_{t=1}^T (X_{tT})^2 - (\bar{X}_T)^2).
\end{aligned}$$

The result follows since $(\bar{X}_T)^2 = O_p(S_T/T)$, Smith (2011, Lemma A.2, p.1219), $S_T = o(T^{1/2})$ by hypothesis and $\sum_{t=1}^T (X_{tT})^2/T \to \Xi$, prob-$\mathcal{P}$, Smith (2011, Lemma A.3, p.1219).

Step 5.

$$\lim_{T\to\infty} \mathcal{P}(\sup_x |\mathcal{P}^*_\omega(\frac{\bar{X}^{\pi*}_{m_T} - \mathrm{E}^*[\bar{X}^{\pi*}_{m_T}]}{\mathrm{var}^*[\bar{X}^{\pi*}_{m_T}]^{1/2}} \le x) - \Phi(x)| \ge \varepsilon) = 0.$$

Applying the Berry-Esséen inequality, Serfling (1980, Theorem 1.9.5, p.33), noting the bootstrap sample observations $\{X^{\pi}_{t^*_s T}\}_{s=1}^{m_T}$ are independently distributed,

$$\sup_x |\mathcal{P}^*_\omega(\frac{m_T^{1/2}(\bar{X}^{\pi*}_{m_T} - \bar{X}^{\pi}_T)}{\mathrm{var}^*[m_T^{1/2}\bar{X}^{\pi*}_{m_T}]^{1/2}} \le x) - \Phi(x)| \le \frac{C}{m_T^{1/2}}\sum\nolimits_{s=1}^{m_T} \mathrm{E}^*[\|X^{\pi}_{t^*_s T} - \bar{X}^{\pi}_T\|^3]\mathrm{var}^*[\sum\nolimits_{s=1}^{m_T} X^{\pi}_{t^*_s T}]^{-3/2}.$$

Now $\mathrm{var}^*[m_T^{1/2}\bar{X}^{\pi*}_{m_T}] \to \Xi > 0$, prob-$\mathcal{P}$, from Step 4. Furthermore, $\mathrm{E}^*[\|X^{\pi}_{t^*_s T} - \bar{X}^{\pi}_T\|^3] = \sum_{t=1}^T \pi_{tT}\|X_{tT} - \bar{X}^{\pi}_T\|^3 = (1 + o_p(1))\sum_{t=1}^T \|X_{tT} - \bar{X}^{\pi}_T\|^3/T$ and

$$\begin{aligned}
\frac{1}{T}\sum\nolimits_{t=1}^T \|X_{tT} - \bar{X}^{\pi}_T\|^3 &\le \max_{1\le t\le T}\|X_{tT} - \bar{X}^{\pi}_T\|\frac{1}{T}\sum\nolimits_{t=1}^T \|X_{tT} - \bar{X}^{\pi}_T\|^2 \\
&= O_p(S_T^{1/2}T^{1/\alpha}).
\end{aligned}$$

The equality follows since

$$\begin{aligned}
\max_{1\le t\le T}\|X_{tT} - \bar{X}^{\pi}_T\| &\le \max_{1\le t\le T}\|X_{tT}\| + \|\bar{X}^{\pi}_T\| \\
&= O_p(S_T^{1/2}T^{1/\alpha}) + O_p((S_T/T)^{1/2}) = O_p(S_T^{1/2}T^{1/\alpha})
\end{aligned}$$

by M and Assumption 2.3(d), cf. Newey and Smith (2004, Proof of Lemma A1, p.239), and $\sum_{t=1}^{T}\|X_{tT}-\bar{X}_T^{\pi}\|^2/T=O_p(1)$, see the Proof of Step 4 above. Therefore

$$\begin{aligned}\sup_x|\mathcal{P}_\omega^*\{\frac{(T/S_T)^{1/2}(\bar{X}_{m_T}^{\pi*}-\bar{X}_T^{\pi})}{\text{var}^*[(T/S_T)^{1/2}\bar{X}_{m_T}^{\pi*}]^{1/2}} &\leq x\}-\Phi(x)|\leq\frac{1}{m_T^{1/2}}O_p(1)O_p(S_T^{1/2}T^{1/\alpha})\\ &= \frac{S_T^{1/2}}{m_T^{1/2}}O_p(T^{1/\alpha})=o_p(1),\end{aligned}$$

by hypothesis, yielding the required conclusion.■

Appendix B: Proofs for GEL Implied Probability GMM-KBB

The indices t_s^* and the consequent bootstrap sample $q_{t_s^*T}^{\pi}(\beta)$, $(s=1,...,m_T)$, denote m_T independent draws with replacement from the index set $\mathcal{T}_T=\{1,...,T\}$ and the bootstrap sample space $\{q_{tT}(\theta)\}_{t=1}^{T}$ with sampling probabilities $\mathcal{P}_\omega^*(q_{t_s^*T}^{\pi}(\theta)=q_{tT}(\theta))=\pi_{tT}$, $(t=1,...,T)$, with $m_T=[T/S_T]$ the integer part of T/S_T.

B.1 Notation

Recall $q_{tT}(\theta)=\sum_{s=t-T}^{t-1}k(s/S_T)q_t(\theta)/(k_2S_T)^{1/2}$, $(t=1,...,T)$, $\hat{q}(\theta)=\sum_{t=1}^{T}q_t(\theta)/T$ and $\hat{q}_T(\theta)=\sum_{t=1}^{T}q_{tT}(\theta)/T$. Also recall $\hat{q}_T^{\pi}(\theta)=\sum_{t=1}^{T}\pi_{tT}q_{tT}(\theta)$ and $\hat{q}_{m_T}^{\pi*}(\theta)=\sum_{s=1}^{m_T}q_{t_s^*T}^{\pi}(\theta)/m_T$.

Additionally, recall the restricted π-GEL GMM and π-GEL GMM-KBB Lagrangeans $\breve{\mathcal{L}}_T^{\pi}(\theta)=\breve{\mathcal{Q}}_T^{\pi}(\theta)/S_T-2\mu'r(\theta)/k$ and $\breve{\mathcal{L}}_{m_T}^{\pi*}(\theta)=\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta)/S_T-2\mu'r(\theta)/k$, where $\breve{\mathcal{Q}}_T^{\pi}(\theta)=\hat{q}_T^{\pi}(\theta)'(W_{qT})^{-1}\hat{q}_T^{\pi}(\theta)$ and $\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta)=\hat{q}_{m_T}^{\pi*}(\theta)'(W_{qT}^{\pi*})^{-1}\hat{q}_{m_T}^{\pi*}(\theta)$, with Lagrange multipliers μ associated with the parametric constraint $r(\theta)=0$. The restricted π-GEL GMM and π-GEL GMM-KBB estimators are $\breve{\theta}_T^{\pi}=\arg\min_{\theta\in\Theta_r}\breve{\mathcal{Q}}_T^{\pi}(\theta)$ and $\breve{\theta}_{m_T}^{\pi*}=\arg\min_{\theta\in\Theta_r}\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta)$ with π-GEL GMM and π-GEL GMM-KBB Lagrange multiplier estimators $\breve{\mu}_T^{\pi}$ and $\breve{\mu}_{m_T}^{\pi*}$. Also recall the GEL criteria corresponding to (2.1) and (2.2) $\hat{\mathcal{P}}_{gT}^{\rho}(\beta,\lambda_g)=\sum_{t=1}^{T}(\rho(\lambda_g'g_{tT}(\beta)/k^{1/2})-\rho_0)/T$ and $\hat{\mathcal{P}}_{qT}^{\rho}(\theta,\lambda_q)=\sum_{t=1}^{T}\rho(\lambda_q'q_{tT}(\theta)/k^{1/2})-\rho_0)/T$ respectively and the corresponding GEL estimators of λ_o, $(o=g,q)$, $\breve{\lambda}_{gT}=\arg\sup_{\lambda_g\in\Lambda_{gT}}\hat{\mathcal{P}}_{gT}^{\rho}(\breve{\beta}_T,\lambda_g)$ and $\breve{\lambda}_{qT}=\arg\sup_{\lambda_q\in\Lambda_{qT}}\hat{\mathcal{P}}_{qT}^{\rho}(\breve{\theta}_T,\lambda_q)$, where the parameter spaces Λ_{oT}, $(o=g,q)$, are defined in Assumption 3.1.

Recall the matrix definitions relevant for restricted GMM estimation and inference $M_{W_q}=(Q'W_q^{-1}Q+R'R)^{-1}$, $K_{W_q}=M_{W_q}-M_{W_q}R'(RM_{W_q}R')^{-1}RM_{W_q}$, $H_{W_q}=K_{W_q}Q'W_q^{-1}$, $J_{W_q}=(RM_{W_q}R')^{-1}RM_{W_q}Q'W_q^{-1}$ and $P_{W_q}=W_q^{-1}-W_q^{-1}QK_{W_q}Q'W_q^{-1}$ and those relevant for unrestricted GMM estimation and inference $K_{W_g}=M_{W_g}=(G'W_g^{-1}G)^{-1}$, $H_{W_g}=K_{W_g}G'W_g^{-1}$, $J_{W_g}=0$ and $P_{W_g}=W_g^{-1}-W_g^{-1}GK_{W_g}G'W_g^{-1}$.

Let $\hat{Q}_T^{\pi}(\theta)=\partial\hat{q}_T^{\pi}(\theta)/\partial\theta'$, $\hat{Q}_{m_T}^{\pi*}(\theta)=\partial\hat{q}_{m_T}^{\pi*}(\theta)/\partial\theta'$ and the (d_g,d_q) selection matrix $S_g=(I_{d_g},0)'$, i.e., $S_g'q_{tT}=g_{tT}$, $(t=1,...,T)$.

B.2 Useful Algebraic Results

$$\textbf{(a) } Q'W_q^{-1}Q=(M_{W_q})^{-1}-R'R;\ \textbf{(b) } H_{W_q}Q=K_{W_q}(M_{W_q})^{-1};$$
$$\textbf{(c) } K_{W_q}R'=0,\ K_{W_q}(M_{W_q})^{-1}K_{W_q}=K_{W_q};\ \textbf{(d) } J_{W_q}Q=((RM_{W_q}R')^{-1}-I_{d_r})R,\ J_{W_q}W_qJ'_{W_q}=(RM_{W_q}R')^{-1}-I_{d_r};$$
$$\textbf{(e) } P_{W_q}Q=J'_{W_q}R.$$

and

$$\textbf{(a) } H_{W_g}G=I_{d_g};\ \textbf{(b) } P_{W_g}G=0.$$

Hence,

$$H_{W_q}QK_\Xi=K_\Xi \text{ and } H_{W_g}GK_\Sigma=K_\Sigma.$$

B.3 GEL Implied Probabilities

Proposition B.1. If Assumptions 2.1, 2.2, 2.3(a)-(d)(f) and 3.1 hold, $\breve{\lambda}_{qT}\to 0$, prob-$\mathcal{P}$, and, if, in addition, Assumption 2.4 also holds, $(T/S_T)^{1/2}\breve{\lambda}_{qT}/k^{1/2}+\Xi^{-1}W_qP_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}$.

Remark B.1. Cf. Smith (2011, Theorem 5.1, p.1210, and eq. (B.16), p.1229) with the GMM estimator $\breve{\theta}_T$ substituting for $\tilde{\theta}_T$. Correspondingly, $(T/S_T)^{1/2}\breve{\lambda}_{gT}/k^{1/2}+\Sigma^{-1}W_gP_{W_g}S'_g(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}$; cf. Smith (2011, Theorems 2.3, p.1201, and eq. (B.2), p.1225).

In this subsection we primarily deal with the restricted implied probabilities $\pi_{tT}=\breve{\pi}_{tT}$ (3.2), $(t=1,...,T)$. Consequently Assumption 3.2 is automatically satisfied as elucidated in the next Lemma.

Lemma B.1. If Assumptions 2.1, 2.2, 2.3(a)-(d)(f), 2.4 and 3.1 are satisfied, then $\max_{1\le t\le T}|T\breve{\pi}_{tT}-1|\to 0$, prob-$\mathcal{P}$, and

$$(T/S_T)^{1/2}(\breve{\pi}_{tT}-\frac{1}{T})=\frac{1}{T}q_{tT}(\breve{\theta}_T)'(T/S_T)^{1/2}\breve{\lambda}_{qT}(1+o_p(1))/k^{1/2}+O_p((S_T/T^3)^{1/2}),\ \text{uniformly } (t=1,...,T). \tag{B.1}$$

Proof. Cf. Smith (2011, eq. (B.4), p.1226).■

Remark B.2. Correspondingly, for unrestricted implied probabilities $\pi_{tT}=\breve{\pi}_{tT}$ (3.2), $(t=1,...,T)$, $\max_{1\le t\le T}|T\breve{\pi}_{tT}-1|\to 0$, prob-$\mathcal{P}$, and

$$(T/S_T)^{1/2}(\breve{\pi}_{tT}-\frac{1}{T})=\frac{1}{T}g_{tT}(\breve{\beta}_T)'(T/S_T)^{1/2}\breve{\lambda}_{gT}(1+o_p(1))/k^{1/2}+O_p((S_T/T^3)^{1/2}),\ \text{uniformly } (t=1,...,T). \tag{B.2}$$

Remark B.3. For the efficient restricted and unrestricted efficient implied probabilities, $\tilde{\pi}_{tT}$ and $\hat{\pi}_{tT}$ (3.3), $(t=1,...,T)$, $\max_{1\le t\le T}|T\tilde{\pi}_{tT}-1|\to 0$, prob-$\mathcal{P}$,

$$(T/S_T)^{1/2}(\tilde{\pi}_t-\frac{1}{T})=\frac{1}{T}q_{tT}(\tilde{\theta}_T)'(T/S_T)^{1/2}\tilde{\lambda}_{qT}(1+o_p(1))/k^{1/2}+O_p((S_T/T^3)^{1/2}),\ \text{uniformly } (t=1,...,T), \tag{B.3}$$

and $\max_{1\leq t\leq T}|T\hat{\pi}_{tT}-1|\to 0$,

$$(T/S_T)^{1/2}(\hat{\pi}_{tT}-\frac{1}{T})=\frac{1}{T}g_{tT}(\hat{\beta}_T)'(T/S_T)^{1/2}\hat{\lambda}_{gT}(1+o_p(1))/k^{1/2}+O_p((S_T/T^3)^{1/2}),\text{ uniformly }(t=1,...,T). \tag{B.4}$$

Let $\hat{q}^{\breve{\pi}}_T(\cdot)=\sum_{t=1}^T\breve{\pi}_{tT}q_{tT}(\cdot)$.

