Technical supplement to "Semiparametric stochastic frontier estimation via profile likelihood"

Carlos Martins-Filho

Department of Economics
University of Colorado
Boulder, CO 80309-0256, USA
email: carlos.martins@colorado.edu
Voice: + 1 303 492 4599

&

IFPRI
2033 K Street NW
Washington, DC 20006-1002, USA
email: c.martins-filho@cgiar.org
Voice: + 1 202 862 8144

and

Feng Yao

Department of Economics
West Virginia University
Morgantown, WV 26505, USA
email: feng.yao@mail.wvu.edu
Voice: +1 304 2937867

December, 2011

Abstract. In this technical supplement to Martins-Filho and Yao (2011) we show that the skew-normal density satisfies all assumptions required in establishing the asymptotic properties of the estimators discussed therein. Also, we provide the proofs for Theorems 1, 2 and Lemma 1.

1 Introduction

Besides this introduction, this technical supplement has four sections. Section 2 provides proofs for Theorems 1, 2 and Lemma 1 in Martins-Filho and Yao (2011). Section 3 verifies that a skew-normal density satisfies all assumptions (A1-A7) required to prove Theorem 2. Section 4 verifies that assumptions PA1, PA2, PB, PC and PD used in Lemma 2 of Martins-Filho and Yao (2011) are also satisfied by the skew-normal density. Section 5 obtains the variance of the asymptotic distribution of $\sqrt{n}(\tilde{\theta}-\theta_0)$ when f_ϵ in Martins-Filho and Yao (2011) is a skew-normal. Throughout the note $\theta=(\sigma_u^2,\sigma_v^2)$ and we have

$$f_\varepsilon(y-g(x);\theta)=\frac{2}{\sqrt{\sigma_u^2+\sigma_v^2}}\phi\left(\frac{y-g(x)}{\sqrt{\sigma_u^2+\sigma_v^2}}\right)\left(1-\Phi\left(\frac{\sqrt{\sigma_u^2/\sigma_v^2}}{\sqrt{\sigma_u^2+\sigma_v^2}}(y-g(x))\right)\right) \tag{1}$$

where $\sigma_u^2,\sigma_v^2>0$, $\lambda=\sqrt{\frac{\sigma_u^2}{\sigma_v^2}}$ and $s^2=\sigma_u^2+\sigma_v^2$. Whenever we need to distinguish the true values of the parameters and the function g we write $\theta_0=\left(\begin{array}{c}\sigma_{u0}^2\\ \sigma_{v0}^2\end{array}\right)$, λ_0, s_0^2 and g_0. C will always denote an arbitrary positive real number.

2 Proofs of Theorems 1, 2 and Lemma 1

Theorem 1: *Proof.* Given A1.1-3, A4 and the definition of $\hat{\theta}$, by Theorem 2.1 in Newey and McFadden (1994) it suffices to prove that $sup_{\theta\in\Theta}\left|\bar{l}_n(\theta,\hat{m}+\gamma(\theta))-E\left(\bar{l}_n(\theta,g_0)\right)\right|=o_p(1)$. We do so by establishing that $sup_{\theta\in\Theta}\left|\bar{l}_n(\theta,\hat{m}+\gamma(\theta))-\bar{l}_n(\theta,g_0)\right|=o_p(1)$ and $sup_{\theta\in\Theta}\left|\bar{l}_n(\theta,g_0)-E\left(\bar{l}_n(\theta,g_0)\right)\right|=o_p(1)$. First, note that by A1.3 and A4.4 $sup_{\theta\in\Theta}\left|\bar{l}_n(\theta,\hat{m}+\gamma(\theta))-\bar{l}_n(\theta,g_0)\right|\leq\frac{1}{n}\sum_{i=1}^n sup_{\theta\in\Theta}b(y_i,x_i,\theta)|\hat{m}(x_i)-m(x_i;\theta,g_0)|$. If $\frac{nh_n^3}{log(n)}\to\infty$ as $n\to\infty$ and given A2.2 and A3, for a compact set G, we have $sup_{x\in G}|\hat{m}(x)-m(x;\theta,g_0)|=o_p(1)$ (Martins-Filho and Yao, 2007). Given that $E\left(sup_{\theta\in\Theta}b(y_i,x_i,\theta)\right)<\infty$ and the fact that (y_i,x_i) are i.i.d, we conclude that $sup_{\theta\in\Theta}\left|\bar{l}_n(\theta,\hat{m}+\gamma(\theta))-\bar{l}_n(\theta,g_0)\right|=o_p(1)$. Second, note that

$$\left|\bar{l}_n(\theta,g_0)-E\left(\bar{l}_n(\theta,g_0)\right)\right|\leq\frac{1}{n}\sum_{i=1}^n\left|logf_\varepsilon(y_i-g_0(x_i);\theta)-E\left(logf_\varepsilon(y_i-g_0(x_i);\theta)\right)\right|.$$

By the Heine-Borel theorem every open covering of Θ contains a finite subcover $\{S_k\}_{k=1}^K$ where $S_k=S(\theta_k,d(\theta_k))$ denotes an open sphere centered at θ_k with radius $d(\theta_k)>0$. Now let

$$\mu(y_i,x_i,\theta,g_0,d(\theta))=sup_{\theta'\in S(\theta,d(\theta))}\left|logf_\varepsilon(y_i-g_0(x_i);\theta)-logf_\varepsilon(y_i-g_0(x_i);\theta')\right|.$$

By A4.2 $\mu(y_i, x_i, \theta, g_0, d(\theta)) \to 0$ for all $\theta \in \Theta$ as $d(\theta) \to 0$ almost everywhere arcording to $f(y, x)$. By the triangle inequality $\mu(y_i, x_i, \theta, g_0, d(\theta)) \leq 2sup_{\theta\in\Theta}|logf_\varepsilon(y_i; \theta, g_0(x_i))|$. By A4.3 and Lebesgue's dominate convergence theorem (LDC) we conclude that for any $\epsilon, d > 0$, $E\left(\mu(y_i, x_i, \theta, g_0, d(\theta))\right) < \epsilon$ whenever $d(\theta) < d$. Letting $d(\theta_k) < d$ for all $k = 1, ..., K$ we have $E(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k))) < \epsilon$ for all k. Also, $|E(logf_\varepsilon(y_i - g_0(x_i); \theta) - E(logf_\varepsilon(y_i - g_0(x_i); \theta_k))| \leq E\left(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k))\right) < \epsilon$. Hence, for $\theta \in S_k$ we have

$$\begin{aligned} \left|\bar{l}_n(\theta, g_0) - E\left(\bar{l}_n(\theta, g_0)\right)\right| &\leq n^{-1}\sum_{i=1}^{n}\left(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k)) - E(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k)))\right) \\ &+ \left|n^{-1}\sum_{i=1}^{n}\left(logf_\varepsilon(y_i - g_0(x_i); \theta_k) - E(logf_\varepsilon(y_i - g_0(x_i); \theta_k)\right)\right| + 2\epsilon \end{aligned}$$

Since $E(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k))) < \infty$ and $E(|logf_\varepsilon(y_i - g_0(x_i); \theta_k)|) < \infty$, we have, by the strong law of large numbers, that there exists $N_{\epsilon,k}$ such that $n > N_{\epsilon,k}$ gives

$$\left|n^{-1}\sum_{i=1}^{n}\mu(y_i, x_i, \theta_k, g_0, d(\theta_k)) - E\left(\mu(y_i, x_i, \theta_k, g_0, d(\theta_k))\right)\right| < \epsilon$$

and $n^{-1}\sum_{i=1}^{n}|logf_\varepsilon(y_i - g_0(x_i); \theta_k) - E(logf_\varepsilon(y_i - g_0(x_i); \theta_k))| < \epsilon$. Given that K is finite, for all $n > max_k N_{k,\epsilon}$ we have $sup_{\theta\in\Theta}|\bar{l}_n(\theta, g_0) - E(\bar{l}_n(\theta, g_0))| = o_p(1)$.

Theorem 2: *Proof.* Given A6.1 and Taylor's Theorem in Graves (1927),

$$\begin{aligned} \bar{l}_n(\theta, \hat{m} + \gamma(\theta)) &= \frac{1}{n}\sum_{i=1}^{n}logf_\varepsilon(y_i - g_0(x_i); \theta) + \frac{1}{n}\sum_{i=1}^{n}\frac{d_F}{dg}logf_\varepsilon(y_i - g_0(x_i), \theta)(\hat{m}(x_i) - m(x_i; \theta, g_0)) \\ &+ \frac{1}{2n}\sum_{i=1}^{n}(\hat{m}(x_i) - m(x_i; \theta, g_0))^2\int_0^1 \frac{d_F^2}{dg^2}logf_\varepsilon(y_i - g_0(x_i) - t(\hat{m}(x_i) - m(x_i; \theta, g_0); \theta) \\ &\times (1 - t)dt. \end{aligned}$$

Denoting the last term in the above inequality by c_n and given that $\frac{d_F}{dg}$ is a bounded linear functional from $\mathcal{G}$ to $\Re$, we have

$$\begin{aligned} |c_n| &\leq \frac{1}{2n}\sum_{i=1}^{n}\left(\sup_{x\in G}|\hat{m}(x) - m(x; \theta, g_0)|\right)^2\int_0^1 C\sup_{x\in G}|g_0(x) + t(\hat{m}(x) - m(x; \theta, g_0))|(1 - t)dt \\ &\leq \frac{1}{2}\left(\sup_{x\in G}|\hat{m}(x) - m(x; \theta, g_0)|\right)^2\int_0^1 C\left(\sup_{x\in G}|g_0(x)| + t\sup_{x\in G}|(\hat{m}(x) - m(x; \theta, g_0))|\right)(1 - t)dt \\ &\leq C\left(\left(\sup_{x\in G}|\hat{m}(x) - m(x; \theta, g_0)|\right)^2 + \left(\sup_{x\in G}|\hat{m}(x) - m(x; \theta, g_0)|\right)^3\right). \end{aligned}$$

Since $\sup_{x\in G}|\hat{m}(x) - m(x; \theta, g_0)| = O_p\left(\left(\frac{log(n)}{nh_n}\right)^{1/2} + h_n^2\right)$, it follows that if $h_n = O(n^{-1/5})$ we have that

$|c_n| = o_p(n^{-1/2})$. Consequently, we can write

$$\bar{l}_n(\theta, \hat{m} + \gamma(\theta)) = \bar{l}_n(\theta, g_0(x_i)) + \frac{1}{n}\sum_{i=1}^{n} \frac{d_F}{dg} logf_\varepsilon(y_i - g_0(x_i), \theta)(\hat{m}(x_i) - m(x_i;\theta, g_0)) + o_p(n^{-1/2}).$$

Since $\frac{\partial}{\partial\theta}\bar{l}_n(\hat{\theta}, \hat{m}+\gamma(\hat{\theta})) = 0$ we have . By A5.1 and the mean value theorem there exists some $\bar{\theta} \in L(\hat{\theta}, \theta_0)$ (the line segment uniting $\hat{\theta}$ and θ_0) such that

$$-\frac{\partial^2}{\partial\theta\partial\theta'}\bar{l}_n(\bar{\theta}, \hat{m} + \gamma(\bar{\theta}))\sqrt{n}(\hat{\theta} - \theta_0) = \sqrt{n}\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0, \hat{m} + \gamma(\theta_0)). \quad (2)$$

We now write

$$\begin{aligned}\sqrt{n}\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0, \hat{m} + \gamma(\theta_0)) &= \sqrt{n}\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0, g_0) + \frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\frac{d_F}{dg}logf_\varepsilon(y_i - g_0(x_i), \theta_0)(\hat{m}(x_i) - m(x_i;\theta_0, g_0)) \\ &+ o_p(1).\end{aligned}$$

Due to the equality of Fréchet and Gateaux differentials, we have

$$\frac{d_F}{dg}logf_\varepsilon(y_i - g_0(x_i), \theta)(\hat{m}(x_i) - m(x_i;\theta, g_0)) = \frac{\partial}{\partial\eta}logf_\varepsilon(y_i - g_0(x_i);\theta)(\hat{m}(x_i) - m(x_i;\theta, g_0)).$$

