Supplemental Appendix: Robust Forecast Superiority Testing with an Application to Assessing Pools of Expert Forecasters *

Valentina Corradi[1], Sainan Jin[2], and Norman R. Swanson[3]

[1]University of Surrey, [2]Singapore Management University, and [3]Rutgers University

September 2020

Abstract

In this supplemental appendix, additional theoretical results are presented for the case of recursive estimation error. Additionally, auxiliary Monte Carlo experiments are reported on, and further empirical results are tabulated based on the empirical illustration in the paper.

Keywords: Robust Forecast Evaluation, Many Moment Inequalities, Bootstrap, Estimation Error, Combination Forecasts, Survey of Professional Forecasters.

———————————————————————

*Valentina Corradi, School of Economics, University of Surrey, Guildford, Surrey, GU2 7XH, UK, v.corradi@surrey.ac.uk; Jin Sainan, School of Economics, Singapore Management University, 90 Stamford Road, #05-30, Singapore 178903, snjin@smu.edu.sg; and Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. We are grateful to Kevin Lee, Patrick Marsh, Luis Martins, Jams Mitchell, Alessia Paccagini, Paulo Parente, Ivan Petrella, Valerio Poti, Barbara Rossi, Simon Van Norden, Claudio Zoli, and to the participants at the 2018 NBER-NSF Times Series Conference, the 2016 European Meeting of the Econometric Society, Conference for 50 years of Keynes College at Kent University, and seminars at Mannheim University, the University of Nottingham, University College Dublin, Instituto Universitário de Lisboa, Universita' di Verona and the Warwick Business School for useful comments and suggestions. Additionally, many thanks are owed to Mingmian Cheng for excellent research assistance.

1 Forecast Superiority Tests in the Presence of Recursive Estimation Error

1.1 The Statistic

This Supplement extends all Lemmas and Theorems in the paper to the case in which there is non vanishing, recursive estimation error.

Let $T = R + n$. At each point in time, $t > R$, update model parameter estimates prior to the construction of each new forecast, using all the available information.[1]

For $j = 1, ..., k$, use the first R observations to compute $\widehat{\theta}_{j,R}$, and construct the first prediction error:

$$\widehat{e}_{j,R+1} = X_{R+1} - \phi_j \left(Z_{j,R}, \widehat{\theta}_{j,R} \right),$$

where $Z_{j,R}$ contains lags of X as well as other regressors. Then, use the first $R+1$ observations to construct

$$\widehat{e}_{j,R+2} = X_{R+2} - \phi_j \left(Z_{j,R+1}, \widehat{\theta}_{j,R+1} \right).$$

Proceed in the same manner until a sequence of n prediction errors has been constructed, defined as:

$$\widehat{e}_{j,t+1} = X_{t+1} - \phi_j \left(Z_{j,t}, \widehat{\theta}_{j,t} \right), \tag{1.1}$$

for $t = R, ..., R+n-1$, where $\widehat{\theta}_{j,t}$ is the estimator computed using observations up to time t. In the sequel, assume that $\widehat{\theta}_{j,t}$ is a recursive $m-$estimator, so that:

$$\widehat{\theta}_{j,t} = \arg \min_{\theta_j \in \Theta_j} \frac{1}{t} \sum_{i=2}^{t} m_j(X_i, Z_{j,i-1}, \theta_j), \quad R \le t \le n-1, \; j = 1, ..., k, \tag{1.2}$$

and

$$\theta_j^{\dagger} = \arg \min_{\theta_j \in \Theta_j} E(m_j(X_i, Z_{j,i-1}, \theta_j)).$$

For $x \geq 0$, define:

$$\widetilde{G}_{j,n}^{+}(x) = \frac{1}{n} \sum_{t=R}^{T-1} \left(1 \left\{ \widehat{e}_{j,t+1} \le x \right\} - 1 \left\{ \widehat{e}_{1,t+1} \le x \right\} \right) = \left(\widetilde{F}_{j,n}(x) - \widetilde{F}_{1,n}(x) \right) \tag{1.3}$$

and

$$\begin{aligned} \widetilde{C}_{j,n}^{+}(x) &= \int_x^{\infty} \left(\widetilde{F}_{j,n}(t) - \widetilde{F}_{1,n}(t) \right) dt \\ &= \frac{1}{n} \sum_{t=R}^{T-1} \left\{ \left[(\widehat{e}_{1,t+1} - x) \right]_+ - \left[(\widehat{e}_{j,t+1} - x) \right]_+ \right\}. \end{aligned} \tag{1.4}$$

Define the following forecast superiority test statistics:

$$\widetilde{S}_n^{G+} = \int_{\mathcal{X}^+} \sum_{j=2}^{k} \left(\max \left\{ 0, \frac{\sqrt{n} \widetilde{G}_{j,n}^{+}(x)}{\widetilde{\sigma}_{jj,n}^{G+}(x) + \varepsilon} \right\} \right)^2 \mathrm{d}Q(x)$$

[1]In the rolling estimation case, we use only the most recent R observations to re-estiamte the forecasdting model, for each $t > R$. The rolling case can be treated analogously, and it is omitted only for brevity.

and

$$\widetilde{S}_n^{C+}=\int_{\mathcal{X}^+}\sum_{j=2}^{k}\left(\max\left\{0,\frac{\sqrt{n}\widetilde{C}_{j,n}^{+}(x)}{\widetilde{\sigma}_{jj,n}^{C+}(x)+\varepsilon}\right\}\right)^2 \mathrm{d}Q(x),$$

where $\widetilde{\sigma}_{jj,n}^{G+}(x)$ and $\widetilde{\sigma}_{jj,n}^{C+}(x)$ include terms accounting for the contribution of parameter estimation error to asymptotic covariance. Here, $\widetilde{\sigma}_{jj,n}^{2,G+}(x)$ is defined as:

$$\begin{aligned}\widetilde{\sigma}_{jj,n}^{2,G+}(x) &= \widehat{\sigma}_{jj,n}^{2,G+}(x)+2\widehat{\Pi}\widehat{f}_{1,n,h}^2(x)\widehat{A}_1\widehat{\Sigma}_{11}\widehat{A}_1'+2\widehat{\Pi}\widehat{f}_{j,n,h}^2(x)\widehat{A}_j\widehat{\Sigma}_{jj}\widehat{A}_j' \\ &\quad -4\widehat{\Pi}\widehat{f}_{1,n,h}(x)\widehat{A}_1\widehat{\Sigma}_{1j}\widehat{A}_j'\widehat{f}_{j,n,h}(x)+2\widehat{\Pi}\widehat{f}_{1,n,h}(x)\widehat{A}_1\widehat{\Sigma}_{u1}(x)-2\widehat{\Pi}\widehat{f}_{j,n,h}(x)\widehat{A}_j\widehat{\Sigma}_{uj}(x),\end{aligned}$$

where $\widehat{\sigma}_{jj,n}^{2,G+}(x)$ is defined as in the text, but using only the last n observations, $\widehat{\Pi}=1-\frac{R}{n}\ln\left(1+\frac{n}{R}\right)$,

$$\widehat{f}_{j,n,h}(x)=\frac{1}{nh}\sum_{t=R+1}^{n}K\left(\frac{\widehat{e}_{j,t}-x}{h}\right),$$