COROLLARY B.1. Under Assumptions 2.1, 2.2, 2.3(a)-(d)(f), 2.4 and 3.1, then

$$(T/S_T)^{1/2}\hat{q}^{\breve{\pi}}_T(\theta_0)=(T/S_T)^{1/2}\hat{q}_T(\theta_0)-\sigma W_qP_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0)+o_p(1),\sigma\in\{0,1\}. \tag{B.5}$$

PROOF. Follows directly from Proposition B.1 and Lemma B.1.■

REMARK B.4. For unrestricted implied probabilities $\pi_{tT}=\breve{\pi}_{tT}$ (3.2), $(t=1,...,T)$,

$$(T/S_T)^{1/2}\hat{q}^{\breve{\pi}}_T(\theta_0)=(T/S_T)^{1/2}\hat{q}_T(\theta_0)-\sigma S_gW_gP_{W_g}S_g'(T/S_T)^{1/2}\hat{q}_T(\theta_0)+o_p(1),\sigma\in\{0,1\}. \tag{B.6}$$

REMARK B.5. Setting $\sigma=0$ in eqs. (B.5) and (B.6) gives Corollary B.1 and Remark B.4 for standard GMM weighting T^{-1}.

Let $a_t(\theta)=a(z_t,\theta)$, $a_{tT}(\theta)=\sum_{s=t-T}^{t-1}k(s/S_T)a_t(\theta)/(k_2S_T)^{1/2}$, $(t=1,...,T)$, and define $\hat{a}^{\breve{\pi}}_T(\theta)=\sum_{t=1}^T\breve{\pi}_{tT}a_{tT}(\theta)$, $\hat{a}_T(\theta)=\sum_{t=1}^Ta_{tT}(\theta)/T$ and $\hat{a}(\theta)=\sum_{t=1}^Ta_t(\theta)/T$.

Lemma B.2 below mirrors Smith (2011, Theorem 3.1, p.1206).

LEMMA B.2. Let $T^{1/2}(\dot{\theta}_T-\theta_0)\to 0$, prob-$\mathcal{P}$. Let Assumptions 2.1, 2.2, 2.3(a)-(d)(f), 2.4 and 3.1 suitably modified be satisfied for $(a_t(\theta)',q_t(\theta)')'$ jointly. Then, **(a)** if $\mathrm{E}[a(z_t,\theta_0)]\neq 0$,

$$\hat{a}^{\breve{\pi}}_T(\dot{\theta}_T)/S_T^{1/2}=\hat{a}(\dot{\theta}_T)/k^{1/2}-\sigma B_{aq}\Xi^{-1}W_qP_{W_q}\hat{q}(\theta_0)/k^{1/2}+o_p(1),\sigma\in\{0,1\},\text{ or} \tag{B.7}$$

(b) if $\mathrm{E}[a(z_t,\theta_0)]=0$,

$$(T/S_T)^{1/2}\hat{a}^{\breve{\pi}}_T(\dot{\theta}_T)=T^{1/2}\hat{a}(\dot{\theta}_T)/k^{1/2}-\sigma B_{aq}\Xi^{-1}W_qP_{W_q}T^{1/2}\hat{q}(\theta_0)/k^{1/2}+o_p(1),\sigma\in\{0,1\}, \tag{B.8}$$

where $B_{aq}=\sum_{s=-\infty}^{\infty}\mathrm{E}[a_t(\theta_0)q_{t-s}(\theta_0)']$. Additionally, if $a_t(\theta)=q_t(\theta)$, $(t=1,...,T)$,

$$(T/S_T)^{1/2}\hat{q}^{\breve{\pi}}_T(\dot{\theta}_T)=T^{1/2}\hat{q}(\dot{\theta}_T)/k^{1/2}-\sigma W_qP_{W_q}T^{1/2}\hat{q}(\theta_0)/k^{1/2}+o_p(1),\sigma\in\{0,1\}. \tag{B.9}$$

PROOF. First, using Lemma B.1,

$$\begin{aligned}(T/S_T)^{1/2}\hat{a}^{\breve{\pi}}_T(\dot{\theta}_T) &= (T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T)+(1+o_p(1))\sigma\frac{1}{T}\sum_{t=1}^Ta_{tT}(\dot{\theta}_T)q_{tT}(\breve{\theta}_T)'(T/S_T)^{1/2}\breve{\lambda}_{qT}/k^{1/2}\\ &\quad+O_p((S_T/T^2))(T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T).\end{aligned}$$

Secondly, by Proposition B.1,

$$\begin{aligned}(T/S_T)^{1/2}\hat{a}^{\breve{\pi}}_T(\dot{\theta}_T) &= (T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T)\\ &\quad-(1+o_p(1))\sigma\frac{1}{T}\sum_{t=1}^Ta_{tT}(\dot{\theta}_T)q_{tT}(\breve{\theta}_T)'(\Xi^{-1}W_qP_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0)+o_p(1))\\ &\quad+O_p((S_T/T^2))(T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T).\end{aligned}$$

By UWL Lemma A.1, $\sum_{t=1}^{T} a_{tT}(\dot{\theta}_T) q_{tT}(\breve{\theta}_T)'/T \rightarrow B_{aq}$, prob-$\mathcal{P}$, cf. Smith (2011, Lemma A.7, p.1223). Now, as **(a)** $\hat{a}_T(\dot{\theta}_T)/S_T^{1/2} = O_p(1)$ if $\mathrm{E}[a(z_t, \theta_0)] \neq 0$, cf. Smith (2011, Lemma A.1, p.1217),

$$O_p(S_T/T^{1/2}))\hat{a}_T(\dot{\theta}_T)/S_T^{1/2} = O_p(S_T/T^{1/2}) = o_p(1),$$

or **(b)** $(T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T) = O_p(1)$ if $\mathrm{E}[a(z_t, \theta_0)] = 0$, cf. Smith (2011, Lemma A.2, p.1219),

$$O_p(S_T/T)(T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T) = O_p(S_T/T) = o_p(1).$$

Thus, as **(a)** $\hat{a}_T(\dot{\theta}_T)/S_T^{1/2} = (1 + o(1))\hat{a}(\dot{\theta}_T)/k^{1/2} + O_p((S_T/T)^{1/2})$ or **(b)** $(T/S_T)^{1/2}\hat{a}_T(\dot{\theta}_T) = (1 + o(1))T^{1/2}\hat{a}(\dot{\theta}_T)/k^{1/2} + O_p((S_T/T)^{1/2})$, cf. Smith (2011, eq. (A.11), p.1219), either

$$\textbf{(a)}\ \hat{a}_T^{\breve{\pi}}(\dot{\theta}_T)/S_T^{1/2} = \hat{a}(\dot{\theta}_T)/k^{1/2} - \sigma B_{aq}\Xi^{-1}W_q P_{W_q}\hat{q}(\theta_0)/k^{1/2} + o_p(1) \text{ or}$$

$$\textbf{(b)}\ (T/S_T)^{1/2}\hat{a}_T^{\breve{\pi}}(\dot{\theta}_T) = T^{1/2}\hat{a}(\dot{\theta}_T)/k^{1/2} - \sigma B_{aq}\Xi^{-1}W_q P_{W_q}T^{1/2}\hat{q}(\theta_0)/k^{1/2} + o_p(1).$$

Finally, if $a_t(\theta) = q_t(\theta)$, $(t = 1, ..., T)$, then $\hat{a}_T^{\breve{\pi}}(\theta) = \hat{q}_T^{\breve{\pi}}(\theta)$, $\hat{a}(\theta) = \hat{q}(\theta)$ and $B_{aq} = \Xi$. Hence, since $(T/S_T)^{1/2}\hat{q}_T(\theta_0) = (1 + o(1))T^{1/2}\hat{q}(\theta_0)/k^{1/2} + O_p((S_T/T)^{1/2})$ from above,

$$(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\dot{\theta}_T) = T^{1/2}\hat{q}(\dot{\theta}_T)/k^{1/2} - \sigma W_q P_{W_q}T^{1/2}\hat{q}(\theta_0)/k^{1/2} + o_p(1).\blacksquare$$

Remark B.6. For unrestricted implied probabilities $\pi_{tT} = \breve{\pi}_{tT}$ (3.2), $(t = 1, ..., T)$,

(a) if $\mathrm{E}[a(z_t, \theta_0)] \neq 0$, $\hat{a}_T^{\breve{\pi}}(\dot{\theta}_T)/S_T^{1/2} = \hat{a}(\dot{\theta}_T)/k^{1/2} - \sigma B_{aq}S_g\Sigma^{-1}W_g P_{W_g}S_g' T^{1/2}\hat{q}(\theta_0)/k^{1/2} + o_p(1)$ or

(b) if $\mathrm{E}[a(z_t, \theta_0)] = 0$, $(T/S_T)^{1/2}\hat{a}_T^{\breve{\pi}}(\dot{\theta}_T) = T^{1/2}\hat{a}(\dot{\theta}_T)/k^{1/2} - \sigma B_{aq}S_g\Sigma^{-1}W_g P_{W_g}S_g' T^{1/2}\hat{q}(\theta_0)/k^{1/2} + o_p(1)$,

and

$$(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\dot{\theta}_T) = T^{1/2}\hat{q}(\dot{\theta}_T)/k^{1/2} - \sigma\Xi S_g\Sigma^{-1}W_g P_{W_g}S_g' T^{1/2}\hat{q}(\theta_0)/k^{1/2} + o_p(1).$$

Thus, if $T^{1/2}(\ddot{\beta}_T - \beta_0) \rightarrow 0$, prob-$\mathcal{P}$,

$$(T/S_T)^{1/2}\hat{g}_T^{\breve{\pi}}(\ddot{\beta}_T) = T^{1/2}\hat{g}(\ddot{\beta}_T)/k^{1/2} - \sigma W_g P_{W_g}T^{1/2}\hat{g}(\beta_0)/k^{1/2} + o_p(1).$$

Remark B.7. The above results are straightforwardly specialised for efficient unrestricted and restricted implied probabilities (3.3) by the substitution of P_Σ (Σ) for $\Sigma^{-1}W_g P_{W_g}$ (W_g) and P_Ξ (Ξ) for $\Xi^{-1}W_q P_{W_q}$ (W_q) respectively.

B.4 GEL Implied Probability GMM Estimation

Proof of Proposition 3.2. **(a)** By T and CS, $|\breve{\mathcal{Q}}_T^{\pi}(\theta) - \breve{\mathcal{Q}}_T(\theta)| \leq \|\hat{q}_T^{\pi}(\theta) - \hat{q}_T(\theta)\|^2/S_T \cdot \|W_{qT}\|^{-1} + 2\|\hat{q}_T^{\pi}(\theta) - \hat{q}_T(\theta)\|/S_T^{1/2} \cdot \|\hat{q}_T(\theta)\|/S_T^{1/2} \cdot \|W_{qT}\|^{-1}$. Assumptions 2.3(a)(b) imply the compactness of the restricted parameter space Θ_r. Now $\sup_{\theta\in\Theta_r}\|\hat{q}_T^{\pi}(\theta) - \hat{q}_T(\theta)\|/S_T^{1/2} \rightarrow 0$, prob-$\mathcal{P}$, by Assumption 3.2(b) and Proof of UWL Lemma A.1. Therefore $\sup_{\theta\in\Theta_r}|\breve{\mathcal{Q}}_T^{\pi}(\theta) - \breve{\mathcal{Q}}_T(\theta)| \rightarrow 0$, prob-$\mathcal{P}$, as $\|W_{qT} - W_q\| \rightarrow 0$, prob-$\mathcal{P}$. The result follows from Proposition 2.1(a).

(b) From the first order conditions, employing UWL Lemma A.1 and GEL-KBB Global UWL Lemma B.3, $Q'W_q^{-1}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\breve{\theta}_T^{\pi}) - R'T^{1/2}\breve{\mu}_T^{\pi}/k^{1/2} \to 0$, prob-$\mathcal{P}$. So, pre-multiplying by RM_{W_q}, $J_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\breve{\theta}_T^{\pi}) - T^{1/2}\breve{\mu}_T^{\pi}/k^{1/2} \to 0$ and, thus, $H_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\breve{\theta}_T^{\pi}) \to 0$, prob-$\mathcal{P}$. After substituting $(T/S_T)^{1/2}\hat{q}_T^{\pi}(\breve{\theta}_T^{\pi}) - (T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) - (Q/k^{1/2})T^{1/2}(\breve{\theta}_T^{\pi} - \theta_0) \to 0$ and since $J_{W_q}Q = ((RM_{W_q}R')^{-1} - I_{d_r})R$, noting $RT^{1/2}(\breve{\theta}_T^{\pi} - \theta_0) \to 0$, prob-$\mathcal{P}$, we conclude $T^{1/2}(\breve{\theta}_T^{\pi} - \theta_0)/k^{1/2} + H_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) \to 0$, prob-$\mathcal{P}$, cf. Smith (2011, eq. (B.15), p.1229), and $T^{1/2}\breve{\mu}_T^{\pi}/k^{1/2} - J_{W_q}(T/S_T)^{1/2}\hat{q}_T^{\pi}(\theta_0) \to 0$, prob-$\mathcal{P}$, cf. Smith (2011, eq. (B.17), p.1229).■

Proof of Corollary 3.1. Substituting $(T/S_T)^{1/2}\hat{q}_T^{\breve{\pi}}(\theta_0) - QH_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$, see Lemma B.2(b) with $\sigma = 1$, and noting $H_{W_q}QK_{W_q} = K_{W_q}$, $T^{1/2}(\breve{\theta}_T^{\breve{\pi}} - \theta_0)/k^{1/2} + H_{W_q}(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}$. The second result follows since $RH_{W_q} = 0$.■

B.5 GEL Implied Probability GMM-KBB Estimation

Proof of Theorem 3.1. The proof verifies the conditions of Gonçalves and White (2004, Lemma A.2, p.212). To do so, replace n by T, $Q_T(\cdot, \theta)$ by $\breve{\mathcal{Q}}_T(\theta)$ and $Q_T^*(\cdot, \omega, \theta)$ by $\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta)$. Conditions (a1) and (a2) hold under Assumptions 2.1 and 2.3(a)(b). Condition (a3) $\sup_{\theta\in\Theta_r} |\breve{\mathcal{Q}}_T(\theta) - \breve{\mathcal{Q}}_0(\theta)/k|/S_T \to 0$, prob-$\mathcal{P}$, where the population GMM criterion $\breve{\mathcal{Q}}_0(\theta) = \mathrm{E}[q_t(\theta)]'W_q\mathrm{E}[q_t(\theta)]$, follows as in Newey and McFadden (1994, Proof of Theorem 2.6, p.2132) from the UWL Lemma A.1 $\sup_{\theta\in\Theta_r} \|\hat{q}_T(\theta)/S_T^{1/2} - \hat{q}(\theta)/k^{1/2}\| \to 0$, prob-$\mathcal{P}$, $\mathrm{E}[\sup_{\theta\in\Theta_r} \|q_t(\theta)\|] < \infty$ from Assumption 2.3(d), $\|W_{qT} - W_q\| \to 0$, prob-$\mathcal{P}$, W_q p.d. by hypothesis and $\breve{\mathcal{Q}}_0(\theta)$ uniquely minimised at θ_0 by Assumption 2.3(c). Hence $\breve{\theta}_T - \theta_0 \to 0$, prob-$\mathcal{P}$.