Given A5.6 we have from Lemma 1 in Martins-Filho and Yao (2006),

$$\begin{aligned}\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\frac{d_F}{dg}logf_\varepsilon(y_i - g_0(x_i);\theta)(\hat{m}(x_i) - m(x_i;\theta, g_0)) &= \sqrt{n}\frac{\partial}{\partial\theta}\left(\frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i;\theta, g_0))\times\right. \\ \int\frac{\partial}{\partial\eta}logf_\varepsilon(y - g_0(x_i);\theta)f_\varepsilon(y - g_0(x_i);\theta_0)dy &+ \left.\frac{1}{2}h_n^2\sigma_K^2 E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}logf_\varepsilon(y_i - g_0(x_i);\theta)\right)\right) \\ &+o_p(1) + \sqrt{n}\ o_p(h_n^2).\end{aligned}$$

We observe that the term $y_i - m(x_i;\theta, g_0)$ is a function of θ. Hence, when passing the $\frac{\partial}{\partial\theta}$ operator and evaluating at θ_0 we obtain,

$$\begin{aligned}&\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\frac{d_F}{dg}logf_\varepsilon(y_i-, g_0(x_i);\theta_0)(\hat{m}(x_i) - m(x_i;\theta_0, g_0)) \\ =\ &\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\gamma(\theta_0)\int\frac{\partial}{\partial\eta}logf_\varepsilon(y - g_0(x_i);\theta_0)f_\varepsilon(y - g_0(x_i);\theta_0)dy + \frac{1}{n}\sum_{i=1}^{n}(y_i - m(x_i;\theta_0, g_0))\right. \\ \times\ &\left.\frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\eta}logf_\varepsilon(y - g_0(x_i);\theta_0)f_\varepsilon(y - g_0(x_i);\theta_0)dy + \frac{1}{2}h_n^2\sigma_K^2\frac{\partial}{\partial\theta}E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}logf_\varepsilon(y_i - g_0(x_i);\theta_0)\right)\right) \\ +\ &o_p(1) + \sqrt{n}\ o_p(h_n^2).\end{aligned}$$

Given assumption A5.3, we have

$$\frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\eta}f_\varepsilon(y - g_0(x_i);\theta_0)f_\varepsilon(y - g_0(x_i);\theta_0)dy = \int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y - g_0(x_i);\theta_0)f_\varepsilon(y - g_0(x_i);\theta_0)dy$$

and

$$\frac{\partial}{\partial\theta}E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)=E\left(g_0^{(2)}(x_i)\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right).$$

In addition, by A5.5

$$\begin{aligned}\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)f_\varepsilon(y-g_0(x_i);\theta_0)dy &= \frac{\partial}{\partial\eta}\int f_\varepsilon(y-g_0(x_i);\theta_0)f_\varepsilon(y-g_0(x_i);\theta_0)dy\\ &= \frac{\partial}{\partial\eta}E(f_\varepsilon(y_i-g_0(x_i);\theta_0)=0.\end{aligned}$$

Hence, we can write

$$\begin{aligned}\sqrt{n}\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0,\hat{m}+\gamma(\theta_0)) &= \sqrt{n}\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0,g_0)+\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n(y_i-m(x_i;\theta_0,g_0))\int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right.\\ &\times\left. f_\varepsilon(y)dy+\frac{1}{2}h_n^2\sigma_K^2E\left(g_0^{(2)}(x_i)\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)\right)\\ &+ \sqrt{n}\ o_p(h_n^2)+o_p(1)\\ &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n\frac{\partial}{\partial\theta}f_\varepsilon(y_i-g_0(x_i);\theta_0)+(y_i-m(x_i;\theta_0,g_0))\right.\\ &\times\left.\int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)f_{y|x}(y)dy\right)\\ &+ \sqrt{n}\left(\frac{1}{2}h_n^2\sigma_K^2E\left(g_0^{(2)}(x_i)\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)+o_p(h_n^2)\right)+o_p(1).\end{aligned}$$

Let $Z_i=\frac{\partial}{\partial\theta}f_\varepsilon(y_i-g_0(x_i);\theta_0)+(y_i-m(x_i;\theta_0,g_0))\int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)f_\varepsilon(y)dy$ and observe that $E(Z_i)=0$ since $E(y_i-m(x_i;\theta_0,g_0)|x_i)=0$ and $E(f_\varepsilon(y-g_0(x_i);\theta)$ has an unique maximum at θ_0. Let $\sigma_F^2=E(Z_iZ_i')$, which exists as a positive definite matrix by A6.3 with

$$\begin{aligned}\sigma_F^2 &= E\left(\left(\frac{\partial}{\partial\theta}f_\varepsilon(y_i-g_0(x_i);\theta_0)+(y_i-m(x_i;\theta_0,g_0))\int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)f_\varepsilon(y-g_0(x_i);\theta_0)dy\right)\right.\\ &\times\left.\left(\frac{\partial}{\partial\theta}f_\varepsilon(y_i-g_0(x_i);\theta_0)+(y_i-m(x_i;\theta_0,g_0))\int\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)f_\varepsilon(y-g_0(x_i);\theta_0)dy\right)'\right).\end{aligned}$$

Since Z_i is a continuous (measurable) function of $\begin{pmatrix}y_i\\x_i\end{pmatrix}$, and given that the sequence $\left\{\begin{pmatrix}y_i\\x_i\end{pmatrix}\right\}_{i=1,2,\cdots}$ is i.i.d., by the Cramer-Wold device and Lévy's central limit theorem, we have

$$\sqrt{n}\left(\frac{\partial}{\partial\theta}\bar{l}_n(\theta_0,\hat{m}+\gamma(\theta_0))-B_{1n}\right)\xrightarrow{d}N(0,\sigma_F^2)\tag{3}$$

where $B_{1n}=\frac{1}{2}h_n^2\sigma_K^2E\left(g_0^{(2)}(x_i)\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)+o_p(h_n^2)$.

We now study the asymptotic behavior of $\frac{\partial^2}{\partial\theta\partial\theta'}\bar{l}_n(\bar{\theta},\hat{m}+\gamma(\bar{\theta}))$. Note that,

$$\begin{aligned}\frac{\partial^2}{\partial\theta\partial\theta'}\bar{l}_n(\bar{\theta},\hat{m}+\gamma(\bar{\theta})) &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta\partial\theta'}f_\varepsilon(y_i-g_0(x_i);\bar{\theta}) \\ &+ \frac{\partial^2}{\partial\theta\partial\theta'}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{d_F}{dg}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})(\hat{m}(x_i)-m(x_i;\bar{\theta},g_0))\right) \\ &+ o_p(n^{-1/2}).\end{aligned}$$

Since $\hat{\theta}-\theta_0=o_p(1)$ and $\bar{\theta}\in L(\hat{\theta},\theta_0)$ we have that $\bar{\theta}-\theta_0=o_p(1)$, that is, for sufficiently large n, $\bar{\theta}\in S_0$. Denote the (i,j) element of $\frac{\partial^2}{\partial\theta\partial\theta'}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})$ by $\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})$ and note that by A5.1 it is continuous on S_0. Furthermore, by A5.2 and Theorem 1 we have that $E\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\theta)\right)$ is continuous at θ_0 and

$$\sup_{\theta\in S_0}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\theta)-E\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\theta)\right)\right|=o_p(1).$$

By Theorem 21.6 in Davidson (1994) we conclude that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})\xrightarrow{p}E\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)\text{ for all }(i,j).$$

Now, from earlier in the proof we have that

$$\begin{aligned}&\frac{\partial}{\partial\theta}\frac{1}{n}\sum_{i=1}^{n}\frac{d_F}{dg}f_\varepsilon(y_i-g_0(x_i);\theta)(\hat{m}(x_i)-m(x_i;\theta,g_0)) \\ =\ & \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\gamma(\theta)\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)f_\varepsilon(y-g_0(x_i);\theta_0)dy+\frac{1}{n}\sum_{i=1}^{n}(y_i-m(x_i;\theta,g_0)) \\ \times\ & \frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)f_\varepsilon(y-g_0(x_i);\theta_0)dy+\frac{1}{2}h_n^2\sigma_K^2\frac{\partial}{\partial\theta}E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}f_\varepsilon(y_i-g_0(x_i);\theta)\right) \\ +\ & o_p(n^{-1/2})+o_p(h_n^2)\text{ and therefore we write}\end{aligned}$$

$\frac{\partial^2}{\partial\theta\partial\theta'}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{d_F}{dg}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})(\hat{m}(x_i)-m(x_i;\bar{\theta},g_0))\right)=\sum_{j=1}^{5}I_{4j}+o_p(n^{-1/2})+o_p(h_n^2)$ where

$$\begin{aligned}I_{41} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta\partial\theta'}\gamma(\bar{\theta})\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\bar{\theta})f_\varepsilon(y-g_0(x_i);\theta_0)dy \\ I_{42} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\bar{\theta})f_\varepsilon(y-g_0(x_i);\theta_0)dy\frac{\partial}{\partial\theta}\gamma(\bar{\theta}) \\ I_{43} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\gamma(\bar{\theta})\frac{\partial}{\partial\theta}\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\bar{\theta})f_\varepsilon(y-g_0(x_i);\theta_0)dy \\ I_{44} &= \frac{1}{n}\sum_{i=1}^{n}(y_i-m(x_i;\theta,g_0))\frac{\partial^2}{\partial\theta\partial\theta'}\int\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\bar{\theta})f_\varepsilon(y-g_0(x_i);\theta_0)dy \\ I_{45} &= \frac{1}{2}h_n^2\sigma_K^2\frac{\partial^2}{\partial\theta\partial\theta'}E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}f_\varepsilon(y_i-g_0(x_i);\bar{\theta})\right)\end{aligned}$$

The order of each term can be obtained by repeated use of Theorem 1 and Theorem 21.6 in Davidson (1994). We obtain,

$$
\begin{aligned}
I_{41} &\xrightarrow{p} \frac{\partial^2}{\partial\theta\partial\theta'}\gamma(\theta_0)E\left(\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right) = 0 \text{ given A5.3, A5.6 and A7.1,}\\
I_{42} = I'_{43} &\xrightarrow{p} \frac{\partial}{\partial\theta}\gamma(\theta_0)E\left(\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right) \text{ given A5.3, A5.6, A7.2, A7.3}\\
I_{44} &\xrightarrow{p} 0 \text{ given A7.2 , A7.4, A7.5 since } E(y_i - m(x_i;\theta_0,g_0)|x_i) = 0\\
\frac{1}{h_n^2}I_{45} &\xrightarrow{p} \frac{1}{2}\sigma_K^2\frac{\partial^2}{\partial\theta\partial\theta'}E\left(g^{(2)}(x_i)\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right) \text{ given A7.2 , A7.4, A7.6}
\end{aligned}
$$

Combining all terms, we have

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta\partial\theta'}\bar{l}_n(\bar\theta,\hat m+\gamma(\bar\theta)) &= E\left(\frac{\partial^2}{\partial\theta\partial\theta}f_\varepsilon(y_i-g_0(x_i);\theta_0)\right)\\
&+ \frac{\partial}{\partial\theta}\gamma(\theta_0)E\left(\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right)\\
&+ E\left(\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right)\frac{\partial}{\partial\theta}\gamma(\theta_0)'\\
&+ h_n^2O_p(1) + o_p(h_n^2) + o_p(n^{-1/2}).
\end{aligned}
$$

Using the last equation together with the results in equations (2) and (3) we conclude that,

$$
\sqrt{n}(\hat\theta-\theta_0-B_{2n}) \xrightarrow{d} N(0,\bar H^{-1}\sigma_F^2\bar H^{-1}) \tag{4}
$$

where $B_{2n} = -\bar H^{-1}\left(\frac{1}{2}h_n^2\sigma_K^2E\left(g_0^{(2))}(x_i)\frac{\partial^2}{\partial\theta\partial\eta}f_\varepsilon(y-g_0(x_i);\theta_0)\right)\right) + o_p(h_n^2)$.

