Online Supplement to "Robust Forecast Comparison"

Sainan Jin[1] , Valentina Corradi[2] and Norman R. Swanson[3]
[1]Singapore Management University
[2]University of Surrey
[3]Rutgers University

September 2016

1 Introduction

The material contained in this supplement includes: (i) tabulated findings from Monte Carlo experiments that are omitted from the aforementioned paper for the sake of brevity; and (ii) proofs of all lemmas used in the aforementioned paper.

2 Additional Monte Carlo Findings

Pairwise comparisons: stationary case

In this part of the supplemental paper, we first re-examine three data generating processes (DGPs) discussed in the aforementioned paper. In the three DGPs, we allow the two forecast errors to be dependent on each other with non-independent observations and generate e_{kt} according to

$$e_{kt} = (1-\lambda)(\sqrt{\rho}\widetilde{e}_{0t} + \sqrt{1-\rho}\widetilde{e}_{kt}) + \lambda e_{k,t-1},$$

where $(\widetilde{e}_{0t}, \widetilde{e}_{1t}, \widetilde{e}_{2t})$ are *i.i.d.* but have different marginals in different DGPs. The parameters λ and ρ determine the mutual dependence of e_{1t} and e_{2t} and their autocorrelations. The three DGPs are presented here for your reference:
DGP2: $(\widetilde{e}_{0t}, \widetilde{e}_{1t}, \widetilde{e}_{2t})$ are $i.i.d$ $N(0, I_3)$;
DGP4: $\widetilde{e}_{1t} \sim i.i.d.N(0, 1.5)$ and $\widetilde{e}_{kt} \sim i.i.d.N(0, 1)$, for $k = 0$ and 2;
DGP6: $\widetilde{e}_{0t} \sim$ *i.i.d.* Beta(1,1), $\widetilde{e}_{1t} \sim$ *i.i.d.* Beta(1,2), and $\widetilde{e}_{2t} \sim$ *i.i.d.* Beta(2,4); where all are recentered around their population means, i.e., 1/2, 1/3 and 1/3, respectively.

The results for the case where $\rho = \lambda = 0.3$ are shown in Table 1 in the aforementioned paper. Simulation results for the DGPs with different values of ρ and λ are reported in Tables 1-4 in this supplement. The main entries in the table are rejection frequencies. Qualitatively similar results obtain when DGPs are specified using other moderate values of ρ and λ (e.g., $\rho = \lambda = 0.5$, $\rho = 0.3$, $\lambda = 0.5$, and $\rho = 0.5$, $\lambda = 0.3$). The tests perform poorly when the forecast errors are highly dependent and strongly autocorrelated (e.g., $\rho = \lambda = 0.8$). In most cases, our tests have worse power performance than the DM test, particularly for small sample sizes; but our tests have better size properties, especially when both ρ and λ are large.

We also conduct Monte Carlo simulations for DGPs with parameter estimation error. In these experiments, DGPs (DGP PEE1 - DGP PEE 16, as presented in the table below) are the same as those examined in Corradi and Swanson (2007). In the setup, the benchmark model (denoted by DGP PEE1 below) is an AR(1). The benchmark model is also called the "small" model. Note that the benchmark or "small" model in our test statistic calculations is always estimated as $y_t = \alpha + \beta y_{t-1} + \epsilon_t$; and the "big" model is the same, but with x_{t-1} added as an additional regressor. The null hypothesis is that the smaller (size) model outperforms the "big" alternative model. DGPs PEE1 - PEE4 satisfy the null hypothesis whereas DGPs PEE 5 - PEE 16 satisfy the alternative hypothesis. Note that only in DGP PEE7 and DGP PEE13 is the alternative model "correct", and all other alternative models are clearly misspecified. Thus the power in these cases might be low. We set $P = 0.5T$, and $T = 600$ as in Corradi and Swanson (2007).

Tables 5-6 show that the results from these simulations are qualitatively similar to those discussed in the paper with no parameter estimation error, when the nulls are least favorable to the alternatives, while the tests are mostly under-sized when the nulls are not least favorable to the alternatives. This verifies our theory, which predicts that the stationary bootstrap works well for least favorable nulls. Our tests in general have comparable power performance relative to the DM test; and the DM test has more conservative sizes compared to our tests in all DGPs.

Data Generating Processes with Parameter Estimation Errors

$u_{1,t} \sim iidN(0,1)$, $u_{2,t} \sim iidN(0,1)$

$x_t = 1 + 0.3x_{t-1} + u_{1,t}$

DGP PEE1: $y_t = 1 + 0.3y_{t-1} + u_{2,t}$

DGP PEE2: $y_t = 1 + 0.3y_{t-1} + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE3: $y_t = 1 + 0.6y_{t-1} + u_{2,t}$

DGP PEE4: $y_t = 1 + 0.6y_{t-1} + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE5 : $y_t = 1 + 0.3y_{t-1} + \exp(\tan^{-1}(x_{t-1}/2)) + u_{3,t}$

DGP PEE6 : $y_t = 1 + 0.3y_{t-1} + \exp(\tan^{-1}(x_{t-1}/2)) + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE7 : $y_t = 1 + 0.3y_{t-1} + x_{t-1} + u_{3,t}$

DGP PEE8 : $y_t = 1 + 0.3y_{t-1} + x_{t-1} + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE9 : $y_t = 1 + 0.3y_{t-1} + x_{t-1}1\{x_{t-1} > 1/(1-0.3)\} + u_{3,t}$

DGP PEE10 : $y_t = 1 + 0.3y_{t-1} + x_{t-1}1\{x_{t-1} > 1/(1-0.3)\} + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE11 : $y_t = 1 + 0.6y_{t-1} + \exp(\tan^{-1}(x_{t-1}/2)) + u_{3,t}$

DGP PEE12 : $y_t = 1 + 0.6y_{t-1} + \exp(\tan^{-1}(x_{t-1}/2)) + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE13 : $y_t = 1 + 0.6y_{t-1} + x_{t-1} + u_{3,t}$

DGP PEE14: $y_t = 1 + 0.6y_{t-1} + x_{t-1} + 0.3u_{3,t-1} + u_{3,t}$

DGP PEE15: $y_t = 1 + 0.6y_{t-1} + x_{t-1}1\{x_{t-1} > 1/(1-0.3)\} + u_{3,t}$

DGP PEE16: $y_t = 1 + 0.6y_{t-1} + x_{t-1}1\{x_{t-1} > 1/(1-0.3)\} + 0.3u_{3,t-1} + u_{3,t}$.

3 Proofs of the Lemmas Used in the Paper

Lemma A.1: *Suppose that Assumptions A.2 and A.4 hold and let* $\alpha \in [0, 0.5)$. *Then, for* $k = 1, ..., l$,

(a) $sup_t \left\| n^{\alpha} H_k(t) \right\| \xrightarrow{p} 0$;

(b) $sup_t \left\| n^{\alpha}(\widehat{\beta}_{k,t} - \beta_{k,0}) \right\| \xrightarrow{p} 0$;

(c) $sup_t \left\| n^{1/2} H_k(t) \right\| = O_p(1)$.