Conditions (b1) and (b2) follow from Assumptions 2.1 and 2.3(a)(b). To prove (b3) $\sup_{\theta\in\Theta_r} |\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta) - \breve{\mathcal{Q}}_T(\theta)|/S_T \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, from T and CS,

$$\begin{aligned}|\breve{\mathcal{Q}}_{m_T}^{\pi*}(\theta) - \breve{\mathcal{Q}}_T(\theta)|/S_T \quad \le \quad & \|\hat{q}_{m_T}^{\pi*}(\beta) - \hat{q}_T(\theta)\|^2/S_T \cdot \|W_{qm_T}^{\pi*}\|^{-1} \\ & +2\|\hat{q}_T(\theta)\|/S_T^{1/2} \cdot \|\hat{q}_{m_T}^{\pi*}(\theta) - \hat{q}_T(\theta)\|/S_T^{1/2} \cdot \|W_{qm_T}^{\pi*}\|^{-1} \\ & +\|\hat{q}_T(\theta)\|^2/S_T \cdot \|W_{qT}\|^{-1}\|W_{qm_T}^{\pi*} - W_{qT}\|\|W_{qm_T}^{\pi*}\|^{-1}.\end{aligned}$$

Now by GEL-KBB UWL Lemma A.3 $\sup_{\theta\in\Theta_r} \|\hat{q}_{m_T}^{\pi*}(\theta) - \hat{q}_T(\theta)\|/S_T^{1/2} \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$. Also, $\sup_{\theta\in\Theta_r} \|\hat{q}_T(\theta)\|/S_T^{1/2} \le \sup_{\theta\in\Theta_r} \|\hat{q}_T(\theta)/S_T^{1/2} - \mathrm{E}[q_t(\theta)]/k^{1/2}\| + \sup_{\theta\in\Theta_r} \|\mathrm{E}[q_t(\theta)]\|/k^{1/2} = O_p(1)$ by UWL Lemma A.1 and Assumption 2.3(d). The result then follows since $\|W_{qm_T}^{\pi*} - W_{qT}\| \to 0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, by hypothesis.■

Proof of Theorem 3.2. Extensive use is made of GEL-KBB Local UWL Lemma A.4 and Assumption 2.3(b), i.e., $\hat{Q}_{m_T}^{\pi*}(\bar{\theta}_{m_T}^{\pi*})/S_T^{1/2} \to Q/k^{1/2}$ and $R(\bar{\theta}_{m_T}^{\pi*}) \to R$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, for any $\bar{\theta}_{m_T}^{\pi*} \to \theta_0$, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$.

From Theorem 3.1, by Assumptions 2.3(c) and 2.4(c), $\breve{\theta}_{m_T}^{\pi*} \in \mathcal{N}$ w.p.a.1, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$. Thus, the first order conditions for $\breve{\theta}_{m_T}^{\pi*}$ are satisfied with equality w.p.a.1, prob-$\mathcal{P}_\omega^*$, prob-$\mathcal{P}$, i.e., $\hat{Q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})'(W_{qm_T}^{\pi*})^{-1}\hat{q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})/S_T - R(\breve{\theta}_{m_T}^{\pi*})'\breve{\mu}_{m_T}^{\pi*}/k = 0$. Pre-multiplying by RM_{W_q}, $RM_{W_q}(\hat{Q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*})/S_T^{1/2})'(W_{qm_T}^{\pi*})^{-1}(T/S_T)^{1/2}\hat{q}_{m_T}^{\pi*}(\breve{\theta}_{m_T}^{\pi*}) -$

$RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})'T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k = 0$. Since $R(\breve{\theta}^{\pi*}_{m_T})-R \to 0$, $RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})'$ is p.d. w.p.a.1, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and, thus, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$,

$$T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k - (RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})')^{-1}RM_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T}) \to 0. \quad \text{(B.10)}$$

Substituting back, $(M_{W_q}-M_{W_q}R(\breve{\theta}^{\pi*}_{m_T})'(RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})')^{-1}RM_{W_q})(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T}) = 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, which, together with the Taylor expansion $(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T}) = (T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T})+$ $(\hat{Q}^{\pi*}_{m_T}(\bar{\theta}^{\pi*}_{m_T})/S_T^{1/2})T^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T})$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, where $\bar{\theta}^{\pi*}_{m_T}$ is on the line segment joining $\breve{\theta}^{\pi*}_{m_T}$ and $\breve{\theta}^{\pi}_{T}$, and $RM_{W_q}Q'(W_q)^{-1}Q = (I_{d_r} - RM_{W_q}R')R$, yields

$$K_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T}) + T^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T})/k \to 0, \quad \text{(B.11)}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, noting $RT^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T}) \to 0$ from the first order conditions, $W^{\pi*}_{qm_T} \to W_q$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, by hypothesis, and extensive use of GEL-KBB Local UWL Lemma A.4.

Now

$$\begin{aligned}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T}) &= (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_{T}(\theta_0)) \\ &\quad + (T/S_T)^{1/2}\hat{q}^{\pi}_{T}(\theta_0) + (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T}) - \hat{q}^{\pi*}_{m_T}(\theta_0)).\end{aligned} \quad \text{(B.12)}$$

By another Taylor expansion, of $(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T})$ about θ_0,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T}) - \hat{q}^{\pi*}_{m_T}(\theta_0)) = (\hat{Q}^{\pi*}_{m_T}(\dot{\theta}^{\pi}_{T})/S_T^{1/2})T^{1/2}(\breve{\theta}^{\pi}_{T}-\theta_0), \quad \text{(B.13)}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, where $\dot{\theta}^{\pi}_{T}$ lies on the line segment joining $\breve{\theta}^{\pi}_{T}$ and θ_0. Since $H_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi}_{T}(\breve{\theta}^{\pi}_{T}) \to 0$ and $(T/S_T)^{1/2}(\hat{q}^{\pi}_{T}(\breve{\theta}^{\pi}_{T}) - \hat{q}^{\pi}_{T}(\theta_0)) - QT^{1/2}(\breve{\theta}^{\pi}_{T}-\theta_0)/k^{1/2} \to 0$, prob-$\mathcal{P}$, cf. Proof of Proposition 3.2(b), back-substitution using GEL-KBB Local UWL Lemma A.4 yields

$$K_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}((T/S_T)^{1/2}\hat{q}^{\pi}_{T}(\theta_0) + (\hat{Q}^{\pi*}_{m_T}(\dot{\theta}^{\pi}_{T})/S_T^{1/2})T^{1/2}(\breve{\theta}^{\pi}_{T}-\theta_0)) \to 0, \quad \text{(B.14)}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. So, substituting eq. (B.13) into eq. (B.12) and thence into eq. (B.11), noting eq. (B.14),

$$T^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T})/k + K_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_{T}(\theta_0)) \to 0, \quad \text{(B.15)}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. By GEL-KBB CLT Lemma A.5, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_{T}(\theta_0)) - (T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Hence, since $\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2} \to Q/k^{1/2}$ from GEL-KBB Local UWL Lemma A.4, $W^{\pi*}_{qm_T} \to W_q$ by hypothesis, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and $(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to^{d_{\mathcal{P}}} N(0, \Xi/k)$, prob-$\mathcal{P}$, $T^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T})$ converges in distribution to $N(0, H_{W_q}\Xi H'_{W_q})$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$.

Substitution from the Taylor expansion for $(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})$ below eq. (B.10) yields $T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k-$ $(RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})')^{-1}RM_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_{T}) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, noting from the first order conditions $J_{W_q}QT^{1/2}(\breve{\theta}^{\pi*}_{m_T}-\breve{\theta}^{\pi}_{T}) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Similarly, from the Taylor expansion eq. (B.13), $T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k-(RM_{W_q}R(\breve{\theta}^{\pi*}_{m_T})')^{-1}RM_{W_q}(\hat{Q}^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(W^{\pi*}_{qm_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\theta_0)\to 0$ or $T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k^{1/2}-J_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. As $T^{1/2}\breve{\mu}^{\pi}_T/k^{1/2}-J_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi}_T(\theta_0)\to 0$, prob-$\mathcal{P}$, cf. Proof of Proposition 3.2(b), again recalling GEL-KBB CLT Lemma A.5, $T^{1/2}(\breve{\mu}^{\pi*}_{m_T}-\breve{\mu}^{\pi}_T)$ converges in distribution to $N(0,J_{W_q}\Xi J'_{W_q})$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$.

Theorem 3.2 then follows by Pólya's Theorem, Serfling (1980, Theorem 1.5.3, p.18), and continuity of the normal distribution c.d.f. ■

Proof of Corollary 3.2. Immediate from eq. (B.11) and first order condition $H_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)\to 0$, prob-$\mathcal{P}$, and $T^{1/2}\breve{\mu}^{\pi*}_{m_T}/k^{1/2}-J_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\breve{\theta}^{\pi}_T)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, cf. Proof of Theorem 3.2, and first order condition $T^{1/2}\breve{\mu}^{\pi}_T/k^{1/2}-J_{W_q}(T/S_T)^{1/2}\hat{q}^{\pi}_T(\breve{\theta}^{\pi}_T)\to 0$, prob-$\mathcal{P}$, cf. Proof of Proposition 3.2(b). ■.

Proof of Corollary 3.3. From eq. (B.13) and, below, $(T/S_T)^{1/2}(\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)-\hat{q}^{\pi}_T(\theta_0))-QT^{1/2}(\tilde{\theta}^{\pi}_T-\theta_0)/k^{1/2}\to 0$, prob-$\mathcal{P}$,

$$Q'\Xi^{-1}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T)-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))-Q'\Xi^{-1}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))\to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P}.$$

Since, from GEL-KBB CLT Lemma A.5, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))-(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$, as Q f.c.r., $(Q'\Xi^{-1}Q)^{-1/2}Q'\Xi^{-1}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))$ converges in distribution to $N(0,I_{d_\theta}/k)$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Therefore,

$$\begin{aligned}\sup_{x\in\mathcal{R}^{d_\theta}}|\mathcal{P}^*_\omega((Q'\Xi^{-1}Q)^{-1/2}Q'\Xi^{-1}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T)-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) &\leq x/k^{1/2})\\ -\mathcal{P}((Q'\Xi^{-1}Q)^{-1/2}Q'\Xi^{-1}(T/S_T)^{1/2}\hat{q}_T(\theta_0) &\leq x/k^{1/2})|\to 0,\ \text{prob-}\mathcal{P},\end{aligned}$$

follows by Pólya's Theorem, Serfling (1980, Theorem 1.5.3, p.18), and the continuity of the normal c.d.f. $\Phi(\cdot)$ recalling $\sup_x\left|\mathcal{P}(\Xi^{-1/2}(T/S_T)^{1/2}\hat{q}_T(\theta_0)\leq x/k^{1/2})-\Phi(x)\right|\to 0$ from Step 3 of the Proof of GEL-KBB CLT Lemma A.5. ■

B.6 GEL Implied Probability GMM-KBB Inference

Lemma B.3. Let Assumptions 2.1-2.4 and 3.2 hold. Then, if $\Xi^{\pi*}_{m_T}-\Xi_T\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Xi^{\pi*}_{m_T}$ p.s.d., and $\Xi_T-\Xi\to 0$, prob-$\mathcal{P}$, Ξ_T p.s.d.,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))-\Xi P_\Xi(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P}.$$

Proof. Using GEL-KBB Local UWL A.4, setting $W^{\pi*}_{qm_T}=\Xi^{\pi*}_{m_T}$ in (B.11), $T^{1/2}(\tilde{\theta}^{\pi*}_{m_T}-\tilde{\theta}^{\pi}_T)/k+K_\Xi(\hat{Q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(\Xi^{\pi*}_{m_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Also, by a Taylor expansion and GEL-KBB local UWL Lemma A.4,

$$\begin{aligned}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T)) &= (\hat{Q}^{\pi*}_{m_T}(\bar{\theta}^{\pi*}_T)/S_T^{1/2})T^{1/2}(\tilde{\theta}^{\pi*}_{m_T}-\tilde{\theta}^{\pi}_T)\\ &= QT^{1/2}(\tilde{\theta}^{\pi*}_{m_T}-\tilde{\theta}^{\pi}_T)/k^{1/2},\end{aligned}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, where $\bar{\theta}^{\pi*}_T$ is on the line segment joining $\tilde{\theta}^{\pi*}_T$ and $\tilde{\theta}^{\pi}_T$. Hence, substituting for $T^{1/2}(\tilde{\theta}^{\pi*}_{m_T}-\tilde{\theta}^{\pi}_T)$,

$$(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \Xi P_\Xi (T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T) \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}. \tag{B.16}$$

Now, by Taylor expansions, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T) - \hat{q}^{\pi*}_{m_T}(\theta_0)) - (\hat{Q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi*}_T)/S_T^{1/2})T^{1/2}(\tilde{\theta}^{\pi}_T-\theta_0) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and $(T/S_T)^{1/2}(T/S_T)^{1/2}(\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\theta_0)) - (\hat{Q}^{\pi}_T(\dot{\theta}_T)/S_T^{1/2})T^{1/2}(\tilde{\theta}^{\pi}_T-\theta_0) \to 0$, prob-$\mathcal{P}$, where $\ddot{\theta}^{\pi*}_T$ and $\dot{\theta}_T$ lie on the line segment joining $\tilde{\theta}^{\pi}_T$ and θ_0. Thus,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi}_T) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}, \tag{B.17}$$

since $\hat{Q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi*}_T)/S_T^{1/2}$, $\hat{Q}^{\pi}_T(\dot{\theta}_T))/S_T^{1/2} \to Q/k^{1/2}$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, by GEL-KBB Local UWL Lemma A.4 and UWL Lemma A.1. Therefore, combining eqs. (B.16) and (B.17), since $H_\Xi(T/S_T)^{1/2}\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T) \to 0$, prob-$\mathcal{P}$,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi P_\Xi (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) \to 0,$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. By GEL-KBB CLT Lemma A.5 $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) - (T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and the conclusion of the Lemma follows.■

Recall the alternative restricted π-GEL GMM estimator $\dot{\theta}^{\pi}_T$ defined by $\dot{\theta}^{\pi}_T = \arg\min_{\theta\in\Theta_r} \tilde{\mathcal{Q}}^{\pi}_T(\theta)$, with associated π-GEL GMM Lagrangean $\dot{\mathcal{L}}^{\pi}_T(\theta) = \tilde{\mathcal{Q}}^{\pi}_T(\theta)/S_T - 2\mu'(r(\theta) - r(\alpha, \hat{\beta}^{\pi}_T))/k$, Lagrange multiplier estimator $\dot{\mu}^{\pi}_T$, and π-GEL GMM criterion $\tilde{\mathcal{Q}}^{\pi}_T(\theta) = \hat{q}^{\pi}_T(\theta)'(\Xi_T)^{-1}\hat{q}^{\pi}_T(\theta)$. Additionally, recall $\ddot{\theta}^{\pi}_T = (\dot{\alpha}^{\pi\prime}_T, \hat{\beta}^{\pi\prime}_T)'$ and the corresponding restricted π-GEL GMM-KBB estimator $\dot{\theta}^{\pi*}_{m_T} = \arg\min_{\theta\in\Theta_r} \tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta)$, with π-GEL GMM-KBB Lagrangean $\tilde{\mathcal{L}}^{\pi*}_{m_T}(\theta) = \tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta)/S_T - 2\mu'(r(\theta) - r(\ddot{\theta}^{\pi}_T)/k$, Lagrange multiplier estimator $\ddot{\mu}^{\pi*}_{m_T}$, and π-GEL GMM-KBB criterion $\tilde{\mathcal{Q}}^{\pi*}_{m_T}(\theta) = \hat{q}^{\pi*}_{m_T}(\theta)'(\Xi^{\pi*}_{qm_T})^{-1}\hat{q}^{\pi*}_{m_T}(\theta)$.