Lemma 1: *Proof.* Let $S_0(\theta,\eta,x) = \frac{1}{n}\sum_{i=1}^n\frac{1}{h_n}\frac{\partial}{\partial\eta}log f_\varepsilon(y_i-\eta;\theta)K\left(\frac{x_i-x}{h_n}\right)$. For $x\in G$, since G is compact there exists a sphere $B(a;r)$ such that $G\subset B(a,r)$. For fixed n, by the Heine-Borel theorem there exist $\{B(x_k,\delta_n)\}_{k=1}^{l_n}$ for $x_k\in G$ such that $G\subseteq\cup_{k=1}^{l_n}B(x_k,\delta_n)$ with $l_n<r/\delta_n$. Similarly, there exists $\{B(\eta_k,\delta_n)\}_{k=1}^{l_{\eta n}}$ for $\eta_k\in\mathcal{H}$ such that $\mathcal{H}\subseteq\cup_{k=1}^{l_{\eta n}}B(\eta_k,\delta_n)$ with $l_{\eta n}<r'/\delta_n$ and $\{B(\theta_k,\delta_n)\}_{k=1}^{l_{\theta n}}$ for $\theta_k\in\Theta$ such that $\Theta\subseteq\cup_{k=1}^{l_{\theta n}}B(\theta_k,\delta_n)$ with $l_{\theta n}<r^P/\delta_n^P$.

$$
\begin{aligned}
|S_0(\theta,\eta,x)-E(S_0(\theta,\eta,x))| \le\ & |S_0(\theta,\eta,x)-S_0(\theta_{k_1},\eta,x)| + |S_0(\theta_{k_1},\eta,x)-S_0(\theta_{k_1},\eta_{k_2},x)|\\
&+ |S_0(\theta_{k_1},\eta_{k_2},x)-S_0(\theta_{k_1},\eta_{k_2},x_{k_3})| + |E(S_0(\theta,\eta,x))-E(S_0(\theta_{k_1},\eta,x))|\\
&+ |E(S_0(\theta_{k_1},\eta,x))-E(S_0(\theta_{k_1},\eta_{k_2},x))| + |E(S_0(\theta,\eta,x))-E(S_0(\theta,\eta,x))|\\
&+ |S_0(\theta_{k_1},\eta_{k_2},x_{k_3})-E(S_0(\theta_{k_1},\eta_{k_2},x_{k_3}))|
\end{aligned}
$$

where $k_1\in\{1,\cdots,l_{\theta n}\}$, $k_2\in\{1,\cdots,l_{\eta n}\}$, $k_3\in\{1,\cdots,l_n\}$. By construction of the open balls, for all $x\in G$,$\theta\in\Theta$, $\eta\in\mathcal{H}$ we can always find k_1,h_2,k_3 such that $||\theta-\theta_{k_1}||<\delta_n$,$|\eta-\eta_{k_2}|<\delta_n$ and

$|x - x_{k_3}|| < \delta_n$.

$|S_0(\theta,\eta,x) - S_0(\theta_{k_1},\eta,x)| \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}K\left(\frac{x_i - x}{h_n}\right)\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta^*)\right|\right| ||(\theta-\theta_{k_1})||$ for $\theta^* \in L(\theta,\theta_{k_1})$.

By the c_r-inequality $\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta^*)\right|\right| \leq \sum_{k=1}^{P}\left|\frac{\partial^2}{\partial\eta\partial\theta_k}logf_\varepsilon(y_i-\eta;\theta^*)\right|$, and since by assumption $E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta)\right|^2\right) < \infty$, we have $E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta^*)\right|\right|\right) \leq \sum_{k=1}^{P}$ $E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta\partial\theta_k}logf_\varepsilon(y_i-\eta;\theta^*)\right|^2\right) < \infty$. Since $||\theta-\theta_{k_1}|| < \delta_n$ and $K(x) < C$

$$\begin{aligned}|S_0(\theta,\eta,x) - S_0(\theta_{k_1},\eta,x) &\leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}K\left(\frac{x_i-x}{h_n}\right)\delta_n\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta^*)\right|\right| \\ &\leq C\delta_n\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta)\right|\right|\end{aligned}$$

where $\frac{1}{n}\sum_{i=1}^{n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta^*)\right|\right| = O_p(1)$. By similar manipulations we have,

$$|S_0(\theta_{k_1},\eta,x) - S_0(\theta_{k_1},\eta_{k_2},x) \leq C\delta_n\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y_i-\eta;\theta)\right|$$

where $\frac{1}{n}\sum_{i=1}^{n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y_i-\eta;\theta)\right| = O_p(1)$ using assumption PA2 with $s = 2$, $r = 0$. Similarly, we also obtain $|S_0(\theta_{k_1},\eta_{k_2},x) - S_0(\theta_{k_1},\eta_{k_2},x_{k_3}) \leq C\frac{1}{h_n^2}\delta_n\frac{1}{n}\sum_{i=1}^{n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y_i-\eta;\theta)\right|$ where $\frac{1}{n}\sum_{i=1}^{n}\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y_i-\eta;\theta)\right| = O_p(1)$ using PA2 with $s = 1$, $r = 0$. In an analogous fashion we obtain, $|E(S_0(\theta,\eta,x)) - E(S_0(\theta_{k_1},\eta,x))| \leq C\frac{1}{h_n}\delta_nE\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta\partial\theta}logf_\varepsilon(y_i-\eta;\theta)\right|\right)$, $|E(S_0(\theta_{k_1},\eta,x)) - E(S_0(\theta_{k_1},\eta_{k_2},x))| \leq C\frac{1}{h_n}\delta_nE\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y_i-\eta;\theta)\right|\right)$ and $|E(S_0(\theta_{k_1},\eta_{k_2},x)) - E(S_0(\theta_{k_1},\eta_{k_2},x_{k_3}))|$ $\leq C\frac{1}{h_n^2}\delta_nE\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y_i-\eta;\theta)\right|\right)$. Hence, we can find k_1, k_2, k_3 such that

$$|S_0(\theta,\eta,x) - E(S_0(\theta,\eta,x))| \leq |E(S_0(\theta_{k_1},\eta_{k_2},x)) - E(S_0(\theta_{k_1},\eta_{k_2},x_{k_3}))| + 3C\delta_nh_n^{-2}\frac{1}{n}\sum_{i=1}^{n}M_i$$

where $M_i = \left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\theta\partial\eta}logf_\varepsilon(y_i-\eta;\theta)\right| + \sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y_i-\eta;\theta)\right| + \sup_{\theta\in\Theta,\eta\in\mathcal{H}}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y_i-\eta;\theta)\right|\right)$ and $E|M_i| < \infty$. Hence,

$$\begin{aligned}\sup_{\theta\in\Theta,\eta\in\mathcal{H},x\in G}|S_0(\theta,\eta,x) - E(S_0(\theta,\eta,x))| &\leq \max_{1\leq k_1\leq l_{\theta n},1\leq k_2\leq l_{\eta n},1\leq k_3\leq l_n}|S_0(\theta_{k_1},\eta_{k_2},x_{k_3}) - E(S_0(\theta_{k_1},\eta_{k_2},x_{k_3}))| \\ &+ 3C\delta_nh_n^{-2}\frac{1}{n}\sum_{i=1}^{n}M_i = I_1 + I_2.\end{aligned}$$

Note that $P(I_1 \geq \epsilon/2) \leq \sum_{k_1=1}^{l_{\theta n}} \sum_{k_2=1}^{l_{\eta n}} \sum_{k_3=1}^{l_n} P\left(|S_0(\theta_{k_1}, \eta_{k_2}, x_{k_3}) - E(S_0(\theta_{k_1}, \eta_{k_2}, x_{k_3}))|\right)$ and put

$$|S_0(\theta_{k_1}, \eta_{k_2}, x_{k_3}) - E(S_0(\theta_{k_1}, \eta_{k_2}, x_{k_3}))| = \left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{h_n} \frac{\partial}{\partial \eta} log f_\varepsilon(y_i - \eta_{k_2}; \theta_{k_1}) K\left(\frac{x_i - x_{k_3}}{h_n}\right) \right.\right.$$
$$\left.\left. - E\left(\frac{1}{h_n} \frac{\partial}{\partial \eta} log f_\varepsilon(y_i - \eta_{k_2}; \theta_{k_1}) K\left(\frac{x_i - x_{k_3}}{h_n}\right)\right)\right)\right| = \left| \frac{1}{n} \sum_{i=1}^{n} W_{in} \right|$$

For fixed n, $E(W_{in}) = 0$. Furthermore, given that $K(x) < C$, $\frac{\partial}{\partial \eta} log f_\varepsilon(y - \eta; \theta)$ exists almost surely with $E(\frac{\partial}{\partial \eta} log f_\varepsilon(y - \eta; \theta)) < \infty$ by PA2 we have $|W_{in}| < c/h_n$. Furthermore, $\{W_{in}\}_{i \geq 1}$ forms an independent sequence since K and $\frac{\partial}{\partial \eta} log f_\varepsilon(y - \eta; \theta))$ are continuous (measurable) functions of $\{(y_i, x_i)\}_{i \geq 1}$, an independent sequence. By Bernstein's inequality we obtain, $P(I_1 \geq \epsilon/2) \leq 2 l_{\theta n} l_{\eta n} l_n exp\left(\frac{-n h_n (\epsilon^2/4)}{2 h_n \bar{\sigma}^2 + (1/3) C\epsilon}\right)$ where $\bar{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} Var(W_{in})$ and $h_n \bar{\sigma}^2 \to B_{\bar{\sigma}^2} = \int \left(\frac{\partial}{\partial \eta} log f_\varepsilon(y - \eta_{k_2}, \theta_{k_1})\right)^2 f(y, x_{k_3}) dy \int K^2(x) dx$. Since $l_{\theta n} < \frac{r^P}{\delta_n^P}$, $l_{\eta n} < \frac{r}{\delta_n}$, $l_n < \frac{r}{\delta_n}$ we have that $P(I_1 \geq \epsilon/2) < \frac{2r^{P+2}}{\delta_n^{P+2}} exp\left(\frac{-n h_n (\epsilon^2/4)}{2 h_n \bar{\sigma}^2 + (1/3) C\epsilon}\right)$. Now, by Markov's Inequality $P(I_2 \geq \epsilon/2) = P(\frac{1}{n} \sum_{i=1}^{n} M_i \geq \frac{h_n^2 \epsilon}{6\delta_n}) \leq \frac{E|M_i|}{\epsilon h_n^2} 6\delta_n$ and we have

$$P\left(\sup_{\theta \in \Theta, \eta \in \mathcal{H}, x \in G} |S_0(\theta, \eta, x) - E(S_0(\theta, \eta, x))| > \epsilon\right) \leq \frac{2r^{P+2}}{\delta_n^{P+2}} exp\left(\frac{-n h_n (\epsilon^2/4)}{2 h_n \bar{\sigma}^2 + (1/3) C\epsilon}\right) + \frac{E|M_i|}{\epsilon h_n^2} 6\delta_n.$$

The denominator in the first term converges as $n \to \infty$, so we set the two terms on the right hand side of the inequality to be of equal magnitide and solve for δ_n as a function of ϵ. As such, $\delta_n = O\left(\epsilon^{1/(p+3)} h_n^{2/(P+3)} exp(-n h_n \epsilon^2/C)\right)$. Since we need $P\left(\sup_{\theta \in \Theta, \eta \in \mathcal{H}, x \in G} |S_0(\theta, \eta, x) - E(S_0(\theta, \eta, x))| > \epsilon\right) \to 0$, we set $\epsilon = \left(\frac{log(n)}{n h_n}\right)^{1/2} \Delta$ for some constant Δ. It is easy to verify that if $n h_n^{3m/(2-m)} \to \infty$ for $3/4 \leq m \leq 1$ then we have the desired convergence. Hence, if $n h_n^3 \to \infty$ $(m = 1)$ we have that

$$\sup_{\theta \in \Theta, \eta \in \mathcal{H}, x \in G} |S_0(\theta, \eta, x) - E(S_0(\theta, \eta, x))| = O_p((log(n)/n h_n)^{1/2}).$$

Now, let $S_1(\theta, \eta, x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n} \frac{\partial}{\partial \eta} log f_\varepsilon(y_i - \eta; \theta) K\left(\frac{x_i - x}{h_n}\right)\left(\frac{x_i - x}{h_n}\right)$ and since the kernel K has bounded support we immediately obtain, $\sup_{\theta \in \Theta, \eta \in \mathcal{H}, x \in G} |S_1(\theta, \eta, x) - E(S_1(\theta, \eta, x))| = O_p((log(n)/n h_n)^{1/2})$ which completes the proof.

