Supplement to

NONPARAMETRIC INFERENCE BASED ON CONDITIONAL MOMENT INEQUALITIES

By

Donald W. K. Andrews and Xiaoxia Shi

December 2011

COWLES FOUNDATION DISCUSSION PAPER NO. 1840

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Appendix 2

to

Nonparametric Inference Based on Conditional Moment Inequalities

Donald W. K. Andrews
Cowles Foundation for Research in Economics
Yale University

Xiaoxia Shi
Department of Economics
University of Wisconsin, Madison

June 2010
Revised: December 2011

10 Appendix 2

This Appendix provides proofs of Theorems N2 and N3 of the paper "Nonparametric inference based on conditional moment inequalities." It also provides some additional simulation results to those given in that paper. We let AS1 and AS2 abbreviate Andrews and Shi (2007a) and Andrews and Shi (2007b), respectively.

10.1 Proofs of Theorems N2 and N3

Proof of Theorem N2. Theorem N2 is analogous to Theorem 3 of AS1. The proof of Theorem 3 of AS1 that is given in Section 14.2 in Appendix C of AS2 goes through with a few changes in the present context. First, $E_{F_0}(\cdot)$ is replaced by $E_{F_0}(\cdot|Z_i = z_0)$ in $m^*(g)$ and elsewhere. Second, $n^{1/2}\beta(g_0)$ is replaced throughout by $(nb^{d_z})^{1/2}\beta(g_0)$. Third, Assumption NFA(a) is used in place of Assumption FA(a) to obtain the inequality in (14.28) of AS2. Fourth, the proof uses Lemma AN3, which employs Assumptions NFA(b) and NFA(c), in place of Lemma A1 of AS2.

Fifth, the second equality of (14.33) of AS2 does not hold. It relies on $n^{-1/2}h_{1,n,F_0}(\theta_*, g) = m^*(g)$, which in the present context is replaced by $(nb^{d_z})^{-1/2}h_{1,n,F_0}(\theta_*, g, b) = m^*(g)$, which does not hold. However, we have

$$\begin{aligned}(nb^{d_z})^{-1/2}h_{1,n,F_0}(\theta_*, g, b) &= D_{F_0}^{-1/2}(\theta_*, b)b^{-d_z/2}E_{F_0}m(W_i, \theta_*, g, b)\\ &= D_{F_0}^{-1/2}(\theta_*, z_0)E_{F_0}m(\theta_*, X_i, z_0)g(X_i) + O(b^2)\\ &= D_{F_0}^{-1/2}(\theta_*, z_0)E_{F_0}(m(W_i, \theta_*, g)|Z_i = z_0)f(z_0) + O(b^2)\\ &= m^*(g) + o(1), \qquad (10.34)\end{aligned}$$

where the second equality holds by Lemma AN1(b) and (9.12) (which holds for $(\theta_*, F_0) \in \mathcal{F}_+$), the third equality holds by the same argument as in the proof of Lemma AN4 with $m(y, x, z, \theta, g)m(y, x, z, \theta, g^*)'$ replaced by $m(y, x, z, \theta, g)$ throughout, and the fourth equality holds by the definition of $m^*(g)$ and Assumption B(a).

Using (10.34), the second equality of (14.33) of AS2 holds with $m^*(g)/\beta(g_0)$ replaced by $m^*(g)/\beta(g_0) + o(1)$.

These are the only changes needed to the proof of Theorem 3 of AS1. □

Proof of Theorem N3. Theorem N3 is analogous to Theorem 4 of AS1. First, we give an analogue of (14.37) in the proof of Theorem 4 of AS1 given in Section 14.3 of

Appendix C in AS2. We have

$$
\begin{aligned}
& h_{1,n,F_n}(\theta_{n,*}, g, b) \\
= & \; n^{1/2} D_{F_n}^{-1/2}(\theta_{n,*}, b) E_{F_n} m(W_i, \theta_{n,*}, g, b) \\
= & \; (nb^{d_z})^{1/2}(I_k + o(1)) D_{F_n}^{-1/2}(\theta_{n,*}, z_0) E_{F_n} m(\theta_{n,*}, X_i, z_0) g(X_i) + o(1) \qquad (10.35) \\
= & \; (nb^{d_z})^{1/2}(I_k + o(1)) D_{F_n}^{-1/2}(\theta_{n,*}, z_0) E_{F_n}(m(W_i, \theta_{n,*}, g) | Z_i = z_0) f_n(z_0) + o(1),
\end{aligned}
$$

where the first equality holds by (9.2), the second equality holds by Lemma AN1(b) and (9.12) because $n^{1/2} b^{2+d_z/2} \to 0$ if $b = o(n^{-1/(4+d_z)})$, and the third equality holds by the same argument as in the proof of Lemma AN4 above.