Lemma B.4. Let Assumptions 2.1-2.4 and 3.2 be satisfied. Then, if $\Xi^{\pi*}_{m_T} - \Xi_T \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, $\Xi^{\pi*}_{m_T}$ p.s.d., and $\Xi_T - \Xi \to 0$, prob-$\mathcal{P}$, Ξ_T p.s.d.,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\dot{\theta}^{\pi*}_{m_T}) - \Xi P_\Xi \hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T)) - \Xi P_\Xi (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

Proof. The Proof replicates the steps in the Proof of Lemma B.3 above. Noting $RT^{1/2}(\dot{\theta}^{\pi*}_{m_T} - \ddot{\theta}^{\pi}_T) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, from the first order conditions, $T^{1/2}(\dot{\theta}^{\pi*}_{m_T} - \ddot{\theta}^{\pi}_T)/k + K_\Xi(\hat{Q}^{\pi*}_{m_T}(\dot{\theta}^{\pi*}_{m_T})/S_T^{1/2})'(\Xi^{\pi*}_{m_T})^{-1}(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Also, by a Taylor expansion and GEL-KBB Local UWL Lemma A.4,

$$\begin{aligned}(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\dot{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T)) &= (\hat{Q}^{\pi*}_{m_T}(\bar{\theta}^{\pi*}_T)/S_T^{1/2})T^{1/2}(\dot{\theta}^{\pi*}_{m_T} - \ddot{\theta}^{\pi}_T) \\ &= QT^{1/2}(\dot{\theta}^{\pi*}_{m_T} - \ddot{\theta}^{\pi}_T)/k^{1/2},\end{aligned}$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, where $\bar{\theta}^{\pi*}_T$ is on the line segment joining $\dot{\theta}^{\pi*}_T$ and $\ddot{\theta}^{\pi}_T$. Hence, substituting for $QT^{1/2}(\dot{\theta}^{\pi*}_{m_T} - \ddot{\theta}^{\pi}_T)$,

$$(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\dot{\theta}^{\pi*}_{m_T}) - \Xi P_\Xi (T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T) \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P},$$

cf. eq. (B.16). Similar Taylor expansions to those in the Proof of Lemma B.3 yield $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T)-\hat{q}^{\pi*}_{m_T}(\theta_0))-(\hat{Q}^{\pi*}_{m_T}(\theta^{\pi*}_T)/S_T^{1/2})T^{1/2}(\ddot{\theta}^{\pi}_T-\theta_0)\to 0$ and $(T/S_T)^{1/2}(\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T)-\hat{q}^{\pi}_T(\theta_0))-(\hat{Q}^{\pi}_T(\acute{\theta}_T)/S_T^{1/2})T^{1/2}(\ddot{\theta}^{\pi}_T-\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, where $\theta^{\pi*}_T$ and $\acute{\theta}_T$ lie on the line segment joining $\ddot{\theta}^{\pi}_T$ and θ_0. Hence,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T)-\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))-(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))\to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P},$$

cf. (B.17), since $\hat{Q}^{\pi*}_{m_T}(\theta^{\pi*}_T)/S_T^{1/2}$, $\hat{Q}^{\pi}_T(\acute{\theta}_T))/S_T^{1/2}\to Q/k^{1/2}$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, by GEL-KBB local UWL Lemma A.4. Therefore, substituting for $(T/S_T)^{1/2}\hat{q}^{\pi*}_{m_T}(\ddot{\theta}^{\pi}_T)$, by a Taylor expansion,

$$(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\dot{\theta}^{\pi*}_{m_T})-\Xi P_\Xi\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T))-\Xi P_\Xi(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))\to 0,\ \text{prob-}\mathcal{P}^*_\omega,\ \text{prob-}\mathcal{P},$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. The result is then immediate by GEL-KBB CLT Lemma A.5, i.e., $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0)-\hat{q}^{\pi}_T(\theta_0))-(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$.■

B.6.1 Overidentification Tests

Proof of Theorem 3.3. Replacing Ξ, P_Ξ by Σ, P_Σ, $\hat{q}^{\pi*}_{m_T}(\cdot)$, $\hat{q}^{\pi}_T(\cdot)$, $\hat{q}_T(\cdot)$ by $\hat{g}^{\pi*}_{m_T}(\cdot)$, $\hat{g}^{\pi}_T(\cdot)$, $\hat{g}_T(\cdot)$ and $\tilde{\theta}^{\pi*}_{m_T}$, $\tilde{\theta}^{\pi}_T$ by $\hat{\beta}^{\pi*}_{m_T}$, $\hat{\beta}^{\pi}_T$, in Lemma B.3, $(T/S_T)^{1/2}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))-\Sigma P_\Sigma(T/S_T)^{1/2}\hat{g}_T(\beta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Likewise $(T/S_T)^{1/2}\hat{g}_T(\hat{\beta}_T)-\Sigma P^\Sigma(T/S_T)^{1/2}\hat{g}_T(\beta_0)\to 0$, prob-$\mathcal{P}$. Therefore, as $(T/S_T)^{1/2}\hat{g}_T(\beta_0)\to^{d_\mathcal{P}} N(0,\Sigma/k)$, prob-$\mathcal{P}$, $\mathcal{J}^{\pi*}_{m_T}$ converges in distribution to $N(0,\Sigma P_\Sigma\Sigma)'\times\Sigma^{-1}\times N(0,\Sigma P_\Sigma\Sigma)=N(0,\Sigma)'\times P_\Sigma\times N(0,\Sigma)\sim\chi^2(d_g-d_\beta)$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and the conclusion follows by Polya's Theorem, Serfling (1980, Theorem 1.5.3, p.18), noting $\mathcal{J}_T\to^{d_\mathcal{P}}\chi^2(d_g-d_\beta)$, prob-$\mathcal{P}$, and the continuity of the $\chi^2(d_g-d_\beta)$ c.d.f.■

B.6.2 Specification Tests

Proof of Theorem 3.4. $\mathcal{LR}^{\pi*}_{m_T}$: From Lemma B.3 $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))-\Xi P_\Xi(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$ and $(T/S_T)^{1/2}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))-\Sigma P_\Sigma(T/S_T)^{1/2}\hat{g}_T(\beta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Now, as $\Xi^{\pi*}_{m_T}\to\Xi$, $\Sigma^{\pi*}_{m_T}\to\Sigma$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, by hypothesis, $\mathcal{LR}^{\pi*}_{m_T}/k-(T/S_T)\hat{q}_T(\theta_0)'(P_\Xi-S_gP_\Sigma S_g')\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, noting $P_\Xi\Xi S_gP_\Sigma=S_gP_\Sigma$. Since $\mathcal{LR}_T/k-(T/S_T)\hat{q}_T(\theta_0)'(P_\Xi-S_gP_\Sigma S_g')\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}$, and $(T/S_T)\hat{q}_T(\theta_0)'(P_\Xi-S_gP_\Sigma S_g')\hat{q}_T(\theta_0)\to N(0,\Xi/k)'\times(P^\Xi-S_gP^\Sigma S_g')\times N(0,\Xi/k)\sim\chi^2(d_r+(d_q-d_g)-(d_\theta-d_\beta))/k$, prob-$\mathcal{P}$, see Rao and Mitra (1971, Theorem 9.2.1, p.171); cf. Smith (2011, p.1229). Hence, cf. Smith (2011, eq. (B.18), p.1229), the claim follows by Polya's Theorem and the continuity of the χ^2 c.d.f.■

$\mathcal{D}^{\pi*}_{m_T}$: From the Proof of Theorem 3.4 for $\mathcal{LR}^{\pi*}_{m_T}$, $(T/S_T)(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))'(\Xi^{\pi*}_{m_T}))^{-1}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))-(T/S_T)\hat{q}_T(\theta_0)'P_\Xi\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Similarly, $(T/S_T)(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))'(\Sigma^{\pi*}_{m_T})^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})-\hat{g}^{\pi}_T(\hat{\beta}^{\pi}_T))-(T/S_T)\hat{q}_T(\theta_0)'S_gP_\Sigma S_g'\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. The equivalence with $\mathcal{LR}^{\pi*}_{m_T}$ then follows.■

$\mathcal{S}^{\pi*}_{m_T}$: Similarly, from Lemma B.3, $((\Xi^{\pi*}_{m_T})^{-1}-S_gP^{\pi*}_{\Sigma m_T}S_g')(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})-\hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T))-(P_\Xi-S_gP_\Sigma S_g')(T/S_T)^{1/2}\hat{q}_T(\theta_0)\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. The conclusion is shown since, cf. $\mathcal{LR}^{\pi*}_{m_T}$, $(T/S_T)\hat{q}_T(\theta_0)'(P_\Xi-$

$S_g P_\Sigma S_g')\hat{q}_T(\theta_0) \to \chi^2(d_r + (d_q - d_g) - (d_\theta - d_\beta))/k$, prob-$\mathcal{P}$.■

$\mathcal{LM}^{\pi*}_{m_T}$: Let

$$\Upsilon_\Xi = \begin{pmatrix} \Xi & 0 & Q \\ 0 & 0 & R \\ Q' & R' & 0 \end{pmatrix}.$$

Define the $(d_q + d_r + d_\theta, d_h + d_r)$ selection matrix $S^\theta_{h,\mu} = (S'_{h,\mu}, 0')'$. Then $(S^\theta_{h,\mu})'(\Upsilon_\Xi)^{-1}(S^\theta_{h,\mu}) = S'_{h,\mu}\Psi_\Xi S_{h,\mu}$, see Smith (2011, p.1230). From a Taylor expansion of $\breve{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T})$ about $\tilde{\theta}^{\pi*}_{m_T}$, $(T/S_T)^{1/2}\breve{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - (T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - S_g\Sigma^{-1}\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})) - QT^{1/2}(\hat{\theta}^{\pi*}_{m_T} - \tilde{\theta}^{\pi*}_{m_T})/k^{1/2} \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Then, from the first order conditions for $\tilde{\theta}^{\pi*}_{m_T}$, noting $\breve{q}^{\pi*}_{m_T}(\cdot) = (I_{d_q} - \Xi^{\pi*}_{m_T}S_g(\Sigma^{\pi*}_{m_T})^{-1}S_g')\hat{q}^{\pi*}_{m_T}(\cdot)$,

$$(\Upsilon_\Xi)^{-1}(S^\theta_{h,\mu})T^{1/2}\begin{pmatrix} \breve{h}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T})/S_T^{1/2} \\ r(\hat{\theta}^{\pi*}_{m_T})/k^{1/2} \end{pmatrix} - T^{1/2}\begin{pmatrix} \Xi^{-1}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})/S_T^{1/2} - S_g\Sigma^{-1}\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T})/S_T^{1/2} \\ \tilde{\mu}^{\pi*}_{m_T}/k^{1/2} \\ (\hat{\theta}^{\pi*}_{m_T} - \tilde{\theta}^{\pi*}_{m_T})/k^{1/2} \end{pmatrix} \to 0,$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Hence,

$$(S^\theta_{h,\mu})'(\Upsilon_\Xi)^{-1}(S^\theta_{h,\mu})T^{1/2}\begin{pmatrix} \breve{h}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T})/S_T^{1/2} \\ r(\hat{\theta}^{\pi*}_{m_T})/k^{1/2} \end{pmatrix} - S_{h,\mu}T^{1/2}\begin{pmatrix} \Xi^{-1}\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})/S_T^{1/2} \\ \tilde{\mu}^{\pi*}_{m_T}/k^{1/2} \end{pmatrix} \to 0,$$

prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$; cf. Smith (2011, eq. (B.22), p.1230). Therefore, together with the similar result

$$(S^\theta_{h,\mu})(\Upsilon_\Xi)^{-1}(S^\theta_{h,\mu})T^{1/2}\begin{pmatrix} \breve{h}^{\pi}_{T}(\hat{\theta}^{\pi}_{T})/S_T^{1/2} \\ r(\hat{\theta}^{\pi}_{T})/k^{1/2} \end{pmatrix} - S_{h,\mu}T^{1/2}\begin{pmatrix} \Xi^{-1}\hat{q}^{\pi}_{Y}(\tilde{\theta}^{\pi}_{T})/S_T^{1/2} \\ \tilde{\mu}^{\pi}_{T}/k^{1/2} \end{pmatrix} \to 0, \text{ prob-}\mathcal{P}.$$

cf. $\mathcal{GW}^{\pi*}_{m_T}$ (3.13), prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$,

$$\mathcal{LM}^{\pi*}_{m_T}/k - T\begin{pmatrix} (\breve{h}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - \breve{h}^{\pi}_{T}(\hat{\theta}^{\pi}_{T}))/S_T^{1/2} \\ (r(\hat{\theta}^{\pi*}_{m_T}) - r(\hat{\theta}^{\pi}_{T}))/k^{1/2} \end{pmatrix}' S'_{h,\mu}\Psi_\Xi S_{h,\mu}\begin{pmatrix} (\breve{h}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - \breve{h}^{\pi}_{T}(\hat{\theta}^{\pi}_{T}))/S_T^{1/2} \\ (r(\hat{\theta}^{\pi*}_{m_T}) - r(\hat{\theta}^{\pi}_{T}))/k^{1/2} \end{pmatrix} \to 0.■$$