3 Verification of assumptions for Theorem 2

Since assumption A1 has been verified in Martins-Filho and Yao (2011) and assumptions A2 and A3 are unrelated to f_ε we will verify A4, A5, A6 and A7.

A4.1: For all $\theta \in \Theta$ we have that if $\theta \neq \theta_0$ then $f_\varepsilon(y - g_0(x); \theta) \neq f_\varepsilon(y - g_0(x); \theta_0)$ for all (y, x).

Letting $e = y - g_0(x)$ for fixed x. Then, $f_\varepsilon(y - g_0(x);\theta) = f_e(e;\theta) = \frac{2}{s}\phi\left(e/s\right)\Phi\left(-\frac{\lambda}{s}e\right)$. Hence, f_e is a skew normal density with scale parameter s and skewness parameter λ. Identification of (s,λ) follows directly from Azzalini (1985). Identification of (σ_u^2, σ_v^2) follows from the uniqueness of the reparametrization given by $\lambda = \sqrt{\frac{\sigma_u^2}{\sigma_v^2}}$ and $s^2 = \sigma_u^2 + \sigma_v^2$. We also note that from property H and Lemma 2 in Azzalini (1985), all even moments of a skew-normal density with scale parameter $s = 1$ coincide with those of a standard normal, and all odd moments can be obtained by the moment generating function $M(t) = 2exp(t^2/2)\Phi(\frac{\lambda}{\sqrt{1+\lambda^2}}t)$.

A4.2: If $\{\theta_i\}_{i=1,2,\cdots}$ is a sequence in Θ such that $\theta_i \to \theta$ as $i \to \infty$, then

$$logf_\varepsilon(y;\theta_i, g_0(x)) \to logf_\varepsilon(y - g_0(x);\theta) \text{ as } i \to \infty \text{ for all } \theta \in \Theta.$$

This assumption requires continuity of $logf_\varepsilon(y - g_0(x);\theta)$ with respect to θ. This is apparent from the structure of f_ε as a function of θ and continuity of the logarithm function.

A4.3: $E\left(sup_{\theta\in\Theta}\left|logf_\varepsilon(y - g_0(x);\theta)\right|\right) < \infty$.

Since $\theta \in \Theta$ a compact set, $\sigma_u^2, \sigma_v^2 > 0$, then $sup_{\theta\in\Theta}|log(\sigma_u^2+\sigma_v^2)| < C$ and $E\left(sup_{\theta\in\Theta}(y - g_0(x))^2/(\sigma_u^2 + \sigma_v^2)\right) < C(\frac{\pi-2}{\pi}\sigma_u^2 + \sigma_v^2) < C$. Now, let $z = (-\lambda/s)e$ and note that by the mean value theorem

$$\begin{aligned}
|log\Phi(z)| &\leq |log\Phi(0)| + \frac{\phi(z')}{\Phi(z')}|z| \text{ for } z' = \delta z \text{ where } \delta \in (0,1). \\
&\leq |log\Phi(0)| + C(1 + \delta|\lambda/s||e|)|e| \text{ and} \\
sup_{\theta\in\Theta}|log\Phi(z)| &\leq |log\Phi(0)| + C(|e| + e^2) \text{ which gives} \\
E(sup_{\theta\in\Theta}|log\Phi(z)|) &\leq |log\Phi(0)| + CE((|e| + e^2)) < C.
\end{aligned}$$

Combining the last inequality with the fact that $E\left(sup_{\theta\in\Theta}(y - g_0(x))^2/(\sigma_u^2 + \sigma_v^2)\right) < C$ and $sup_{\theta\in\Theta}|log(\sigma_u^2 + \sigma_v^2)| < C$ verifies A4.3.

A4.4: $E\left(sup_{\theta\in\Theta}\left|logf_\varepsilon(y - g_0(x);\theta)\right|\right) < \infty$.For all (y,x), $g \in \mathcal{G}$ and $\theta \in \Theta$, $|logf_\varepsilon(y;\theta, g(x)) - logf_\varepsilon(y - g_0(x);\theta)| \leq b(y,x,\theta)|g(x) - g_0(x)| = b(y,x,\theta)|m(x;\theta,g) - m(x;\theta,g_0)|$ with $b(y,x,\theta) > 0$, and $E\left(sup_{\theta\in\Theta}b(y,x,\theta)\right) < \infty$. We first observe that

$$\begin{aligned}
|logf_\varepsilon(y;\theta,g(x)) - logf_\varepsilon(y - g_0(x);\theta)| &\leq \frac{1}{2s}|(y - g(x))^2 - (y - g_0(x))^2| \\
&+ \left|log\left(\Phi(-\frac{\lambda}{s}(y - g(x)))\right) - \Phi(-\frac{\lambda}{s}(y - g_0(x)))\right| \\
&= I_1 + I_2.
\end{aligned}$$

By Taylor's Theorem in Graves (1927), we have that $I_1 = \frac{1}{s^2}|y - g(x)||g(x) - g_0(x)| = b_1(y, x, \theta)|g(x) - g_0(x)|$. In addition, $sup_{\theta\in\Theta} b_1(y, x, \theta) \leq C|y - g(x)|$ and given that σ_u^2, $\sigma_v^2 > 0$ and Θ is compact we have $E(sup_{\theta\in\Theta} b_1(y, x, \theta)) < \infty$. Again, by Taylor's Theorem $I_2 \leq \frac{\lambda}{s}\int_0^1 \frac{\phi(w_1)}{\Phi(w_1)} dt|g(x) - g_0(x)| = b_2(y, x, \theta)|g(x) - g_0(x)|$ where $w_1 = y - g(x) + t(g(x) - g_0(x))$. Now,

$$\begin{aligned} b_2(y, x, \theta) &\leq \frac{\lambda}{s}\int_0^1 C(1 + w_1)dt \\ &\leq C\frac{\lambda}{s} + C\frac{\lambda^2}{s^2}(|e| + |g(x) - g_0(x)|). \end{aligned}$$

Since $E(|e|) < C$ we have given σ_u^2, $\sigma_v^2 > 0$, Θ compact and the results for I_1 that,

$$|logf_\varepsilon(y - g(x); \theta) - logf_\varepsilon(y - g_0(x); \theta)| \leq (b_1(y, x\theta) + b_2(y, x, \theta))|g(x) - g_0(x)|$$

where $E(sup_{\theta\in\Theta}(b_1(y, x\theta) + b_2(y, x, \theta))) < C$.

A5.1 : For all $\eta = g(x) \in \mathcal{H}$, $logf_\varepsilon(y - \eta; \theta)$ is twice continuously differentiable with respect to θ and $f_\varepsilon(y - \eta; \theta) > 0$ on some open ball $S_{0,\theta} = S(\theta_0, d(\theta_0))$ of θ_0 with $S_{0,\theta} \subset \Theta$ and $d(\theta_0)$ the radius of the ball.

For fixed $\eta = g(x)$ and given the structure of $log\ f_\varepsilon$ as a function of θ, routine partial differentiation with respect to σ_u^2 and σ_v^2 show the existence of $\frac{\partial^3}{\partial(\sigma_u^2)^r \partial(\sigma_v^2)^s} log\ f_\varepsilon(y - \eta; \theta)$ for $r + s \leq 3$ and $r, s = 0, 1, 2, 3$ provided that $\sigma_u^2 + \sigma_v^2 > 0$. That $f_\varepsilon(y - \eta; \theta) > 0$ on an open ball around θ_0 follows from the fact that $\sigma_u^2, \sigma_v^2 > 0$, $exp(x) > 0$ and $\Phi(x) \in (0, 1)$.

A5.2 : $E\left(sup_{\theta\in S_{0,\theta}}\left|\frac{\partial^2}{\partial\theta_j\partial\theta_k} logf_\varepsilon(y - g_0(x); \theta)\right|\right) < \infty$ for $k, j = 1, \cdots, P$

Routine partial differentiation of $logf_\varepsilon(y - g(x); \theta)$ with respect to σ_u^2 gives

$$\begin{aligned} \left|\frac{\partial^2 logf_\varepsilon(y - g(x); \theta)}{\partial(\sigma_u^2)^2}\right| &\leq \frac{1}{2s^2} + \frac{e^2}{s^6} + (1 - \Phi(e\lambda/s))^{-2}(\phi(e\lambda/s)ew)^2 + \frac{1}{2}(1 - \Phi(e\lambda/s))^{-1}\phi(e\lambda/s)|e^3|\frac{|w|}{s^4} \\ &+ (1 - \Phi(e\lambda/s))^{-1}\phi(e\lambda/s)|e|\left|\frac{\partial w}{\partial\sigma_u^2}\right| = I_1 + I_2 + I_3 + I_4 + I_5 \end{aligned}$$

where $w = \frac{1}{2\lambda s^3}$. $I_1 < C$ since $\sigma_u^2, \sigma_v^2 > 0$, $E(sup_{\theta\in\Theta} I_2) < C$ since $1/s_6 < C$ and $E(e^2) < C$. Note that $I_3 < C(1 - \Phi(e\lambda/s))^{-2}\phi(e\lambda/s)^2|e|^2$. Let $I = (1 - \Phi(e\lambda/s))^{-i}\phi(e\lambda/s)^i|e|^j$ and given the convexity of $\phi(x)/\Phi(x)$ and $lim_{x\to\infty}(\phi(x)/\Phi(x)) = 0$, $lim_{x\to-\infty}(\phi(x)/\Phi(x)) = -x$ we have that $|I| \leq C(|e|^j + |e|^{i+j}(\lambda/s)^i)$. In the case of I_3, $i = j = 2$ and since $E(e^3), E(e^4) < C$ we have $E(sup_{\theta\in\Theta} I_3) < C$. Using similar arguments $E(sup_{\theta\in\Theta} I_4) < C$, $E(sup_{\theta\in\Theta} I_5) < C$. The same arguments give the following bounds $E\left(sup_{\theta\in\Theta}\left|\frac{\partial^2 logf_\varepsilon(y - g(x); \theta)}{\partial(\sigma_v^2)^2}\right|\right) < C$ and $E\left(sup_{\theta\in\Theta}\left|\frac{\partial^2 logf_\varepsilon(y - g(x); \theta)}{\partial\sigma_v^2\partial\sigma_u^2}\right|\right) < C$.

A5.3 : $\frac{\partial}{\partial \eta} f_\varepsilon(y - g_0(x);\theta)$ is continuously differentiable in $S_{0,\theta}$. Furthermore,

$$\int \sup_{\theta \in S_{0,\theta}} ||\frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x);\theta)||_E f_\varepsilon(y - g_0(x);\theta_0)dy < \infty$$

and

$$E\left(\sup_{\theta \in S_{0,\theta}} ||\frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x);\theta)||_E |g_0^{(2)}(x)|\right) < \infty.$$

Taking partial derivatives we have $\frac{\partial}{\partial\eta} log\ f_\varepsilon(y-\eta;\theta) = \frac{e}{s^2} + \left(1 - \Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\frac{\lambda}{s}$. For fixed $\eta = g(x)$ and given the structure of $\frac{\partial}{\partial\eta} log f_\varepsilon$ as a function of θ, routine partial differentiation with respect to σ_u^2 and σ_v^2 show the existence of $\frac{\partial^3}{\partial\eta\partial(\sigma_u^2)^r(\sigma_v^2)^s} log\ f_\varepsilon(y;\theta,\eta)$ for $r+s \le 2$ and $r,s = 0,1,2$ provided that $\sigma_u^2 + \sigma_v^2 > 0$. For $x \in \Re^2$, $||x||_E \le x_1^2 + x_2^2$. Hence, we first show that $E\left(sup_{\theta\in\Theta}\left|\frac{\partial^2}{\partial\eta\partial\sigma_u^2} log\ f_\varepsilon(y-\eta;\theta)\right| |x\right) < C$. Using the structure of $log\ f_\varepsilon(y-\eta;\theta)$ and taking partial derivatives shows that the components of $\frac{\partial^2}{\partial\eta\partial\sigma_u^2} log\ f_\varepsilon(y-\eta;\theta)$ take the form $C\left(1-\Phi(e\lambda/s)\right)^{-i}\phi(e\lambda/s)^i|e|^j \le C(|e|^j + |e|^{i+j}C)$ for $(i,j) = (0,1),(2,1),(1,2),$ or $(1,0)$. Since all finite order moments of skew normal densities exist (Azzalini(1985)), it follows that $E\left(sup_{\theta\in\Theta}\left|\frac{\partial^2}{\partial\eta\partial\sigma_u^2} log\ f_\varepsilon(y-\eta;\theta)\right|\right) < C$. The same argument provides

$$E\left(sup_{\theta\in\Theta}\left|\frac{\partial^2}{\partial\eta\partial\sigma_v^2} log\ f_\varepsilon(y-\eta;\theta)\right|\right) < C.$$

Finally, with the above given bounds $E\left(\sup_{\theta \in S_{0,\theta}} ||\frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x);\theta)||_E |g_0^{(2)}(x)|\right) < \infty$. provided that $|g_0^2(x)| < \infty$.