$$\widehat{A}_j=\frac{1}{n}\sum_{t=R+1}^{T}\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\widehat{\theta}_{j,R}\right)'\left(\frac{1}{R}\sum_{t=1}^{R}\nabla_{\theta_j}^2 m_j(X_t,Z_{j,t-1},\widehat{\theta}_{j,R})\right)^{-1},$$

$$\begin{aligned}\widehat{\Sigma}_{jj} &= \frac{1}{n}\sum_{t=R+1}^{T}\nabla_{\theta_j}m_j(X_t,Z_{t,i-1},\widehat{\theta}_{j,R})\nabla_{\theta_j}m_j(X_t,Z_{t,i-1},\widehat{\theta}_{j,R})' \\ &\quad +2\frac{1}{n}\sum_{\tau=1}^{l_n}\sum_{t=R+\tau+1}^{T}w_\tau\nabla_{\theta_j}m_j(X_t,Z_{t,i-1},\widehat{\theta}_{j,R})\nabla_{\theta_j}m_j(X_{t-\tau},Z_{t-\tau,i-1},\widehat{\theta}_{j,R})',\end{aligned}$$

and

$$\begin{aligned}\widehat{\Sigma}_{uj}(x) &= \frac{1}{n}\sum_{t=R+1}^{T}\nabla_{\theta_j}m_j(X_t,Z_{t,i-1},\widehat{\theta}_{j,R})\left(\left(1\left\{\widehat{e}_{j,t}\le x\right\}-\frac{1}{n}\sum_{t=1}^{n}1\left\{\widehat{e}_{j,t}\le x\right\}\right)\right. \\ &\quad \left.-\left(1\left\{\widehat{e}_{1,t}\le x\right\}-\frac{1}{n}\sum_{t=1}^{n}1\left\{\widehat{e}_{1,t}\le x\right\}\right)\right) \\ &\quad +2\frac{1}{n}\sum_{\tau=1}^{l_n}\sum_{t=R+\tau+1}^{T}w_\tau\nabla_{\theta_j}m_j(X_t,Z_{t,i-1},\widehat{\theta}_{j,R})\left(\left(1\left\{\widehat{e}_{j,t-\tau}\le x\right\}-\frac{1}{n}\sum_{t=1}^{n}1\left\{\widehat{e}_{j,t-\tau}\le x\right\}\right)\right. \\ &\quad \left.-\left(1\left\{\widehat{e}_{1,t-\tau}\le x\right\}-\frac{1}{n}\sum_{t=1}^{n}1\left\{\widehat{e}_{1,t-\tau}\le x\right\}\right)\right),\end{aligned}$$

By noting that

$$
\begin{aligned}
& \widetilde{C}^{+}_{j,n}(x) \\
= \; & \frac{1}{n}\sum_{t=R}^{T-1}\Big([(e_{1,t+1}-x)]_{+}-[(e_{j,t+1}-x)]_{+}\Big) \\
& +\frac{1}{n}\sum_{t=R}^{T-1}\left((\widehat{e}_{1,t}-e_{1,t})\,1\{e_{1,t}\geq x\}-(\widehat{e}_{j,t}-e_{j,t})\,1\{e_{j,t}\geq x\}\right) \\
& +\frac{1}{n}\sum_{t=R}^{T-1}\left((e_{1,t}-x)\left(1\{\widehat{e}_{1,t}\geq x\}-1\{\widehat{e}_{1,t}\geq x\}\right)-(e_{j,t}-x)\left(1\{\widehat{e}_{j,t}\geq x\}-1\{e_{j,t}\geq x\}\right)\right) \qquad (1.5) \\
& +\frac{1}{n}\sum_{t=R}^{T-1}\left((\widehat{e}_{1,t}-e_{1,t})\left(1\{\widehat{e}_{1,t}\geq x\}-1\{e_{1,t}\geq x\}\right)-(\widehat{e}_{j,t}-e_{j,t})\left(1\{\widehat{e}_{j,t}\geq x\}-1\{e_{j,t}\geq x\}\right)\right)
\end{aligned}
$$

we see that $\widetilde{\sigma}^{2,C+}_{jj,n}(x)$ is defined as:

$$
\begin{aligned}
& \widetilde{\sigma}^{2,G+}_{jj,n}(x) \\
= \; & \widehat{\sigma}^{2,C+}_{jj,n}(x)+2\widehat{\Pi}\widehat{f}^{2}_{1,n,h}(x)\widetilde{A}_1(x)\widehat{\Sigma}_{11}\widetilde{A}_1'(x)+2\widehat{\Pi}\widehat{f}^{2}_{j,n,h}(x)\widetilde{A}_j(x)\widehat{\Sigma}_{jj}\widetilde{A}_j'(x) \\
& -4\widehat{\Pi}\widehat{f}_{1,n,h}(x)\widetilde{A}_1(x)\widehat{\Sigma}_{1j}\widetilde{A}_j'(x)\widehat{f}_{j,n,h}(x)+2\widehat{\Pi}\widehat{f}_{1,n,h}(x)\widetilde{A}_1(x)\widehat{\Sigma}_{u1}(x)-2\widehat{\Pi}\widehat{f}_{j,n,h}(x)\widetilde{A}_j(x)\widehat{\Sigma}_{uj}(x) \\
& +2\widehat{\Pi}\widetilde{B}_1(x)\widehat{\Sigma}_{11}\widetilde{B}_1'(x)+2\widehat{\Pi}\widetilde{B}_j'(x)\widehat{\Sigma}_{jj}\widetilde{B}_j'(x)-4\widehat{\Pi}\widetilde{B}_1(x)\widehat{\Sigma}_{1j}\widetilde{B}_j'(x) \\
& +2\widehat{\Pi}\widetilde{B}_1(x)\widehat{\Sigma}_{u1}(x)-2\widehat{\Pi}\widetilde{B}_j(x)'\widehat{\Sigma}_{uj}(x) \\
& +2\widehat{\Pi}\widehat{f}_{1,n,h}(x)\widetilde{A}_1(x)\widehat{\Sigma}_{11}\widetilde{B}_1'(x)+2\widehat{\Pi}\widehat{f}_{j,n,h}(x)\widetilde{A}_j\widehat{\Sigma}_{jj}\widetilde{B}_j'(x)-2\widehat{\Pi}\widetilde{B}_1(x)\widehat{\Sigma}_{1j}\widetilde{A}_j'(x)\widehat{f}_{j,n,h}(x) \\
& -2\widehat{\Pi}\widetilde{B}_j(x)\widehat{\Sigma}_{1j}\widetilde{A}_1'(x)\widehat{f}_{1,n,h}(x),
\end{aligned}
$$

where $\widehat{\sigma}^{2,C+}_{jj,n}(x)$ is defined as in the statement of Lemma 1, but computed using only the last n observations. Also,

$$
\widetilde{A}_j(x)=\frac{1}{n}\sum_{t=R+1}^{T}(\widehat{e}_{t+1,j}-x)\,\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\widehat{\theta}_{j,R}\right)'\left(\frac{1}{R}\sum_{t=1}^{R}\nabla^2_{\theta_j}m_j(X_t,Z_{j,t-1},\widehat{\theta}_{j,R})\right)^{-1}
$$

and

$$
\widetilde{B}_j(x)=\frac{1}{n}\sum_{t=R+1}^{T}1\{\widehat{e}_{t+1,j}>x\}\,\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\widehat{\theta}_{j,R}\right)'\left(\frac{1}{R}\sum_{t=1}^{R}\nabla^2_{\theta_j}m_j(X_t,Z_{j,t-1},\widehat{\theta}_{j,R})\right)^{-1}.
$$

In order to formalize the case of asymptotically non-vanishing parameter estimation error, we require the following assumptions.