Next, by element-by-element mean-value expansions about θ_n, we have

$$
\begin{aligned}
& D_{F_n}^{-1/2}(\theta_{n,*}, z_0) E_{F_n}(m(W_i, \theta_{n,*}, g) | Z_i = z_0) f_n(z_0) \\
= & \; D_{F_n}^{-1/2}(\theta_n, z_0) E_{F_n}(m(W_i, \theta_n, g) | Z_i = z_0) f_n(z_0) \\
& + \Pi_{F_n}(\theta_{n,g}, g)(\theta_{n,*} - \theta_n), \qquad (10.36)
\end{aligned}
$$

using Assumption NLA2, where $\theta_{n,g}$ may differ across rows of $\Pi_{F_n}(\theta_{n,g}, g)$, $\theta_{n,g}$ lies between $\theta_{n,*}$ and θ_n, and $\theta_{n,g} \to \theta_0$.

Combining (10.35) and (10.36) gives the analogue of (14.37) of AS2:

$$
\begin{aligned}
& h_{1,n,F_n}(\theta_{n,*}, g, b) \\
= & \; (nb^{d_z})^{1/2}(I_k + o(1)) D_{F_n}^{-1/2}(\theta_n, z_0) E_{F_n}(m(W_i, \theta_n, g) | Z_i = z_0) f_n(z_0) \\
& + (I_k + o(1)) \Pi_{F_n}(\theta_{n,g}, g)(nb^{d_z})^{1/2}(\theta_{n,*} - \theta_n) \\
\to & \; h_1(g) + \Pi_0(g)\lambda, \qquad (10.37)
\end{aligned}
$$

where $h_1(g)$ and $\Pi_0(g)$ are defined in (7.2) and the convergence uses Assumptions NLA1(a), NLA1(b), and NLA2.

Now, the proof of Theorem N3 is similar to the proof of Theorem 4 of AS1 given in AS2 with the following changes:

(i) $\{(\theta_{n,*}, F_n) \in \mathcal{F} : n \geq 1\} \in SubSeq(h_2)$, where $h_2 = h_{2,F_0}(\theta_0) \in \mathcal{H}_{2,+}$ by Assumptions NLA1(a) and NLA1(c)-(e),

(ii) part (i) and Assumptions B and MN imply that the results of Lemma AN3 hold under $\{(\theta_{n,*}, F_n) \in \mathcal{F} : n \geq 1\}$ and these results are used in place of Lemma A1 of AS2,

(iii) equation (14.38) of AS2 is replaced by

$$\begin{aligned}
&\kappa_n^{-1}\overline{D}_{F_n}^{-1/2}(\theta_{n,*}, g, b)D_{F_n}^{1/2}(\theta_{n,*}, b)h_{1,n,F_n}(\theta_{n,*}, g, b) \\
&= (I_k + o(1))\kappa_n^{-1}(nb^{d_z})^{1/2}\overline{D}_{F_n}^{-1/2}(\theta_n, g, z_0)E_{F_n}(m(W_i, \theta_n, g)|Z_i = z_0)f_n(z_0) \\
&\quad +\kappa_n^{-1}\overline{D}_{F_0}^{-1/2}(\theta_0, g, z_0)D_{F_0}^{1/2}(\theta_0, z_0)(I_k + o(1))\Pi_{F_n}(\theta_{n,g}, g)(nb^{d_z})^{1/2}(\theta_{n,*} - \theta_n)] \\
&= \pi_1(g) + o(1), \qquad (10.38)
\end{aligned}$$

where the first equality holds by the equality in (10.37) and Lemma AN1(b) and the second equality holds because (a) the first term on the rhs of the first equality is $\pi_1(g) + o(1)$ by Assumption NLA4 and (b) the second term on the rhs of the first equality is $o(1)$ by the convergence of the second term in (10.37) plus $\kappa_n^{-1} \to 0$, and