A5.4 : The matrix

$$\begin{aligned}\bar{H} &= E\left(\frac{\partial^2}{\partial\theta\partial\theta'} log f_\varepsilon(y - g_0(x);\theta_0)\right) + \frac{\partial}{\partial\theta}\gamma(\theta_0) E\left(\frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x);\theta_0)\right)' \\ &+ E\left(\frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x);\theta_0)\right)\frac{\partial}{\partial\theta}\gamma(\theta_0)'\end{aligned}$$

exists and is nonsingular.

Given the structure of f_ε as a function of θ and the fact that $\gamma(\theta) = \sqrt{\frac{2}{\pi}\sigma_u^2}$, we have by routine partial differentiation that $\bar{H} = \begin{pmatrix} \bar{H}_{11} & \bar{H}_{12} \\ \bar{H}_{21} & \bar{H}_{22}\end{pmatrix}$ where $\bar{H}_{11} = -\frac{1}{2s^4} - w^2 I_1 + \sqrt{\frac{2}{\pi\sigma_u^2}} C_1$, $\bar{H}_{12} = -\frac{1}{2s^4} - ww_1 I_1 + \sqrt{\frac{2}{\pi\sigma_u^2}} C_2$, and $\bar{H}_{22} = -\frac{1}{2s^4} - w_1^2 I_1$. Here, $w = \frac{1}{2\lambda s^3}$, $w_1 = -\frac{1}{2\lambda}\frac{(\sigma_u^2 + 2\sigma_v^2)\sigma_u^2}{(\sigma_v^2)^2 s^3}$, $I = \int \frac{e\frac{\sqrt{2}}{s\pi^{3/2}}}{1-erf(e\frac{\lambda}{s\sqrt{2}})} exp(-e^2(\frac{\lambda^2}{s^2} + \frac{1}{2s^2}))de$, $I_1 = \int \frac{e^2\frac{\sqrt{2}}{s\pi^{3/2}}}{1-erf(e\frac{\lambda}{s\sqrt{2}})} exp(-e^2(\frac{\lambda^2}{s^2} + \frac{1}{2s^2}))de$, $C_1 = \frac{\gamma(\theta)}{s^4} + \frac{\lambda}{s} wI - (\frac{\lambda}{s})^2 w\sqrt{\frac{2}{\pi(\lambda^2+1)}}\frac{s^2}{\lambda^2+1} + w\sqrt{\frac{2}{\pi(\lambda^2+1)}}$, $C_2 = \frac{\gamma(\theta)}{s^4} + \frac{\lambda}{s} w_1 I - (\frac{\lambda}{s})^2 w_1\sqrt{\frac{2}{\pi(\lambda^2+1)}}\frac{s^2}{\lambda^2+1} + w_1\sqrt{\frac{2}{\pi(\lambda^2+1)}}$ where erf is the Gaussian error function. The determinant of $\bar{H}$ denoted by $det(\bar{H}) = \bar{H}_{11}\bar{H}_{22} - \bar{H}_{21}\bar{H}_{12} = \frac{I_1}{2s^4}(w_1 - w)^2 + \frac{1}{\sqrt{2}s^4}\sqrt{\frac{1}{\pi\sigma_u^2}}(C_2 - C_1) + ww_1^2 I_1^2(w_1 - w) + w_1 I_1\sqrt{\frac{2}{\pi\sigma_u^2}}(wC_2 - w_1 C_1) - \frac{C_2^2}{2\pi\sigma_u^2} \neq 0$ given the structure of C_1, C_2 and $w \neq w_1$, $C_1 \neq C_2$.

A5.5 : Let $\eta_0 = g_0(x)$ for any $x \in G$ and denote by $S_{0,\eta} = S(\eta_0, d(\eta_0))$. $f_\varepsilon(y-\eta;\theta)$ is continuously differentiable on $S_{0,\eta}$, an open interval of $\mathcal{H}$, $E\left(\sup_{\eta \in S_{0,\eta}} \left|\frac{\partial}{\partial \eta} f_\varepsilon(y-\eta;\theta)\right|\right) < \infty$ and for all $x \in G$, $\frac{\partial}{\partial \eta} E\left(f_\varepsilon(y-g_0(x);\theta_0)|x\right) = 0$.

From A5.3 we have $\frac{\partial}{\partial \eta} log\ f_\varepsilon(y-\eta;\theta) = \frac{e}{s^2} + \left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\frac{\lambda}{s}$ and taking a second partial derivative with respect to η gives $\frac{\partial^2}{\partial \eta^2} log\ f_\varepsilon(y-\eta;\theta) = -\frac{1}{s^2} - \left(1-\Phi(e\frac{\lambda}{s})\right)^{-2}\phi(e\frac{\lambda}{s})^2(\frac{\lambda}{s})^2 + \left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})(\frac{\lambda}{s})^3 e$ establishing continuous differentiability of $log\ f_\varepsilon(y-\eta;\theta)$ with respect to the argument η. Now,

$$E\left(sup_{\eta \in S_{0,\eta}} |f_\varepsilon(y-\eta;\theta)|\right) \leq \frac{1}{s^2} E(sup_{\eta \in S_{0,\eta}}|y-\eta|) + (\lambda/s) E\left(sup_{\eta \in S_{0,\eta}} \left|\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\right|\right).$$

The first term on the right hand side of the inequality is bounded since $\eta \in \mathcal{H}$ a compact set in $\Re$ and $E(|y|) \leq E(|g_0(x)|) + E(|e|) < C$. For the second term, note that

$$E\left(sup_{\eta \in S_{0,\eta}} \left|\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\right|\right) \leq C\left(1 + (\lambda/s) E\left(sup_{\eta \in S_{0,\eta}}|y-\eta|\right)\right) < C.$$

Lastly, $\frac{\partial}{\partial \eta} E\left(f_\varepsilon(y-g_0(x);\theta_0)\right) = E\left(\frac{\partial}{\partial \eta} f_\varepsilon(y-g_0(x);\theta_0)\right)$ and

$$\begin{aligned} E\left(\frac{\partial}{\partial \eta} f_\varepsilon(y-g_0(x);\theta_0)\right) &= -\frac{1}{s^2}\sqrt{\frac{2}{\pi}\sigma_u^2} + \frac{\lambda}{s} E\left(\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})|x\right) \\ &= -\frac{1}{s^2}\sqrt{\frac{2}{\pi}\sigma_u^2} + \frac{\lambda}{s}\sqrt{\frac{2}{\pi(\lambda^2+1)}} = 0 \end{aligned}$$

given the definition of s and λ.

A5.6 : For all $\theta \in \Theta$, $\frac{\partial}{\partial \eta} f_\varepsilon(y-g_0(x);\theta)$ is continuous at x,

$$E\left(\sup_{\theta \in \Theta}\left(\frac{\partial}{\partial \eta} f_\varepsilon(y-\eta;\theta)\right)^2\right) < \infty \text{ and } E\left(\sup_{\theta \in \Theta}\left(\frac{\partial}{\partial \eta} f_\varepsilon(y-\eta;\theta)\right)^2 y^2\right) < \infty.$$

Continuity of $\frac{\partial}{\partial \eta} f_\varepsilon(y-g_0(x);\theta)$ at x follows directly from the existence $\frac{\partial^2}{\partial \eta \partial x} f_\varepsilon(y-g_0(x);\theta)$, which follows given that $\sigma_u^2, \sigma_v^2 > 0$ and differentiability of $g_0(x)$ with respect to x. Now,

$$\begin{aligned} E\left(\sup_{\theta \in \Theta}\left(\frac{\partial}{\partial \eta} f_\varepsilon(y-\eta;\theta)\right)^2\right) &= E\left(\sup_{\theta \in \Theta}\left(\frac{1}{s^2} + \left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})(\frac{\lambda}{s})\right)^2\right) \\ &\leq CE(e^2) + CE\left(\sup_{\theta \in \Theta}\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})|e|\right) \\ &+ CE\left(\sup_{\theta \in \Theta}\left(1-\Phi(e\frac{\lambda}{s})\right)^{-2}\phi(e\frac{\lambda}{s})^2\right) < C \end{aligned}$$

given that $\sigma_u^2, \sigma_v^2 > 0$ and the arguments made to verify A5.2. Given that $E(e^2y^2) \leq C(E(e^2 g_0(x)^2) + E(e^4)) < C(E(e^2) + E(e^4)) < C$ from Azzalini(1985) and the fact that $g_0(x) \in \mathcal{H}$ a compact set, we immediately have $E\left(\sup_{\theta\in\Theta}\left(\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)\right)^2 y^2\right)<\infty$ using the same arguments that gave the boundedness of $E\left(\sup_{\theta\in\Theta}\left(\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)\right)^2\right)$.

A6.1 : $logf_\varepsilon(y-g_0(x);\theta)$ is twice Fréchet differentiable at g_0 with increment $h(x)=g(x)-g_0(x)$ and denote the Fréchet derivatives of order $i=1,2$ at g_0 by $\frac{d_F^i}{dg^i}logf_\varepsilon(y-g_0(x);\theta)$.

Let $T(g_0)=f_\varepsilon(y-g_0(x);\theta)$. It follows directly from the definition of the first order Gateaux differential with increment $h=g-g_0$ that $\delta_G T(g_0,h)=\frac{d}{d\alpha}T(g_0+\alpha(g-g_0))\,|_{\alpha=0}=\frac{e}{s^2}(g(x)-g_0(x))+\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\frac{\lambda}{s}(g(x)-g_0(x))$. We will establish that $\delta_G T(g_0,h)=\delta_F T(g_0,h)$. First, observe that

$$\begin{aligned}|T(g)-T(g_0)-\delta_G T(g_0,h)| &\leq \left|-\frac{1}{2s^2}(y-g(x))^2+\frac{1}{2s^2}(y-g_0(x))^2-\frac{e}{s^2}(g(x)-g_0(x))\right|\\ &+ \left|log\left(1-\Phi((y-g(x))\frac{\lambda}{s})\right)-log\left(1-\Phi((y-g_0(x))\frac{\lambda}{s})\right)\right.\\ &- \left.log\left(1-\Phi((y-g(x))\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\frac{\lambda}{s}(g(x)-g_0(x))\right|\\ &= I_1+I_2.\end{aligned}$$

Let $T_1(g)=-\frac{1}{2s^2}(y-g(x))^2$, then $\delta_G T_1(g_0,h)=\frac{e}{s^2}(g(x)-g_0(x))$ and $\delta_G^2 T_1(g_0,h)=-\frac{1}{s^2}(g(x)-g_0(x))^2$. By Taylor's Theorem $T_1(g)=T_1(g_0)+\delta_G T_1(g_0,h)+\int_0^1\delta_G^2 T_1(g_0,h)(1-t)dt$, therefore $I_1\leq\frac{1}{2s^2}||g-g_0||^2$. Setting $T_2(g)=log\left(1-\Phi((y-g(x))\frac{\lambda}{s})\right)$ and taking Gateaux differentials of order 2 and using Taylor's Theorem as in the case for T_1 gives

$$\begin{aligned}I_2 &= \left|\int_0^1(1-t)\left(-\left(1-\Phi(e^*\frac{\lambda}{s})\right)^{-2}\phi(e^*\frac{\lambda}{s})^2(\frac{\lambda}{s})^2\right.\right.\\ &+ \left.\left.\left(1-\Phi(e^*\frac{\lambda}{s})\right)^{-1}\phi(e^*\frac{\lambda}{s})^2e^*(\frac{\lambda}{s})^3\right)dt\right|(g(x)-g_0(x))\\ &\leq \int_0^1(1-t)\left(C(1+|e^*|\frac{\lambda}{s})^2+C(1+|e^*|\frac{\lambda}{s})|e^*|\right)dt||g(x)-g_0(x)||^2\end{aligned}$$

where $e^*=y-(g_0+th)(x)$. Since $|e^*|\leq|e|+t|g(x)-g_0(x)|$, $E(e^2)<C$ and $g\in\mathcal{G}$ we have that $I_2=O_p(1)||g-g_0||^2$. Since $I_1=O(1)||g-g_0||^2$ we have combining the two orders that $\delta_G T(g_0,h)=\delta_F T(g_0,h)=\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta)(g(x)-g_0(x))$ where $\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta)=\frac{e}{s^2}+\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})\frac{\lambda}{s}$ is the first order Fréchet derivative. The same arguments show the existence and equivalence of second order Gateaux and Fréchet differentials. In this case,

$$\begin{aligned}\delta_F^2 T(g_0,h) &= -\frac{1}{s^2}(g(x)-g_0(x))^2+\left(-\left(1-\Phi(e\frac{\lambda}{s})\right)^{-2}\phi(e\frac{\lambda}{s})^2(\frac{\lambda}{s})^2\right.\\ &+ \left.\left(1-\Phi(e\frac{\lambda}{s})\right)^{-1}\phi(e\frac{\lambda}{s})^2e(\frac{\lambda}{s})^3\right)(g(x)-g_0(x)).\end{aligned}$$

A6.2 : $\frac{d_F}{dg} log f_\varepsilon(y - g_0(x); \theta_0)$ is continuous at every $x \in G$.