$\mathcal{GW}^{\pi*}_{m_T}$: Substituting the expansions $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\hat{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T})) - QT^{1/2}(\hat{\theta}^{\pi*}_{m_T} - \tilde{\theta}^{\pi*}_{m_T})/k^{1/2} \to 0$ and $T^{1/2}r(\hat{\theta}^{\pi*}_{m_T}) - RT^{1/2}(\hat{\theta}^{\pi*}_{m_T} - \tilde{\theta}^{\pi*}_{m_T}) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, and, similarly, those for $\hat{q}^{\pi}_T(\hat{\theta}^{\pi}_T)$ and $r(\hat{\theta}^{\pi}_T)$, as $P_\Xi Q = J_\Xi' R$ and $J_\Xi Q = J_\Xi \Xi J_\Xi' R$, $\mathcal{GW}^{\pi*}_{m_T}$ may be expressed as $(T/S_T)((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi S_g\Sigma^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}_T(\hat{\beta}^{\pi}_T))'P_\Xi((\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi S_g\Sigma^{-1}(\hat{g}^{\pi*}_{m_T}(\hat{\beta}^{\pi*}_{m_T}) - \hat{g}_T(\hat{\beta}^{\pi}_T))$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Cf. $\mathcal{LR}^{\pi*}_{m_T}$ eq. (3.9). Note that, from the first order conditions for $\tilde{\theta}^{\pi*}_{m_T}$, $\tilde{\theta}^{\pi}_T$, $\hat{\beta}^{\pi*}_{m_T}$ and $\hat{\beta}^{\pi}_T$, P_Ξ may be replaced by Ξ^{-1}.■

Proof of Theorem 3.5. From Lemma B.4, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\mathring{\theta}^{\pi*}_{m_T}) - \Xi^{\pi*}_{m_T}P^{\pi*}_{\Xi m_T}\hat{q}^{\pi}_T(\ddot{\theta}^{\pi}_T)) - \Xi P_\Xi(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Since, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\tilde{\theta}^{\pi*}_{m_T}) - \hat{q}^{\pi}_T(\tilde{\theta}^{\pi}_T)) - \Xi P_\Xi(T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, from Lemma B.3, and, by GEL-KBB CLT Lemma A.5, $(T/S_T)^{1/2}(\hat{q}^{\pi*}_{m_T}(\theta_0) - \hat{q}^{\pi}_T(\theta_0)) - (T/S_T)^{1/2}\hat{q}_T(\theta_0) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, the results follow immediately. Cf. Proof of Theorem 3.2 for $\mathcal{LR}^{\pi*}_{m_T}$ and $\mathcal{D}^{\pi*}_{m_T}$.■

Appendix C: GEL Implied Probability GMM-KBB HAC Variance Matrix Estimation

In the following $X_t(\theta)$ and $Y_t(\theta)$ substitute, where appropriate, for $q_t(\theta)$, $(t = 1, 2, ...)$, in Assumptions 2.1, 2.3 and 2.4.

Let $X_{tT}(\theta) = \sum_{s=t-T}^{t-1} k(s/S_T)X_t(\theta)/(k_2 S_T)^{1/2}$, $Y_{tT}(\theta) = \sum_{s=t-T}^{t-1} k(s/S_T)Y_t(\theta)/(k_2 S_T)^{1/2}$, $(t = 1, ..., T)$, $\bar{X}(\theta) = \sum_{t=1}^{T} X_t(\theta)/T$, $\bar{Y}(\theta) = \sum_{t=1}^{T} Y_t(\theta)/T$ and $\bar{X}_T(\theta) = \sum_{t=1}^{T} X_{tT}(\theta)/T$, $\bar{Y}_T(\theta) = \sum_{t=1}^{T} Y_{tT}(\theta)/T$. Also let $\bar{X}_T^{\pi}(\theta) = \sum_{t=1}^{T} \pi_{tT} X_{tT}(\theta)$, $\bar{Y}_T^{\pi}(\theta) = \sum_{t=1}^{T} \pi_{tT} Y_{tT}(\theta)$ and $\bar{X}_{m_T}^{\pi*}(\theta) = \sum_{s=1}^{m_T} X_{t_s^* T}^{\pi}(\theta)/m_T$, $\bar{Y}_{m_T}^{\pi*}(\theta) = \sum_{s=1}^{m_T} Y_{t_s^* T}^{\pi}(\theta)/m_T$ where the indices t_s^* and the consequent bootstrap sample $(X_{t_s^* T}^{\pi}(\theta), Y_{t_s^* T}^{\pi}(\theta))$, $(s = 1, ..., m_T)$, denote m_T independent draws with replacement from the index set $\mathcal{T}_T = \{1, ..., T\}$ and the bootstrap sample space $\{X_{tT}(\theta), Y_{tT}(\theta)\}_{t=1}^{T}$ with sampling probabilities $\mathcal{P}_{\omega}^*(X_{t_s^* T}(\theta) = X_{tT}(\theta), Y_{t_s^* T}^{\pi}(\theta) = Y_{tT}(\theta)) = \pi_{tT}$, $(t = 1, ..., T)$, with $m_T = [T/S_T]$ the integer part of T/S_T.

The argument θ is suppressed in the proofs of the following lemmas without loss of generality.

Lemma C.1. (GEL-KBB HAC Variance Pointwise WLLN.) Let $E[X_t(\theta)] = 0$. Also let $\{X_t(\theta)\}_{t=1}^{\infty}$ satisfy Assumptions 2.1(b) and 2.3(d). If Assumptions 2.2 and 3.2 hold, then

$$\textbf{(a)}\ \frac{1}{m_T}\sum_{s=1}^{m_T}(X_{t_s^* T}^{\pi}(\theta))^2 - \frac{1}{T}\sum_{t=1}^{T}(X_{tT}(\theta))^2 \to 0,\ \text{prob-}\mathcal{P}_{\omega}^*,\ \text{prob-}\mathcal{P};$$

$$\textbf{(b)}\ \frac{1}{m_T}\sum_{s=1}^{m_T}(X_{t_s^* T}^{\pi}(\theta))^2 - \sum_{t=1}^{T}\pi_{tT}(X_{tT}(\theta))^2 \to 0,\ \text{prob-}\mathcal{P}_{\omega}^*,\ \text{prob-}\mathcal{P}.$$

Proof. By Assumption 3.2(b) $\sum_{t=1}^{T}(X_{tT})^2/T - \sum_{t=1}^{T}\pi_{tT}(X_{tT})^2 \to 0$, prob-$\mathcal{P}$. Hence, as $\sum_{t=1}^{T}(X_{tT})^2/T = O_p(1)$, cf. Smith (2011, Lemma A.3, p.1219), the Lemma follows by T if **(b)** is proven. The proof offered is similar to that in Gonçalves and White (2004, Proof of Lemma B.1, pp.217-218).

First, since $\mathrm{E}^*[(X_{t_s^* T}^{\pi})^2] = \sum_{t=1}^{T}\pi_{tT}(X_{tT})^2$, by M,

$$\mathcal{P}_{\omega}^*(|\frac{1}{m_T}\sum_{s=1}^{m_T}(X_{t_s^* T}^{\pi})^2 - \sum_{t=1}^{T}\pi_{tT}(X_{tT})^2| > \varepsilon) \leq \frac{1}{\varepsilon^p}\mathrm{E}^*[|\frac{1}{m_T}\sum_{s=1}^{m_T}(X_{t_s^* T}^{\pi})^2 - \sum_{t=1}^{T}\pi_t(X_{tT})^2|^p] \tag{C.1}$$

for some $p > 1$. Now

$$\begin{aligned}\mathrm{E}^*[|\frac{1}{m_T}\sum_{s=1}^{m_T}(X_{t_s^* T}^{\pi})^2 - \sum_{t=1}^{T}\pi_{tT}(X_{tT})^2|^p] &= \frac{1}{m_T^p}\mathrm{E}^*[|\sum_{s=1}^{m_T}((X_{t_s^* T}^{\pi})^2 - \mathrm{E}^*[(X_{t_s^* T}^{\pi})^2])|^p] \\ &\leq \frac{1}{m_T^p}C\mathrm{E}^*[(\sum_{s=1}^{m_T}|(X_{t_s^* T}^{\pi})^2 - \mathrm{E}^*[(X_{t_s^* T}^{\pi})^2]|^2)^{p/2}]\end{aligned}$$

for some $C < \infty$ by the extension to the Burkholder inequality in White and Chen (1996, Lemma A.2(iv), p.299) as $(X_{t_s^* T}^{\pi})^2 - \mathrm{E}^*[(X_{t_s^* T}^{\pi})^2]$, $(s = 1, ..., m_T)$, are `i.i.d.` zero mean. For $1 < p \leq 2$, by Jensen's inequality, White (1984, Proposition 2.38, p.27), $\mathrm{E}^*[(\sum_{s=1}^{m_T}|(X_{t_s^* T}^{\pi})^2 - \mathrm{E}^*[(X_{t_s^* T}^{\pi})^2]|^2)^{p/2}] \leq m_T\mathrm{E}^*[|(X_{t_s^* T}^{\pi})^2 - \mathrm{E}^*[(X_{t_s^* T}^{\pi})^2]|^p]$. Invoking the c_r-inequality, White (1984, Proposition 3.8, p.33), $\mathrm{E}^*[|(X_{t_s^* T}^{\pi})^2$

$-\mathrm{E}^*[(X^\pi_{t_s^*T})^2]|^p] \leq 2^p\mathrm{E}^*[|X^\pi_{t_s^*T}|^{2p}]$. Hence, substituting in eq. (C.1),

$$\begin{aligned}\mathcal{P}^*_\omega(|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(X^\pi_{t_s^*T})^2-\sum\nolimits_{t=1}^{T}\pi_t(X_{tT})^2| &> \varepsilon) \leq \frac{1}{\varepsilon^p m_T^{p-1}}2^pC\mathrm{E}^*[|X^\pi_{t_s^*T}|^{2p}] \\ &= \frac{1}{\varepsilon^p m_T^{p-1}}2^pC\sum\nolimits_{t=1}^{T}\pi_{tT}|X_{tT}|^{2p} \\ &= \frac{1}{\varepsilon^p m_T^{p-1}}2^pC(1+o_p(1))\frac{1}{T}\sum\nolimits_{t=1}^{T}|X_{tT}|^{2p}.\end{aligned} \tag{C.2}$$

Now, by M, cf. Newey and Smith (2004, Proof of Lemma A1, p.239),

$$\begin{aligned}\frac{1}{T}\sum\nolimits_{t=1}^{T}|X_{tT}|^{2p} &\leq \frac{1}{T}\sum\nolimits_{t=1}^{T}|X_{tT}|^2(\max_{1\leq t\leq T}|X_{tT}|)^{2(p-1)} \\ &= O_p(T^{2(p-1)/\alpha}),\end{aligned}$$

noting $\sum_{t=1}^T|X_{tT}|^2/T = O_p(1)$, cf. Smith (2011, Lemma A.3, p.1219). Thus, from eq. (C.2),

$$\begin{aligned}\mathcal{P}^*_\omega(|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(X^\pi_{t_s^*T})^2-\sum\nolimits_{t=1}^{T}\pi_t(X_{tT})^2| &> \varepsilon) \leq \frac{1}{\varepsilon^p m_T^{p-1}}2^pC(1+o_p(1))O_p(T^{2(p-1)/\alpha}) \\ &= O_p(T^{(p-1)(2/\alpha-\eta-1/2)}) = o_p(1).\end{aligned}$$

The second equality and, thus, the result, follow from Assumptions 2.2(a) and 2.3(d) since $\alpha > 4$.■

Lemma C.2. (GEL-KBB Outer Product Estimation.) Let $\{(X_t, Y_t\}_{t=1}^\infty$ satisfy Assumption 2.1(b) and $\mathrm{E}[|X_t|^{dp}], \mathrm{E}[|Y_t|^{\frac{dp}{d-1}}] < \Delta$, $0 < \Delta < \infty$, for some $1 < p \leq 2$ and $d > 1$. If Assumptions 2.2 and 3.2 hold, then

$$\frac{1}{T^{1/2}m_T}\sum\nolimits_{s=1}^{m_T}X^\pi_{t_s^*T}Y^\pi_{t_s^*T} \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

Proof. Cf. Gonçalves and White (2004, Proof of Lemma B.2, p.218). By M, for some $1 < p \leq 2$, using the c_r-inequality with $r = p$,

$$\begin{aligned}\mathcal{P}^*_\omega(|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}X^\pi_{t_s^*T}Y^\pi_{t_s^*T}| > T^{1/2}\varepsilon) &\leq \frac{1}{\varepsilon^pT^{p/2}}\mathrm{E}^*[|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}X^\pi_{t_s^*T}Y^\pi_{t_s^*T}|^p] \\ &\leq \frac{1}{\varepsilon^pT^{p/2}}2^{p-1}(\mathrm{E}^*[|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(X^\pi_{t_s^*T}Y^\pi_{t_s^*T}-\mathrm{E}^*[X^\pi_{t_s^*T}Y^\pi_{t_s^*T}])|^p] \\ &+|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}\mathrm{E}^*[X^\pi_{t_s^*T}Y^\pi_{t_s^*T}]|^p) \\ &= \frac{1}{\varepsilon^pT^{p/2}}2^{p-1}(F_1+F_2).\end{aligned}$$

Since $X^\pi_{t_s^*T}Y^\pi_{t_s^*T} - \mathrm{E}^*[X^\pi_{t_s^*T}Y^\pi_{t_s^*T}]$, $(s = 1, ..., m_T)$, are independent zero mean,

$$F_1 \leq \frac{1}{m_T^p}C\mathrm{E}^*[(\sum\nolimits_{s=1}^{m_T}|X^\pi_{t_s^*T}Y^\pi_{t_s^*T}-\mathrm{E}^*[X^\pi_{t_s^*T}Y^\pi_{t_s^*T}]|^2)^{p/2}]$$

for some $C < \infty$ by the extension to the Burkholder inequality White and Chen (1996, Lemma A.2(iv), p.299). Hence, for $1 < p \leq 2$, by Jensen's inequality and the c_r-inequality,

$$\begin{aligned} F_1 &\leq \frac{1}{m_T^{p-1}} C\mathrm{E}^*[|X^\pi_{t^*_s T} Y^\pi_{t^*_s T} - \mathrm{E}^*[X^\pi_{t^*_s T} Y^\pi_{t^*_s T}]|^p] \\ &\leq \frac{1}{m_T^{p-1}} 2^p C\mathrm{E}^*[|X^\pi_{t^*_s T} Y^\pi_{t^*_s T}|^p] \\ &= \frac{1}{m_T^{p-1}} 2^p C \sum\nolimits_{t=1}^{T} \pi_{tT} |X_{tT} Y_{tT}|^p = \frac{1}{m_T^{p-1}} 2^p C(1 + o_p(1)) \frac{1}{T} \sum\nolimits_{t=1}^{T} |X_{tT} Y_{tT}|^p. \end{aligned}$$