Proof of Lemma A.1: The results follow from Lemma A.1 and the proof of Lemma 2.3.2 of McCracken (2000).

The following lemma holds for all $k = 1, ..., l$.

Lemma A.2: *(a) Suppose that Assumption A.1 holds. Then, for each* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for all* $x, \dot{x} \in \mathcal{X}^-$ or $x, \dot{x} \in \mathcal{X}^+$,

$$\overline{\lim_{T \to \infty}} \left\| \sup_{\rho_g^*((x,\beta_k),(\dot{x},\dot{\beta}_k))<\delta} \left| \nu_{k,n}^g (x, \beta_k) - \nu_{k,n}^g \left(\dot{x}, \dot{\beta}_k\right) \right| \right\|_q < \varepsilon, \tag{3.1}$$

where

$$\rho_g^* \left((x, \beta_k), \left(\dot{x}, \dot{\beta}_k\right)\right) = \left\{ E \left[\left(1(e_{kt}(\beta_k) \leq x) - 1\left(e_{kt}\left(\dot{\beta}_k\right) \leq \dot{x} \right) \right) \right]^2 \right\}^{1/2}. \tag{3.2}$$

(b) Suppose that Assumption A.1 holds. Then, for each* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for all* $x, \dot{x} \in \mathcal{X}^-$ or $x, \dot{x} \in \mathcal{X}^+$,

$$\overline{\lim_{T \to \infty}} \left\| \sup_{\rho_c^*((x,\beta_k),(\dot{x},\dot{\beta}_k))<\delta} \left| \nu_{k,n}^c (x, \beta_k) - \nu_{k,n}^c \left(\dot{x}, \dot{\beta}_k\right) \right| \right\|_q < \varepsilon, \tag{3.3}$$

where

$$\rho_c^*((x, \beta_k), (\dot{x}, \dot{\beta}_k)) = \left\{ E \left| \int_{-\infty}^{x} 1(e_{k,t}(\beta_k) \leq s) ds - \int_{-\infty}^{\dot{x}} 1(e_{k,t}((\dot{\beta}_k) \leq s) ds \right|^r \right\}^{1/r} 1(x < 0,\ \dot{x} < 0)$$

$$+ \left\{ E \left| \int_{x}^{\infty} 1(e_{k,t}(\beta_k) > s) ds - \int_{\dot{x}}^{\infty} 1(e_{k,t}(\dot{\beta}_k) > s) ds \right|^r \right\}^{1/r} 1(x \geq 0,\ \dot{x} \geq 0). \tag{3.4}$$

Proof of Lemma A.2: We first prove part (a). Without loss of generality (WLOG), we verify the conditions of Theorem 2.2 in Andrews and Pollard (1994) hold with $Q = q$ and $\gamma = 1$ for the case when $x, \dot{x} \in \mathcal{X}^+$, which is bounded on the real line. The mixing condition is implied by Assumption A.1(i). The bracketing condition also holds by the following argument. Let

$$\mathcal{F}_k^{g+} = \{1(e_{k,t}(\beta_k) \leq x) : (x, \beta_k) \in \mathcal{X}^+ \times \Theta_{k0}\}.$$

We now show $\mathcal{F}_k^{g+}$ is a class of uniformly bounded functions satisfying the L^2-continuity conditions. Let $\sup_{(\dot{x},\dot{\beta}_k)}$ denote $\sup_{\{(\dot{x},\dot{\beta}_k)\in\mathcal{X}^+\times\Theta_{k0},\ |\dot{x}-x|\le r_1, ||\dot{\beta}_k-\beta_k||\le r_2, \sqrt{r_1^2+r_2^2}\le\tilde{r}\}}$, we have

$$\begin{aligned}
& \sup_t E \sup_{(\dot{x},\dot{\beta}_k)} \left|1(e_{k,t+\tau}(\dot{\beta}_k)\le \dot{x}) - 1(e_{k,t+\tau}(\beta_k)\le x)\right|^2 \\
= \ & E \sup_{(\dot{x},\dot{\beta}_k)} \Big|1(e_{k,t+\tau}\le m_k\left(Z_{k,t+\tau},\dot{\beta}_k\right) - m_k\left(Z_{k,t+\tau},\beta_{k0}\right)+\dot{x}) \\
& -1(e_{k,t+\tau}\le m_k(Z_{k,t+\tau},\beta_k) - m_k(Z_{k,t+\tau},\beta_{k0})+x)| \\
\le \ & E \sup_{(\dot{x},\dot{\beta}_k)} 1\{|e_{k,t+\tau}-m_k(Z_{k,t+\tau},\beta_k)+m_k(Z_{k,t+\tau},\beta_{k0})-x|\le \\
& |m_k(Z_{k,t+\tau},\dot{\beta}_k)-m_k(Z_{k,t+\tau},\beta_k)+\dot{x}-x|\Big\} \\
\le \ & E \sup_{(\dot{x},\dot{\beta}_k)} 1\{|e_{k,t+\tau}-m_k(Z_{k,t+\tau},\beta_k)+m_k(Z_{k,t+\tau},\beta_{k0})-x|\le ||M_k(Z_{k,t+\tau},\beta_k^*)||r_2+r_1\} \\
\le \ & C \sup_{\beta_k\in\Theta_{k,}} E||M_k(Z_{k,t},\beta_k)||r_2+r_1 \\
\le \ & \widetilde{C}\widetilde{r}. \qquad (3.5)
\end{aligned}$$

where β_k^* lies between $\dot{\beta}_k$ and β_k. The first inequality is due to the fact $|1(z\le t)-1(z\le 0)|\le 1(|z|\le|t|)$ for any scalars z and t. The second inequality follows from Assumption A.1(ii), the triangle inequality and the Cauchy-Schwartz inequality. The third inequality holds by Assumptions A.1(ii) and (iii), and $\widetilde{C}=\sqrt{2}C(\sup_{\beta_k\in\Theta_{k0}} E||M_k(Z_{k,t},\beta_k)||\vee 1)$ is finite by Assumption A.1(ii). The desired bracketing condition holds because the L^2-continuity condition implies the bracketing number satisfies

$$N(\varepsilon,\mathcal{F}_k^{g+})\le C(1/\varepsilon)^{L_k+1}.$$

The other cases can be done in the same fashion.