From A6.1 $\frac{d_F}{dg} log f_\varepsilon(y - g_0(x); \theta_0) = \frac{e}{s^2} + \left(1 - \Phi(e\frac{\lambda}{s})\right)^{-1} \phi(e\frac{\lambda}{s})\frac{\lambda}{s}$ which is clearly continuous in x given differentiability of $g(x)$ and the structure of $log f_\varepsilon(y - g_0(x); \theta_0)$ as a function of x.

A6.3 : The matrix

$$
\begin{aligned}
\sigma_F^2 \;=\;& E\left(\left(\frac{\partial}{\partial\theta} f_\varepsilon(y_i - g_0(x_i); \theta_0) + (y_i - m(x_i; \theta_0, g_0)) \int \frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x_i); \theta_0) f_\varepsilon(y - g_0(x_i); \theta_0) dy\right)\right. \\
\times& \left.\left(\frac{\partial}{\partial\theta} f_\varepsilon(y_i - g_0(x_i); \theta_0) + (y_i - m(x_i; \theta_0, g_0)) \int \frac{\partial^2}{\partial\theta\partial\eta} f_\varepsilon(y - g_0(x_i); \theta_0) f_\varepsilon(y - g_0(x_i); \theta_0) dy\right)'\right).
\end{aligned}
$$

exists and is positive definite.

We start by noting that routine partial differentiation with respect to σ_u^2 gives, $\frac{\partial f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2} = -\frac{1}{2s^2} + \frac{e^2}{2s^4} - \Phi\left(-e\frac{\lambda}{s}\right)^{-1} \phi(-e\frac{\lambda}{s})ew$ and consequently $E\left(\left(\frac{\partial f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2}\right)^2\right) = \frac{1}{4s^4} + \frac{1}{4s^6} E(e^4) + w^2 E(\Phi\left(-e\frac{\lambda}{s}\right)^{-2} \phi(-e\frac{\lambda}{s})^2$ $e^2) < C$ by the arguments used in verifying A5.2. Now,

$$
\begin{aligned}
E\left(\frac{\partial^2 f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2\partial\eta}\right) \;=\;& -e\frac{1}{4s^4} - \Phi\left(-e\frac{\lambda}{s}\right)^{-2} \phi(-e\frac{\lambda}{s})^2(-\frac{\lambda}{s})ew \\
-\;& \Phi\left(-e\frac{\lambda}{s}\right)^{-1} \phi(-e\frac{\lambda}{s})(\frac{\lambda}{s})e^2w + \Phi\left(-e\frac{\lambda}{s}\right)^{-1} \phi(-e\frac{\lambda}{s})(\frac{\lambda}{s})w
\end{aligned}
$$

and by the Cauchy-Schwarz inequality

$$
\begin{aligned}
E\left(\frac{\partial f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2} eE\left(\frac{\partial^2 f_\varepsilon(y - g_0(x); \theta))}{\partial\sigma_u^2\partial\eta}|x\right)\right) \;\le\;& \left(E\left(e^2\left(\frac{\partial f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2}\right)^2\right)\right)^{1/2} \\
\times& \left(E\left(\left(\frac{\partial^2 f_\varepsilon(y - g_0(x); \theta))}{\partial\sigma_u^2\partial\eta}\right)^2\right)\right)^{1/2}.
\end{aligned}
$$

We observe that

$$
\begin{aligned}
E\left(e^2\left(\frac{\partial f_\varepsilon(y - g_0(x); \theta)}{\partial\sigma_u^2}\right)^2\right) \;=\;& E\left(\frac{1}{4s^4} + \frac{e^4}{4s^8} + \Phi\left(-e\frac{\lambda}{s}\right)^{-2} \phi(-e\frac{\lambda}{s})^2e^2w^2 - \frac{e^2}{2s^6}\right. \\
+\;& \left.\frac{1}{s^2}\Phi\left(-e\frac{\lambda}{s}\right)^{-1} \phi(-e\frac{\lambda}{s})ew - \frac{1}{s^4}\Phi\left(-e\frac{\lambda}{s}\right)^{-1} \phi(-e\frac{\lambda}{s})e^3w\right) \\
\times\;& (e^2 + 2e\gamma(\theta) + \gamma(\theta)^2)
\end{aligned}
$$

and note that the leading expectation takes the form

$$
E\left(C(s^2)^{-i}\Phi\left(-e\frac{\lambda}{s}\right)^{-j} \phi(-e\frac{\lambda}{s})^j e^q w^h\right)
$$

where $(s^2)^{-i}w^h < C$ for all i, h nonnegative and (j, q) is either (0,2), (0,1), (0,0), (0,6), (0,5), (0,4), (2,4), (2,3), (0,2), (2,2), (0,3), (0,2), (1,3), (1,2), (1,1), (1,5), (1,4), (1,3). In all cases, given the existence of all moments for a skew normal we have $E\left(e^2\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2}\right)^2\right) < C$. Similarly, $E\left(\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}\right)^2\right) < C$ and consequently $E\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2} eE\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}|x\right)\right) < C$. By the definition of conditional expectation,

$$\begin{aligned}
E\left(e^2\left(E\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}|x\right)\right)^2\right) &= E(e^2)E\left(E\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}|x\right)^2\right)\\
&\leq E\left(e^2+2e\gamma(\theta)+\gamma(\theta)^2\right)E\left(\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}\right)^2\right)\\
&\leq C
\end{aligned}$$

given the existence of all moments for a skew normal. Hence, we conclude that

$$E\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2} + eE\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}|x\right)\right)^2 < C.$$

Repeating the same sequence of arguments for σ_v^2 gives

$$E\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_v^2} + eE\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_v^2\partial\eta}|x\right)\right)^2 < C.$$

Lastly, by the Cauchy-Schwarz inequality

$$\begin{aligned}
&E\left(\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2} + eE\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_u^2\partial\eta}|x\right)\right)\times\right.\\
&\left.\left(\frac{\partial f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_v^2} + eE\left(\frac{\partial^2 f_\varepsilon(y-g_0(x);\theta)}{\partial\sigma_v^2\partial\eta}|x\right)\right)\right) \leq C
\end{aligned}$$

by the bound obtained above. In addition, σ_F^2 is positive definite.

A7.1 : $\sup_{\theta\in S_{0,\theta}}\left|E\left(\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)\right| < \infty$

Using the same steps taken to verify A5.3 we have that

$$\begin{aligned}
\sup_{\theta\in S_{0,\theta}}\left|E\left(\frac{\partial}{\partial\eta}f_\varepsilon(y;\theta,g_0(x_i))|x_i\right)\right| &\leq \sup_{\theta\in S_{0,\theta}}\frac{1}{s^2}E(|e|) + \sup_{\theta\in S_{0,\theta}}\frac{\lambda}{s}E\left(\left|1-\Phi\left(e\frac{\lambda}{s}\right)^{-1}\phi(e\frac{\lambda}{s})\right|\right)\\
&= I_1 + I_2 < C
\end{aligned}$$

by the same arguments we have used above.

A7.2 : $\frac{\partial^2}{\partial\theta_i\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)$ is continuously differentiable in $S_{0,\theta}$ and

$$E\left(\sup_{\theta\in S_{0,\theta}}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)\right||x_i\right) < \infty$$

for all $x_i \in G$ almost surely.

We verify that $\frac{\partial^2}{\partial\theta_i\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)$ is continuously differentiable in $S_{0,\theta}$ by taking routine partial derivatives of $\frac{\partial^2}{\partial\theta_i\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)$ and establishing that $\frac{\partial^4}{\partial\theta_i^k\partial\theta_j^l\partial\eta}f_\varepsilon(y-g_0(x_i);\theta) < C$ for all $k,l = 0,1,2,3$ and $k+l \leq 3$. The fact that

$$E\left(\sup_{\theta\in S_{0,\theta}}\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)\right||x_i\right) < \infty$$

follows directly from the arguments used to verify A5.3.

A7.3 : $\sup_{\theta\in S_{0,\theta}}\left|E\left(\frac{\partial^2}{\partial\theta_i\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)\right| < \infty.$

$$\begin{aligned}\sup_{\theta\in S_{0,\theta}}\left|E\left(\frac{\partial^2}{\partial\sigma_u^2\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)\right| &\leq \sup_{\theta\in S_{0,\theta}}\frac{1}{s^4}|E(e)| + \sup_{\theta\in S_{0,\theta}}|w\frac{\lambda}{s}|\left|E\left(1-\Phi\left(e\frac{\lambda}{s}\right)^{-2}\phi(e\frac{\lambda}{s})^2e\right)\right| \\ &+ \sup_{\theta\in S_{0,\theta}}|\frac{\lambda}{2s^5}|E\left(1-\Phi\left(e\frac{\lambda}{s}\right)^{-1}\phi(e\frac{\lambda}{s})\right) \\ &+ \sup_{\theta\in S_{0,\theta}}|w|E\left(1-\Phi\left(e\frac{\lambda}{s}\right)^{-1}\phi(e\frac{\lambda}{s})\right) < C\end{aligned}$$

by the arguments repeatedly used above. In a similar fashion we have $\sup_{\theta\in S_{0,\theta}}\left|E\left(\frac{\partial^2}{\partial\sigma_v^2\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)\right| < C$.

A7.4 : $E\left(\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)$ is continuous in $S_{0,\theta}$ almost surely.

In verifying A7.2 we obtained the existence of $\frac{\partial^4}{\partial\theta_i^k\partial\theta_j^l\partial\eta}f_\varepsilon(y-g_0(x_i);\theta) < C$ for all $k,l = 0,1,2,3$ and $k+l\leq 3$. Hence following the argument in A7.3 we verify that $E\left(\frac{\partial^4}{\partial\theta_i^k\partial\theta_j^l\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)\right) < C$ given the existence of moments of a skew normal. This implies continuity of $E\left(\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)|x_i\right)$ in $S_{0,\theta}$.

A7.5 : Follows from Azzalini (1985).

A7.6 : $\frac{\partial^2}{\partial\theta_i\partial\theta_j}E\left(g_0^{(2)}(x_i)\frac{\partial}{\partial\eta}f_\varepsilon(y-g_0(x_i);\theta)\right)$ is continuous in $S_{0,\theta}$ almost surely.

This follows directly from A7.2 and A7.4.