Also, by Jensen's inequality and Assumption 3.2(b),

$$\begin{aligned} F_2 &= \frac{1}{m_T^p} |\sum\nolimits_{s=1}^{m_T} \mathrm{E}^*[X^\pi_{t^*_s T} Y^\pi_{t^*_s T}]|^p \\ &\leq \mathrm{E}^*[|X^\pi_{t^*_s T} Y^\pi_{t^*_s T}|^p] \\ &= \sum\nolimits_{t=1}^{T} \pi_{tT} |X_{tT} Y_{tT}|^p = (1 + o_p(1)) \frac{1}{T} \sum\nolimits_{t=1}^{T} |X_{tT} Y_{tT}|^p. \end{aligned}$$

By M and Hölder inequality, White (1984, Proposition 3.4, p.30),

$$\begin{aligned} \mathcal{P}(\frac{1}{T} \sum\nolimits_{t=1}^{T} |X_{tT} Y_{tT}|^p &> \delta) \leq \frac{1}{T\delta} \sum\nolimits_{t=1}^{T} \mathrm{E}[|X_{tT} Y_{tT}|^p] \\ &\leq \frac{1}{T\delta} \sum\nolimits_{t=1}^{T} (\mathrm{E}[|X_{tT}|^{dp}])^{1/d} (\mathrm{E}[|Y_{tT}|^{\frac{dp}{d-1}}])^{(d-1)/d}. \end{aligned}$$

Then, by T and Minkowski inequality, White (1984, Proposition 3.11, p.34),

$$\begin{aligned} \mathrm{E}[|(k_2)^{1/2} X_{tT} / S_T^{1/2}|^{dp}] &= \mathrm{E}[|\frac{1}{S_T} \sum\nolimits_{s=t-T}^{t-1} k(\frac{s}{S_T}) X_{t-s}|^{dp}] \\ &\leq \mathrm{E}[(\frac{1}{S_T} \sum\nolimits_{s=t-T}^{t-1} |k(\frac{s}{S_T})| |X_{t-s}|)^{dp}] \\ &\leq (\frac{1}{S_T} \sum\nolimits_{s=t-T}^{t-1} |k(\frac{s}{S_T})| \mathrm{E}[|X_{t-s}|^{dp}]^{1/dp})^{dp} = O(1) \end{aligned}$$

as $\mathrm{E}[|X_t|^{dp}]$ is bounded by hypothesis and $\sum\nolimits_{s=1-T}^{T-1} |k(s/S_T)|/S_T = O(1)$. By the same reasoning $\mathrm{E}[|(k_2)^{1/2} Y_{tT} / S_T^{1/2}|^{\frac{dp}{d-1}}] = O(1)$. The result follows since $S_T/T^{1/2} = o(1)$ by Assumption 2.2(a).■

Let

$$\Xi^{\pi*}_{m_T}(\theta) = \frac{1}{m_T} \sum\nolimits_{s=1}^{m_T} q^\pi_{t^*_s T}(\theta) q^\pi_{t^*_s T}(\theta)'.$$

Lemma C.3. (GEL-GMM HAC Variance Estimation.) Under Assumptions 2.1-2.4 and 3.2, if $(T/S_T)^{1/2} \hat{q}^\pi_T(\theta_0) = O_p(1)$,

$$\Xi^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T}) \to \Xi, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

Proof. Adopting a proof strategy similar to that of Gonçalves and White (2004, Proof of Theorem 3.1, pp.216-217), first consider the infeasible estimator $\Xi^{\pi*}_{m_T}(\theta_0)$ of Ξ. Fix any $\xi \in \mathcal{R}^{d_q}$, $0 < \|\xi\| < \infty$. Now $\xi' \Xi^{\pi*}_{m_T}(\theta_0) \xi = \sum_{t=1}^{m_T} (\xi' q^\pi_{t^*_s T}(\theta_0))^2 / m_T$. Applying Lemma C.1 with $X_t(\theta_0) = \xi' q_t(\theta_0)$, $(t = 1, ..., T)$,

$$\frac{1}{m_T} \sum\nolimits_{s=1}^{m_T} (\xi' q^\pi_{t^*_s T}(\theta_0))^2 - \frac{1}{T} \sum\nolimits_{t=1}^{T} (\xi' q_{tT}(\theta_0))^2 \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

Thus, by Smith (2011, Lemma A.3, p.1219),

$$\frac{1}{m_T}\sum\nolimits_{t=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\theta_0))^2 - \xi'\Xi\xi \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

It remains to prove that $\xi'\Xi^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})\xi - \xi'\Xi^{\pi*}_{m_T}(\theta_0)\xi \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. By a first order Taylor expansion of $(\xi' q^{\pi}_{t^*_s T}(\breve{\theta}^{\pi*}_{m_T}))^2$ around θ_0

$$(\xi' q^{\pi}_{t^*_s T}(\breve{\theta}^{\pi*}_{m_T}))^2 = (\xi' q^{\pi}_{t^*_s T}(\theta_0))^2 + 2(\xi' q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\xi' Q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\breve{\theta}^{\pi*}_{m_T} - \theta_0))$$

where $\bar{\theta}^{\pi*}_{m_T}$ is on the line segment joining $\breve{\theta}^{\pi*}_{m_T}$ and θ_0 and $Q^{\pi}_{t^*_s T}(\theta) = \partial q^{\pi}_{t^*_s T}(\theta)/\partial\theta'$, $(s = 1, ..., m_T)$. Substituting

$$\xi'\Xi^{\pi*}_{m_T}(\breve{\theta}^{\pi*}_{m_T})\xi = \frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\theta_0))^2 + \frac{2}{m_T}\sum\nolimits_{s=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\xi' Q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\breve{\theta}^{\pi*}_{m_T} - \theta_0).$$

The first term is $\xi'\Xi^{\pi*}_{m_T}(\theta_0)\xi$. For the second term, denoting the jth column of $Q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T})$ by $Q^{\pi}_{t^*_s T,j}(\bar{\theta}^{\pi*}_{m_T})$, $(\xi' Q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\breve{\theta}^{\pi*}_{m_T} - \theta_0) = \sum_{j=1}^{d_\theta}(\xi' Q^{\pi}_{t^*_s T,j}(\bar{\theta}^{\pi*}_{m_T}))(\breve{\theta}^{\pi*}_{m_T,j} - \theta_{0,j})$, and, thus,

$$\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\xi' Q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\breve{\theta}^{\pi*}_{m_T} - \theta_0) = \sum\nolimits_{j=1}^{d_\theta}(\breve{\theta}^{\pi*}_{m_T,j} - \theta_{0,j})\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\xi' Q^{\pi}_{t^*_s T,j}(\bar{\theta}^{\pi*}_{m_T})).$$

Now, by T and CS,

$$\left|\frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}(\xi' q^{\pi}_{t^*_s T}(\bar{\theta}^{\pi*}_{m_T}))(\xi' Q^{\pi}_{t^*_s T,j}(\bar{\theta}^{\pi*}_{m_T}))\right| \leq \frac{1}{m_T}\sum\nolimits_{s=1}^{m_T}\sup_{\theta\in\Theta_r}|(\xi' q^{\pi}_{t^*_s T}(\theta))|\sup_{\theta\in\Theta_r}|(\xi' Q^{\pi}_{t^*_s T,j}(\theta))|.$$

Define $X^{\pi}_{t^*_s T}(\theta) = \sup_{\theta\in\Theta_r}|(\xi' q^{\pi}_{t^*_s T}(\theta))|$ and $Y^{\pi}_{t^*_s T}(\theta) = \sup_{\theta\in\Theta_r}|(\xi' Q^{\pi}_{t^*_s T,j}(\theta))|$. Applying Lemma C.2 with $p = 1 + \varepsilon$ and $d = \alpha/(1 + \varepsilon)$ for some $\varepsilon > 0$,

$$\frac{1}{T^{1/2}m_T}\sum\nolimits_{s=1}^{m_T}\sup_{\theta\in\Theta_r}|(\xi' q^{\pi}_{t^*_s T}(\theta))|\sup_{\theta\in\Theta_r}|(\xi' Q^{\pi}_{t^*_s T,j}(\theta))| \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P},$$

by Assumption 2.4(b). The result then follows from Proposition 3.2 and Theorem 3.2 writing $T^{1/2}(\breve{\theta}^{\pi*}_{m_T} - \theta_0) = T^{1/2}(\breve{\theta}^{\pi*}_{m_T} - \breve{\theta}^{\pi}_{T}) + T^{1/2}(\breve{\theta}^{\pi}_{T} - \theta_0)$. ■

References

Allen, J., Gregory, A.W. and K. Shimotsu (2011): "Empirical likelihood block bootstrapping," *Journal of Econometrics*, 161, 110-121.

Andrews, D.W.K (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, 59, 817-58.

Andrews, D.W.K., (2002): "Equivalence of the higher order asymptotic efficiency of k-step and extremum statistics," *Econometric Theory*, 18, 1040-1085.

Andrews, D.W.K. and J.C. Monahan (1992): "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator," *Econometrica*, 60, 953-966.

Antoine, B., Bonnal, H., and E. Renault, (2007): "On the efficient use of the informational content of estimating equations: Implied probabilities and Euclidean empirical likelihood," *Journal of Econometrics*, 138(2), 461-487.

Back, K. and D.P. Brown (1993): "Implied probabilities in GMM estimators", *Econometrica*, 61, 971-975.

Bravo, F. and F. Crudu (2012): "Efficient bootstrap with weakly dependent processes," *Computational Statistics and Data Analysis*, 56, 3444-3458.

Brown, B.W. and W.K. Newey (1992): "Bootstrapping for GMM," Working Paper, M.I.T.

Brown, B.W. and W.K. Newey (1998): "Efficient semiparametric estimation of expectations", *Econometrica* 66, 453-464.

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

Burnside, C. and M.S. Eichenbaum (1996): "Small-sample properties of GMM-based Wald tests," *Journal of Business and Economic Statistics*, 14, 294-308.

Camponovo, L. (2016): "Asymptotic refinements of nonparametric bootstrap for quasi-likelihood ratio tests for classes of extremum estimators," *Econometrics Journal*, 19, 33-54.

Camponovo, L., Scaillet, O. and F. Trojani (2012): "Robust subsampling," *Journal of Econometrics*, 167, 197-210.

Carlstein, E. (1986): "The use of subseries values for estimating the variance of a general statistic from a stationary sequence," *Annals of Statistics*, 14, 1171-1179.

Christiano, L. J. and W.J. den Haan (1996): "Small-sample properties of GMM for business-cycle analysis," *Journal of Business and Economic Statistics*, 14, 309-327.

Cressie, N., and T. Read (1984): "Multinomial goodness-of-fit tests", *Journal of the Royal Statistical Society, Series B*, 46, 440-464.

D'Antona, G. and A. Ferrero (2005): *Digital Signal Processing for Measurement Systems: Theory and Applications*. New York, U.S.A.: Springer.

Davidson, R. and J. G. MacKinnon (1999): "Bootstrap testing in nonlinear models," *International Economic Review*, 40, 487-508.

Dufour, J.-M. (2003): "Identification, weak instruments, and statistical inference in econometrics," *Canadian Journal of Economics*, 36, 767-808.

Efron B. (1979): "Bootstrap methods: another look at the jackknife," *Annals of Statistics,* 7, 1–26.

Eichenbaum, M.S., Hansen, L.P. and K.J. Singleton (1988): "A time series analysis of representative agent models of consumption and leisure choice under uncertainty," *Quarterly Journal of Economics*, 103, 51-78.

Gallant, R. (1987): *Nonlinear Statistical Models*. New York, U.S.A.: Wiley.

Gonçalves, S., and White, H. (2000): "Maximum likelihood and the bootstrap for nonlinear dynamic models," University of California at San Diego, Economics Working Paper Series 2000-32, Department of Economics, UC San Diego.

Gonçalves, S. and H. White (2004): "Maximum likelihood and the bootstrap for nonlinear dynamic models," *Journal of Econometrics*, 119, 199-219.

Gouriéroux, C., and A. Monfort (1989): "A general framework for testing a null hypothesis in a "mixed form", *Econometric Theory*, 5, 63-82.

Gradshteyn, I.S., and I.M. Ryzhik (1980): *Tables of Integrals, Series and Products.* New York: Academic Press.

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

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

Hall, P. and J. L. Horowitz (1996): "Bootstrap critical values for tests based on generalized method of moments estimators," *Econometrica*, 891-916.

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

Hansen, L.P., Heaton, J. and A. Yaron (1996): "Finite-sample properties of some alternative GMM estimators", *Journal of Business and Economic Statistics*, 14, 262-280.

Hong, H. and O. Scaillet (2006): "A fast subsampling method for nonlinear dynamic models," *Journal of Econometrics*, 133, 557-578.

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

Kitamura, Y. and M. Stutzer (1997): "An information-theoretic alternative to generalized method of moments estimation," *Econometrica*, 65, 861-874.

Kreiss, J.P. and Paparoditis, E. (2011). "Bootstrap methods for dependent data: a review," *Journal of the Korean Statistical Society*, 40, 357-378.

Künsch, H. (1989): "The jacknife and the bootstrap for general stationary observations," *Annals of Statistics,* 17, 1217–1241.

La Vecchia, D., Moor A. and Scaillet, O. (2023): "A higher-order correct fast moving-average bootstrap for dependent data," *Journal of Econometrics*, 235, 65-81.

Lahiri, S.N. (2003): *Resampling Methods for Dependent Data*. New York: Springer.

Liu, R. and K. Singh (1992): "Moving blocks jacknife and bootstrap capture weak dependence," in *Exploring the Limits of Bootstrap*, Wiley, New York, eds. Lepage, R. and Billard, L., 225–248.

Newey, W.K. (1985): "Generalized method of moments specification testing", *Journal of Econometrics*, 29, 229-256.

Newey, W.K. and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, Volume 4, R.F. Engle and D.L. McFadden (eds.), 2111-2245.

Newey, W.K. and R.J. Smith (2004): "Higher order properties of GMM and generalized empirical likelihood estimators," *Econometrica*, 72, 219-255.

Newey, W.K. and K. D.West, (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703-08.

Ng, S. and P. Perron (1996): "The exact error in estimating the spectral density at the origin," *Journal of Time Series Analysis* 17, 379–408.