To prove part (b), WLOG, we only verify the case for $x_1\ge 0$ and $x_2\ge 0$. We show that the result follows from Theorem 3 of Hansen (1996) with $a=L_{\max}+1$, $\lambda=1$. Let

$$\mathcal{F}_k^{c+}=\{\int_x^\infty 1(e_{k,t}(\beta_k)>s)ds:(x,\beta_k)\in\mathcal{X}^+\times\Theta_{k0}\}.$$

Then the functions in $\mathcal{F}_k^{c+}$ satisfy the Lipschitz condition:

$$\begin{aligned}
& \left|\int_{\dot{x}}^\infty 1(e_{k,t+\tau}(\dot{\beta}_k)>s)ds - \int_x^\infty 1(e_{k,t+\tau}(\beta_k)>s)ds\right| \\
= \ & \Big|\max\left\{e_{k,t+\tau}+m_k\left(Z_{k,t+\tau},\beta_{k0}\right)-m_k\left(Z_{k,t+\tau},\dot{\beta}_k\right)-\dot{x},0\right\} \\
& -\max\{e_{k,t+\tau}+m_k(Z_{k,t+\tau},\beta_{k0})-m_k(Z_{k,t+\tau},\beta_k)-x,0\}| \\
\le \ & \left|m_k(Z_{k,t+\tau},\dot{\beta}_k)-m_k(Z_{k,t+\tau},\beta_k)\right|+\left|\dot{x}-x\right| \\
\le \ & \sqrt{2}(\sup_{\beta_k\in\Theta_{k0}}||M_k(Z_{k,t+\tau},\beta_k)||\vee 1)(||\dot{\beta}_k-\beta_k||^2+(\dot{x}-x)^2)^{1/2}
\end{aligned}$$

where the first inequality follows from the fact that $|\max\{z_1,0\}-\max\{z_2,0\}|\le|z_1-z_2|$ and the triangle inequality, and the second inequality holds by Assumption A.1*(ii) and the Cauchy-Schwartz inequality.

We have $\max_k \sup_{\beta_k \in \Theta_{k0}} ||M_k(Z_{k,t+\tau}, \beta_k)||_r < \infty$ by Assumption A.1*(ii) which yields the conditions (12) and (13) of Hansen (1996). Finally, the mixing condition (11) in Hansen (1996) holds by Assumption A.1*(i).

Lemma A.3: *Suppose that Assumptions A.1, A.1*, and A.4 hold. Denote* $\zeta^i_{k,t+\tau}(x,\beta) = f^i_{k,t+\tau}(x,\beta) - Ef^i_{k,t+\tau}(x,\beta) - f^i_{k,t+\tau}(x,\beta_0) + Ef^i_{k,t+\tau}(x,\beta_0)$, $i = g, c$. Then, for $k = 2, ..., l$,

(a)

$$\begin{aligned} \sup_t E \sup_{\{\beta\} \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} [\zeta^i_{k,t+\tau}(x,\beta)]^2 &\leq Cn^{-\alpha}\varepsilon, \\ \sup_t E \sup_{\{\beta\} \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^-} [\zeta^i_{k,t+\tau}(x,\beta)]^2 &\leq Cn^{-\alpha}\varepsilon, \ i = g, c \end{aligned}$$

(b)

$$\begin{aligned} \sup_t |E \sup_{\{\beta,\ \dot{\beta}\} \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^i_{k,t+\tau}(x,\beta)\zeta^i_{k,t+\tau+j}(x,\dot{\beta})| &\leq \widetilde{C}\alpha(j)^d (n^{-\alpha}\varepsilon)^2, \\ \sup_t |E \sup_{\{\beta,\ \dot{\beta}\} \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^-} \zeta^i_{k,t+\tau}(x,\beta)\zeta^i_{k,t+\tau+j}(x,\dot{\beta})| &\leq \widetilde{C}\alpha(j)^d (n^{-\alpha}\varepsilon)^2, \end{aligned}$$

where $d = 1$ and $\delta/(2+\delta)$ for $i = g$ and c, respectively.

Proof of Lemma A.3: Part (a) holds directly from the proof of Lemma A.2 by taking $\dot{x} = x$ and $q = 2$ and applying the Cauchy-Schwartz inequality.
For part (b), WLOG, we consider the case $x \geq 0$.

Define $\{x^*, \gamma_1^*, \gamma_2^*\} = \operatorname{argsup}_{\{x \in \mathcal{X}^+,\ \{\gamma_1,\ \gamma_2\} \in N(n^{-\alpha}\varepsilon)\}} \zeta^i_{k,t+\tau}(x,\gamma_1)\zeta^i_{k,t+\tau+j}(x,\gamma_2)$, where we suppress the dependence of $(x^*, \gamma_1^*, \gamma_2^*)$ on $i = g$ or c. By the proof of Lemma A.2, it is easy to verify $\left\|\zeta^i_{k,t+\tau}(x^*,\gamma_1^*)\right\|_{2+\delta} \leq \left\|\sup_{\beta \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^i_{k,t+\tau}(x,\beta)\right\|_{2+\delta} = Cn^{-\alpha}\varepsilon$. By Assumptions A.1, A.1* and Corollary 1.1 of Bosq (1996),

$$\begin{aligned} & |cov(\zeta^g_{k,t+\tau}(x^*,\gamma_1^*),\ \zeta^g_{k,t+\tau+j}(x^*,\gamma_2^*))| \\ \leq\ & 4\alpha(j) \left\| \sup_{\beta \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^g_{k,t+\tau}(x,\beta) \right\|_\infty \left\| \sup_{\beta \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^g_{k,t+\tau+j}(x,\beta) \right\|_\infty \\ \leq\ & C\alpha(j)(n^{-\alpha}\varepsilon)^2, \end{aligned}$$

and

$$\begin{aligned} & |cov(\zeta^c_{k,t+\tau}(x^*,\gamma_1^*),\ \zeta^c_{k,t+\tau+j}(x^*,\gamma_2^*))| \\ \leq\ & 2(1+2/\delta)(2\alpha(j))^{\delta/(2+\delta)} \left\| \sup_{\beta \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^c_{k,t+\tau}(x,\beta) \right\|_{2+\delta} \left\| \sup_{\beta \in N(n^{-\alpha}\varepsilon)} \sup_{x \in \mathcal{X}^+} \zeta^c_{k,t+\tau+j}(x,\beta) \right\|_{2+\delta} \\ \leq\ & C\alpha(j)^{\delta/(2+\delta)}(n^{-\alpha}\varepsilon)^2. \end{aligned}$$

This completes the proof.

Lemma A.4: *(a) Suppose that Assumptions A.1-A.4 hold. Then, we have for* $k = 1, ..., l$,

$$\sup_{x\in\mathcal{X}^+} |\xi^g_{k1}(x) - \nu^g_{k,n}(x, \beta_{k0}) + \nu^g_{1,n}(x, \beta_{1,0})| \xrightarrow{p} 0, \tag{3.6}$$
$$\sup_{x\in\mathcal{X}^-} |\xi^g_{k1}(x) - \nu^g_{k,n}(x, \beta_{k0}) + \nu^g_{1,n}(x, \beta_{1,0})| \xrightarrow{p} 0$$

(b) Suppose that Assumptions A.1, A.2, A.3* and A.4 hold. Then, we have for* $k = 1, ..., l$,

$$\sup_{x\in\mathcal{X}^+} |\xi^c_{k1}(x) - \nu^c_{k,n}(x, \beta_{k0}) + \nu^c_{1,n}(x, \beta_{10})| \xrightarrow{p} 0, \tag{3.7}$$
$$\sup_{x\in\mathcal{X}^-} |\xi^c_{k1}(x) - \nu^c_{k,n}(x, \beta_{k0}) + \nu^c_{1,n}(x, \beta_{10})| \xrightarrow{p} 0.$$

