Appendix 5a Proofs for Part 5a

Proof of Lemma 5.a.1: First we consider the case in which x is an interior value in the support C = (a,b). As a precursory note, throughout this appendix we will be using GS to designate the Gaussian distribution, and often use $\mathbf{\beta}$ to designate $\hat{\boldsymbol{\beta}}$, and f_{β} to designate $\frac{\partial f}{\partial \boldsymbol{\beta}}$. Observe that since $h''(x) = 0$, then Theorems 1a and 2 give us that for x in the interior of C,

$$(nh)^{.5}\sigma_{0,1}(x,K,U)^{-1}\left[\hat{g}_{LL} - \theta\right] \xrightarrow{D} GS(0,1) \qquad \text{(A5.A.1)}$$

where

$$\sigma^2_{0,1}(x,K,U) = v(Y|X = x)(d(x))^{-1}\int_U K_{0,1}(z,U)^2\,dz\,,$$

$$U = [-1,1],$$

$$N_1(U) = \begin{bmatrix} u_0 & u_1 \\ u_1 & u_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & .2 \end{bmatrix},$$

$$M_{0,1}(z,U) = \begin{bmatrix} 1 & 0 \\ z & .2 \end{bmatrix}, \text{ and}$$

$$K_{0,1}(z,U) = \{\left|M_{0,1}(z,U)\right| / \left|N_1(U)\right|\}K(z) = K(z)\,.$$

So that A5.A.1 becomes

$$(nh)^{.5}(.6\theta(1-\theta))^{-.5}\left[\hat{g}_{LL} - \theta\right] \xrightarrow{D} GS(0,1) \qquad \text{(A5.A.2)}$$

Next, note that since C has finite length, and θ is bounded,

$$\inf_{x\in C}((.6\theta(1-\theta))^{-.5}) = C^{-.5} > 0$$

Hence,

$$\sup_{x\in C}(.6\theta(1-\theta)) = C < \infty$$

So A5.A.2 becomes

$$(nh)^{.5}C^{-.5}\left[\hat{g}_{LL}(x_i) - \theta(x_i)\right] \xrightarrow{D} GS(0,1)\,. \qquad \text{(A5.A.3)}$$

Then from Bishop, Feinberg and Holland (1975) we have that a sequence that converges in distribution is bounded in probability (p. 477). So that for x in the interior of C,

$$(nh)^{.5}C^{-.5}\left[\hat{g}_{LL}(x_i)-q(x_i)\right]=O_P(1).$$

Then

$$\left[\hat{g}_{LL}(x_i)-q(x_i)\right]=O_P(n^{-.5}t_n^{.5}). \qquad \text{(A5.A.4)}$$

Next suppose that there are k interior values of x, (where $k = np$ and $0 < p < 1$). Then

$$\left\|\hat{g}_{LL}-\mathbf{q}\right\|^2=\frac{\sum_{i=1}^{k}(\hat{g}_{LL}(x_i)-q(x_i))^2}{k}.$$

Note that by A5.A.4, for each of the k interior values

$$=\frac{(\hat{g}_{LL}(x_i)-q(x_i))^2}{k}=O_P(n^{-1}t_n^{1}k^{-1}).$$

So that

$$\left\|\hat{g}_{LL}-\mathbf{q}\right\|^2= kO_P(n^{-1}t_n^{1}k^{-1})=O_P(t_n n^{-1}) \qquad \text{(A5.A.5)}$$

for the k interior values as desired.

Next let x be a boundary value of the support C = (a,b), and when we say boundary value, we mean that $D_{xh}\neq[-1,1]$, where

$$D_{xh}=\{z{:}\,x-hz\in[a,b]\}\cap[-1,1]$$

as defined by Fan, Heckman and Wand (1995). Notice that $D_{xh}\subseteq[-1,1]=U$. According to Theorems 1a and 2 of Fan, Heckman and Wand (1995), A5.A.1 becomes

$$(nh)^{.5}\,s_{0,1}(x,K,D_{xh})^{-1}\left[\hat{g}_{LL}-q\right]\xrightarrow{D}GS(0,1) \qquad \text{(A5.A.6)}$$

Notice that we can rewrite D_{xh} as $[e,f]$, where $e\in[-1,0],\ and\ \ f\in[0,1]$. Note also that because of the assumption in the statement of the Lemma we have that either

$|e|$ = 1, or $|f|$ = 1.