Assumption A5: ϕ_j is twice continuously differentiable on the interior of Θ_j and the elements of $\nabla_{\theta_j}\phi_j(Z_{j,i-1},\theta_i)$ and $\nabla^2_{\theta_j}\phi_j(Z_{j,i-1},\theta_i)$ are $p-$dominated on Θ_i, for $j=1,...,k$, with $p>4$.

Assumption A6: For $j=1,...,k$: (i) $\theta_j^{\dagger}$ is uniquely identified (i.e. $E(m_j(X_t,Z_{j,t-1},\theta_j))>E(m_j(X_t,Z_{j,t-1},\theta_j^{\dagger}))$, for any $\theta_j\neq\theta_j^{\dagger}$); (ii) m_j is twice continuously differentiable on the interior of Θ_j; (iii) the elements of $\nabla_{\theta_j}m_j$ and $\nabla^2_{\theta_j}m_j$ are $p-$dominated on Θ_j, with $p>4$; and (iii) $E\left(-\nabla^2_{\theta_j}m_j(\theta_j)\right)$ is positive definite, uniformly on Θ_j.[2]

Assumption A7: $T = R + n$, and as $T \to \infty$, $n/R \to \pi$, with $0 \leq \pi < \infty$.

As explained earlier, it is crucial to have a consistent estimator of the variance of the moment conditions. Otherwise, bootstrap critical values are not scale invariant. Hence, we need to construct estimators which properly capture parameters estimation error, regardless the fact that we rely on bootstrap critical values. GMS tests in the presence of non-vanishing estimation error have been considered in Coroneo, Corradi and Santos-Monteiro (2019). We have the following result.

Lemma 3: *Let Assumptions A1-A3, and A5-A7 hold. If* $l_n \approx n^{\delta}$ $\delta < \frac{1}{2}$, *as defined in Assumption A1, then:*

(i) $\sup_{x \in \mathcal{X}^+} \left| \widetilde{\sigma}^{2,G+}_{jj,n}(x) - \omega^{2,G+}_{jj}(x) \right| = o_p(1)$, *with* $\omega^{2G+}_{jj}(x) = avar\left(\sqrt{n}\widetilde{G}^{+}_{j,n}(x)\right)$; *and*

(ii) $\sup_{x \in \mathcal{X}^+} \left| \widetilde{\sigma}^{2,C+}_{jj,n}(x) - \omega^{2,C+}_{jj}(x) \right| = o_p(1)$, *with* $\omega^{2C+}_{jj}(x) = avar\left(\sqrt{n}\widetilde{C}^{+}_{j,n}(x)\right)$.

Lemma 3 mirrors Lemma 1 for the case of non-vanishing estimation error. In order to provide the analog of Theorem 1, we need define the counterparts of $S^{\dagger G+}_n$ and $S^{\dagger C+}_n$ which take into account of parameter estimation error. Let $\overline{\Omega}^{G+}(x,x) = \Omega^{G+}(x,x) + \varepsilon I_{k-1}$, where $\Omega^{G+}(x,x) = [\omega^{2,G+}_{ij,n}(x)]$. Also,

$$\mathcal{D}^{G+}(x) = \text{diag}\Omega^{G+}(x,x),$$

$$h^{G+}_{1,n}(x) = \mathcal{D}^{G+}(x)^{-1/2}\left(\sqrt{n}G^{+}_{2}(x), ..., \sqrt{n}G^{+}_{k}(x)\right)',$$

$$h^{G+}_{2}(x,x') = \mathcal{D}^{G+}(x)^{-1/2}\overline{\Omega}^{G+}(x,x')\,\mathcal{D}^{G+}(x')^{-1/2},$$

and

$$v^{G+}(.) = (v^{G+}_{2}(.), ..., v^{G+}_{k}(.))'.$$

Here, $v^{G+}(.)$ is a $(k-1)-$dimensional zero mean Gaussian process with correlation $h^{G+}_{2}(x,x')$. Also, let $\mathcal{D}^{C+}(x), h^{C+}_{1,n}(x), h^{C+}_{2}(x,x')$, and $v^{C+}(.)$ be defined analogously by replacing $\Omega^{G+}(x,x), G^{+}_{2}(x), ..., G^{+}_{k}(x)$ with $\Omega^{C+}(x,x), C^{+}_{2}(x), ..., C^{+}_{k}(x)$.

Finally, define

$$S^{\ddagger G+}_n = \int_{\mathcal{X}^+} \sum_{j=2}^{k} \left(\max\left\{ 0, \frac{v^{G+}_{j}(x) + h^{G+}_{j,1,n}(x)}{\sqrt{h^{G+}_{jj,2}(x)}} \right\} \right)^2 \mathrm{d}Q(x),$$

where $h^{G+}_{jj,2}(x)$ is the $jj-$th element of $h^{G+}_{2}(x,x)$, and let

$$S^{\ddagger C+}_n = \int_{\mathcal{X}^+} \sum_{j=2}^{k} \left(\max\left\{ 0, \frac{v^{C+}_{j}(x) + h^{C+}_{j,1,n}(x)}{\sqrt{h^{C+}_{jj,2}(x)}} \right\} \right)^2 \mathrm{d}Q(x),$$

which is defined analogously, by replacing $v^{G+}_{j}(x), h^{G+}_{j,1,n}(x)$, and $h^{G+}_{jj,2}(x)$ with $v^{C+}_{j}(x), h^{C+}_{j,1,n}(x)$, and $h^{C+}_{jj,2}(x)$. The following result holds.

Theorem 5: *Let Assumptions A1-A7 hold.*

[2] We say that $\nabla_{\theta_j} \ln f_j(y_t, Z^{t-1}, \theta_j)$ is $2r-$dominated on Θ_j if its $v-th$ element, $v = 1, ..., \varrho(j)$, is such that $\left|\nabla_{\theta_j} \ln f_j(y_t, Z^{t-1}, \theta_j)\right|_v \leq D_t$, and $E(|D_t|^{2r}) < \infty$. For more details on domination conditions, see Gallant and White (1988, pp. 33).