Proof of Lemma A.4: WLOG, we consider the case $x \geq 0$. Denote $\zeta^i_{k,t+\tau}\left(x, \widehat{\beta}_t\right) = f^i_{k,t+\tau}\left(x, \widehat{\beta}_t\right) - Ef^i_{k,t+\tau}(x, \beta)|_{\beta=\widehat{\beta}_t} - f^i_{k,t+\tau}(x, \beta_0) + Ef^i_{k,t+\tau}(x, \beta_0)$, $i = g$, c, then

$$\xi^i_{k1}(x) - \nu^i_{1,n}(x, \beta_{10}) + \nu^i_{k,n}(x, \beta_{k0}) = n^{-1/2} \sum_t \zeta^i_{k,t+\tau}\left(x, \widehat{\beta}_t\right).$$

Fix ε_0, $\delta > 0$. By Lemma A.1 (b), for all $\varepsilon > 0$, there exists T_0 such that for all $T > T_0$, $P\left(\sup_k \sup_t n^\alpha \left\|\widehat{\beta}_{k,t} - \beta_{k0}\right\| > \varepsilon\right) < \delta/2$. It is useful then to note that for all $T > T_0$ and $\varepsilon_0 > 0$,

$$\begin{aligned} & P\left(\sup_{x\in\mathcal{X}^+} n^{-1/2} \left|\sum_t \zeta^i_{k,t+\tau}\left(x, \widehat{\beta}_t\right)\right| > \varepsilon_0\right) \\ \leq\ & P\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1/2} \left|\sum_t \zeta^i_{k,t+\tau}(x, \beta_t)\right| > \varepsilon_0\right) + P\left(\sup_k \sup_t n^\alpha \left\|\widehat{\beta}_{k,t} - \beta_{k0}\right\| > \varepsilon\right) \\ \leq\ & P\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1/2} \left|\sum_t \zeta^i_{k,t+\tau}(x, \beta_t)\right| > \varepsilon_0\right) + \delta/2 \end{aligned} \tag{3.8}$$

where $\{\beta_t\} \equiv \{\beta_t\}^T_{t=R}$ is a nonrandom sequence. Now we show that there exists $T_1 > T_0$ such that for all $T > T_1$, the first term on the right hand side (r.h.s.) of (3.8) is less than $\delta/2$. For the remainder of this proof only, let $\sum_j$ denote the summation $\sum_{-n+1\leq j\neq 0\leq n-1}$. Applying the Chebyshev's inequality, we have

$$
\begin{aligned}
& \varepsilon_0^2 P\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1/2}\left|\sum_t \zeta^i_{k,t+\tau}(x,\beta_t)\right| > \varepsilon_0\right) \\
\leq\ & E\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1/2}\left|\sum_t \zeta^i_{k,t+\tau}(x,\beta_t)\right|\right)^2 \\
=\ & E\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1}\sum_t \left[\zeta^i_{k,t+\tau}(x,\beta_t)\right]^2\right) \\
& +E\left(\sup_{\{\beta_t,\beta_{t+j}\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} \sum_j \left[n^{-1}\sum_{t=R}^{T-|j|} \zeta^i_{k,t+\tau}(x,\beta_t)\zeta^i_{k,t+\tau+j}\left(x,\beta_{t+j}\right)\right]\right) \\
\leq\ & E\left(\sup_{\{\beta_t\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} n^{-1}\sum_t \left[\zeta^i_{k,t+\tau}(x,\beta_t)\right]^2\right) \\
& +\sum_j\left\{n^{-1}\sum_{t=R}^{T-|j|}\left|E\left[\sup_{\{\beta_t,\beta_{t+j}\}\in N(n^{-\alpha}\varepsilon)} \sup_{x\in\mathcal{X}^+} \zeta^i_{k,t+\tau}(x,\beta_t)\zeta^i_{k,t+\tau+j}(x,\beta_{t+j})\right]\right|\right\}. \qquad (3.9)
\end{aligned}
$$

For part (a), substituting the results of Lemma A.3 into (3.9), the r.h.s. of (3.9) is less than or equal to

$$
\begin{aligned}
& \widetilde{C}(n^{-\alpha}\varepsilon) + \sum_j (1-|j|/n)\widetilde{C}\alpha(j)(n^{-\alpha}\varepsilon)^2 \\
\leq\ & \widetilde{C}n^{-\alpha}\varepsilon\left\{1+2\sum_{j=1}^{n-1}\alpha(j)\right\} \\
\leq\ & Cn^{-\alpha}\varepsilon, \text{ say,} \qquad (3.10)
\end{aligned}
$$

provided $0 < n^{-\alpha}\varepsilon < 1$. Where $0 < C \equiv \left\{1 + 2\sum_{j=1}^{\infty}\alpha(j)\right\}\widetilde{C} < \infty$. Thus we can choose T_1 and ε such that for all $T > T_1 > T_0$, $\varepsilon < (\delta\varepsilon_0^2 n^{\alpha}/2C)$ and $0 < n^{-\alpha}\varepsilon < 1$, the result follows.

Similarly, for part (b), (3.10) holds by Lemma A.3 if we replace $\alpha(j)$ by $\alpha(j)^{\delta/(2+\delta)}$. In this case, $0 < C \equiv \left\{1 + 2\sum_{j=0}^{\infty}\alpha(j)^{\delta/(2+\delta)}\right\}\widetilde{C} < \infty$ by Assumption A.1* (see Eq. (14.6) in Davidson, 1994). Then the result follows analogously.

Lemma A.5: *(a) Suppose that Assumptions A.1-A.4 hold. Then, we have for $k = 1, ..., l$,*

$$
\begin{aligned}
\sup_{x\in\mathcal{X}^+}\left|\xi^g_{k2}(x) - \sqrt{n}\Delta'_{k0}(x)B_k\overline{H}_{k,n} + \sqrt{n}\Delta'_{10}(x)B_1\overline{H}_{1,n}\right| &= o_p(1), \qquad (3.11) \\
\sup_{x\in\mathcal{X}^-}\left|\xi^g_{k2}(x) - \sqrt{n}\Delta'_{k0}(x)B_k\overline{H}_{k,n} + \sqrt{n}\Delta'_{10}(x)B_1\overline{H}_{1,n}\right| &= o_p(1). \qquad (3.12)
\end{aligned}
$$

(b) Suppose that Assumptions A.1*, A.2, A.3* and A.4 hold. Then, we have for $k = 1, ..., l$,

$$
\begin{aligned}
\sup_{x\in\mathcal{X}^+}\left|\xi^c_{k2}(x) - \sqrt{n}\Lambda'_{k0}(x)B_k\overline{H}_{k,n} + \sqrt{n}\Lambda'_{10}(x)B_1\overline{H}_{1,n}\right| &= o_p(1), \qquad (3.13) \\
\sup_{x\in\mathcal{X}^-}\left|\xi^c_{k2}(x) - \sqrt{n}\Lambda'_{k0}(x)B_k\overline{H}_{k,n} + \sqrt{n}\Lambda'_{10}(x)B_1\overline{H}_{1,n}\right| &= o_p(1). \qquad (3.14)
\end{aligned}
$$

