Newey and West, "Automatic Lag Selection in Covariance Matrix Estimation"

July 1992, revised September 1993

Additional Appendix

This additional appendix contains material omitted from the published paper to conserve space: (1)a sketch of the proof of Theorem 2; (2)additional simulation results.

(1)SKETCH OF PROOF OF THEOREM 2

Lemma A9: Let $y=\hat{\gamma}T^{1/(2q+1)}$. Let N be an arbitrary integer ≥ 1. Then if $\hat{\gamma} \xrightarrow{p} \gamma \neq 0$, $|([y]+1)^{-N}-y^{-N}| = T^{-(N+1)/(2q+1)}O_p(1)$.

Proof: We have

$$
\begin{aligned}
|([y]+1)^{-N}-y^{-N})| &= ([y]+1)^{-N}|1 - \{([y]+1)/y\}^N| \\
&= y^{-N-1}\{y/([y]+1)\}^N y|1 - \{([y]+1)/y\}^N| \\
&= y^{-N-1}O_p(1)|y||1 - \{([y]+1)/y\}^N|,
\end{aligned}
$$

where the final equality follows since $0 < y/([y]+1) \leq 1$ and $y \geq 0$. The result will now follow if

$$
\text{(A-11)} \qquad |y||1 - \{([y]+1)/y\}^N| = O_p(1)
$$

Now, $([y]+1)/y = 1 + d_T/y$ for some $0 < d_T \leq 1$, and the "T" subscript is present to emphasize that d_T is random. Thus

$$
1 - \{([y]+1)/y\}^N = \Sigma_{j=1}^N c_j (d_T/y)^j
$$

where the c_j are the negative of the binomial coefficients, $c_j = -N!/[j!(N-j)!]$. Upon multiplying the summation in the preceding expression by y and taking absolute values, we obtain

$$
\text{(A-12)} \qquad |c_1 d_T + \Sigma_{j=2}^N c_j (d_T^j/y^{j-1})|
$$

We have $0 < c_1 d_T \leq c_1 \Rightarrow c_1 d_T = O_p(1)$. For $j\geq 2$ $(c_j d_T^j/y^{j-1}) = (c_j d_T^j/\hat{\gamma}^{j-1})(1/T^{(j-1)/(2q+1)}) \xrightarrow{p} 0$ since $\hat{\gamma}^{j-1} \xrightarrow{p} \gamma^{j-1} \neq 0$, $0 < c_j d_T^j \leq c_j \Rightarrow c_j d_T^j = O_p(1)$. It follows that (A-12), which is just (A-11) rewritten, is $O_p(1)$, which is the desired result.

Lemma A10: Under the assumptions of Theorem 2, $T^{q/(2q+1)}[\Sigma_{j=-T+1}^{T-1}(\bar{k}_j - \hat{k}_j)(\tilde{\sigma}_j - E\tilde{\sigma}_j)] \xrightarrow{p} 0$.

Proof: Choose v and define a so that

$$
1 + 1/(2b-2) < v < 5/4, \quad a = [T^{v/(2q+1)}];
$$

b>3 guarantees that this can be done. We need to show that

$$
\begin{aligned}
&2T^{q/(2q+1)}\Sigma_{j=1}^a(\bar{k}_j - \hat{k}_j)(\tilde{\sigma}_j - E\tilde{\sigma}_j) \\
&\quad + 2T^{(q/2q+1)}\Sigma_{j=a+1}^{T-1}\bar{k}_j(\tilde{\sigma}_j - E\tilde{\sigma}_j) - 2T^{(q/2q+1)}\Sigma_{j=a+1}^{T-1}\hat{k}_j(\tilde{\sigma}_j - E\tilde{\sigma}_j) \\
&\quad = 2F_1 + 2F_2 + 2F_3 \xrightarrow{p} 0.
\end{aligned}
$$

The Lipschitz condition and Lemma A9 imply that

$$
\begin{aligned}
|F_1| &\leq cT^{q/(2q+1)}\Sigma_{j=1}^a |j/([\hat{\gamma}T^{1/(2q+1)}]+1) - (j/\hat{\gamma}T^{1/(2q+1)})||\tilde{\sigma}_j - E\tilde{\sigma}_j| \\
&= cO_p(1)T^{(q-2)/(2q+1)}\Sigma_{j=1}^a j|\tilde{\sigma}_j - E\tilde{\sigma}_j|.
\end{aligned}
$$

It follows from the proof of Lemma A6 that the right hand side of this expression converges in probability to zero if $T^{(q-2)/(2q+1)}T^{-1/2}T^{2v/(2q+1)} \to 0$, which in turn follows since $v < 5/4$.

That $F_3 \stackrel{p}{\to} 0$ was already established in the proof of Lemma A6. That $F_2 \stackrel{p}{\to} 0$ follows from a similar argument.

