Supplemental appendix for "An asymptotically normal out-of-sample test based on mixed estimation windows"

Gray Calhoun*
Iowa State University

January 9th, 2015

This appendix contains mathematical proofs and some supporting Lemmas for the paper, "An asymptotically normal out-of-sample test based on mixed estimation windows" (Calhoun, 2015). Define the following additional terms:

$$F_t(\beta) = 2(x_t'\beta - \hat{y}_{t+1})x_t',$$

$F_t = F_t(\beta_0)$, $\hat{F}_t = F_t(\hat{\beta}_t)$, $F = \mathrm{E}\, F_t$, $B = (\mathrm{E}\, x_t x_t')^{-1}$, $B_t = (\sum_{s=1}^{t-1} x_s x_s'/(t-1))^{-1}$, and $H_t = \sum_{s=1}^{t-1} x_s \varepsilon_{s+1}/(t-1)$. And let $\|\cdot\|$ denote the L_2 norm in $\mathbb{R}^k$.

Note that Assumptions 1 and 2 imply that f_t, g_t, and F_t are all strong mixing of size $-r/(r-2)$ or uniform mixing of size $-r/(2r-2)$ and are stationary with bounded rth moments.

Theorem 1. *If Assumptions 1–3 hold then*

$$\sqrt{P}(\bar{f} - \mathrm{E}\,\bar{f}^*) \to^d N(0, \sigma^2),$$

with $\sigma^2 = s_1 + 2(s_2 + s_3)$ *and*

$$s_1 = \lim \mathrm{var}(\sqrt{P}\,\bar{f}^*), \qquad s_2 = \lim \mathrm{cov}(\sqrt{P}\,\bar{f}^*, \sqrt{P}\,\bar{g}^*), \qquad s_3 = \lim \mathrm{var}(\sqrt{P}\,\bar{g}^*).$$

Proof. Let R' be a new sequence such that $R' \to \infty$ as $T \to \infty$ and $R' = o(\sqrt{P})$, and then rewrite the centered OOS average as

$$\sqrt{P}(\bar{f} - \mathrm{E}\,\bar{f}^*) = \tfrac{1}{\sqrt{P}} \sum_{t=R'}^{T-1} ((f_t - \mathrm{E}\, f_t) + (\hat{f}_t - f_t)) + \tfrac{1}{\sqrt{P}} \sum_{t=R}^{R'-1} (\hat{f}_t - \mathrm{E}\, f_t). \tag{1}$$

*Economics Department; Iowa State University; Ames, IA 50011. Telephone: (515) 294-6271. Email: «gcalhoun@iastate.edu», web: «www.econ.iastate.edu/~gcalhoun».

Lemma A.1 ensures that the second summation is $o_p(1)$, so we can use a Taylor expansion to rewrite (1) as

$$\begin{aligned}\sqrt{P}(\bar{f} - \mathrm{E}\bar{f}^*) = {} & \tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(f_t - \mathrm{E} f_t) + FB\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1} H_t \\ & + \tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)BH_t + \tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1} F(B_t - B)H_t \\ & + \tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)(B_t - B)H_t + \tfrac{1}{\sqrt{P}}\sum_{t=R}^{T-1} w_t + o_p(1)\end{aligned}$$

where w_t equals $2(\hat{\beta}_t - \beta_0)' x_t x_t' (\hat{\beta}_t - \beta_0)$. Lemma A.3 shows that

$$\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)BH_t \to^p 0 \tag{2}$$

$$\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1} F(B_t - B)H_t \to^p 0 \tag{3}$$

and

$$\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)(B_t - B)H_t \to^p 0 \tag{4}$$

and Lemma A.2 along with the CLT ensures that $\frac{1}{\sqrt{P}}\sum_{t=R}^{T-1} w_t = o_p(1)$. The proof that

$$\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(f_t - \mathrm{E} f_t) + FB\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1} H_t \to N(0, \sigma^2).$$

follows the same argument as in West (1996) and McCracken (2000). □

Lemma 2. *If Assumptions 1–4 hold then*

$$\hat{\sigma}_1^2 \to^p \sigma^2.$$

If Assumptions 1–3 hold and $\{\varepsilon_t, \mathscr{F}_t\}$ *is an MDS then*

$$\hat{\sigma}_2^2 \to^p \sigma^2.$$

We will only prove $\hat{\sigma}_2 \to^p \sigma$. The result for $\hat{\sigma}_1$ is essentially the same and uses de Jong and Davidson's (2000) Theorem 2.1 for the HAC equivalent of Equations (5)–(7).