Proof of Lemma A. 5: We first prove part (a). Recall that $\Delta_{k0}(x) = (\partial F_k(x, \beta_{k0})/\partial\beta) sgn(x)$ and $\xi^g_{k2}(x) = n^{-1/2}\sum_{t=R}^{T}\left[F_k\left(x, \widehat{\beta}_{k,t}\right) - F_k(x,\beta_{k0}) - F_1\left(x,\widehat{\beta}_{1,t}\right) + F_1(x,\beta_{10})\right] sgn(x)$, WLOG, we consider the case $x \geq 0$ and prove

$$\sup_{x\in\mathcal{X}^+}\left|n^{-1/2}\sum_t\left(F_k\left(x,\widehat{\beta}_{k,t}\right) - F_k(x,\beta_{k0})\right) - \sqrt{n}\left(\frac{\partial F_k(x,\beta_{k0})}{\partial\beta_k'}\right)B_k\overline{H}_{k,n}\right| = o_p(1). \tag{3.15}$$

By Assumption A.3 (i) and the mean value theorem,

$$n^{-1/2}\sum_t\left\{F_k\left(x,\widehat{\beta}_{k,t}\right) - F_k(x,\beta_{k0})\right\} = n^{-1/2}\sum_t\left(\frac{\partial F_k(x,\beta^*_{k,t}(x))}{\partial\beta_k'}\right)\left(\widehat{\beta}_{k,t} - \beta_{k0}\right),$$

where $\beta^*_{k,t}(x)$ lies between $\widehat{\beta}_{k,t}$ and β_{k0}. By Lemma A.1 (b), for all $\alpha \in [0, 1/2)$ and all $\varepsilon > 0$, there exists δ, such that $P(\sup_t \sup_{x\in\mathcal{X}^+} n^\alpha||\beta^*_{k,t}(x) - \beta_{k0}|| \leq \varepsilon) < \delta/2$ for sufficiently large n. Let

$$A_{1n} = \sup_{x\in\mathcal{X}^+}\sup_{\{\beta_k\}\in N_k(n^{-\alpha}\varepsilon)}\left\|\frac{\partial F_k(x,\beta_k)}{\partial\beta_k} - \frac{\partial F_k(x,\beta_{k0})}{\partial\beta_k}\right\|.$$

Then $A_{1n} = O(n^{-\eta\alpha})$ by Assumption A.3(ii).

$$\begin{aligned}
A_{2n} &\equiv \sup_{x\in\mathcal{X}^+}\left\|n^{-1}\sum_t\frac{\partial F_k(x,\beta^*_{k,t}(x))}{\partial\beta_k} - \frac{\partial F_k(x,\beta_{k0})}{\partial\beta_k}\right\| \\
&\leq \sup_{x\in\mathcal{X}^+}\sup_t\left\|\frac{\partial F_k(x,\beta^*_{k,t}(x))}{\partial\beta_k} - \frac{\partial F_k(x,\beta_{k0})}{\partial\beta_k}\right\| = O_p(n^{-\eta\alpha}).
\end{aligned}$$

where the last equality holds because $P(A_{2n} \leq A_{1n}) \to 1$ as $n \to \infty$ by construction. Now we have the desired result

$$\begin{aligned}
&\sup_{x\in\mathcal{X}^+}\left|n^{-1/2}\sum_t\left(F_k(x,\widehat{\beta}_{k,t}) - F_k(x,\beta_{k0})\right) - n^{1/2}\left(\frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right)B_k\overline{H}_{k,n}\right| \\
= &\sup_{x\in\mathcal{X}^+}\left|n^{-1/2}\sum_t\left(\frac{\partial F_k(x,\beta^*_{k,t}(x))}{\partial\beta_k'}\right)\left(\widehat{\beta}_{k,t} - \beta_{k0}\right) - n^{1/2}\left(\frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right)B_k\overline{H}_{k,n}\right| \\
\leq &\sup_{x\in\mathcal{X}^+}\left|n^{-1/2}\sum_t\left(\frac{\partial F_k(x,\beta^*_{k,t}(x))}{\partial\beta_k'} - \frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right)\left(\widehat{\beta}_{k,t} - \beta_{k0}\right)\right| \\
&+\sqrt{n}\sup_{x\in\mathcal{X}^+}\left|\left(\frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right)n^{-1}\sum_t\left(\widehat{\beta}_{k,t} - \beta_{k0}\right) - \left(\frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right)B_k\overline{H}_{k,n}\right| \\
\leq &A_{2n}\sup_t\left\|\sqrt{n}\left(\widehat{\beta}_{k,t} - \beta_{k0}\right)\right\| + \sup_{x\in\mathcal{X}^+}\left\|\frac{\partial F_k(x,\beta_{k0})}{\partial\beta'}\right\|\left\|n^{-1/2}\sum_t\left(\widehat{\beta}_{k,t} - \beta_{k0}\right) - B_k\sqrt{n}\overline{H}_{k,n}\right\| \\
= &o_p(1) + o_p(1) = o_p(1)
\end{aligned}$$

where the first $o_p(1)$ follows from the fact that $A_{2n}\sup_{t=R,\ldots,T}\left\|\sqrt{n}\left(\widehat{\beta}_{k,t} - \beta_{k0}\right)\right\| = O_p(n^{-\alpha(1+\eta)+1/2}) = o_p(1)$ for all $\alpha \in (1/2(1+\eta), 1/2)$ by Lemma A.1(b), and the second $o_p(1)$ holds by Assumption A.3(iii), Lemma A.1(c) and the following argument

$$\begin{aligned}\left\| n^{-1/2}\sum_{t=R}^{T}\left(\widehat{\beta}_{k,t}-\beta_{k0}\right)-B_k\sqrt{n}\overline{H}_{k,n}\right\| &= \left\| n^{-1/2}\sum_{t=R}^{T}B_k(t)H_k(t)-B_k n^{-1/2}\sum_{t=R}^{T}H_k(t)\right\| \\ &= \left\| n^{-1/2}\sum_{t=R}^{T}(B_k(t)-B_k)H_k(t)\right\| \\ &\leq \sup_t \|B_k(t)-B_k\|\sup_t n^{1/2}\|H_k(t)\| \\ &= o_p(1)O_p(1)=o_p(1).\end{aligned}$$

The proof of part (b) is similar and thus omitted.