Using the Epanechnikov Kernel we obtain the following results

$$u_l(D_{xh}) = \int_e^f z^l K(z)dz$$

$$= \int_e^f z^l .75(1 - z^2)dz$$

So that

$$u_0 = .75\left(z - \frac{z^3}{3} \right]_e^f = .75((f - e) - \frac{f^3 - e^3}{3}) = .75(v_1 - v_3) \text{ (say)}$$

$$u_1 = .75\left(\frac{z^2}{2} - \frac{z^4}{4} \right]_e^f = .75(\frac{f^2 - e^2}{2} - \frac{f^4 - e^4}{4}) = .75(v_2 - v_4) \text{ (say)}$$

$$u_2 = .75\left(\frac{z^3}{3} - \frac{z^5}{5} \right]_e^f = .75(\frac{f^3 - e^3}{3} - \frac{f^5 - e^5}{5}) = .75(v_3 - v_5) \text{ (say).}$$

And

$$N_1(D_{xh}) = .75\begin{bmatrix} v_1 - v_3 & v_2 - v_4 \\ v_2 - v_4 & v_3 - v_5 \end{bmatrix}$$

$$M_{0,1}(D_{xh}) = \begin{bmatrix} 1 & .75(v_2 - v_4) \\ z & .75(v_3 - v_5) \end{bmatrix},$$

so that

$$\left| N_1(D_{xh}) \right| = .75((v_1 - v_3)(v_3 - v_5) - (v_2 - v_4)^2)$$

$$\left| M_{0,1}(D_{xh}) \right| = .75((v_3 - v_5) - z(v_2 - v_4)) .$$

Next, observe that the following inequalities hold.

$$For\ n \in \aleph,\ n\ odd,\ v_n \geq \frac{1}{n}.$$

$$For\ n \in \aleph,\ n\ even,\ \left| v_n \right| \leq \frac{1}{n}.$$

$$For\ n \in \aleph,\ n\ odd,\ v_n > v_{n+2}\,.$$

$$For\ n \in \aleph,\ n\ even,\ \left|v_n\right| > \left|v_{n+2}\right|.$$

$$For\ n \in \aleph,\ n\ odd,\ v_n - \frac{1}{n} \geq v_{n+2} - \frac{1}{n+2}\,.$$

$$For\ n \in \aleph,\ n\ even,\ \frac{1}{n} - \left|v_n\right| \geq \frac{1}{n+2} - \left|v_{n+2}\right|.$$

$$For\ n \in \aleph,\ n\ even,\ v_n v_{n+2} \geq 0\,.$$

From these we obtain $\left|N_1(D_{xh})\right| = .75((v_1 - v_3)(v_3 - v_5) - (v_2 - v_4)^2)$, so that

$$.234 \cong .75\left(\left(\frac{4}{3}\right)\left(\frac{4}{15}\right) - 0\right) \geq \left|N_1(D_{xh})\right| \geq .75\left(\left(\frac{2}{3}\right)\left(\frac{2}{15}\right) - \left(\frac{1}{4}\right)^2\right) \cong .022 > 0\,.$$

Since we know this determinant is always positive and bounded we will simply refer to the determinant as Dn, as it is used as a denominator in the expressions that follow.

Next recall that by Fan, Heckman and Wand (1995) Theorem 2,

$$s^2{}_{0,1}(x, K, D_{xh}) = v(Y|X = x)(f(x))^{-1} \int_{D_{xh}} K_{0,1}(z, D_{xh})^2\, dz$$

$$= (q(1-q))\frac{.75^2}{Dn}\int_e^f (1-z^2)^2((v_3 - v_5) - z(v_2 - v_4))^2\, dz \qquad \text{(A5.A.7)}$$