4 Verification of assumptions PA1, PA2, PB, PC and PD for Lemma 2

Before we verify that the assumptions in Lemma 2 are met by the density in (1) we note that in the statement of Lemma 2 $\alpha_\theta(x)$ is assumed to be the unique maximizer of $E(f_\varepsilon(y-\eta;\theta))$ for fixed x and θ and satisfies $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\alpha_\theta(x);\theta)|x) = 0$. This must be verified for the density in (1). For this purpose, we will show that for given $\theta \in \Theta$ and $x \in G$, there exists a unique maximizer of $E(logf_\varepsilon(y-\eta;\theta)|x)$ with respect to η given by $\alpha_\theta(x)$ such that

$$(i)\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\alpha_\theta(x);\theta)|x)=0, \quad (ii)\inf_{\theta\in\Theta,x\in G} -\frac{\partial^2}{\partial\eta^2}E(logf_\varepsilon(y-\alpha_\theta(x);\theta)|x)>0.$$

(i) for fixed θ, $logf_\varepsilon(y-\eta;\theta)=\frac{1}{2}log(\frac{2}{\pi})-\frac{1}{2}log(\sigma_u^2+\sigma_v^2)-\frac{1}{2}\frac{e^2}{\sigma_u^2+\sigma_v^2}+log(\Phi(-\frac{\lambda e}{s}))$, and only the last two terms depend on η. If each is a strictly concave function e then $logf_\varepsilon(y-\eta;\theta)$ is a strictly concave function of η since η enters e linearly. Note that $\frac{e^2}{2(\sigma_u^2+\sigma_v^2)}$ is strictly concave in e, hence we only need to show that $log(\Phi(-\frac{\lambda e}{s}))$ is a strictly concave function of e. Let $u=-\frac{\lambda e}{s}$ and note that for fixed $-\frac{\lambda}{s}$, e enters u linearly, i.e., $\frac{\partial u}{\partial e}=-\frac{\lambda}{s}$ does not depend on e. Consequently, if we show $log(\Phi(-\frac{\lambda e}{s}))$ is strictly concave in u, then it is also strictly concave in e since $\frac{\partial^2}{\partial u^2}log(\Phi(u))=\frac{1}{\Phi^2(u)}[\phi'(u)\Phi(u)-\phi^2(u)]$. Inspired by Burridge (1981), we consider three cases:

(a) if $\phi'(u)<0$, then it is clear that $\frac{\partial^2}{\partial u^2}log(\Phi(u))<0$;

(b) if $\phi'(u)=0$, then $\frac{\partial^2}{\partial u^2}log(\Phi(u))<0$;

(c) if $\phi'(u)>0$, then we show that $\frac{\partial^2}{\partial u^2}log(\Phi(u))<0$.

For $x\neq y$, since $\phi(x)$ is log-concave, we have $log(\phi(x))<log(\phi(y))+h(y)(x-y)$, where $h(y)=\frac{\phi'(y)}{\phi(y)}$. So $\phi(x)<\phi(y)exp(h(y)(x-y))$. Integrating with respect to x from $-\infty$ to z, we obtain

$$\Phi(z)<\int_{-\infty}^{z}\phi(y)exp(h(y)(x-y))dx=\frac{\phi(y)}{h(y)}[exp(h(y)(z-y))-exp(h(y)(x-y))|_{x=-\infty}]$$

If we let $y=u$, and $z=u$, then $h(y)=g(u)=\frac{\phi'(u)}{\phi(u)}>0$, so $\Phi(u)<\frac{\phi(u)}{h(u)}=\frac{\phi^2(u)}{\phi'(u)}$ hence $\frac{\partial^2}{\partial u^2}log(\Phi(u))<0$. In all, $logf_\varepsilon(y-\eta;\theta)$ is strictly concave in η, so $E(logf_\varepsilon(y-\eta;\theta))$ is strictly concave in η. Given PA2, we have $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\eta;\theta)|x)=E(\frac{\partial}{\partial\eta}logf_\varepsilon(y-\eta;\theta)|x)=\frac{E(y-\eta)}{\sigma_u^2+\sigma_v^2}+E[\Phi(-\frac{\lambda e}{s})]^{-1}\phi(-\frac{\lambda e}{s})\frac{\lambda}{s}$. Since the second term is greater than zero, and the first term could be less than zero for large value of η, $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\eta;\theta)|x)$ could take positive or negative values. Since $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\eta;\theta)|x)$ is continuous and decreasing with η, there exists a unique value η^* such that $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\eta^*;\theta)|x)=0$. Since $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\eta;\theta)|x)$ is continuous with respect to x,θ,η in a neighborhood of x,θ, and η^* and $|\frac{\partial^2}{\partial\eta^2}E(logf_\varepsilon(y-\eta^*;\theta)|x)|\neq 0$ (see (ii) below). By the Implicit Function Theorem, there exists a neighborhood of x,θ and η^* such that for every x,θ in it, there is a unique $\alpha_\theta(x)$ in the neighborhood of η^* such that $\frac{\partial}{\partial\eta}E(logf_\varepsilon(y-\alpha_\theta(x);\theta)|x)=0$.

To verify (ii) note that as in (i) $E(logf_\varepsilon(y-\eta;\theta)|x)$ is strictly concave in η, $logf_\varepsilon(y-\eta;\theta)$ does not depend on x and $\sigma_u^2,\sigma_v^2>0$, so $\inf_{\theta\in\Theta,x\in G}-\frac{\partial^2}{\partial\eta^2}E(logf_\varepsilon(y-\alpha_\theta(x);\theta)|x)>\inf_{\theta\in\Theta,x\in G,\eta\in\mathcal{H}}-\frac{\partial^2}{\partial\eta^2}E(logf_\varepsilon(y-\eta;\theta)|x)>0$.

We now turn to the verification of PA1, PA2, PB, PC and PD.

PA1: 1. For fixed (but arbitrary) $\theta' \in \Theta$ and $\eta' \in \mathcal{H}$ and for all $\theta \in \Theta$ and $\eta \in \mathcal{H}$, let $\rho(\theta,\eta) = \int log f_\varepsilon(y-\eta;\theta) f_\varepsilon(y-\eta';\theta')dy$. If $\theta \neq \theta'$ then $\rho(\theta,\eta) < \rho(\theta',\eta')$; 2. $\tilde{I}_\theta(\theta,\eta) > 0$ for all $\theta \in \Theta$ and $\eta \in \mathcal{H}$ where

$$\begin{aligned}\tilde{I}_\theta(\theta,\eta) &= E\left(\frac{\partial}{\partial\theta} f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta} f_\varepsilon(y-\eta;\theta)'\right) \\ &- \frac{E(\frac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta} f_\varepsilon(y-\eta;\theta))E(\frac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta'} f_\varepsilon(y-\eta;\theta))}{E((\frac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta)^2)}\end{aligned}$$

Since $f_\varepsilon(y-\eta;\theta)$ is a density function, $\int f_\varepsilon(y-\eta;\theta)dy = 1$. By the strict version of Jensen's inequality, for any nonconstant positive random variable a, $E(-ln(a)) > -lnE(a)$. If $\theta \neq \theta'$ implies $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$, then for $a = \frac{f_\varepsilon(y-\eta;\theta)}{f_\varepsilon(y-\eta';\theta')}$ and $\theta \neq \theta'$,

$$\begin{aligned}&\int(-log(a))f_\varepsilon(y-\eta';\theta'dy) = \int log(f_\varepsilon(y-\eta';\theta'))f_\varepsilon(y-\eta';\theta')dy - \int log(f_\varepsilon(y-\eta;\theta))f_\varepsilon(y-\eta';\theta')dy \\ &> -log\int a f_\varepsilon(y-\eta';\theta')dy = -log\int \frac{f_\varepsilon(y-\eta;\theta)}{f_\varepsilon(y-\eta';\theta')} f_\varepsilon(y-\eta';\theta')dy = 0.\end{aligned}$$

So we only need to show if $\theta \neq \theta'$ then $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$.

We consider two cases: (i) If $\eta = \eta'$, and if further $E(y-\eta)^2 > 0$, then as in the verification of A4.1, we have $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$.

(ii) If $\eta \neq \eta'$ and if $\frac{\lambda}{s} \neq \frac{\lambda'}{s'}$, then $(y-\eta)\frac{\lambda}{s} \neq (y-\eta')\frac{\lambda'}{s'}$ and since $\Phi(x)$ is strictly increasing, $1-\Phi((y-\eta)\frac{\lambda}{s}) \neq 1-\Phi((y-\eta')\frac{\lambda'}{s'})$. We consider three subcases:

(1) If $\sigma_v^2 = \sigma_v^{2'}$, then $\sigma_u^2 = \sigma_u^{2'}$ since $\theta \neq \theta'$. We also have $1-\Phi((y-\eta)\frac{\lambda}{s}) \neq 1-\Phi((y-\eta')\frac{\lambda'}{s'})$, $\frac{2}{\sqrt{\sigma_u^2+\sigma_v^2}}\phi\left(\frac{y-\eta}{\sqrt{\sigma_u^2+\sigma_v^2}}\right) \neq \frac{2}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\phi\left(\frac{y-\eta'}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\right)$. So $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$ except on the set

$$\begin{aligned}A &= \left\{\frac{2}{\sqrt{\sigma_u^2+\sigma_v^2}}\phi\left(\frac{y-\eta}{\sqrt{\sigma_u^2+\sigma_v^2}}\right)\left(1-\Phi\left(\frac{\sqrt{\sigma_u^2/\sigma_v^2}}{\sqrt{\sigma_u^2+\sigma_v^2}}(y-\eta)\right)\right) = \frac{2}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\phi\left(\frac{y-\eta'}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\right)\left(1-\Phi\left(\frac{\sqrt{\sigma_u^{2'}/\sigma_v^{2'}}}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}(y-\eta')\right)\right)\right\} \\ &= \left\{-\frac{1}{2}\frac{(y-\eta)^2}{\sigma_u^2+\sigma_v^2} + log[(\sigma_u^2+\sigma_v^2)^{-\frac{1}{2}}(1-\Phi(\frac{\lambda}{s}(y-\eta)))] = -\frac{1}{2}\frac{(y-\eta')^2}{\sigma_u^{2'}+\sigma_v^{2'}} + log[(\sigma_u^{2'}+\sigma_v^{2'})^{-\frac{1}{2}}(1-\Phi(\frac{\lambda'}{s'}(y-\eta')))]\right\}\end{aligned}$$

Since the two quadratic functions of y are not equal, $log[(\sigma_u^2+\sigma_v^2)^{-\frac{1}{2}}(1-\Phi(\frac{\lambda}{s}(y-\eta)))]$ is a monotonic function of y, set A contains at most finite discrete points which make the equality to be true and $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$ almost everywhere.

(2) If $\sigma_v^2 \neq \sigma_v^{2'}$, and $\frac{\sigma_u^2}{\sigma_v^2} = \frac{\sigma_u^{2'}}{\sigma_v^{2'}}$, then $\sigma_u^2 \neq \sigma_u^{2'}$. We consider two subcases:

(a) If $\sigma_u^2+\sigma_v^2 = \sigma_u^{2'}+\sigma_v^{2'}$, then $\frac{2}{\sqrt{\sigma_u^2+\sigma_v^2}}\phi\left(\frac{y-\eta}{\sqrt{\sigma_u^2+\sigma_v^2}}\right) \neq \frac{2}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\phi\left(\frac{y-\eta'}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\right)$. Since $\frac{\lambda}{s} \neq \frac{\lambda'}{s'}$, $1-\Phi((y-\eta)\frac{\lambda}{s}) \neq 1-\Phi((y-\eta')\frac{\lambda'}{s'})$. So following similar arguments as in (1), $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$.

(b) If $\sigma_u^2+\sigma_v^2 \neq \sigma_u^{2'}+\sigma_v^{2'}$, similar arguments in (a) give $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$.

(3) $\sigma_v^2 \neq \sigma_v^{2'}$, and $\frac{\sigma_u^2}{\sigma_v^2} \neq \frac{\sigma_u^{2'}}{\sigma_v^{2'}}$. We again consider three subcases:

(a) $\sigma_u^2 + \sigma_v^2 = \sigma_u^{2'} + \sigma_v^{2'}$. So $\frac{\lambda}{s} \neq \frac{\lambda'}{s'}$, and $1 - \Phi((y-\eta)\frac{\lambda}{s}) \neq 1 - \Phi((y-\eta')\frac{\lambda'}{s'})$. So as in (2)(a), $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$.

(b) $\sigma_u^2 + \sigma_v^2 \neq \sigma_u^{2'} + \sigma_v^{2'}$ and $\frac{\lambda}{s} = \frac{\lambda'}{s'}$. Since $1-\Phi((y-\eta)\frac{\lambda}{s}) \neq 1-\Phi((y-\eta')\frac{\lambda'}{s'})$, $\frac{2}{\sqrt{\sigma_u^2+\sigma_v^2}}\phi\left(\frac{y-\eta}{\sqrt{\sigma_u^2+\sigma_v^2}}\right) \neq \frac{2}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\phi\left(\frac{y-\eta'}{\sqrt{\sigma_u^{2'}+\sigma_v^{2'}}}\right)$. $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$ as in (1).