Lemma A.6: *(a) Suppose that Assumptions A.1-A.4 hold. Then, we have for* $k=2,...,l$,

$$\begin{pmatrix} v^g_{k,n}(\cdot,\beta_{k,0})-v^g_{1,n}(\cdot,\beta_{1,0}) \\ \sqrt{n}\overline{H}_{k,n} \\ \sqrt{n}\overline{H}_{1,n}\end{pmatrix} \Rightarrow \begin{pmatrix}\widetilde{g}_k(\cdot) \\ v_{k0} \\ v_{10}\end{pmatrix}$$

and except at zero, the sample paths of $\widetilde{g}_k(\cdot)$ *are uniformly continuous with respect to a pseudometric* ρ_g *on* $\mathcal{X}$ *with probability one, where for* $x_1,x_2\in\mathcal{X}^+$ *or* $x_1,x_2\in\mathcal{X}^-$,

$$\rho_g(x_1,x_2)=\left\{E[(1(e_{1,t}\leq x_1)-1(e_{k,t}\leq x_1))-(1(e_{1,t}\leq x_2)-1(e_{k,t}\leq x_2))]^2\right\}^{1/2}.$$

(b) Suppose Assumptions A.1, A.2, A.3* and A.4 hold. Then, we have for* $k=2,...,l$,

$$\begin{pmatrix} v^c_{k,n}(\cdot,\beta_{k0})-v^c_{1,n}(\cdot,\beta_{10}) \\ \sqrt{n}\overline{H}_{k,n} \\ \sqrt{n}\overline{H}_{1,n}\end{pmatrix} \Rightarrow \begin{pmatrix}\widetilde{c}_k(\cdot) \\ v_{k0} \\ v_{10}\end{pmatrix}$$

and except at zero, the sample paths of $\widetilde{c}_k(\cdot)$ *are uniformly continuous with respect to a pseudometric* ρ_c *on* $\mathcal{X}$ *with probability one, where for* $x_1,x_2\in\mathcal{X}^+$ *or* $x_1,x_2\in\mathcal{X}^-$,

$$\rho_c(x_1,x_2)=\left\{E\left|\int_{-\infty}^{x_1}(1(e_{1,t}\leq s)-1(e_{k,t}\leq s))ds-\int_{-\infty}^{x_2}(1(e_{1,t}\leq s)-1(e_{k,t}\leq s))ds\right|^r\right\}^{1/r}1(x_1<0,x_2<0)$$

$$+\left\{E\left|\int_{x_1}^{\infty}(1(e_{1,t}>s)--1(e_{k,t}>s)))ds-\int_{x_2}^{\infty}(1(e_{1,t}>s)-1(e_{k,t}>s)))ds\right|^r\right\}^{1/r}1(x_1\geq 0,x_2\geq 0).$$

Proof of Lemma A.6: We first prove (a). By Theorem 10.2 of Pollard (1990), the results hold if we have (i) total boundedness of the pseudometric space $\left(\mathcal{X},\rho_g\right)$, (ii) stochastic equicontinuity of $\left\{v^g_{k,n}(\cdot,\beta_{k0})-v^g_{1,n}(\cdot,\beta_{10}):n\geq 1\right\}$ and (iii) finite dimensional (fidi) convergence. The first two conditions follow from Lemma A.2. We now verify condition (iii), i.e., we need to show that $(v^g_{k,n}(x_1,\beta_{k0})$ $-v^g_{1,n}(x_1,\beta_{10}),...,v^g_{k,n}(x_J,\beta_{10})-v^g_{1,n}(x_J,\beta_{k0}),\ \sqrt{n}\overline{H}'_{k,n},\sqrt{n}\overline{H}'_{1,n})'$ converges in distribution to $(\widetilde{g}_k(x_1),$ $...,\widetilde{g}_k(x_J),\ v'_{k0},v'_{10})'\ \ \forall x_1,...,x_J\in\mathcal{X}^+$ or $x_1,...,x_J\in\mathcal{X}^-$, and $\forall J\geq 1$. The central limit theorem (CLT) holds for $\sqrt{n}\overline{H}_{k,n}$ by Lemma 4.1 in West (1996). A CLT for bounded random variables under $\alpha-$mixing conditions (see Hall and Heyde, 1980) hold for $v^g_{k,n}(x_j,\beta_{k0})-v^g_{1,n}(x_j,\beta_{10})$, $j=1,...,J$. Then one obtains the above weak convergence result by the Cramer-Wold device. This establishes part (a).

For part (b), we need to verify the fidi convergence again. Note that the moment condition of Hall and Heyde (1980, Corollary 5.1) holds since (WLOG), for $x > 0$,

$$E\left|\int_x^\infty (1(e_{1,t} > s)ds - 1(e_{k,t} > s))ds\right|^{2+\delta} \leq E\left|e_{1,t} - e_{k,t}\right|^{2+\delta} < \infty.$$

The mixing condition also holds since we have $\sum \alpha(j)^{\delta/(2+\delta)} \leq C \sum j^{-M\delta/(2+\delta)} < \infty$ by Assumption A.1*.

Lemma HA.1: *(a) Suppose that Assumption HA.1 holds. Then, for each* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for all* x, $\dot{x} \in \mathcal{X}^+$ or x, $\dot{x} \in \mathcal{X}^-$,

$$\overline{\lim_{T\to\infty}} \left\| \sup_{\rho^*_{hg}(x,\dot{x}))<\delta} |\nu^{hg}_{k,n}(x) - \nu^{hg}_{k,n}(\dot{x})| \right\|_q < \varepsilon, \tag{3.16}$$

where

$$\rho^*_{hg}(x,\dot{x}) = \{E[1(e_{k,t+\tau} \leq x) - 1(e_{k,t+\tau} \leq \dot{x})]^2\}^{1/2}. \tag{3.17}$$

(b) Suppose that Assumption HA.1 holds. Then, for each* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that for all* x, $\dot{x} \in \mathcal{X}^+$ or x, $\dot{x} \in \mathcal{X}^-$,

$$\overline{\lim_{T\to\infty}} \left\| \sup_{\rho^*_{hc}(x,\dot{x})<\delta} |\nu^{hc}_{k,n}(x) - \nu^{hc}_{k,n}(\dot{x})| \right\|_q < \varepsilon, \tag{3.18}$$

where

$$\begin{aligned}\rho^*_{hc}(x,\dot{x}) = &\left\{E\left|\int_{-\infty}^x 1(e_{k,t+\tau} \leq s)ds - \int_{-\infty}^{\dot{x}} 1(e_{k,t+\tau} \leq s)ds\right|^r\right\}^{1/r} 1(x<0,\ \dot{x}<0) \\ &+ \left\{E\left|\int_x^\infty 1(e_{k,t+\tau} > s)ds - \int_{\dot{x}}^\infty 1(e_{k,t+\tau} > s)ds\right|^r\right\}^{1/r} 1(x \geq 0,\ \dot{x} \geq 0). \end{aligned} \tag{3.19}$$

Proof of Lemma HA.1: Assumptions HA.4 and HA.1 (i) (resp. HA.1*(i)) imply that $\{e_{k,t+\tau} : t \geq R\}$ is an $\alpha-$mixing sequence with mixing coefficients $\alpha(l) = O(l^{-C_0})$, where C_0 is as defined in HA.1 (resp. HA.1*). Note that Theorem 2.2 in Andrews and Pollard (1994) and Theorem 3 in Hansen (1996) do not require any stationarity assumption, the proof is analogous to that of Lemma A.2. For example, for part (b), Eq. (12) of Hansen (1996) is satisfied with our mixing coefficient $C_0 = 1/q - 1/r$, Eq. (12) is true by Assumption HA.1 (ii) and Equation (13) is satisfied with the dominating function b =1. Then theorem 3 in Hansen (1996) follows by taking $a = 1$ and $\lambda = 1$.