(i) Under H_0^{G+}, there exist $\delta > 0$ such that:

$$\lim_{n\to\infty}\sup \sup_{P\in\mathcal{P}_0^{G+}} \left[P\left(\widetilde{S}_n^{G+} > a_{h_{A,n}}^{G+}\right) - P\left(S_n^{\ddagger G+} + \delta > a_{h_{A,n}}^{G+}\right)\right] \leq 0$$

and

$$\lim_{n\to\infty}\inf \inf_{P\in\mathcal{P}_0^{G+}} \left[P\left(\widetilde{S}_n^{G+} > a_{h_{A,n}}^{G+}\right) - P\left(S_n^{\ddagger G+} - \delta > a_{h_{A,n}}^{G+}\right)\right] \geq 0.$$

(ii) Under H_0^{C+}, there exist $\delta > 0$ such that:

$$\lim_{n\to\infty}\sup \sup_{P\in\mathcal{P}_0^{C+}} \left[P\left(\widetilde{S}_n^{C+} > a_{h_{A,n}}^{C+}\right) - P\left(S_n^{\ddagger C+} + \delta > a_{h_{A,n}}^{C+}\right)\right] \leq 0$$

and

$$\lim_{n\to\infty}\inf \inf_{P\in\mathcal{P}_0^{C+}} \left[P\left(\widetilde{S}_n^{C+} > a_{h_{A,n}}^{C+}\right) - P\left(S_n^{\ddagger C+} - \delta > a_{h_{A,n}}^{C+}\right)\right] \geq 0.$$

Theorem 5 provides upper and lower bounds for $P\left(\widetilde{S}_n^{G+} > a_{h_{A,n}}^{G+}\right)$ and $P\left(\widetilde{S}_n^{C+} > a_{h_{A,n}}^{C+}\right)$, uniformly, over the probabilities under the null H_0^{G+} and H_0^{C+}, respectively.

1.2 Bootstrap Estimators

When computing recursive $m-$estimators, it is important to note that earlier observations are used more frequently than temporally subsequent observations. On the other hand, in the standard block bootstrap, all blocks from the original sample have the same probability of being selected, regardless of the dates of the observations in the blocks. Thus, the bootstrap estimator, say $\widehat{\theta}_{j,t}^*$, which is constructed as a direct analog of $\widehat{\theta}_{j,t}$ in (1.2), is characterized by a location bias that can be either positive or negative, depending on the sample that we observe. In order to circumvent this problem, Corradi and Swanson (2007) suggest a re-centering of the bootstrap score which ensures that the new bootstrap estimator is asymptotically unbiased. Also, assume that $T = R + n = b_T l_T$, with $b_T = b_n \frac{T}{n}$ and $l_T = l_n \frac{T}{n}$, and define:

$$\widetilde{\theta}_{j,t}^* = \arg\min_{\theta_j\in\Theta_j} \frac{1}{t}\sum_{i=1}^{t}\left(m_j(X_i^*, Z_{j,i-1}^*, \theta_j) - \theta_j' \left(\frac{1}{T}\sum_{k=1}^{T-1} \nabla_{\theta_j} m_j(X_k, Z_{j,k-1}, \widehat{\theta}_{j,t})\right)\right),$$

where $X_i^*, Z_{j,i-1}^*$ are resampled via the "standard" block bootstrap outlined in the previous section, but with block length l_T. Theorem 1 in Corradi and Swanson (2007) establish that $\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widetilde{\theta}_{j,t}^* - \widehat{\theta}_{j,t}\right)$ has the same limiting distribution as $\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{\theta}_{j,t} - \theta_j^\dagger\right)$, conditional of the sample.

With a slight abuse of notation, let $u_{j,t}^*(x) = 1\{e_{j,t}^* \leq x\} - \frac{1}{T}\sum_{t=1}^{T} 1\{\widehat{e}_{j,t} \leq x\}$ and $\eta_{j,t}^*(x) = \left[e_{j,t}^* - x\right]_+ - \frac{1}{T}\sum_{t=1}^{n}\left[\widehat{e}_{j,t} - x\right]_+$, with $e_{j,t+1}^* = X_{t+1}^* - \phi_j\left(Z_{j,t}^*, \widehat{\theta}_{j,t}\right)$, and let $\widehat{u}_{j,t}^*(x) = 1\{\widehat{e}_{j,t}^* \leq x\} - \frac{1}{T}\sum_{t=1}^{T} 1\{\widehat{e}_{j,t} \leq x\}$ and $\widehat{\eta}_{j,t}^*(x) = \left[\widehat{e}_{j,t}^* - x\right]_+ - \frac{1}{T}\sum_{t=1}^{n}\left[\widehat{e}_{j,t} - x\right]_+$, with $\widehat{e}_{j,t+1}^* = X_{t+1}^* - \phi_j\left(Z_{j,t}^*, \widetilde{\theta}_{j,t}^*\right)$. Our first goal is to construct the bootstrap counterparts of $\widetilde{\sigma}_{jj,n}^{2,G+}(x)$ and $\widetilde{\sigma}_{jj,n}^{2,C+}(x)$, called $\widetilde{\sigma}_{jj,n}^{*2,G+}(x)$ and $\widetilde{\sigma}_{jj,n}^{*2,C+}(x)$. Define:

$$
\begin{aligned}
&\widetilde{\sigma}^{*2,G+}_{jj,n}(x)\\
=\;& \widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{u}^*_{j,t}(x)-\widehat{u}^*_{1,t}(x)\right)\right)\\
=\;& \widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(u^*_{j,t}(x)-u^*_{1,t}(x)\right)\right)+\widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{f}^*_{j,n,h}(x)\widehat{PEE^*}_{j,t}-\widehat{f}^*_{1,n,h}(x)\widehat{PEE^*}_{1,t}\right)\right)\\
&-2\widehat{\text{acov}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(u^*_{j,t}(x)-u^*_{1,t}(x)\right),\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{f}^*_{j,n,h}(x)\widehat{PEE^*}_{j,t}-\widehat{f}^*_{1,n,h}(x)\widehat{PEE^*}_{1,t}\right)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
&\widetilde{\sigma}^{*2,C+}_{jj,n}(x)\\
=\;& \widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{\eta}^*_{j,t}(x)-\widehat{\eta}^*_{1,t}(x)\right)\right)\\
=\;& \widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\eta^*_{j,t}(x)-\eta^*_{1,t}(x)\right)\right)\\
&+\widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\left[\widehat{f}^*_{j,n,h}\widehat{PEE}^*_{j,t}-x\right]_+-\left[\widehat{f}^*_{1,n,h}\widehat{PEE}^*_{1,t}-x\right]_+\right)\right)+\\
&-2\widehat{\text{acov}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\eta^*_{j,t}(x)-\eta^*_{1,t}(x)\right),\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\left[\widehat{f}^*_{j,n,h}\widehat{PEE}^*_{j,t}-x\right]_+-\left[\widehat{f}^*_{1,n,h}\widehat{PEE}^*_{1,t}-x\right]_+\right)\right),
\end{aligned}
$$

where avar* and acov* denote asymptotic variances and covariances, with respect to the bootstrap probability measure, $\widehat{f}^*_{j,n,h}$ is an estimator of the density of e_j based on the resampled observations, and $\widehat{PEE^*}_{j,t}$ is an estimator of:

$$
\begin{aligned}
PEE^*_{j,t} \;=\;& \mathrm{E}^*\left(\nabla_{\theta_j}\phi_j\left(Z^*_{j,t},\widetilde{\theta}^*_{j,t}\right)\right)\mathrm{E}^*\left(\nabla^2_{\theta}m_j\left(X^*_i,Z^*_{j,i-1},\widetilde{\theta}^*_{j,t}\right)\right)\\
&\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta}m_j\left(X^*_i,Z^*_{j,i-1},\widehat{\theta}_{j,t}\right)-\frac{1}{T}\sum_{i=1}^{T}\nabla_{\theta_j}m_j(X_k,Z_{j,k-1},\widehat{\theta}_{j,t})\right). \qquad (1.6)
\end{aligned}
$$

Closed form expressions for $\widehat{PEE^*}_{j,t}$, $\widehat{\text{avar}^*}$, and $\widehat{\text{acov}^*}$ are given in the proof of Lemma 4.