Proof. First, we can rewrite the components of the variance estimator as

$$\hat{s}_{21} = \tfrac{1}{P}\sum_{t=R}^{T-1}\big[(f_t - \mathrm{E}\, f_t) + (\hat{f}_t - f_t) - (\bar{f} - \mathrm{E}\, f_t)\big]^2$$
$$\hat{s}_{22} = \tfrac{1}{P}\sum_{t=R}^{T-1}\big[(f_t - \mathrm{E}\, f_t) + (\hat{f}_t - f_t) - (\bar{f} - \mathrm{E}\, f_t)\big]\big[(g_t - \mathrm{E}\, g_t) + (\hat{g}_t - g_t) - (\bar{g} - \mathrm{E}\, g_t)\big]$$

and

$$\hat{s}_{23} = \tfrac{1}{P}\sum_{t=R}^{T-1}\big[(g_t - \mathrm{E}\, g_t) + (\hat{g}_t - g_t) - (\bar{g} + \mathrm{E}\, g_t)\big]^2$$

so $\hat{\sigma}_2 \to^p \sigma$ as long as the following hold: $\bar{f} - \mathrm{E}\,\bar{f}^* \to^p 0$, $\bar{g} - \mathrm{E}\,\bar{g}^* \to^p 0$,

$$\tfrac{1}{P}\sum_{t=R}^{T-1}(f_t - \mathrm{E}\, f_t)^2 \to^p \lim \mathrm{var}(\sqrt{P}\bar{f}^*) \tag{5}$$
$$\tfrac{1}{P}\sum_{t=R}^{T-1}(g_t - \mathrm{E}\, g_t)^2 \to^p \lim \mathrm{var}(\sqrt{P}\bar{g}^*) \tag{6}$$
$$\tfrac{1}{P}\sum_{t=R}^{T-1}(f_t - \mathrm{E}\, f_t)(g_t - \mathrm{E}\, g_t) \to^p \lim \mathrm{cov}(\sqrt{P}\bar{f}^*, \sqrt{P}\bar{g}^*) \tag{7}$$
$$\tfrac{1}{P}\sum_{t=R}^{T-1}(\hat{f}_t - f_t)^2 \to^p 0, \tag{8}$$

and

$$\tfrac{1}{P}\sum_{t=R}^{T-1}(\hat{g}_t - g_t)^2 \to^p 0. \tag{9}$$

The first two results are implied by the proof of Theorem 1 and (5), (6), and (7) follow from the LLN, since each summand is an L_1-mixingale of size -1 (see, for example Davidson, 1994, Theorem 17.5), so it suffices to prove (8) and (9).

As in the proof of Theorem 1, let R' be a new sequence such that $R' \to \infty$ as $T \to \infty$ and $R' = o(\sqrt{P})$. Straightforward algebra reveals that (8) holds if

$$\tfrac{1}{P}\sum_{t=R}^{T-1}((\hat{\beta}_t - \beta_0)' x_t)^4 \to^p 0 \tag{10}$$

and

$$\tfrac{1}{P}\sum_{t=R}^{T-1}(x_t'(\hat{\beta}_t - \beta_0))^2(2x_t'\beta_0 - y_{t+1} - \hat{y}_{t+1})^2 \to^p 0. \tag{11}$$

The LHS of (10) is bounded by

$$
\begin{aligned}
&\tfrac{1}{P}\sum_{t=R}^{T-1}\|\hat{\beta}_t-\beta_0\|^4\|x_t\|^4 \\
&\quad= \tfrac{1}{P}\sum_{t=R}^{R'-1}\|\hat{\beta}_t-\beta_0\|^4\|x_t\|^4+\tfrac{1}{P}\sum_{t=R'}^{T-1}\|\hat{\beta}_t-\beta_0\|^4\|x_t\|^4 \\
&\quad\leq \max_{t=R,\dots,R'-1}\|\hat{\beta}_t-\beta_0\|^4\,\tfrac{1}{P}\sum_{t=R}^{R'-1}\|x_t\|^4+\max_{t=R',\dots,T-1}\|\hat{\beta}_t-\beta_0\|^4\,\tfrac{1}{P}\sum_{t=R'}^{T-1}\|x_t\|^4 \\
&\quad= O_p(R'/P)+o_p(1)
\end{aligned}
$$

by Lemma A.2 and the LLN. A similar argument holds for the second term:

$$
\begin{aligned}
&\tfrac{1}{P}\sum_{t=R}^{T-1}(x_t'(\hat{\beta}_t-\beta_0))^2(2x_t'\beta_0-y_{t+1}-\hat{y}_{t+1})^2 \\
&\qquad= \tfrac{1}{P}\sum_{t=R}^{R'-1}(x_t'(\hat{\beta}_t-\beta_0))^2(2x_t'\beta_0-y_{t+1}-\hat{y}_{t+1})^2 \\
&\qquad\quad+\tfrac{1}{P}\sum_{t=R'}^{T-1}(x_t'(\hat{\beta}_t-\beta_0))^2(2x_t'\beta_0-y_{t+1}-\hat{y}_{t+1})^2 \\
&\qquad\leq \max_{t=R,\dots,R'-1}\|\hat{\beta}_t-\beta_0\|^2\tfrac{1}{P}\sum_{t=R}^{R'-1}\|x_t(2x_t'\beta_0-y_{t+1}-\hat{y}_{t+1})\|^2 \\
&\qquad\quad+\max_{t=R',\dots,T-1}\|\hat{\beta}_t-\beta_0\|^2\tfrac{1}{P}\sum_{t=R'}^{T-1}\|x_t(2x_t'\beta_0-y_{t+1}-\hat{y}_{t+1})\|^2 \\
&\qquad= O_p(R'/P)+o_p(1)
\end{aligned}
$$

again by Lemma A.2 and the LLN. Both terms converge to zero in probability by construction. The proof of (9) is similar. □

Supporting Results

Lemma A.1. *Suppose the conditions of Theorem 1 hold, and define R' to be a sequence that satisfies $R'\to\infty$ as $T\to\infty$ and $R'=o(\sqrt{P})$. Then*

$$\tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(\hat{f}_t-\mathrm{E}\,f_t)\to^p 0.$$

Proof. We can rewrite this summation as

$$\tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(\hat{f}_t - \mathrm{E} f_t) = \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(f_t - \mathrm{E} f_t) + \\ \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(4x_t'\beta_0 - 2y_{t+1} - 2\hat{y}_{i,t+1})x_t'(\hat{\beta}_t - \beta_0) + \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(x_t'\hat{\beta}_t - x_t'\beta_0)^2.$$

Each of these individual summations can be shown to converge to zero in probability. First,

$$\mathrm{E}\Big|\tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(f_t - \mathrm{E} f_t)\Big| \leq \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}\mathrm{E}|f_t - \mathrm{E} f_t| = O(R'/\sqrt{P}).$$

Also,

$$\begin{aligned}&\Big|\tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(4x_t'\beta_0 - 2y_{t+1} - 2\hat{y}_{i,t+1})x_t'(\hat{\beta}_t - \beta_0)\Big| \\ &\qquad \leq \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}\|(4x_t'\beta_0 - 2y_{t+1} - 2\hat{y}_{i,t+1})x_t\| \max_{t=R,\dots,R'-1}\|\hat{\beta}_t - \beta_0\| \\ &\qquad = O_p(R'/\sqrt{P})\end{aligned}$$

and

$$\Big|\tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}(x_t'\hat{\beta}_t - x_t'\beta_0)^2\Big| \leq \tfrac{1}{\sqrt{P}}\sum_{t=R}^{R'-1}\|x_t\|^2 \max_{t=R,\dots,R'-1}\|\hat{\beta}_t - \beta_0\|^2 = O_p(R'/\sqrt{P}).$$

by Lemma A.2 and the LLN. Since $R'/\sqrt{P} \to 0$ by construction, this completes the proof. □

Lemma A.2. *Suppose $a \in [0, 1/2)$ and Assumptions 1 – 3 hold, and let R' be a sequence such that $R' \to \infty$ as $T \to \infty$ and $R' = o(\sqrt{P})$. Then*

1. $\max_{t=R',\dots,T-1}|(t-1)^a H_t| \to^p 0$,
2. $\max_{t=R,\dots,R'-1}|(t-1)^a H_t| = O_p(1)$,
3. $\max_{t=R',\dots,T-1}|B_t - B| \to^p 0$,
4. $\max_{t=R,\dots,R'-1}|B_t - B| = O_p(1)$,
5. $\max_{t=R',\dots,T-1}|(t-1)^a(\hat{\beta}_t - \beta_0)| \to^p 0$, *and*
6. $\max_{t=R,\dots,R'-1}|(t-1)^a(\hat{\beta}_t - \beta_0)| = O_p(1)$,

where the absolute value is taken as the largest of the element-by-element absolute values.

To streamline the presentation, we'll assume in these proofs that x_t is a scalar.

Proof. We will prove each part in order.