(iv) in the verification of (14.23) in part (ix) of the proof of Theorem 4 of AS1 given in Section 14.3 of Appendix C in AS2, (10.37) is used in place of (14.37) of AS2. This completes the proof. □

10.2 Additional Simulation Results

In this section, we provide some additional simulation results. Tables A1 and A2 report the robustness results for the CvM/Max and KS/Max test statistics in the kinked and the peaked bound cases, respectively, for the quantile selection model. As in Tables I-III, the results in Tables A1 and A2 are for the lower endpoints of the identified intervals. Tables A3 and A4 report the robustness results for the CvM and KS test statistics in the kinked and tilted bound cases, respectively, for the conditional treatment effect model.

Both Tables A1 and A2 show that there is little sensitivity to r_1, ε, the GMS tuning parameters, and the kernel bandwidth in terms of coverage probabilities. There is some sensitivity in terms of the FCP's. The FCP decreases (gets better) with the sample size for the KS/MAX-GMS/Asy pair and is stable for the CvM/Max-GMS/Asy pair. The FCP is smaller (better) with (κ_n, B_n) halved and bigger with (κ_n, B_n) doubled.

There is quite a bit sensitivity to the kernel bandwidth. With both the kinked and the peaked bound, doubling the bandwidth reduces the FCP's for tests with the KS/Max statistics. The same is true with the kinked bound and the CvM/Max statistic. However, with the peaked bound, both doubling and halving the bandwidth increases the FCP's.

Tables A1 and A2 show that 0.50 CI's cover the true value with probability noticeably higher than 0.50. This indicates that the lower boundary point of the 0.50 CI as an estimator for the lower end point of the identified set is not median unbiased, but does not have an inward bias which has been a concern in the literature.

Table A1. Nonparametric Quantile Selection Model with Kinked Bound: Variations on the Base Case

		(a) Coverage Probabilities		(b) False Cov Probs (CPcor)	
	Statistic:	CvM/Max	KS/Max	CvM/Max	KS/Max
Case	Crit Val:	GMS/Asy	GMS/Asy	GMS/Asy	GMS/Asy
Base Case: $(n = 250, r_1 = 3,$ $\varepsilon = 0.05, b = b^0 n^{-2/7})$		.989	.987	.49	.57
$n = 100$		.988	.991	.48	.59
$n = 500$		.989	.991	.45	.54
$r_1 = 2$		.988	.987	.50	.53
$r_1 = 4$		.990	.989	.48	.60
$(\kappa_n, B_n) = 1/2(\kappa_{n,bc}, B_{n,bc})$		.991	.987	.49	.55
$(\kappa_n, B_n) = 2(\kappa_{n,bc}, B_{n,bc})$		.993	.991	.56	.61
$\varepsilon = 1/100$		.989	.987	.47	.57
$b = 0.5 b^0 n^{-2/7}$		.986	.987	.69	.77
$b = 2 b^0 n^{-2/7}$		.997	.995	.35	.45
$\alpha = .5$		.771	.739	.05	.06
$\alpha = .5$ & $n = 500$		.787	.753	.05	.06

Table A2. Nonparametric Quantile Selection Model with Peaked Bound: Variations on the Base Case

		(a) Coverage Probabilities		(b) False Cov Probs (CPcor)	
	Statistic:	CvM/Max	KS/Max	CvM/Max	KS/Max
Case	Crit Val:	GMS/Asy	GMS/Asy	GMS/Asy	GMS/Asy
Base Case: ($n = 250, r_1 = 3,$ $\varepsilon = 0.05, b = b^0 n^{-2/7}$)		.991	.991	.49	.53
$n = 100$		.989	.990	.56	.65
$n = 500$		.994	.995	.50	.45
$r_1 = 2$		.990	.990	.51	.50
$r_1 = 4$		.992	.991	.48	.58
$(\kappa_n, B_n) = 1/2(\kappa_{n,bc}, B_{n,bc})$		.992	.990	.47	.52
$(\kappa_n, B_n) = 2(\kappa_{n,bc}, B_{n,bc})$		.994	.994	.54	.56
$\varepsilon = 1/100$		.991	.991	.47	.53
$b = 0.5 b^0 n^{-2/7}$		.988	.989	.62	.70
$b = 2 b^0 n^{-2/7}$		.997	.996	.53	.47
$\alpha = .5$		.803	.761	.04	.05
$\alpha = .5$ & $n = 500$		.836	.795	.04	.04

Tables A3 and A4 show the sensitivity results for the nonparametric conditional treatment effect model with kinked bound and tilted bound, respectively.