Owen, A. (1988): "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika* 75, 237-249.

Paparoditis, E. and D. N. Politis (2001): "Tapered block bootstrap," *Biometrika*, vol. 88, No. 4, 1105-1119.

Parente, P.M.D.C. and R.J. Smith (2018): "Kernel block bootstrap," CWP 48/18, Centre for Microdata Methods and Practice, U.C.L and I.F.S.

Parente, P.M.D.C. and R.J. Smith (2021): "Quasi-maximum likelihood and the kernel block bootstrap for nonlinear dynamic models," *Journal of Time Series Analysis*, 42, 377-405.

Politis, D.N. and J.P. Romano (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.

Politis, D.N. and J.P. Romano (1995): "Bias-corrected nonparametric spectral estimation". *Journal of Time Series Analysis* 16, 67–103.

Rao, C. (1973): *Linear Statistical Inference and its Applications.* New York: Wiley.

Rao, C. R. and S. K. Mitra, (1971): *Generalized Inverse of Matrices and Its Applications.* New York: Wiley.

Ruud, P. (2000): *An Introduction to Classical Econometric Theory.* Oxford: Oxford University Press.

Serfling, R. (1980): *Approximation Theorems of Mathematical Statistics.* New York: Wiley.

Shao, J. and D. Tu (1995). *The Jackknife and Bootstrap.* New York: Springer.

Smith, R.J., (1997): "Alternative semi-parametric likelihood approaches to generalised method of moments estimation," *Economic Journal*, 107, 503-519.

Smith, R.J. (2005): "Automatic positive semidefinite HAC covariance matrix and GMM estimation," *Econometric Theory*, 21, 158-170.

Smith. R.J. (2011): "GEL criteria for moment condition models," *Econometric Theory*, 27, 1192-1235.

White, H. (1984): *Asymptotic Theory for Econometricians.* New York: Academic Press.

White, H. and X. Chen (1996): "Laws of large numbers for Hilbert space-valued mixingales with applications," *Econometric Theory*, 12, 284-304.

Table 1. Model 1: Empirical Rejection Probabilities: Overidentifying Moment Restrictions Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	0.40	**4.04**	**9.64**	0.96	5.34	12.22	0.68	4.48	**10.04**	0.88	5.12	10.74
KBB_{TR}	0.28	3.28	8.46	0.60	3.90	9.60	0.58	4.16	9.64	0.80	4.60	9.46
KBB^{π}_{TR}	0.30	3.34	8.80	0.72	3.98	**9.96**	0.82	4.36	9.94	0.80	4.78	**9.76**
KBB_{BT}	0.28	2.78	7.28	0.70	3.32	7.90	0.62	4.02	9.32	0.84	4.24	8.50
KBB^{π}_{BT}	0.34	3.34	8.14	0.80	3.94	8.96	0.70	4.40	9.64	0.76	4.12	8.68
KBB_{PP}	0.34	2.98	6.88	1.22	3.56	7.10	0.64	4.00	8.82	1.48	4.50	8.16
KBB^{π}_{PP}	**0.70**	3.68	8.00	1.78	4.86	9.42	**1.00**	**4.64**	9.82	1.56	4.90	9.10
KBB_{QS}	0.22	3.08	7.62	0.82	3.72	8.50	0.62	4.02	9.44	**1.00**	4.30	8.68
KBB^{π}_{QS}	0.40	3.46	8.36	**1.02**	4.08	9.28	0.80	4.38	9.70	0.88	4.52	9.18
MBB	0.24	2.86	6.86	1.14	3.86	8.56	0.64	3.82	9.02	1.28	4.52	8.70
MBB^{π}	0.54	3.34	7.98	1.48	**4.90**	9.72	0.60	4.26	9.48	1.44	**4.94**	9.42

Table 2. Model 2: Empirical Rejection Probabilities: Overidentifying Moment Restrictions Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	0.42	**3.94**	**9.40**	0.76	**4.90**	11.52	0.52	4.86	10.16	0.76	5.36	**10.52**
KBB_{TR}	0.28	2.92	7.96	0.54	3.42	8.20	0.46	3.92	9.06	0.56	4.24	9.22
$\text{KBB}^{\pi}_{\text{TR}}$	0.34	3.32	8.20	0.74	3.66	**8.70**	0.48	4.54	**9.72**	0.70	4.80	9.34
KBB_{BT}	0.24	2.40	6.50	0.60	2.84	6.88	0.42	3.86	8.60	0.74	3.90	7.98
$\text{KBB}^{\pi}_{\text{BT}}$	0.42	3.44	8.06	0.72	3.48	7.98	0.50	4.36	9.26	0.68	4.28	8.54
KBB_{PP}	0.44	2.58	6.00	**1.00**	2.96	6.02	0.54	3.66	8.12	1.26	4.26	7.42
$\text{KBB}^{\pi}_{\text{PP}}$	**0.70**	3.92	8.00	1.58	4.54	8.38	**0.80**	**4.94**	9.34	1.32	5.16	9.02
KBB_{QS}	0.28	2.58	6.92	0.70	2.98	7.38	0.48	3.88	8.86	0.76	4.20	8.22
$\text{KBB}^{\pi}_{\text{QS}}$	0.40	3.52	8.12	0.74	3.76	8.26	0.54	4.46	9.50	0.84	4.54	9.00
MBB	0.32	2.60	6.42	1.04	3.18	7.16	0.46	3.76	8.42	**1.00**	4.44	8.02
MBB^{π}	0.48	3.44	7.94	1.28	4.28	8.68	0.58	4.46	9.30	1.02	**5.12**	9.28

Table 3. Model 1: Empirical Rejection Probabilities: Parametric Restrictions *t*-Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	6.10	13.54	20.34	20.38	31.98	39.84	3.14	9.08	15.14	12.54	23.50	30.58
KBB_{TR}	3.34	9.70	15.78	9.44	20.18	28.32	1.90	7.16	12.72	5.42	14.66	22.64
$\text{KBB}^{\pi}_{\text{TR}}$	2.48	8.12	13.54	7.42	15.34	22.12	1.58	6.34	11.50	4.22	11.10	17.88
$\text{KBB}^{\pi_r}_{\text{TR}}$	3.20	9.46	15.58	6.86	16.64	24.74	1.90	7.26	12.50	4.52	13.06	20.80
KBB_{BT}	2.20	7.36	12.90	6.04	15.00	21.78	1.38	5.90	11.04	3.48	10.40	17.06
$\text{KBB}^{\pi}_{\text{BT}}$	1.78	6.46	11.88	4.78	11.26	16.74	1.22	5.28	10.58	2.54	7.88	13.22
$\text{KBB}^{\pi_r}_{\text{BT}}$	1.94	6.84	12.60	3.52	10.94	16.80	1.34	6.04	11.12	2.38	8.10	14.98
KBB_{PP}	**0.96**	4.60	9.14	3.00	9.68	16.32	0.80	4.06	8.44	1.52	**5.52**	**10.86**
$\text{KBB}^{\pi}_{\text{PP}}$	0.88	4.16	8.48	2.30	8.14	13.54	0.64	3.92	8.12	**1.04**	4.46	9.12
$\text{KBB}^{\pi_r}_{\text{PP}}$	0.64	4.40	8.84	**1.50**	**6.54**	**13.04**	0.70	4.08	8.54	0.82	4.22	9.06
KBB_{QS}	2.30	7.50	13.36	6.46	14.98	21.92	1.50	6.02	11.42	3.40	10.50	17.52
$\text{KBB}^{\pi}_{\text{QS}}$	2.04	6.78	12.40	5.04	11.76	17.40	1.46	5.58	10.74	2.74	8.20	13.98
$\text{KBB}^{\pi_r}_{\text{QS}}$	2.24	7.44	12.86	4.26	12.36	19.02	1.46	6.02	11.32	2.72	9.30	15.98
MBB	1.78	5.92	11.42	4.78	12.24	19.58	**1.02**	4.86	**9.94**	2.32	8.02	14.42
MBB^{π}	1.52	**5.16**	**10.66**	3.78	9.50	15.20	0.86	4.84	9.40	1.96	6.56	11.72
MBB^{π_r}	1.30	5.84	11.14	2.68	9.02	16.06	1.08	**4.88**	9.88	1.76	6.30	12.86

Table 4. Model 2: Empirical Rejection Probabilities: Parametric Restrictions *t*-Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	5.40	13.92	20.44	19.00	30.44	38.22	3.16	10.46	16.54	12.50	22.58	30.36
KBB_{TR}	2.64	9.04	15.58	8.62	18.30	26.72	1.72	7.22	13.50	5.12	13.56	20.46
$\text{KBB}^{\pi}_{\text{TR}}$	2.20	7.38	13.02	6.32	13.78	20.08	1.52	6.46	12.48	3.74	10.54	16.10
$\text{KBB}^{\pi_r}_{\text{TR}}$	2.28	8.98	14.94	5.82	14.82	22.58	1.94	7.46	13.60	4.20	12.62	19.56
KBB_{BT}	1.50	6.34	12.16	5.16	12.94	20.20	1.12	5.94	11.34	2.94	9.40	15.18
$\text{KBB}^{\pi}_{\text{BT}}$	1.20	5.82	10.94	3.72	10.06	15.34	1.08	5.56	10.88	2.16	7.36	11.78
$\text{KBB}^{\pi_r}_{\text{BT}}$	1.44	6.46	12.40	3.22	9.34	15.86	1.20	5.96	11.64	2.20	8.18	14.20
KBB_{PP}	0.78	3.80	8.06	2.30	8.70	14.66	0.62	3.82	8.62	1.24	**5.08**	**10.08**
$\text{KBB}^{\pi}_{\text{PP}}$	0.70	3.38	7.86	1.90	6.90	12.34	0.56	3.64	8.22	**0.80**	4.26	8.16
$\text{KBB}^{\pi_r}_{\text{PP}}$	0.46	3.54	8.32	**1.48**	**5.94**	**11.82**	0.74	4.04	9.10	0.60	4.42	9.28
KBB_{QS}	1.94	6.88	12.70	5.42	12.98	20.40	1.20	5.98	11.62	3.32	9.64	15.80
$\text{KBB}^{\pi}_{\text{QS}}$	1.40	5.98	11.40	4.18	10.44	16.08	1.06	5.66	10.92	2.38	7.86	12.34
$\text{KBB}^{\pi_r}_{\text{QS}}$	1.72	6.82	13.04	3.80	10.78	17.82	1.24	6.32	12.08	2.64	8.98	15.28
MBB	1.24	5.16	10.42	3.94	11.06	17.38	0.88	**5.06**	**10.06**	2.18	7.40	13.28
MBB^{π}	0.90	4.76	9.40	2.82	8.26	13.58	0.74	4.58	9.82	1.80	5.92	10.88
MBB^{π_r}	**0.94**	**5.08**	**10.56**	14.30	7.94	2.24	**1.00**	5.40	10.74	1.38	6.66	12.48

Table 5. Model 1: Empirical Rejection Probabilities: Parametric Restrictions LR-GMM Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	6.10	13.54	20.34	20.38	31.98	39.84	3.14	9.08	15.14	12.54	23.50	30.58
KBB_{TR}	1.36	5.90	11.82	1.74	8.78	16.00	**0.98**	5.30	10.72	1.32	7.18	13.94
$\text{KBB}^{\text{C}}_{\text{TR}}$	3.34	9.70	15.78	9.44	20.18	28.32	1.90	7.16	12.72	5.42	14.66	22.64
$\text{KBB}^{\pi}_{\text{TR}}$	1.26	5.60	11.40	1.30	7.24	14.30	0.94	**5.04**	10.64	0.92	6.82	13.58
$\text{KBB}^{\pi,\text{C}}_{\text{TR}}$	2.48	8.12	13.54	7.42	15.34	22.12	1.58	6.34	11.50	4.22	11.10	17.88
$\text{KBB}^{\pi_r}_{\text{TR}}$	3.20	9.46	15.58	6.86	16.64	24.74	1.90	7.26	12.50	4.52	13.06	20.80
$\text{KBB}^{\pi_r,\text{C}}_{\text{TR}}$	2.30	7.80	13.18	5.14	13.42	19.74	1.52	6.32	11.18	3.66	11.16	17.90
KBB_{BT}	**1.00**	4.36	9.06	**0.74**	**4.84**	10.88	0.48	4.10	9.16	0.48	4.24	9.74
$\text{KBB}^{\text{C}}_{\text{BT}}$	2.20	7.36	12.90	6.04	15.00	21.78	1.38	5.90	11.04	3.48	10.40	17.06
$\text{KBB}^{\pi}_{\text{BT}}$	0.62	3.84	8.66	0.48	3.96	9.06	0.40	4.18	9.04	0.52	3.68	8.88
$\text{KBB}^{\pi,\text{C}}_{\text{BT}}$	1.78	6.46	11.88	4.78	11.26	16.74	1.22	5.28	10.58	2.54	7.88	13.22
$\text{KBB}^{\pi_r}_{\text{BT}}$	1.94	6.84	12.60	3.52	10.94	16.80	1.34	6.04	11.12	2.38	8.10	14.98
$\text{KBB}^{\pi_r,\text{C}}_{\text{BT}}$	1.22	**5.06**	**10.16**	2.52	8.26	14.24	0.96	4.66	9.38	1.78	6.50	12.16
KBB_{PP}	0.18	2.20	5.18	0.10	1.34	4.76	0.24	2.54	6.20	0.08	1.18	4.16
$\text{KBB}^{\text{C}}_{\text{PP}}$	0.96	4.60	9.14	3.00	9.68	16.32	0.80	4.06	8.44	1.52	**5.52**	10.86
$\text{KBB}^{\pi}_{\text{PP}}$	0.14	1.92	5.02	0.08	1.06	3.84	0.24	2.42	6.28	0.08	1.06	3.62
$\text{KBB}^{\pi,\text{C}}_{\text{PP}}$	0.88	4.16	8.48	2.30	8.14	13.54	0.64	3.92	8.12	**1.04**	4.46	9.12
$\text{KBB}^{\pi_r}_{\text{PP}}$	0.64	4.40	8.84	1.50	6.54	13.04	0.70	4.08	8.54	0.82	4.22	9.06
$\text{KBB}^{\pi_r,\text{C}}_{\text{PP}}$	0.34	2.72	6.24	0.80	3.60	7.92	0.46	3.18	7.10	0.60	2.70	6.62
KBB_{QS}	0.88	4.50	9.50	0.72	5.24	11.42	0.66	4.34	9.50	0.54	4.42	**9.94**
$\text{KBB}^{\text{C}}_{\text{QS}}$	2.30	7.50	13.36	6.46	14.98	21.92	1.50	6.02	11.42	3.40	10.50	17.52
$\text{KBB}^{\pi}_{\text{QS}}$	0.72	4.22	8.86	0.54	4.42	**9.98**	0.58	4.08	9.18	0.60	4.16	9.62
$\text{KBB}^{\pi,\text{C}}_{\text{QS}}$	2.04	6.78	12.40	5.04	11.76	17.40	1.46	5.58	10.74	2.74	8.20	13.98
$\text{KBB}^{\pi_r}_{\text{QS}}$	2.24	7.44	12.86	4.26	12.36	19.02	1.46	6.02	11.32	2.72	9.30	15.98
$\text{KBB}^{\pi_r,\text{C}}_{\text{QS}}$	1.52	5.90	11.16	2.86	9.02	15.36	1.04	5.14	**9.98**	2.10	7.62	14.00
MBB	0.56	3.40	7.30	0.28	3.02	8.06	0.38	3.26	7.72	0.20	2.54	6.88
MBB^{C}	1.78	5.92	11.42	4.78	12.24	19.58	**1.02**	4.86	9.94	2.32	8.02	14.42
MBB^{π}	0.44	3.24	7.00	0.20	2.34	6.66	0.26	3.26	7.86	0.26	2.12	6.48
$\text{MBB}^{\pi,\text{C}}$	1.52	5.16	10.66	3.78	9.50	15.20	0.86	4.84	9.40	1.96	6.56	11.72
MBB^{π_r}	1.30	5.84	11.14	2.68	9.02	16.06	1.08	4.88	9.88	1.76	6.30	12.86
$\text{MBB}^{\pi_r,\text{C}}$	0.82	4.14	8.64	1.70	6.60	11.94	0.86	3.98	8.66	1.32	5.38	10.38

Table 6. Model 2: Empirical Rejection Probabilities: Parametric Restrictions LR-GMM Tests.