We will demonstrate that A5.A.7 is a positive bounded quantity. First, observe that the coefficient of the integral is positive and bounded (since C is finite, and q is bounded and using the reasoning above for A5.A.3, $Dn > 0$, and $.75^2 > 0$). Second, observe that on the finite interval [e,f], the integrand is nonnegative and not identically zero. To see this note that

$$(1-z^2)^2 = 0,\ iff\ \ z = \pm 1$$

$$(v_3 - v_5) > 0,\ and\ \ (v_2 - v_4) \neq 0\,.$$

Then note that

$$(1-z^2)^2((v_3-v_5)-z(v_2-v_4))^2 \geq 0$$

and

$$(1-z^2)^2((v_3-v_5)-z(v_2-v_4))^2 = 0, iff \;\; z = \pm\frac{(v_3-v_5)}{(v_2-v_4)}.$$

Third, recall that $f-e \geq 1$. Finally, with the information above, we can observe that

$$.54 \geq \int_e^f (1-z^2)^2((v_3-v_5)-z(v_2-v_4))^2\,dz \geq .003 > 0$$

so that A5.A.7 is a positive bounded quantity, which we shall again call C. Then A5.A.6 becomes A5.A.3 and using the resoning from A5.A.3 through A5.A.5 (replacing the term k interior values with k boundary values) we have the desired result (A5.A.5) holds for the boundary values as well.

One may also argue that since C is an open interval, for every x, A5.A.6 becomes A5.A.1 for sufficiently large n. In other words, for every x in C, as the bandwidth approaches zero, x will eventually become an “interior” value and A5.A.7 will transform into $s^2_{0,1}(x,K,U)$. So that for our case (C open) the interior asymptotic result automatically holds for the entire support C. Thus, A5.A.5 holds for the mean response vectors (both true and estimated) based on the entire data set (n observation) and the proof of the lemma is completed.//.

Proof of Lemma 5.a.2: $S_{Gn} = n^{-1}\sum_{i=1}^{n}\left(w_i \frac{\P f(x_i;\boldsymbol{b})}{\P\boldsymbol{b}} \left(\frac{\P f(x_i;\boldsymbol{b})}{\P\boldsymbol{b}} \right)^T \right)$ (w_i represents the weight)

$$= n^{-1}D^T \boldsymbol{L} D$$

where D is the matrix of derivatives defined earlier, and $\boldsymbol{L}$ is the covariance matrix for $\boldsymbol{Y}$, that is

$$\boldsymbol{L} = \begin{bmatrix} V(y_1) & 0 & \cdots & 0 \\ 0 & V(y_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & V(y_n) \end{bmatrix}.$$

Next observe that for any $n > p+1$, rk(D) = rk(D^T) = $p+1$, and rk($\boldsymbol{L}^{.5}$) = n, by requirements R3 and R1 respectively. Then by the properties of rank (see Myers and Milton (1991)), since $\boldsymbol{L}^{.5}$ is nonsingular,

$$\text{rk}(D^T \boldsymbol{L}^{.5}) = p+1 = \text{rk}(\boldsymbol{L}^{.5} D),$$

and since

$$D^T \boldsymbol{L}^{.5} = (\boldsymbol{L}^{.5} D)^T,$$

then rk($n^{-1}D^T\boldsymbol{L}D$) = $p+1$, and ($n^{-1}D^T\boldsymbol{L}D$) = S_{Gn} is full rank so that it is a nonsingular ($p+1$) x ($p+1$) matrix. We need only show that the determinant does not converge to zero. Notice that because of R1, and R2, $\boldsymbol{L}$ is full rank asymptotically. So that if, in fact, S_G is singular, then by the properties of rank (Myers and Milton (1991)),

$$\text{rk}(D^T\boldsymbol{L}) = \text{rk}(D^T) < p+1.$$

But this is equivalent to

$$\frac{\P f(x;\boldsymbol{b})}{\P\boldsymbol{b}_i} = f_{b_i}(x;\boldsymbol{b}) = \sum_{j \neq i} c_j f_{b_j}(x;\boldsymbol{b}),\ for\ a \le x \le b$$

(for some $i \in \{0,1,\ldots p\}$) with $c_j \neq 0,\ j \in \{\{0,1,\ldots,p\} \setminus i\}$, which violates requirement R3.#.

So S_G is nonsingular asymptotically.