1. Our assumptions ensure that $x_t \varepsilon_{t+1}$ is L_2-mixingale of size $-1/2$ (see Theorem 17.5 of Davidson, 1994); let c_t and ζ_k denote its mixingale constants and coefficients. Note that, for any b, $t^b x_t \varepsilon_{t+1}$ is also an L_2-mixingale array with constants $t^b c_s$ and coefficients ζ_k, since

$$\begin{aligned}\|\mathrm{E}_{t-k}\, t^b x_t \varepsilon_{t+1}\| &= t^b \|\mathrm{E}_{t-k}\, x_t \varepsilon_{t+1}\| \\ &\leq (t^b c_t)\zeta_k\end{aligned}$$

and

$$\begin{aligned}\|t^b x_t \varepsilon_{t+1} - t^b\, \mathrm{E}_{t+k}\, x_t \varepsilon_{t+1}\| &= t^b \|x_t \varepsilon_{t+1} - \mathrm{E}_{t+k}\, x_t \varepsilon_{t+1}\| \\ &\leq (t^b c_t)\zeta_{k+1}.\end{aligned}$$

Let δ be a positive number less than $1/2 - \alpha$, so

$$\begin{aligned}\mathrm{E}\Big[\max_{t=R',\dots,T-1} \Big|(t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big|^2\Big] \\ &\leq (R'-1)^{-2\delta}\,\mathrm{E}\Big[\max_{t=R',\dots,T-1}\Big|\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}(s-1)^{a-1+\delta}\Big|^2\Big] \\ &\leq (R'-1)^{-2\delta} O(1) \sum_{s=1}^{T-1} (s-1)^{2(a-1+\delta)}\end{aligned}$$

where the second inequality follows from McLeish's (1975) maximal inequality (also available as Davidson, 1994, Theorem 16.9 and Corollary 16.10). The summation converges to a constant as $T \to \infty$ and $(R'-1)^{-2\delta} \to 0$, completing the proof.

2. Now $t^{a-1} x_t \varepsilon_{t+1}$ is an L_2-mixingale of size $-1/2$ and we can again use McLeish's (1975) maximal inequality to get

$$\begin{aligned}\mathrm{E}\Big|\max_{t=R,\dots,R'-1}\Big((t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big)^2\Big| &\leq \mathrm{E}\Big|\max_{t=R,\dots,T-1}\Big(\sum_{s=1}^{t-1} s^{a-1} x_s \varepsilon_{s+1}\Big)^2\Big| \\ &= O(1)\sum_{s=1}^{R'-1} s^{2a-2}\end{aligned}$$

which converges to a finite limit.

3. The same argument used in Part 1 implies that $\max_{t=R',\dots,T-1}|B_t^{-1} - B^{-1}| \to^p 0$. Since matrix inversion is continuous, the result follows.

4. Holds by Assumptions 1 and 2.

5. We have

$$\begin{aligned}\max_{t=R',\dots,T-1}|(t-1)^a(\hat{\beta}_t - \beta_0)| \leq & \max_{t=R',\dots,T-1}|\hat{B}_t - B| \max_{t=R',\dots,T-1}\Big|(t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big| \\ & + \max_{t=R',\dots,T-1}\Big|B(t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big|\end{aligned}$$

   and both terms converge to zero in probability by Parts 1 and 3.

6. Similar to the previous argument, we have

$$\begin{aligned}\max_{t=R,\dots,R'-1}|(t-1)^a(\hat{\beta}_t - \beta_0)| \leq & \max_{t=R,\dots,R'-1}|\hat{B}_t - B| \max_{t=R,\dots,R'-1}\Big|(t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big| \\ & + \max_{t=R,\dots,R'-1}\Big|B(t-1)^{a-1}\sum_{s=1}^{t-1} x_s \varepsilon_{s+1}\Big|.\end{aligned}$$

   Both terms are $O_p(1)$ by Parts 2 and 4. □

Lemma A.3. *Under the conditions of Theorem 1, Equations* (2)–(4) *hold.*

Proof. We can write

$$\Big|\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)BH_t\Big| \leq \Big\|\tfrac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F)B\Big\| \max_{t=R',\dots,T-1}\|H_t\|.$$

From Lemma A.2, $\max_{t=R',\dots,T-1}\|H_t\| \to^p 0$. de Jong's (1997) CLT implies that $\frac{1}{\sqrt{P}}\sum_{t=R'}^{T-1}(F_t - F) = O_p(1)$, establishing (2). The proofs of (3) and (4) are similar. □

References

G. Calhoun. An asymptotically normal out-of-sample test based on mixed estimation windows. Available at http://www.econ.iastate.edu/~gcalhoun, 2015.

J. Davidson. *Stochastic Limit Theory: An Introduction for Econometricians*. Advanced Texts in Econometrics. Oxford University Press, 1994.

R. M. de Jong. Central limit theorems for dependent heterogeneous random variables. *Econometric Theory*, 13(3):353–367, 1997.

R. M. de Jong and J. Davidson. Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica*, 68(2):407–423, 2000.

M. W. McCracken. Robust out-of-sample inference. *Journal of Econometrics*, 99(2): 195–223, 2000.

D. McLeish. A maximal inequality and dependent strong laws. *The Annals of Probability*, 3(5):829–839, Oct. 1975.

K. D. West. Asymptotic inference about predictive ability. *Econometrica*, 64(5):1067–1084, Sept. 1996.