Lemma 4: *Let Assumptions A1-A3 and A5-A7 hold. Then, if* $l_n \approx n^{\delta}$ $\delta < \frac{1}{2}$, *and* β *the mixing coefficient in Assumption A1 is such that* $\beta > \frac{6\delta}{1-2\delta}$:

(i) $\sup_{x\in\mathcal{X}^+}\left|\widetilde{\sigma}^{*2,G+}_{jj,n}(x)-\mathrm{E}^*\left(\widetilde{\sigma}^{*2,G+}_{jj,n}(x)\right)\right|=o_p(1)$ *and*

(ii) $\sup_{x\in\mathcal{X}^+}\left|\widetilde{\sigma}^{*2,C+}_{jj,n}(x)-\mathrm{E}^*\left(\widetilde{\sigma}^{*2,C+}_{jj,n}(x)\right)\right|=o_p(1)$.

1.3 Bootstrap Critical Values

The bootstrap statistics in the non-vanishing recursive parameter estimation error case are:

$$\widetilde{S}_n^{*G+} = \int_{\mathcal{X}^+} \sum_{j=2}^{k} \max\left(\left\{0, \frac{\widetilde{v}_{j,n}^{*G+}(x) - \widetilde{\phi}_{j,n}^{G+}(x)}{\sqrt{\widetilde{\overline{h}}_{2,jj}^{*G+}(x)}}\right\}\right)^2 \mathrm{d}Q(x), \tag{1.7}$$

where $\widetilde{\overline{h}}_{2,jj}^{*G+}(x)$ is the jj element of $\widetilde{D}_n^{-1/2,G+}(x)\widetilde{\overline{\Sigma}}_n^{*G+}(x,x)\,\widetilde{D}_n^{-1/2,G+}(x)$, with $\widetilde{D}_n^{G+}(x) = \mathrm{diag}\widetilde{\Sigma}_n^{*G+}(x,x)$, $\widetilde{\Sigma}_n^{*G+}(x,x) = [\widetilde{\sigma}_{ij,n}^{*2,G+}(x)]$ $i,j = 1,...,k$, and $\widetilde{\overline{\Sigma}}_n^{*G+} = \widetilde{\Sigma}_n^{*G+} + \varepsilon I_{k-1}$. Also,

$$\begin{aligned}\widetilde{v}_n^{*G+}(x) &= \sqrt{n}\widetilde{D}_n^{-1/2,G+}(x)\frac{1}{\sqrt{n}}\sum_{i=R+1}^{n}\Big(\left(1\left\{\widehat{e}_{j,i}^{*} \le x\right\} - 1\left\{\widehat{e}_{1,i}^{*} \le x\right\}\right) \\ &\quad \frac{1}{T}\sum_{t=1}^{T}\left(1\left\{\widehat{e}_{j,t} \le x\right\} - 1\left\{\widehat{e}_{1,t} \le x\right\}\right)\Big)\end{aligned}$$

and for $\widetilde{\xi}_{j,n}^{G+}(x) = \kappa_n^{-1}n^{1/2}\widetilde{\overline{D}}_{jj,n}^{-1/2,G+}(x)\widetilde{G}_{j,n}^{+}(x)$,

$$\widetilde{\phi}_{j,n}^{G+}(x) = c_n 1\left\{\widetilde{\xi}_{j,n}^{G+}(x) < -1\right\}. \tag{1.8}$$

Finally, also define

$$\widetilde{S}_n^{*C+} = \int_{\mathcal{X}^+} \sum_{j=2}^{k} \max\left(\left\{0, \frac{\widetilde{v}_{j,n}^{*C+}(x) - \widetilde{\phi}_{j,n}^{C+}(x)}{\sqrt{\widetilde{\overline{h}}_{2,jj}^{*C+}(x)}}\right\}\right)^2 \mathrm{d}Q(x),$$

where $\widetilde{v}_n^{*C+}(x), \widetilde{D}_n^{C^+}(x), \widetilde{\xi}_{j,n}^{C+}(x)$, and $\widetilde{\phi}_{j,n}^{C+}(x)$ are defined analogously to $\widetilde{v}_n^{*G+}(x), \widetilde{D}_n^{G+}(x), \widetilde{\xi}_{j,n}^{G+}(x)$, and $\widetilde{\phi}_{j,n}^{G+}(x)$. It is immediate to see that estimation error contributes to the bootstrap statistics not only as a scaling factor, but also in determining which moment conditions are binding. This is why we need an estimator of the variance, even if inference is based on bootstrap critical values.

We now define the GMS bootstrap critical values for the case of non-vanishing recursive estimation error. Let $\widetilde{c}_{n,B,1-\alpha}^{*G+}\left(\widetilde{\phi}_n^{G+}, \overline{h}_{2,n}^{*G+}\right)$ be the $(1-\alpha)$-th critical value of $\widetilde{S}_n^{*G+}$, based on B bootstrap replications, with $\widetilde{\phi}_n^{G+}$ as in (1.8) and $\widetilde{\overline{h}}_{2,jj}^{*G+}(x)$ as in (1.7). The $(1-\alpha)$-th GMS bootstrap critical value, $\widetilde{c}_{0,n,1-\alpha}^{*G+}\left(\widetilde{\phi}_n^{G+}, \overline{h}_{2,n}^{*G+}\right)$ is defined as:

$$\widetilde{c}_{0,n,1-\alpha}^{*G+}\left(\widetilde{\phi}_n^{G+}, \overline{h}_{B,n}^{*G+}\right) = \lim_{B\to\infty} \widetilde{c}_{n,B,1-\alpha+\eta}^{*G+}\left(\widetilde{\phi}_n^{G+}, \overline{h}_{2,n}^{*G+}\right) + \eta,$$

for arbitrarily small $\eta > 0$. Also, $\widetilde{c}_{n,B,1-\alpha+\eta}^{*C+}\left(\widetilde{\phi}_n^{C+}, \overline{h}_{2,n}^{*C+}\right)$ and $\widetilde{c}_{0,n,1-\alpha}^{*C+}\left(\widetilde{\phi}_n^{C+}, \overline{h}_{B,n}^{*C+}\right)$ are defined analogously. The following result then holds.

Theorem 6: Let Assumptions A1-A7 hold, and let $l_n \to \infty$ and $l_n n^{\frac{1}{3}-\varepsilon} \to 0$, as $n \to \infty$. Under H_0^{G+} :

(i) if as $n \to \infty$, $\kappa_n \to \infty$ and $c_n/\kappa_n \to 0$, then

$$\limsup_{n\to\infty} \sup_{P \in \mathcal{P}_0^{G+}} P\left(\widetilde{S}_n^{G+} \geq \widetilde{c}^{*C+}_{n,B,1-\alpha+\eta}\left(\widetilde{\phi}_n^{C+}, \overline{h}^{*C+}_{2,n}\right)\right) \leq \alpha;$$

and (ii) if as $n \to \infty$, $\kappa_n \to \infty$, $c_n \to \infty$, $\sqrt{n}/\kappa_n \to \infty$ and $Q\left(\mathcal{B}^{G+}\right) > 0$, $\mathcal{B}^{G+}$ as in Eq. (3.13), then