<u>Lemma</u> <u>A11:</u> If $\hat{\gamma} \stackrel{p}{\to} \gamma \neq 0$, $T^{q/(2q+1)}[\Sigma_{j=-T+1}^{T-1}(\bar{k}_j-\hat{k}_j)E\tilde{\sigma}_j] \stackrel{p}{\to} 0$.

<u>Proof</u>: Define

$$\bar{j} = \min((T-1),\ [\underline{x}\gamma T^{1/(2q+1)}]).$$

We need to show that

$$2T^{q/(2q+1)}\Sigma_{j=1}^{\bar{j}}(\bar{k}_j-\hat{k}_j)E\tilde{\sigma}_j + 2T^{q/(2q+1)}\Sigma_{j=\bar{j}+1}^{T-1}\bar{k}_jE\tilde{\sigma}_j - 2T^{q/(2q+1)}\Sigma_{j=\bar{j}+1}^{T-1}\hat{k}_jE\tilde{\sigma}_j$$
$$= 2D_1 + 2D_2 + 2D_3 \stackrel{p}{\to} 0.$$

As in the proof of Lemma A7, for $1 \le j \le \bar{j}$, expand $\bar{k}_j$ and $\hat{k}_j$ around $k(0)$. It follows from the proof of that lemma that

$$D_1=T^{q/(2q+1)}(k^{([q])}(0)/[q]!)\Sigma_{j=1}^{\bar{j}}(j/([\hat{\gamma}T^{1/(2q+1)}]+1)^{[q]}-(j/\hat{\gamma}T^{1/(2q+1)})^{[q]})E\tilde{\sigma}_j$$
$$+ T^{q/(2q+1)}(1/([q]+1)!)\Sigma_{j=1}^{\bar{j}}\bar{k}^{([q]+1)}(\overset{*}{x}_j)j/([\hat{\gamma}T^{1/(2q+1)}]+1)^{[q]+1}E\tilde{\sigma}_j$$
$$+ T^{q/(2q+1)}(1/([q]+1)!)\Sigma_{j=1}^{\bar{j}}\hat{k}^{([q]+1)}(\overset{**}{x}_j)(j/\hat{\gamma}T^{1/(2q+1)})^{[q]+1}E\tilde{\sigma}_j$$
$$= D_{11} + D_{12} + D_{13},$$

where $\overset{*}{x}_j$ lies between 0 and $\bar{x}_j=j/([\hat{\gamma}T^{1/(2q+1)}]+1)$, $\overset{**}{x}_j$ between 0 and $\hat{x}_j=(j/\hat{\gamma}T^{1/(2q+1)})$. By Lemma A9,

$$|D_{11}| \le cT^{(q-[q]-1)/(2q+1)}\Sigma_{j=1}^{\bar{j}}j^{[q]}|E\tilde{\sigma}_j|O_p(1) \stackrel{p}{\to} 0$$

since $0 > q-[q]-1$ and $\Sigma_{j=1}^{\infty}j^q|\sigma_j| < \infty$. That $D_{12} \stackrel{p}{\to} 0$, $D_{13} \stackrel{p}{\to} 0$ follows from the argument used to show $B_{12} \stackrel{p}{\to} 0$, $B_{13} \stackrel{p}{\to} 0$ in the proof of Lemma A7.

That $D_2 \stackrel{p}{\to} 0$, $D_3 \stackrel{p}{\to} 0$ follows from the argument used to show $B_2 \stackrel{p}{\to} 0$, $B_3 \stackrel{p}{\to} 0$ in the proof of Lemma A7.

<u>Lemma</u> <u>A12:</u> If $\hat{\gamma} \stackrel{p}{\to} \gamma$, $T^{q/(2q+1)}[\Sigma_{j=-T+1}^{T-1}(\bar{k}_j-\hat{k}_j)(\hat{\sigma}_j-\tilde{\sigma}_j)] \stackrel{p}{\to} 0$.

<u>Proof</u>: Follows by logic similar to that used in proving Lemmas A8, A10, and A11.

<u>Proof</u> <u>of</u> <u>Theorem</u> <u>2:</u> Follows from Theorem 1 and Lemmas A10, A11 and A12.

(2) ADDITIONAL SIMULATION RESULTS

Here are: (a) a version of Table IV that contains results for additional kernels, (b) a version of Table V that contains results for additional kernels, and (c) sample MSE's relative to QS for various kernels.