One may also argue that since $D^T L^{.5}$ is asymptotically full rank, S_G is positive definite, and by Myers and Milton (1991) Theorem 2.1.1 and properties of rank, S_G is full rank and nonsingular asymptotically. To see that this determinant is bounded asymptotically, observe that S_G

$$= \begin{bmatrix} \int_a^b f_{b_0}^2 w dx & \int_a^b f_{b_0} f_{b_1} w dx & \cdots & \int_a^b f_{b_0} f_{b_k} w dx \\ \int_a^b f_{b_0} f_{b_1} w dx & \int_a^b f_{b_1}^2 w dx & & \\ \vdots & & \ddots & \vdots \\ \int_a^b f_{b_0} f_{b_k} w dx & & \cdots & \int_a^b f_{b_k}^2 w dx \end{bmatrix}$$

where $f_{bi}(x)$ (for $i = 0,1,\ldots p$) is as defined earlier, and $w(x) = V(x) = V(y|x)$. By page 187 of Spence, Insel and Friedberg (2000), $|\det(S_G)|$

$$= \prod_{i=0}^{p} u_{ii} < \infty$$

where u_{ii} represents the ith diagonal element of the upper triangular matrix formed by no more than $p+1$ non-scaling elementary row operations, and is finite since all elements of S_G are finite by requirements R2, and R4. Then from the above arguments and Myers and Milton (1991) p. 38, we have that

$$0 < |\det(S_G)| < \infty .$$

So by Theorem 3.4 of Spence, Insel and Friedberg (2000), S_G^{-1} is defined and finite as desired.//.

Proof of Lemma 5.a.3: First we have from Carroll and Ruppert (1988) Theorem 2.1 that for any $n^{.5}$ consistent starting estimate for $\boldsymbol{b}$ (i.e. one coming from LS) that

$$n^{.5}\left[\hat{\boldsymbol{b}} - \boldsymbol{b}\right] \xrightarrow{D} GS(0, s^2 S_G^{-1}) \qquad \text{(A5.A.8)}$$

(where S_G^{-1} was defined and shown to be finite in Lemma 5.a.2) for any number of cycles of the IRLS algorithm. Note that for finite s^2, $s^2 S_G^{-1}$ is also defined and finite (by Lemma 5.a.2).

What we need is an asymptotic result for the function $f(\mathbf{x}_i;\boldsymbol{b})$. Theorem 14.6-2 of Bishop, Feinberg, and Holland (1975) gives us just that. With T = $p+1$, R = 1 we note that for x_i in C,

$$f(\mathbf{b}) = f(\boldsymbol{b}) + (\mathbf{b} - \boldsymbol{b})^T f_{\boldsymbol{b}} + o\left(\left\|\mathbf{b} - \boldsymbol{b}\right\|\right) \; as \; \mathbf{b} \to \boldsymbol{b}$$

by Taylor Series expansion for vectors, since $\dfrac{\P f}{\P \mathbf{b}}$ exists (see Bishop, Feinberg, and Holland (1975)) by requirements R3, and R4. Then with result A5.A.8, we can invoke Theorem 14.6-2 which says that for each x_i in C

$$n^{.5}\left[\hat{f}(x_i;\hat{\boldsymbol{b}}) - f(x_i;\boldsymbol{b})\right] \xrightarrow{D} GS\left(0, \left(\frac{\P f(x_i;\boldsymbol{b})}{\P \boldsymbol{b}}\right)^T (\boldsymbol{s}^2 S_G^{-1}) \left(\frac{\P f(x_i;\boldsymbol{b})}{\P \boldsymbol{b}}\right)\right) \text{.(A5.A.9)}$$

Then repeating the logical sequence from A5.A.3 to A5.A.5 above, and replacing C with $\left(\dfrac{\P f(x_i;\boldsymbol{b})}{\P \boldsymbol{b}}\right)^T (\boldsymbol{s}^2 S_G^{-1}) \left(\dfrac{\P f(x_i;\boldsymbol{b})}{\P \boldsymbol{b}}\right)$, $\boldsymbol{q}$ with f, h (and subsequently $\boldsymbol{t}_n$) with 1, k with n, $\hat{g}_{LL}$ with $\hat{f}$, and deleting "interior" we have that

$$\left\|\hat{f} - f\right\|^2 = O_P(n^{-1})$$

as desired.//.