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \sup_{P \in \mathcal{P}_0^{G+}} P\left(\widetilde{S}_n^{G+} \geq \widetilde{c}^{*C+}_{n,B,1-\alpha+\eta}\left(\widetilde{\phi}_n^{C+}, \overline{h}^{*C+}_{2,n}\right)\right) = \alpha.$$

Also, under H_0^{C+},
(iii) if as $n \to \infty$, $\kappa_n \to \infty$ and $c_n/\kappa_n \to 0$, then

$$\limsup_{n\to\infty} \sup_{P \in \mathcal{P}_0^{C+}} P\left(\widetilde{S}_n^{C+} \geq \widetilde{c}^{*C+}_{0,n,1-\alpha}\left(\widetilde{\phi}_n^{C+}, \overline{h}^{*C+}_{B,n}\right)\right) \leq \alpha;$$

and (iv) if as $n \to \infty$, $\kappa_n \to \infty$, $c_n \to \infty$, $\sqrt{n}/\kappa_n \to \infty$ and $Q\left(\mathcal{B}^{C+}\right) > 0$, $\mathcal{B}^{C+}$ as in Eq. (3.14), then

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \sup_{P \in \mathcal{P}_0^{C+}} P\left(\widetilde{S}_n^{C+} \geq \widetilde{c}^{*C+}_{0,n,1-\alpha}\left(\widetilde{\phi}_n^{C+}, \overline{h}^{*C+}_{B,n}\right)\right) = \alpha.$$

Statements (i) and (iii) of Theorem 6 establish that inference based on GMS bootstrap critical values has uniform correct size, in the parameter estimation error case. Statements (ii) and (iv) of the theorem establish that inference based on the GMS bootstrap critical values is asymptotically non-conservative, whenever $Q\left(\mathcal{B}^{+}\right) > 0$ or $Q\left(\mathcal{B}^{C+}\right) > 0$.

1.4 Proofs

Proof of Lemma 3: **(i)** Letting $\overline{F}_j(x) = \frac{1}{n}\sum_{t=R}^{T-1} 1\left\{\widehat{e}_{j,t+1} \leq x\right\}$, by an intermediate value expansion, in the case of a recursive estimation scheme, we have that

$$
\begin{aligned}
& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{\widehat{e}_{j,t+1} \leq x\right\} - F_j(x)\right) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) + \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{\widehat{e}_{j,t+1} \leq x\right\} - 1\left\{e_{j,t+1} \leq x\right\}\right) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) + \left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right\}\right.\right. \\
& \left.\left. - F_j\left(x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right)\right) - \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right)\right) \\
& + \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(F_j\left(x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right) - F_j(x)\right) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) + \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(F_j\left(x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right) - F_j(x)\right) \\
& + o_p(1) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) - f_j(x)\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right) + o_p(1) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) \qquad (1.9) \\
& - f_j(x)\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)'\frac{1}{t}\sum_{i=1}^{t}\left(\nabla^2_{\theta_j} m_j(X_i, Z_{j,i-1}, \theta_j^{\dagger})\right)^{-1}\left(\nabla_{\theta_j} m_j(X_i, Z_{j,i-1}, \theta_j^{\dagger})\right) + o_p(1) \\
= \;& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) - f_j(x)\widehat{A}_j\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_j} m_j(X_i, Z_{j,i-1}, \theta_j^{\dagger})\right) + o_p(1)
\end{aligned}
$$

where the $o_p(1)$ term on the RHS of the third equality in (1.9) comes from the fact that

$$
\begin{aligned}
& \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right\}\right. \\
& \left. - F_j\left(x - \nabla_{\theta_j}\phi_j\left(Z_{j,t+1}, \overline{\theta}_{j,t}\right)\left(\widehat{\theta}_{j,t} - \theta_j^{\dagger}\right)\right)\right) - \frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(1\left\{e_{j,t+1} \leq x\right\} - F_j(x)\right) = o_p(1),
\end{aligned}
$$

because of stochastic equicontinuity.

Hence,

$$
\begin{aligned}
& \operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(\left(1\left\{\widehat{e}_{1,t+1}\leq x\right\}-F_1(x)\right)-\left(1\left\{\widehat{e}_{j,t+1}\leq x\right\}-F_j(x)\right)\right)\right) \\
= \; & \operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(\left(1\left\{e_{1,t+1}\leq x\right\}-F_1(x)\right)-\left(1\left\{e_{j,t+1}\leq x\right\}-F_j(x)\right)\right)\right) \\
& +f_1(x)^2\mathrm{E}\left(\nabla_{\theta_1}\phi_1\left(Z_{1,t+1},\theta_1^\dagger\right)\right)'\left(\mathrm{E}\left(\nabla^2_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right)^{-1} \\
& \operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right) \\
& \left(\mathrm{E}\left(\nabla^2_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right)^{-1}\mathrm{E}\left(\nabla_{\theta_1}\phi_1\left(Z_{1,t+1},\theta_1^\dagger\right)\right) \\
& +f_j(x)^2\mathrm{E}\left(\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\theta_j^\dagger\right)\right)'\left(\mathrm{E}\left(\nabla^2_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right)^{-1} \\
& \operatorname{var}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right) \\
& \left(\mathrm{E}\left(\nabla^2_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right)^{-1}\mathrm{E}\left(\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\theta_j^\dagger\right)\right) \\
& -2f_1(x)f_j(x)\mathrm{E}\left(\nabla_{\theta_1}\phi_1\left(Z_{1,t+1},\theta_1^\dagger\right)\right)'\left(\mathrm{E}\left(\nabla^2_{\theta_1}m_j(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right)^{-1} \\
& \operatorname{cov}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right) \\
& \left(\mathrm{E}\left(\nabla^2_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right)^{-1}\mathrm{E}\left(\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\theta_j^\dagger\right)\right) \\
& +2f_1(x)\mathrm{E}\left(\nabla_{\theta_1}\phi_1\left(Z_{1,t+1},\theta_1^\dagger\right)\right)'\left(\mathrm{E}\left(\nabla^2_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right)^{-1} \\
& \operatorname{cov}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(\left(1\left\{e_{1,t+1}\leq x\right\}-F_1(x)\right)-\left(1\left\{e_{j,t+1}\leq x\right\}-F_j(x)\right)\right)\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_1}m_1(X_i,Z_{1,i-1},\theta_1^\dagger)\right)\right) \\
& -2f_j(x)^2\mathrm{E}\left(\nabla_{\theta_j}\phi_j\left(Z_{j,t+1},\theta_j^\dagger\right)\right)'\left(\mathrm{E}\left(\nabla^2_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right)^{-1} \\
& \operatorname{cov}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\left(\left(1\left\{e_{1,t+1}\leq x\right\}-F_1(x)\right)-\left(1\left\{e_{j,t+1}\leq x\right\}-F_j(x)\right)\right)\frac{1}{\sqrt{n}}\sum_{t=R}^{T-1}\frac{1}{t}\sum_{i=1}^{t}\left(\nabla_{\theta_j}m_j(X_i,Z_{j,i-1},\theta_j^\dagger)\right)\right)
\end{aligned}
$$

(ii) Recalling (1.5) by a similar argument as in part (i).

Proof of Theorem 5:

Given Lemma 3, the statement follows by the same argument as in Theorem 1.