(a) Table IV, with Additional Results

Sizes of Nominal 1, 5, and 10 Percent Tests, Experiment A

(1) Kernel	(2) PW?	(3a) bandwidth m or lag selection parameter n A1	(3b) A2	(4a) Experiment A1 Size 1.0	(4b) 5.0	(4c) 10.0	(5a) Experiment A2 Size 1.0	(5b) 5.0	(5c) 10.0
1. Bartlett		4	5	4.5	12.4	18.6	2.7	8.1	14.0
2. QS-AR(1)		n.a.	n.a.	3.6	11.6	17.4	3.0	8.5	15.1
3. QS-AR(1)	y	n.a.	n.a.	3.9	11.2	17.3	2.3	7.3	13.4
* QS-AR(1)	y	n.a.	n.a.	4.3	11.6	18.4	2.5	7.5	13.9
4. Truncated		4	5	9.0	18.3	24.6	4.8	11.8	18.2
5. Population S		n.a.	n.a.	0.3	1.4	3.8	0.5	2.5	5.8
6. Bartlett, population γ		8	13	5.9	14.9	20.9	4.5	10.7	16.7
7. Bartlett		4	5	4.6	13.1	19.7	3.5	9.3	15.0
8. Bartlett		12	15	11.0	19.7	27.0	5.2	11.6	17.9
Bartlett with $n=[4(T/100)^{1/9}]$							3.1	9.1	14.4
Bartlett with $n=[12(T/100)^{1/9}]$							4.5	10.9	17.3
9. Bartlett	y	3	3	4.8	12.9	19.1	3.0	8.5	14.1
Bartlett	y	6	7	6.0	14.7	20.9	3.3	10.1	15.9
* Bartlett	y	3	3	4.9	13.2	19.8	2.9	8.7	14.9
* Bartlett	y	6	7	6.4	15.5	22.1	3.8	10.3	15.7
10. Parzen		4	5	5.5	15.3	21.7	3.3	9.8	15.4
11. Parzen		12	15	17.0	25.6	34.1	7.9	14.7	21.5
12. Parzen	y	3	3	5.1	13.4	20.8	3.3	9.8	15.5
Parzen	y	6	7	8.2	17.3	23.6	4.5	11.3	17.9
* Parzen	y	3	3	5.7	14.5	20.5	3.6	9.8	15.7
* Parzen	y	6	7	8.6	17.9	24.1	4.6	11.6	17.3
13. QS		4	4	5.0	14.0	21.4	3.1	9.1	15.3

A line number in the leftmost column indicates a row that also appears in Table IV in the paper.
"*" in the leftmost column indicates that the OLS estimate of A was used in prewhitening, with no adjustment for singular values.

(b) Table V, with Additional Results

Sizes of Nominal 1, 5, and 10 Percent Tests, Experiment B

(1) Kernel	(2) PW?	(3a) bandwidth or lag selection parameter T=300	(3b) T=1000	(4a) Experiment B1 T=300 Size 1.0	(4b) 5.0	(4c) 10.0	(5a) T=1000 Size 1.0	(5b) 5.0	(5c) 10.0	(6a) Experiment B2 T=300 Size 1.0	(6b) 5.0	(6c) 10.0	(7a) T=1000 Size 1.0	(7b) 5.0	(7c) 10.0
1. Bartlett		4	5	17.8	31.1	41.5	10.7	23.0	30.9	17.3	32.3	42.2	10.7	23.6	33.3
2. QS-AR(1)		n.a.	n.a.	16.2	27.5	36.3	6.3	14.4	24.2	16.5	27.2	36.6	6.6	15.9	24.6
3. QS-AR(1)	y	n.a.	n.a.	10.4	20.0	26.0	4.0	9.8	15.6	10.7	20.7	27.6	5.4	12.2	20.4
* QS-AR(1)	y	n.a.	n.a.	7.8	16.1	23.0	3.1	9.3	15.1	7.7	15.8	22.4	3.0	9.8	16.4
4. Truncated		12	12	14.7	23.9	30.9	3.5	12.2	18.3	14.7	24.7	32.3	4.8	13.0	21.5
5. Pop. S		n.a.	n.a.	2.5	6.8	11.2	1.5	5.9	11.5	0.9	1.7	2.6	0.7	2.3	3.3
6. Bartlett, population γ		16	24	11.5	21.4	30.6	5.0	14.9	22.4	12.6	23.5	31.1	5.5	14.9	23.8
7. Bartlett		5	6	11.1	20.9	30.7	5.0	15.2	22.0	11.5	22.5	31.2	5.5	15.6	23.9
Bartlett		15	20	17.2	27.3	36.4	6.7	17.0	24.8	17.8	28.9	37.6	8.4	18.9	26.6
8. Bartlett		12	12	14.8	24.0	32.5	5.7	15.6	22.9	15.1	26.3	34.4	5.6	16.4	25.0
9. Bartlett	y	3	5	9.2	17.3	24.2	2.5	8.1	14.0	9.2	19.0	24.9	4.1	9.1	16.5
Bartlett	y	7	10	8.4	16.5	23.7	2.5	8.3	14.0	8.6	18.3	24.0	4.4	10.0	17.4
* Bartlett	y	3	5	7.7	15.6	21.9	2.1	7.7	13.2	7.5	15.4	21.4	2.6	8.4	14.7
* Bartlett	y	7	10	7.7	15.4	21.4	2.5	8.0	12.6	7.4	16.2	22.0	3.5	8.7	15.1
10. Parzen		4	5	11.3	20.0	28.8	4.4	14.6	21.0	11.3	21.4	30.9	4.8	13.6	22.5
Parzen		14	17	16.4	25.3	34.0	5.6	14.7	22.2	17.6	28.1	35.7	6.4	15.4	25.5
11. Parzen		12	12	14.4	23.1	31.4	4.7	14.3	21.6	14.4	23.8	33.2	5.6	14.5	23.6
12. Parzen	y	3	4	8.8	16.8	24.0	2.6	7.6	13.6	8.7	18.3	23.6	3.5	8.7	16.0
Parzen	y	7	8	8.8	16.6	23.4	2.4	7.7	13.0	8.3	17.5	24.1	3.6	8.4	16.3
* Parzen	y	3	4	7.9	15.3	21.7	2.0	7.2	12.5	7.2	15.5	21.8	2.4	8.1	14.5
* Parzen	y	7	8	7.8	15.9	22.0	2.4	7.1	12.1	8.5	16.1	22.9	3.1	8.3	13.6
13. QS		12	12	14.1	23.0	31.9	4.8	15.0	21.9	14.4	24.2	33.3	5.6	14.3	23.8