T	64						128					
ρ	0.5			0.9			0.5			0.9		
LEVEL	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00	1.00	5.00	10.00
ASYMP	5.40	13.92	20.44	19.00	30.44	38.22	3.16	10.46	16.54	12.50	22.58	30.36
$\mathrm{KBB}_{\mathrm{TR}}$	0.92	5.34	10.82	1.54	7.08	14.30	1.04	5.40	11.04	**0.82**	6.48	13.02
$\mathrm{KBB}^{\mathrm{C}}_{\mathrm{TR}}$	2.64	9.04	15.58	8.62	18.30	26.72	1.72	7.22	13.50	5.12	13.56	20.46
$\mathrm{KBB}^{\pi}_{\mathrm{TR}}$	0.72	**4.96**	10.60	**1.08**	6.18	12.62	0.94	5.16	11.06	0.68	5.68	12.18
$\mathrm{KBB}^{\pi,\mathrm{C}}_{\mathrm{TR}}$	2.20	7.38	13.02	6.32	13.78	20.08	1.52	6.46	12.48	3.74	10.54	16.10
$\mathrm{KBB}^{\pi_r}_{\mathrm{TR}}$	2.28	8.98	14.94	5.82	14.82	22.58	1.94	7.46	13.60	4.20	12.62	19.56
$\mathrm{KBB}^{\pi_r,\mathrm{C}}_{\mathrm{TR}}$	1.62	6.96	12.48	3.92	11.26	17.96	1.32	6.44	11.98	3.16	9.82	16.10
$\mathrm{KBB}_{\mathrm{BT}}$	0.40	3.48	7.80	0.50	3.84	9.10	0.62	3.94	9.24	0.36	3.70	8.76
$\mathrm{KBB}^{\mathrm{C}}_{\mathrm{BT}}$	1.50	6.34	12.16	5.16	12.94	20.20	1.12	5.94	11.34	2.94	9.40	15.18
$\mathrm{KBB}^{\pi}_{\mathrm{BT}}$	0.44	3.16	7.56	0.40	3.28	7.64	0.56	3.84	8.92	0.28	3.12	8.32
$\mathrm{KBB}^{\pi,\mathrm{C}}_{\mathrm{BT}}$	1.20	5.82	10.94	3.72	10.06	15.34	1.08	5.56	10.88	2.16	7.36	11.78
$\mathrm{KBB}^{\pi_r}_{\mathrm{BT}}$	1.44	6.46	12.40	3.22	9.34	15.88	1.20	5.96	11.64	2.20	8.18	14.20
$\mathrm{KBB}^{\pi_r,\mathrm{C}}_{\mathrm{BT}}$	0.86	4.44	9.22	2.06	7.14	12.18	0.88	4.62	9.76	1.48	6.30	11.30
$\mathrm{KBB}_{\mathrm{PP}}$	0.10	1.66	4.64	0.06	1.26	4.30	0.30	2.20	6.46	0.02	1.02	3.92
$\mathrm{KBB}^{\mathrm{C}}_{\mathrm{PP}}$	0.78	3.80	8.06	2.30	8.70	14.66	0.62	3.82	8.62	1.24	**5.08**	**10.08**
$\mathrm{KBB}^{\pi}_{\mathrm{PP}}$	0.08	1.50	4.02	0.02	0.92	3.28	0.32	2.26	6.40	0.12	0.80	3.32
$\mathrm{KBB}^{\pi,\mathrm{C}}_{\mathrm{PP}}$	0.70	3.38	7.86	1.90	6.90	12.34	0.56	3.64	8.22	0.80	4.26	8.16
$\mathrm{KBB}^{\pi_r}_{\mathrm{PP}}$	0.46	3.54	8.32	1.48	5.94	11.82	0.74	4.04	9.10	0.60	4.42	9.28
$\mathrm{KBB}^{\pi_r,\mathrm{C}}_{\mathrm{PP}}$	0.30	1.96	5.08	0.58	3.30	7.06	0.34	3.14	7.26	0.46	2.64	6.22
$\mathrm{KBB}_{\mathrm{QS}}$	0.50	3.68	8.28	0.66	4.36	9.36	0.52	4.32	9.48	0.42	4.12	9.00
$\mathrm{KBB}^{\mathrm{C}}_{\mathrm{QS}}$	1.94	6.88	12.70	5.42	12.98	20.40	1.20	5.98	11.62	3.32	9.64	15.80
$\mathrm{KBB}^{\pi}_{\mathrm{QS}}$	0.50	3.20	8.02	0.48	3.66	8.66	0.56	3.92	9.18	0.48	3.56	8.86
$\mathrm{KBB}^{\pi,\mathrm{C}}_{\mathrm{QS}}$	1.40	5.98	11.40	4.18	10.44	16.08	1.06	5.66	10.92	2.38	7.86	12.34
$\mathrm{KBB}^{\pi_r}_{\mathrm{QS}}$	1.72	6.82	13.04	3.80	10.78	17.82	1.24	6.32	12.08	2.64	8.98	15.28
$\mathrm{KBB}^{\pi_r,\mathrm{C}}_{\mathrm{QS}}$	1.06	4.84	**10.30**	2.20	7.96	13.38	0.92	5.50	10.64	2.00	6.72	12.44
MBB	0.34	2.64	6.74	0.24	2.64	6.66	0.52	3.38	8.16	0.10	2.14	6.26
$\mathrm{MBB}^{\mathrm{C}}$	1.24	5.16	10.42	3.94	11.06	17.38	0.88	**5.06**	**10.06**	2.18	7.40	13.28
MBB^{π}	0.30	2.56	6.26	0.20	2.02	5.48	0.44	3.14	7.94	0.12	1.70	5.86
$\mathrm{MBB}^{\pi,\mathrm{C}}$	0.90	4.76	9.40	2.82	8.26	13.58	0.74	4.58	9.82	1.80	5.92	10.88
MBB^{π_r}	**0.94**	5.08	10.56	2.24	7.94	14.30	**1.00**	5.40	10.74	1.38	6.66	12.48
$\mathrm{MBB}^{\pi_r,\mathrm{C}}$	0.58	3.42	8.02	1.52	**5.38**	**10.18**	0.70	4.38	8.92	1.14	4.84	9.40

Table 7. Model 1: Percentage of Cases inside Convex Hull.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$\mathrm{ET}_{\mathrm{TR}}$	100.00	99.84	100.00	100.00
$\mathrm{ET}_{\mathrm{BT}}$	99.96	99.82	100.00	99.84
$\mathrm{ET}_{\mathrm{PP}}$	100.00	99.98	100.00	100.00
$\mathrm{ET}_{\mathrm{QS}}$	99.98	100.00	100.00	99.98
$\mathrm{ET}_{\mathrm{MBB}}$	100.00	99.96	100.00	100.00
$\mathrm{ET}_{r,\mathrm{TR}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{BT}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{PP}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{QS}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{MBB}}$	100.00	100.00	100.00	100.00

Table 8. Model 2: Percentage of Cases inside Convex Hull.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$\mathrm{ET}_{\mathrm{TR}}$	99.94	99.86	100.00	100.00
$\mathrm{ET}_{\mathrm{BT}}$	99.92	99.76	99.98	99.88
$\mathrm{ET}_{\mathrm{PP}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{\mathrm{QS}}$	99.96	99.96	99.98	100.00
$\mathrm{ET}_{\mathrm{MBB}}$	100.00	99.98	100.00	100.00
$\mathrm{ET}_{r,\mathrm{TR}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{BT}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{PP}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{QS}}$	100.00	100.00	100.00	100.00
$\mathrm{ET}_{r,\mathrm{MBB}}$	100.00	100.00	100.00	100.00

Table 9. Model 1: Percentage of Cases with Ill Conditioned $\Sigma^*_{m_T}$.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$(\Sigma^*_{m_T})_{\mathrm{TR}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi*}_{m_T})_{\mathrm{TR}}$	0.06	0.32	0.00	0.08
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{TR}}$	0.14	1.86	0.00	0.36
$(\Sigma^*_{m_T})_{\mathrm{BT}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi*}_{m_T})_{\mathrm{BT}}$	0.12	0.36	0.00	0.16
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{BT}}$	**0.38**	**4.34**	0.00	**1.26**
$(\Sigma^*_{m_T})_{\mathrm{PP}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi*}_{m_T})_{\mathrm{PP}}$	0.00	0.02	0.00	0.00
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{PP}}$	0.02	0.92	0.00	0.12
$(\Sigma^*_{m_T})_{\mathrm{QS}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi*}_{m_T})_{\mathrm{QS}}$	0.04	0.06	0.00	0.08
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{QS}}$	0.22	2.44	**0.02**	0.64
$(\Sigma^*_{m_T})_{\mathrm{MBB}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi*}_{m_T})_{\mathrm{MBB}}$	0.02	0.14	0.00	0.00
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{MBB}}$	0.28	2.16	0.00	0.18

Table 10. Model 2: Percentage of Cases with Ill Conditioned $\Sigma^*_{m_T}$.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$(\Sigma^*_{m_T})_{\mathrm{TR}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi *}_{m_T})_{\mathrm{TR}}$	0.20	0.28	0.00	0.02
$(\Sigma^{\pi_r *}_{m_T})_{\mathrm{TR}}$	0.42	1.90	0.02	0.34
$(\Sigma^*_{m_T})_{\mathrm{BT}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi *}_{m_T})_{\mathrm{BT}}$	0.16	0.46	0.02	0.12
$(\Sigma^{\pi_r *}_{m_T})_{\mathrm{BT}}$	**0.62**	**4.24**	**0.16**	**1.16**
$(\Sigma^*_{m_T})_{\mathrm{PP}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi *}_{m_T})_{\mathrm{PP}}$	0.02	0.00	0.00	0.00
$(\Sigma^{\pi_r *}_{m_T})_{\mathrm{PP}}$	0.04	0.82	0.00	0.08
$(\Sigma^*_{m_T})_{\mathrm{QS}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi *}_{m_T})_{\mathrm{QS}}$	0.12	0.06	0.02	0.00
$(\Sigma^{\pi_r *}_{m_T})_{\mathrm{QS}}$	0.32	2.06	0.08	0.50
$(\Sigma^*_{m_T})_{\mathrm{MBB}}$	0.00	0.00	0.00	0.00
$(\Sigma^{\pi *}_{m_T})_{\mathrm{MBB}}$	0.06	0.22	0.00	0.00
$(\Sigma^{\pi_r *}_{m_T})_{\mathrm{MBB}}$	0.26	2.24	0.00	0.18

Table 11. Model 1: Average Number of Extra Bootstrap Replications Required Because of Ill Conditioned $\Sigma^{*}_{m_T}$.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$(\Sigma^{*}_{m_T})_{\text{TR}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\text{TR}}$	0.0	1.9	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\text{TR}}$	0.1	65.0	0.0	3.5
$(\Sigma^{*}_{m_T})_{\text{BT}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\text{BT}}$	0.5	196.4	0.0	4.4
$(\Sigma^{\pi_r*}_{m_T})_{\text{BT}}$	**1.7**	**359.6**	0.0	**8.8**
$(\Sigma^{*}_{m_T})_{\text{PP}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\text{PP}}$	0.0	0.3	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\text{PP}}$	0.0	26.8	0.0	2.9
$(\Sigma^{*}_{m_T})_{\text{QS}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\text{QS}}$	0.3	0.0	0.0	0.1
$(\Sigma^{\pi_r*}_{m_T})_{\text{QS}}$	0.1	183.2	0.0	1.5
$(\Sigma^{*}_{m_T})_{\text{MBB}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\text{MBB}}$	0.0	1.1	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\text{MBB}}$	0.0	95.9	0.0	1.3

Table 12. Model 2: Average Number of Extra Bootstrap Replications Required Because of Ill Conditioned $\Sigma^*_{m_T}$.

T	64		128	
ρ	0.5	0.9	0.5	0.9
$(\Sigma^*_{m_T})_{\mathrm{TR}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\mathrm{TR}}$	1.7	2.9	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{TR}}$	20.2	35.6	0.1	2.8
$(\Sigma^*_{m_T})_{\mathrm{BT}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\mathrm{BT}}$	0.8	13.6	0.2	0.4
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{BT}}$	5.3	**502.4**	0.4	**25.8**
$(\Sigma^*_{m_T})_{\mathrm{PP}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\mathrm{PP}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{PP}}$	2.3	78.6	0.0	0.5
$(\Sigma^*_{m_T})_{\mathrm{QS}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\mathrm{QS}}$	1.5	1.1	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{QS}}$	**30.3**	221.6	**5.8**	8.5
$(\Sigma^*_{m_T})_{\mathrm{MBB}}$	0.0	0.0	0.0	0.0
$(\Sigma^{\pi*}_{m_T})_{\mathrm{MBB}}$	0.1	82.4	0.0	0.0
$(\Sigma^{\pi_r*}_{m_T})_{\mathrm{MBB}}$	6.0	85.2	0.0	5.3