Proof of Lemma 4:

(i) Note that $\widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\eta^*_{j,t}(x)-\eta^*_{1,t}(x)\right)\right)=\widehat{\sigma}^{2*G+}_{jj,n}(x)$ as defined in Eq. (3.1), $\widehat{PEE^*}_{j,t}$ is defined as $PEE^*_{j,t}$ with E^* replaced by an average, also

$$\widehat{\text{avar}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{PEE^*}_{j,t}-\widehat{PEE^*}_{1,t}\right)\right)$$
$$=\frac{1}{b_n}\sum_{k=1}^{b_n}\left(\frac{1}{l_n^{1/2}}\sum_{i=1}^{l_n}\left(\widehat{PEE^*}_{j,(k-1)l_n+i}-\widehat{PEE^*}_{j,(k-1)l_n+i}\right)\right)^2$$

and by Theorem 1 in Corradi and Swanson (2007),

$$\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{PEE^*}_{j,t}-\widehat{PEE^*}_{1,t}\right)$$
$$=\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{PEE}_{j,t}-\widehat{PEE}_{1,t}\right)+o_p(1)^*.$$

$$\widehat{\text{acov}^*}\left(\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(u^*_{j,t}(x)-u^*_{1,t}(x)\right),\frac{1}{\sqrt{n}}\sum_{t=R}^{n-1}\left(\widehat{PEE^*}_{j,t}-\widehat{PEE^*}_{1,t}\right)\right)$$
$$=\frac{1}{b_n}\sum_{k=1}^{b_n}\left(\frac{1}{l_n^{1/2}}\sum_{i=1}^{l_n}\left(\widehat{PEE^*}_{j,(k-1)l_n+i}-\widehat{PEE^*}_{j,(k-1)l_n+i}\right)\right.$$
$$\left.\frac{1}{l_n^{1/2}}\sum_{i=1}^{l_n}\left(u^*_{j,t}(x)-u^*_{1,t}(x)\right)\right)$$

and for $h\to 0$, $nh\to\infty$, $\widehat{f}^*_{j,n,h}(x)=\widehat{f}_{j,n,h}(x)+o_{p^*}(1)=f(x)+o_p(1)+o_{p^*}(1)$. The statement then follow by the same argument as in Lemma 2 and Lemma 3.
(ii) By a similar argument as in Part (i).

Proof of Theorem 6:
(i) By a similar argument as in the proof of Theorem 2 in Corradi and Swanson (2007),

$$\widetilde{S}_n^{*G+}=\int_{\mathcal{X}^+}\sum_{j=2}^{k}\max\left(\left\{0,\frac{\widetilde{v}^{*G+}_{j,n}(x)-\widetilde{\phi}^{G+}_{j,n}(x)}{\sqrt{\widetilde{\overline{h}}^{*G+}_{2,jj}(x)}}\right\}\right)^2\mathrm{d}Q(x)$$
$$=\int_{\mathcal{X}^+}\sum_{j=2}^{k}\max\left(\left\{0,\frac{\widetilde{v}^{G+}_{j,n}(x)-\widetilde{\phi}^{G+}_{j,n}(x)}{\sqrt{\widetilde{\overline{h}}^{G+}_{2,jj}(x)}}\right\}\right)^2\mathrm{d}Q(x)+o_{p^*}(1)$$

The statement then follows from Lemma 4 and Theorem 2.
(ii) By a similar argument as in Part (i).

2 Additional Monte Carlo Experimental Results

In this section, experimental results are tabulated for the following DGPs:

$$\widetilde{e}_{i,t} = \varrho\widetilde{e}_{i,t-1} + (1-\varrho^2)^{1/2}\eta_{i,t}, \text{with } \eta_{kt} \sim i.i.d.N(0,1)\ i = 1, ..., 5$$

DGP11: $e_{1t} = \widetilde{e}_{1,t}$ and $e_{kt} = \widetilde{e}_{k,t}$, $k = 2, 3, 4, 5$.
DGP12: $e_{1t} = \widetilde{e}_{1,t}$, $e_{kt} = \widetilde{e}_{k,t}$, $k = 2, 3$ and $e_{kt} = 1.4\widetilde{e}_{k,t}$, $k = 4, 5$
DGP13: $e_{1t} = \widetilde{e}_{1,t}$, $e_{kt} = 0.8\widetilde{e}_{k,t}$, $k = 2, 3$ and $e_{kt} = 1.2\widetilde{e}_{k,t}$, $k = 4, 5$.
DGP14: $e_{1t} = \widetilde{e}_{1,t}$, $e_{kt} = \widetilde{e}_{k,t}$, $k = 2, 3$ and $e_{kt} = 0.6\widetilde{e}_{k,t}$, $k = 4, 5$.

See Tables S1 and S2 for tabulated results, and Section 4 of the paper for complete details regarding the experiments that were run.

3 Additional Empirical Results

Tables S3 and S4 gather root mean square forecast errors associated with the models reported on in Tables 3 and 4 of the paper. See Section 5 of the paper for a complete discussion.

Table 1: Supplemental S1 – Monte Carlo Results for JCS_n^{G+}, JCS_n^{G-}, JCS_n^{C+}, and JCS_n^{C-} Forecast Superiority Tests*

DGP	n	$J_n = 0.20$	$J_n = 0.35$	$J_n = 0.50$	$J_n = 0.65$	$J_n = 0.20$	$J_n = 0.35$	$J_n = 0.50$	$J_n = 0.65$
		GL Forecast Superiority				CL Forecast Superiority			
		Empirical Size							
DGP11	300	0.125	0.125	0.135	0.150	0.115	0.135	0.140	0.150
	600	0.115	0.130	0.160	0.155	0.115	0.115	0.130	0.145
	900	0.095	0.115	0.110	0.115	0.125	0.135	0.165	0.170
DGP12	300	0.075	0.070	0.075	0.090	0.030	0.030	0.050	0.055
	600	0.085	0.090	0.100	0.110	0.015	0.025	0.030	0.035
	900	0.070	0.065	0.080	0.090	0.020	0.020	0.020	0.025
		Empirical Power							
DGP13	300	0.385	0.430	0.460	0.490	0.650	0.660	0.700	0.745
	600	0.720	0.735	0.745	0.780	0.945	0.950	0.960	0.965
	900	0.900	0.915	0.920	0.935	0.995	1.000	1.000	1.000
DGP14	300	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	600	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

* Notes: Entries denote rejection frequencies of (JCS_n^{G+},JCS_n^{G-}) tests (i.e., GL forecast superiority) and (JCS_n^{C+},JCS_n^{C-}) tests (i.e., CL forecast superiority) under a variety of data generating processes denoted by DGP11-DGP14. In DGP11-DGP12, no alternative outperforms the benchmark model. In DGP13-DGP14, at least one alternative model outperfroms the benchmark model. Sample sizes include n=300, 600, and 900 observations, as indicated in the second column of entries in the table. Nominal test size is 10%, and tests are carried out using critical values constructed for values of J_n including 0.20, 0.35, 0.50, and 0.65. See Section 4 for complete details.