See notes on previous page of additional appendix

(c) EMPIRICAL MSE'S RELATIVE TO QS, VARIOUS KERNELS

	Kernel	PW?	bandwidth m or lag selection parameter n A1	A2	Relative MSE's, Experiment A A1	A2
1.	Bartlett		4	5	0.966	1.016
2.	QS-AR(1)		n.a.	n.a.	1.036	1.061
3.	QS-AR(1)	y	n.a.	n.a.	0.959	0.945
*	QS-AR(1)	y	n.a.	n.a.	1.021	1.028
4.	Truncated		4	5	1.003	0.945
6.	Bartlett,		8	13	1.005	0.942
7.	Bartlett		4	5	1.011	0.989
8.	Bartlett		12	15	1.296	1.166
9.	Bartlett	y	3	10	0.976	0.900
	Bartlett	y	6	13	1.052	0.953
*	Bartlett	y	3	10	0.996	0.916
*	Bartlett	y	6	13	1.057	0.956
10.	Parzen		4	5	1.029	0.986
11.	Parzen		12	15	1.430	1.205
12.	Parzen	y	3	7	0.989	0.901
	Parzen	y	6	14	1.126	0.992
*	Parzen	y	3	7	1.000	0.899
*	Parzen	y	6	14	1.121	0.989
13.	QS		4	4	1.000	1.000

Relative MSE's, Experiment B

	Kernel		bandwidth or lag selection parameter T=300	T=1000	B1 T=300	B1 T=1000	B2 T=300	B2 T=1000
1.	Bartlett		4	5	0.955	1.213	1.062	1.103
2.	QS-AR(1)		n.a.	n.a.	0.982	1.199	1.016	0.999
3.	QS-AR(1)	y	n.a.	n.a.	1.035	0.967	1.025	1.054
*	QS-AR(1)	y	n.a.	n.a.	58.533	4.329	1.277	0.892
4.	Truncated		12	12	1.107	0.987	0.992	0.975
6.	Bartlett,		16	24	0.901	0.937	1.010	1.020
7.	Bartlett		5	6	0.878	0.912	1.010	1.023
	Bartlett		15	20	1.013	1.154	1.031	1.040
8.	Bartlett		12	12	0.981	0.976	1.021	1.017
9.	Bartlett	y	10	21	0.879	0.838	1.009	1.012
	Bartlett	y	13	28	0.893	0.876	0.992	1.005
*	Bartlett	y	10	21	73.748	5.361	1.518	0.918
*	Bartlett	y	13	28	73.762	3.914	1.479	0.911
10.	Parzen		4	5	0.865	0.833	1.005	1.008
	Parzen		14	17	1.052	1.117	1.026	1.013
11.	Parzen		12	12	1.007	1.005	1.003	1.001
12.	Parzen	y	7	9	0.869	0.780	1.000	1.002
	Parzen	y	14	18	0.938	0.854	0.983	0.987
*	Parzen	y	7	9	77.631	5.583	1.593	0.911
*	Parzen	y	14	18	67.302	4.670	1.493	0.924
13.	QS		12	12	1.000	1.000	1.000	1.000

See notes on page 5 of additional appendix.