Table A3 shows that, with the kinked bound, the test has NRP's smaller than 0.05 for all the test configurations and sample sizes that we experimented with. This is expected because with the kinked bound, the conditional moment inequality is only binding at a measure-zero set of the instrumental variable and Assumption GMS2 is not likely to hold. The ARP's are relatively stable as we vary r_1, decrease ε or decrease (κ_n, B_n). Doubling (κ_n, B_n) makes the ARP's smaller (worse). Both doubling and halving the kernel bandwidth reduces ARP's noticeably.

Table A3. Nonparametric Conditional Treatment Effect Model with Kinked Bound: Variations on the Base Case

Case	Statistic: Crit Val:	(a) Null Rejection Probabilities CvM GMS/Asy	KS GMS/Asy	(b) Rej Probs under H_1 (NRP-corrected) CvM GMS/Asy	KS GMS/Asy
Base Case: $(n = 250, r_1 = 3,$ $\varepsilon = 0.05, b = b^0 n^{-2/7})$		.000	.000	.52	.49
$n = 100$		.000	.000	.65	.55
$n = 500$		.000	.000	.33	.40
$r_1 = 2$		.000	.000	.52	.53
$r_1 = 4$		.000	.000	.51	.45
$(\kappa_n, B_n) = 1/2(\kappa_{n,bc}, B_{n,bc})$		.000	.000	.52	.52
$(\kappa_n, B_n) = 2(\kappa_{n,bc}, B_{n,bc})$		.000	.000	.44	.42
$\varepsilon = 1/100$		.000	.000	.52	.44
$b = 0.5b^0 n^{-2/7}$		.000	.000	.38	.30
$b = 2b^0 n^{-2/7}$		.000	.000	.34	.43

Table A4 shows a new aspect of the sensitivity analysis. The NRP for the CvM test in the base case is somewhat bigger than 0.05. Halving the bandwidth reduces NRP's to below 0.05. while doubling the bandwidth increases the NRP's to disastrous level. This is expected because with the tilted bound the unconditional moment formed using the kernel functions has negative expectation for any fixed bandwidth. The negative expectation converges to zero as the bandwidth converges to zero. Thus, letting b converge to zero is central to the theoretical validity of our method. Using a large b deviates from the asymptotic theory.

The ARP's in Table A4 are reasonably stable across different configurations and sample sizes, except that they are somewhat sensitive to the kernel bandwidth.

Table A4. Nonparametric Conditional Treatment Effect Model with Tilted Bound: Variations on the Base Case

		(a) Null Rejection Probabilities		(b) Rej Probs under H_1 (NRP-corrected)	
	Statistic:	CvM	KS	CvM	KS
Case	Crit Val:	GMS/Asy	GMS/Asy	GMS/Asy	GMS/Asy
Base Case: ($n = 250, r_1 = 3,$ $\varepsilon = 0.05, b = b^0 n^{-2/7}$)		.072	.047	.53	.36
$n = 100$		.085	.042	.49	.34
$n = 500$		.072	.050	.53	.40
$r_1 = 2$		.074	.059	.52	.38
$r_1 = 4$		.069	.036	.53	.32
$(\kappa_n, B_n) = 1/2(\kappa_{n,bc}, B_{n,bc})$		.081	.054	.50	.35
$(\kappa_n, B_n) = 2(\kappa_{n,bc}, B_{n,bc})$		.066	.045	.53	.36
$\varepsilon = 1/100$		.071	.040	.52	.31
$b = 0.5b^0 n^{-2/7}$		.044	.023	.29	.14
$b = 2b^0 n^{-2/7}$		.467	.313	.69	.57

References

Andrews, D. W. K., Shi, X., 2007a. Inference based on conditional moment inequalities. Cowles Foundation Discussion Paper No. 1761R, Yale University.

Andrews, D. W. K., Shi, X., 2007b. Supplement to "inference based on conditional moment inequalities." Cowles Foundation Discussion Paper No. 1761R, Yale University.