Table 2: Supplemental S2 – Monte Carlo Results for S_n^{G+}, S_n^{G-}, S_n^{C+}, and S_n^{C-} Forecast Superiority Tests*

DGP	n	$\eta = 0.0015$	$\eta = 0.002$	$\eta = 0.0025$	$\eta = 0.003$	$\eta = 0.0015$	$\eta = 0.002$	$\eta = 0.0025$	$\eta = 0.003$
		GL Forecast Superiority				CL Forecast Superiority			
		Empirical Size							
DGP11	300	0.075	0.075	0.075	0.074	0.082	0.082	0.082	0.082
	600	0.114	0.114	0.113	0.112	0.099	0.099	0.098	0.098
	900	0.123	0.122	0.122	0.120	0.153	0.153	0.152	0.151
DGP12	300	0.034	0.033	0.032	0.031	0.040	0.037	0.037	0.036
	600	0.076	0.076	0.076	0.075	0.102	0.102	0.102	0.102
	900	0.092	0.092	0.091	0.091	0.110	0.110	0.110	0.110
		Empirical Power							
DGP13	300	0.896	0.895	0.895	0.895	0.952	0.952	0.952	0.952
	600	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DGP14	300	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	600	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	900	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

* Notes: Entries denote rejection frequencies of (S_n^{G+},S_n^{G-}) tests (i.e., GL forecast superiority) and (S_n^{C+},S_n^{C-}) tests (i.e., CL forecast superiority) under a variety of data generating processes denoted by DGP11-DGP14. In DGP11-DGP12, no alternative outperforms the benchmark model. In DGP13-DGP14, at least one alternative model outperfroms the benchmark model. Sample sizes include n=300, 600, and 900 observations, as indicated in the second column of entries in the table. Nominal test size is 10%, and tests are carried out using critical values constructed for values of η including 0.0015, 0.002, 0.0025, and 0.0030. See Section 4 for complete details.

Table 3: Supplemental S3 – SPF Forecast Pooling Analysis of Quarterly Nominal GDP Using Mean Benchmark Model and Mean Expert Pool Predictions*

Group	*Model*	*Forecast Horizon*				
		$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Group 1	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007237	0.012054	0.016277	0.020490	0.024470
	alternative 2	0.007264	0.012108	0.016358	0.020575	0.024611
	alternative 3	0.007272	0.012141	0.016328	0.020596	0.025276
Group 2	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.008794	0.012535	0.016267	0.021398	0.026000
	alternative 2	0.007476	0.012562	0.017698	0.021331	0.028178
	alternative 3	0.007602	0.012931	0.018113	0.022797	0.025831
Group 3	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007517	0.012197	0.015320	0.019503	0.023099
	alternative 2	0.007128	0.012595	0.016236	0.021133	0.024353
	alternative 3	0.007110	0.012292	0.016534	0.020799	0.024430
Group 4	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007811	0.011955	0.015533	0.019227	0.022871
	alternative 2	0.006971	0.012656	0.017039	0.021008	0.026196
	alternative 3	0.007306	0.012764	0.017251	0.021353	0.025116
Group 5	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007257	0.012007	0.015367	0.019122	0.023173
	alternative 2	0.007197	0.012401	0.016357	0.020830	0.024536
	alternative 3	0.007219	0.012215	0.016720	0.020553	0.024250
Group 6	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007237	0.012054	0.016277	0.020490	0.024470
	alternative 2	0.008794	0.012535	0.016267	0.021398	0.026000
	alternative 3	0.007517	0.012197	0.015320	0.019503	0.023099
	alternative 4	0.007811	0.011955	0.015533	0.019227	0.022871
	alternative 5	0.007257	0.012007	0.015367	0.019122	0.023173
Group 7	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007264	0.012108	0.016358	0.020575	0.024611
	alternative 2	0.007476	0.012562	0.017698	0.021331	0.028178
	alternative 3	0.007128	0.012595	0.016236	0.021133	0.024353
	alternative 4	0.006971	0.012656	0.017039	0.021008	0.026196
	alternative 5	0.007197	0.012401	0.016357	0.020830	0.024536
Group 8	benchmark	0.007245	0.012063	0.016287	0.020506	0.024494
	alternative 1	0.007272	0.012141	0.016328	0.020596	0.025276
	alternative 2	0.007602	0.012931	0.018113	0.022797	0.025831
	alternative 3	0.007110	0.012292	0.016534	0.020799	0.024430
	alternative 4	0.007306	0.012764	0.017251	0.021353	0.025116
	alternative 5	0.007219	0.012215	0.016720	0.020553	0.024250

* Notes: Entries are root mean square forecast errors (RMSFEs) of benchmark and alternative forecasting models for h =0,1,2,3,4. Rejections of the null of no forecast superiority based on S_n tests at a 10% level are denoted by a superscipt on the benchmark model RMSFE - a 1 denotes rejection based on the GL test, and a 2 denotes rejection based on the CL test. Analogous rejections based on application of the J_n test are denoted by superscipts 3 and 4. See Section 5 for complete details.

Table 4: Supplemental S4 – SPF Forecast Pooling Analysis of Quarterly Nominal GDP Using Median Benchmark Model and Median Expert Pool Predictions*

Group	*Model*	*Forecast Horizon*				
		$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Group 1	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007301	0.012033	0.016464	0.020595	0.024785
	alternative 2	0.007340	0.012049	0.016419	0.020566	0.024728
	alternative 3	0.007364	0.012067	0.016498	0.020632	0.025513
Group 2	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.008794	0.012535	0.016267	0.021398	0.026000
	alternative 2	0.007476	0.012562	0.017698	0.021331	0.028178
	alternative 3	0.007602	0.012931	0.018113	0.022797	0.025831
Group 3	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007825	0.012264	0.016016	0.019494	0.023117
	alternative 2	0.007291	0.012425	0.016649	0.021205	0.024090
	alternative 3	0.007372	0.012410	0.016961	0.021245	0.024526
Group 4	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007893	0.012048	0.015983	0.018890	0.022662
	alternative 2	0.007058	0.012606	0.017110	0.021198	0.026052
	alternative 3	0.007231	0.012771	0.017289	0.021477	0.025041
Group 5	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007459	0.011891	0.015555	0.019435	0.023215
	alternative 2	0.007197	0.012299	0.016555	0.020814	0.024760
	alternative 3	0.007400	0.012164	0.017024	0.020698	0.024653
Group 6	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007301	0.012033	0.016464	0.020595	0.024785
	alternative 2	0.008794	0.012535	0.016267	0.021398	0.026000
	alternative 3	0.007825	0.012264	0.016016	0.019494	0.023117
	alternative 4	0.007893	0.012048	0.015983	0.018890	0.022662
	alternative 5	0.007459	0.011891	0.015555	0.019435	0.023215
Group 7	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007340	0.012049	0.016419	0.020566	0.024728
	alternative 2	0.007476	0.012562	0.017698	0.021331	0.028178
	alternative 3	0.007291	0.012425	0.016649	0.021205	0.024090
	alternative 4	0.007058	0.012606	0.017110	0.021198	0.026052
	alternative 5	0.007197	0.012299	0.016555	0.020814	0.024760
Group 8	benchmark	0.007300	0.012036	0.016398	0.020542	0.024699
	alternative 1	0.007364	0.012067	0.016498	0.020632	0.025513
	alternative 2	0.007602	0.012931	0.018113	0.022797	0.025831
	alternative 3	0.007372	0.012410	0.016961	0.021245	0.024526
	alternative 4	0.007231	0.012771	0.017289	0.021477	0.025041
	alternative 5	0.007400	0.012164	0.017024	0.020698	0.024653

* Notes: See notes to Table Supplemental S3.