Lemma HA.2: *(a) Suppose Assumptions HA.1* and HA.4 hold. Then, we have for* $k = 2, ..., l$,

$$v^{hg}_{k,n}(\cdot) - v^{hg}_{1,n}(\cdot) \Rightarrow \widetilde{hg}_k(\cdot)$$

and except at zero, the sample paths of $\widetilde{hg}_k(\cdot)$ are uniformly continuous with respect to a pseudometric ρ_{hg} on $\mathcal{X}$ with probability one, where for $x_1, x_2 \in \mathcal{X}^+$ or $x_1, x_2 \in \mathcal{X}^-$,

$$\rho_{hg}(x_1,x_2) = \left\{E[(1(e_{1,t+\tau} \leq x_1) - 1(e_{k,t+\tau} \leq x_1)) - (1(e_{1,t+\tau} \leq x_2) - 1(e_{k,t+\tau} \leq x_2))]^2\right\}^{1/2}.$$

(b) Suppose that Assumptions A.1, A.2, A.3* and A.4 hold. Then, we have for* $k = 2, ..., l$,

$$v_{k,n}^{hc}(\cdot) - v_{1,n}^{hc}(\cdot) \Rightarrow \widetilde{hc}_k(\cdot)$$

and except at zero, the sample paths of $\widetilde{hc}_k(\cdot)$ are uniformly continuous with respect to a pseudometric ρ_{hc} on $\mathcal{X}$ with probability one, where for $x_1, x_2 \in \mathcal{X}^+$ or $x_1, x_2 \in \mathcal{X}^-$,

$$\begin{aligned} &\rho_{hc}(x_1, x_2) \\ = \; &\left\{ E \left| \int_{-\infty}^{x_1} (1(e_{1,t+\tau} \le s) - 1(e_{k,t+\tau} \le s))ds - \int_{-\infty}^{x_2} (1(e_{1,t+\tau} \le s) - 1(e_{k,t+\tau} \le s))ds \right|^r \right\}^{1/r} 1(x_1 < 0, x_2 < 0) \\ &+ \left\{ E \left| \int_{x_1}^{\infty} (1(e_{1,t+\tau} > s) - 1(e_{k,t+\tau} > s)))ds - \int_{x_2}^{\infty} (1(e_{1,t+\tau} > s) - 1(e_{k,t+\tau} > s)))ds \right|^r \right\}^{1/r} 1(x_1 \ge 0, x_2 \ge 0). \end{aligned}$$

Proof of Lemma HA.2: The proof is analogous to that of Lemma A.6. The total boundedness of the pseudometric space $(\mathcal{X}, \rho_i)$, $i = hg$ and hc, and the stochastic equicontinuity of $\left\{ v_{k,n}^i(\cdot) - v_{1,n}^i(\cdot) : n \ge 1 \right\}$, $i = hg$ and hc follow from Lemma HA.1. The finite dimensional convergence follows from Hall and Heyde (1980). Then the result follows from Theorem 10.2 of Pollard (1990).

References

[1] Andrews, D. W. K. and D. Pollard (1994), An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes, *International Statistical Review* 62, 119-132.

[2] Corradi, V. and N. R. Swanson (2007), Nonparametric Bootstrap Procedures For Predictive Inference Based On Recursive Estimation Schemes, *International Economic Review* 48, 67–109.

[3] Davidson, J. (1994), *Stochastic Limit Theory*, Oxford University Press, New York.

[4] Hall, P. and C. C. Heyde (1980), Martingale Limit Theory and Its Applications, Academic Press.

[5] Hansen, B. E. (1996), Stochastic Equicontinuity for Unbounded Dependent Heterogeneous Arrays, *Econometric Theory* 12, 347-359.

[6] McCracken, M. W. (2000), Robust Out-of-Sample Inference, *Journal of Econometrics* 99, 195-223.

[7] Pollard, D. (1990), Empirical Processes: Theory and Applications, CBMS Conference Series in Probability and Statistic, Vol.2, Institute of Mathematical Statistics, Hayward.

[8] West, K. D. (1996), Asymptotic Inference about Predictive Ability, *Econometrica* 64, 1067-1084.

Table 1: GL and CL Forecast Superiority Tests (DGPs 2, 4 and 6)

	DGP2	DGP4	DGP6	DGP2	DGP4	DGP6
S_n	GL forecast superiority			CL forecast superiority		
	n=100			n=100		
0.63	0.136	0.582	0.277	0.170	0.710	0.372
0.54	0.130	0.605	0.259	0.169	0.726	0.361
0.44	0.115	0.558	0.247	0.170	0.712	0.345
0.35	0.129	0.559	0.239	0.131	0.684	0.348
0.25	0.134	0.555	0.290	0.151	0.696	0.380
0.16	0.124	0.541	0.270	0.146	0.693	0.363
DM	0.165	0.845	0.506			
	n=500			n=500		
0.54	0.140	0.978	0.516	0.161	0.996	0.718
0.45	0.127	0.974	0.517	0.118	0.997	0.715
0.36	0.126	0.981	0.491	0.142	0.998	0.678
0.27	0.106	0.975	0.482	0.130	0.999	0.679
0.17	0.105	0.971	0.462	0.125	0.995	0.672
0.08	0.109	0.976	0.478	0.109	0.995	0.661
DM	0.146	1.000	0.906			
	n=1000			n=1000		
0.50	0.129	1.000	0.717	0.147	1.000	0.898
0.41	0.130	1.000	0.698	0.121	1.000	0.891
0.33	0.130	1.000	0.697	0.128	1.000	0.877
0.24	0.090	1.000	0.687	0.108	1.000	0.876
0.15	0.098	1.000	0.699	0.088	1.000	0.879
0.06	0.095	1.000	0.696	0.119	1.000	0.859
DM	0.165	1.000	0.986			

Notes: See Notes to Table 1 in the paper. $\rho = \lambda = 0.5$.

Table 2: GL and CL Forecast Superiority Tests (DGPs 2, 4 and 6)

	DGP2	DGP4	DGP6	DGP2	DGP4	DGP6
S_n	GL forecast superiority			CL forecast superiority		
	n=100			n=100		
0.63	0.116	0.580	0.222	0.120	0.751	0.352
0.54	0.125	0.579	0.219	0.146	0.750	0.341
0.44	0.109	0.579	0.208	0.130	0.762	0.355
0.35	0.112	0.589	0.232	0.123	0.735	0.328
0.25	0.120	0.580	0.259	0.118	0.720	0.348
0.16	0.127	0.615	0.241	0.136	0.770	0.361
DM	0.127	0.887	0.494			
	n=500			n=500		
0.54	0.112	0.990	0.506	0.129	1.000	0.709
0.45	0.112	0.988	0.484	0.088	1.000	0.709
0.36	0.105	0.989	0.479	0.121	1.000	0.709
0.27	0.098	0.983	0.495	0.107	1.000	0.698
0.17	0.101	0.984	0.447	0.121	1.000	0.727
0.08	0.124	0.987	0.472	0.108	1.000	0.713
DM	0.130	1.000	0.927			
	n=1000			n=1000		
0.50	0.115	1.000	0.767	0.107	1.000	0.918
0.41	0.119	1.000	0.718	0.106	1.000	0.909
0.33	0.120	1.000	0.715	0.107	1.000	0.914
0.24	0.089	1.000	0.709	0.094	1.000	0.912
0.15	0.104	1.000	0.749	0.090	1.000	0.912
0.06	0.092	1.000	0.726	0.116	1.000	0.899
DM	0.122	1.000	0.995			

Notes: See Notes to Table 1 in the paper. $\rho = 0.5$, $\lambda = 0.3$.

Table 3: GL and CL Forecast Superiority Tests (DGPs 2, 4 and 6)

	DGP2	DGP4	DGP6	DGP2	DGP4	DGP6
S_n	GL forecast superiority			CL forecast superiority		
	n=100			n=100		
0.63	0.146	0.580	0.222	0.120	0.751	0.352
0.54	0.125	0.579	0.219	0.146	0.750	0.341
0.44	0.109	0.579	0.208	0.130	0.762	0.355
0.35	0.112	0.589	0.232	0.123	0.735	0.328
0.25	0.120	0.580	0.259	0.118	0.720	0.348
0.16	0.127	0.615	0.241	0.136	0.770	0.361
DM	0.127	0.887	0.494			
	n=500			n=500		
0.54	0.112	0.990	0.506	0.129	1.000	0.709
0.45	0.112	0.988	0.484	0.088	1.000	0.709
0.36	0.105	0.989	0.479	0.121	1.000	0.709
0.27	0.098	0.983	0.495	0.107	1.000	0.698
0.17	0.101	0.984	0.447	0.121	1.000	0.727
0.08	0.124	0.987	0.472	0.108	1.000	0.713
DM	0.130	1.000	0.927			
	n=1000			n=1000		
0.50	0.115	1.000	0.767	0.107	1.000	0.918
0.41	0.119	1.000	0.718	0.106	1.000	0.909
0.33	0.120	1.000	0.715	0.107	1.000	0.914
0.24	0.089	1.000	0.709	0.094	1.000	0.912
0.15	0.104	1.000	0.749	0.090	1.000	0.912
0.06	0.092	1.000	0.726	0.116	1.000	0.899
DM	0.122	1.000	0.995			

Notes: See Notes to Table 1 in the paper. $\rho = 0.3, \lambda = 0.5$.

Table 4: GL and CL Forecast Superiority Tests (DGPs 2, 4 and 6)

	DGP2	DGP4	DGP6	DGP2	DGP4	DGP6
S_n	GL forecast superiority			CL forecast superiority		
	n=100			n=100		
0.63	0.213	0.399	0.216	0.280	0.522	0.327
0.54	0.185	0.389	0.198	0.286	0.450	0.338
0.44	0.155	0.378	0.183	0.220	0.480	0.299
0.35	0.181	0.368	0.180	0.223	0.435	0.298
0.25	0.170	0.320	0.209	0.205	0.432	0.292
0.16	0.162	0.315	0.195	0.206	0.389	0.281
DM	0.283	0.566	0.404			
	n=500			n=500		
0.54	0.197	0.660	0.276	0.258	0.763	0.439
0.45	0.184	0.632	0.274	0.210	0.753	0.391
0.36	0.166	0.589	0.259	0.181	0.733	0.369
0.27	0.164	0.563	0.248	0.170	0.685	0.358
0.17	0.131	0.514	0.217	0.141	0.663	0.317
0.08	0.128	0.507	0.192	0.144	0.653	0.300
DM	0.276	0.860	0.546			
	n=1000			n=1000		
0.50	0.181	0.817	0.347	0.221	0.908	0.509
0.41	0.178	0.776	0.309	0.185	0.909	0.489
0.33	0.160	0.763	0.294	0.186	0.889	0.466
0.24	0.144	0.753	0.289	0.158	0.854	0.428
0.15	0.151	0.706	0.269	0.128	0.844	0.396
0.06	0.118	0.663	0.286	0.137	0.799	0.369
DM	0.280	0.953	0.650			

Notes: See Notes to Table 1 in the paper. $\rho = \lambda = 0.8$.

Table 5: GL and CL Forecast Superiority Tests (DGPs PEE 1-8)

	DGP PEE1	DGP PEE2	DGP PEE3	DGP PEE4	DGP PEE5	DGP PEE6	DGP PEE7	DGP PEE8
S_n				GL forecast superiority				
0.53	0.045	0.042	0.051	0.057	0.777	0.675	0.999	0.999
0.44	0.045	0.055	0.052	0.043	0.795	0.679	1.000	0.999
0.35	0.056	0.047	0.044	0.047	0.760	0.689	1.000	0.999
0.26	0.052	0.052	0.049	0.048	0.810	0.683	1.000	1.000
0.17	0.053	0.050	0.055	0.059	0.766	0.721	1.000	1.000
0.08	0.063	0.052	0.064	0.056	0.813	0.704	1.000	0.999
				CL forecast superiority				
0.53	0.057	0.055	0.065	0.048	0.990	0.976	1.000	1.000
0.44	0.059	0.053	0.047	0.063	0.992	0.973	1.000	1.000
0.35	0.055	0.050	0.065	0.061	0.992	0.973	1.000	1.000
0.26	0.059	0.054	0.055	0.067	0.991	0.969	1.000	1.000
0.17	0.065	0.067	0.062	0.087	0.981	0.978	1.000	1.000
0.08	0.059	0.082	0.061	0.075	0.990	0.984	1.000	1.000
DM	0.011	0.003	0.008	0.019	0.999	0.997	1.000	1.000

DGPs PEE1 - PEE4 satisfy the null hypothesis whereas the other DGPs satisfy the alternative hypothesis. n=600. Entry numbers are the rejection frequency in 1000 repetitions. The number of bootstrap resamples is 300. S_n is the bootstrap smoothing parameter. The nominal test size is 10%.

Table 6: GL and CL Forecast Superiority Tests (DGPs PEE9- 16)

	DGPPEE9	DGPPEE10	DGPPEE11	DGPPEE12	DGPPEE13	DGPPEE14	DGPPEE15	DGPPEE16
S_n				GL forecast superiority				
0.53	1.000	1.000	0.776	0.711	1.000	1.000	0.839	0.799
0.44	1.000	1.000	0.780	0.672	1.000	1.000	0.827	0.762
0.35	1.000	1.000	0.775	0.692	1.000	0.999	0.837	0.761
0.26	1.000	1.000	0.778	0.706	1.000	1.000	0.848	0.759
0.17	1.000	1.000	0.793	0.719	1.000	1.000	0.838	0.788
0.08	1.000	1.000	0.813	0.736	1.000	1.000	0.859	0.778
				CL forecast superiority				
0.53	1.000	1.000	0.990	0.978	1.000	1.000	0.989	0.979
0.44	1.000	1.000	0.986	0.974	1.000	1.000	0.981	0.985
0.35	1.000	1.000	0.995	0.979	1.000	1.000	0.992	0.968
0.26	1.000	1.000	0.991	0.983	1.000	1.000	0.985	0.978
0.17	1.000	1.000	0.994	0.985	1.000	1.000	0.988	0.981
0.08	1.000	1.000	0.991	0.987	1.000	1.000	0.986	0.977
DM	1.000	1.000	0.998	0.997	1.000	1.000	0.988	0.984

Notes: See Notes to Table 5. DGPs PEE9 - PEE16 satisfy the alternative hypothesis.