(c) $\sigma_u^2 + \sigma_v^2 \neq \sigma_u^{2'} + \sigma_v^{2'}$ and $\frac{\lambda}{s} \neq \frac{\lambda'}{s'}$, $f_\varepsilon(y-\eta;\theta) \neq f_\varepsilon(y-\eta';\theta')$ similarly.

PA1.2 : $\tilde{I}_\theta(\theta,\eta) > 0$ for all $\theta \in \Theta$ and $\eta \in \mathcal{H}$. The Fisher information for the parametric submodel $\alpha_\theta(x)$ is

$$\begin{aligned}\mathcal{I}_0\left(\frac{\partial}{\partial\theta}\alpha_{\theta_0}(x)\right) &= E\left(\frac{\partial}{\partial\theta}f_\varepsilon(y-g_0(x);\theta_0) + \frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\theta}\alpha_{\theta_0(x)}\right)\left(\frac{\partial}{\partial\theta}f_\varepsilon(y-g_0(x);\theta_0)\right.\\ &+ \left.\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\theta}\alpha_{\theta_0}(x)\right)'\end{aligned}$$

When we use the least favorable direction,

$$\begin{aligned}\mathcal{I}_0\left(\tfrac{\partial}{\partial\theta}\alpha_{\theta_0}(x)^*\right) &= E\left(\tfrac{\partial}{\partial\theta}f_\varepsilon(y-g_0(x);\theta_0)\tfrac{\partial}{\partial\theta}f_\varepsilon(y-g_0(x);\theta_0)'\right)\\ &-E\left\{\frac{E(\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\theta}f_\varepsilon(y-g_0(x);\theta_0)|x)E(\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\theta'}f_\varepsilon(y-g_0(x);\theta_0)|x)}{E((\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)^2|x)}\right\}\end{aligned}$$

$$\begin{aligned}\tilde{I}_\theta(\theta,\eta) &= E\left(\tfrac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta)\tfrac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta)'\right)\\ &-\frac{E(\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta))E(\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta'}f_\varepsilon(y-\eta;\theta))}{E((\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)^2)}\\ &= E\left(\tfrac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta) + \tfrac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\tfrac{\partial}{\partial\theta}\alpha_\theta\right)\left(\tfrac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta) + \tfrac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\tfrac{\partial}{\partial\theta}\alpha_\theta\right)'\end{aligned}$$

where $\frac{\partial}{\partial\theta}\alpha_\theta = -\frac{E(\frac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta))}{E((\frac{d_F}{dg}f_\varepsilon(y-\eta;\theta)^2)}$. Since $\tilde{I}_\theta(\theta,\eta)$ is in a quadratic form, we only need to show $\frac{\partial}{\partial\theta}f_\varepsilon(y-\eta;\theta) + \frac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\frac{\partial}{\partial\theta}\alpha_\theta \neq 0$. Let w, w_1, I, I_1, C_1, C_2 be as in the verification of A5.4 and $I_2 = \int \frac{\frac{\sqrt{2}}{s\pi^{3/2}}}{1-erf(e\frac{\lambda}{s\sqrt{2}})}exp(-e^2(\frac{\lambda^2}{s^2}+\frac{1}{2\sigma^2}))de$. Also, $\alpha'_{\sigma_u^2} = C_1(\frac{1}{s^2}+(\frac{\lambda}{s})^2I_2)^{-1}$, $\alpha'_{\sigma_v^2} = C_2(\frac{1}{s^2}+(\frac{\lambda}{s})^2I_2)^{-1}$. If $e = y-\eta$ then,

$$\begin{aligned}&\tfrac{\partial}{\partial\sigma_u^2}f_\varepsilon(y-\eta;\theta) + \tfrac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\tfrac{\partial}{\partial\sigma_u^2}\alpha_\theta\\ &= -\tfrac{1}{2s^2} + \tfrac{e^2}{2s^4} - [1-\Phi(\tfrac{\lambda e}{s})]^{-1}\phi(\tfrac{\lambda e}{s})ew + [\tfrac{e}{s^2} + [1-\Phi(\tfrac{\lambda e}{s})]^{-1}\phi(\tfrac{\lambda e}{s})(\tfrac{\lambda}{s})]C_1(\tfrac{1}{s^2}+(\tfrac{\lambda}{s})^2I_2)^{-1} \neq 0\end{aligned}$$

since it is linear combination of constant, quadratic function of e, $[1-\Phi(\frac{\lambda e}{s})]^{-1}\phi(\frac{\lambda e}{s})$ and $[1-\Phi(\frac{\lambda e}{s})]^{-1}\phi(\frac{\lambda e}{s})e$. Note $\frac{\phi(x)}{\Phi(x)}$ is convex, asymptotes to $-x$ as $x \to -\infty$ and to zero as $x \to \infty$. $\frac{\phi(x)x}{\Phi(x)}$ is nonlinear, with nonlinear second order derivative.

$$\begin{aligned}&\tfrac{\partial}{\partial\sigma_v^2}f_\varepsilon(y-\eta;\theta) + \tfrac{d_F}{dg}f_\varepsilon(y-\eta;\theta)\tfrac{\partial}{\partial\sigma_v^2}\alpha_\theta\\ &= -\tfrac{1}{2s^2} + \tfrac{e^2}{2s^4} - [1-\Phi(\tfrac{\lambda e}{s})]^{-1}\phi(\tfrac{\lambda e}{s})ew_1 + [\tfrac{e}{s^2} + [1-\Phi(\tfrac{\lambda e}{s})]^{-1}\phi(\tfrac{\lambda e}{s})(\tfrac{\lambda}{s})]C_2(\tfrac{1}{s^2}+(\tfrac{\lambda}{s})^2I_2)^{-1} \neq 0 \text{ similarly.}\end{aligned}$$

PA2 : For $r, s = 0, 1, 2, 3, 4$ and $r + s \leq 4$, $\frac{\partial^{r+s}}{\partial\theta_p^r \partial\eta^s} log f_\varepsilon(y - \eta; \theta)$ exists for $p = 1, \cdots, P$ and

$$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{\partial^{r+s}}{\partial\theta_p^r \partial\eta^s} log f_\varepsilon(y-\eta;\theta)\right|^2\right) < \infty.$$

Consider $r + s = 0$.

$$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{\partial^{r+s}}{\partial\theta_p^r \partial\eta^s} log f_\varepsilon(y-\eta;\theta)\right|^2\right) = E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} |log f_\varepsilon(y-\eta;\theta)|^2\right)$$
$$= E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{1}{2}log(\frac{2}{\pi}) - \frac{1}{2}log(\sigma_u^2+\sigma_v^2) - \frac{1}{2}\frac{e^2}{\sigma_u^2+\sigma_v^2} + log(\Phi(-\frac{\lambda e}{s}))\right|^2\right) < \infty.$$

Since $\sigma_u^2, \sigma_v^2 > 0$, $E(\sup_{\theta\in\Theta}|-\frac{1}{2}log(\sigma_u^2+\sigma_v^2)|^2) < \infty$.
$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} |-\frac{1}{2}\frac{e^2}{\sigma_u^2+\sigma_v^2}|^2\right) < \infty$ since $E\sup_{\eta\in\mathcal{H}}|e|^4 \leq Ey^4 + \sup_{\eta\in\mathcal{H}}\eta^4 \leq E(g_0(x))^4 + E\epsilon^4 + c < \infty$ as $g_0(x) \in \mathcal{H}$, a compact subset of $\Re$. Here $\epsilon = y - g_0(x)$ and $E\epsilon^i < \infty$ for any finite positive integer i. By Taylor expansion of $log(\Phi(-\frac{\lambda e}{s})$ around $e = 0$, for δ between 0 and 1,

$$|log(\Phi(-\frac{\lambda e}{s})| \leq |log(\Phi(0) + \frac{\phi(-\frac{\lambda\delta e}{s})}{\Phi(-\frac{\lambda\delta e}{s})}(-\frac{\lambda e}{s}) \leq |log(\Phi(0)| + \frac{\phi(-\frac{\lambda\delta e}{s})}{\Phi(-\frac{\lambda\delta e}{s})}|-\frac{\lambda e}{s}|$$
$$\leq |log(\Phi(0)| + c(1+\delta|e||\frac{\lambda}{s}|)|\frac{\lambda}{s}||e| \leq c + c|e| + c|e|^2$$

since $\frac{\phi(x)}{\Phi(x)}$ is convex, asymptotes to $-x$ as $x \to -\infty$ and to zero as $x \to \infty$, so $\frac{\phi(x)}{\Phi(x)} < c(1 + |x|)$. So $E\sup_{\theta\in\Theta,\eta\in\mathcal{H}}|log(\Phi(-\frac{\lambda e}{s})|^2 < c + E\sup_{\eta\in\mathcal{H}}|y-\eta|^2 + E\sup_{\eta\in\mathcal{H}}|y-\eta|^4 < \infty$. For $r + s = 1$,

$$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{\partial}{\partial\sigma_u^2} log f_\varepsilon(y-\eta;\theta)\right|^2\right) = E(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}|-\frac{1}{\sigma_u^2+\sigma_v^2} + \frac{1}{2}\frac{e^2}{(\sigma_u^2+\sigma_v^2)^2} - [\Phi(-\frac{\lambda e}{s})]^{-1}\phi(-\frac{\lambda e}{s})we|^2) < \infty$$
$$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{\partial}{\partial\sigma_v^2} log f_\varepsilon(y-\eta;\theta)\right|^2\right) = E(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}|-\frac{1}{\sigma_u^2+\sigma_v^2} + \frac{1}{2}\frac{e^2}{(\sigma_u^2+\sigma_v^2)^2} - [\Phi(-\frac{\lambda e}{s})]^{-1}\phi(-\frac{\lambda e}{s})w_1e|^2) < \infty$$
$$E\left(\sup_{\theta\in\Theta,\eta\in\mathcal{H}} \left|\frac{\partial}{\partial\eta} log f_\varepsilon(y-\eta;\theta)\right|^2\right) = E(\sup_{\theta\in\Theta,\eta\in\mathcal{H}}|\frac{1}{2}\frac{e}{\sigma_u^2+\sigma_v^2} - [\Phi(-\frac{\lambda e}{s})]^{-1}\phi(-\frac{\lambda e}{s})\frac{\lambda}{s}|^2) < \infty$$

Above are true since the term $|[\Phi(-\frac{\lambda e}{s})]^{-i}\phi^i(-\frac{\lambda e}{s})e^j| \leq c(1+\frac{\lambda}{s}|e|^i)|e|^j \leq c|e|^j + c|e|^{i+j}$ and by consequence $|[\Phi(-\frac{\lambda e}{s})]^{-i}\phi^i(-\frac{\lambda e}{s})e^j|^2 \leq c|e|^{2j} + c|e|^{2(i+j)}$. For $(i, j) = (1, 1)$ and $(1, 0)$, we only need $E\sup_{\theta\in\Theta,\eta\in\mathcal{H}}|e|^4 < \infty$, which is true since $\mathcal{H}$ is a compact subset of $\Re$ and $E|\epsilon|^4 < \infty$. For $r+s = 2, 3$ and 4, the partial derivatives and bounded moments are obtained in a similar fashion.

PA3 : This condition has been verified in Lemma 2 in Martins-Filho and Yao (2011).

PB : 1. $\sup_{x\in G} f_x(x) < C$; 2. $\sup_{x\in G} f_x^{(1)}(x) < C$; 3. $\sup_{x\in G} f_x^{(2)}(x) < C$; 4. $\inf_{x\in G} f_x(x) > 0$; 5. $\sup_{x\in G,\theta\in\Theta}|\frac{\partial}{\partial x}\alpha_\theta(x)| < C$.

PB1-4 are all conditions on the marginal density of x. PB5 is implied by the fact that $\frac{\partial}{\partial x}\alpha_\theta(x)$ is continuous at $x \in G$ and $\theta \in \Theta$. G and Θ are all compact sets.

PC : $\int \sup_{x\in G}\left|\frac{\partial}{\partial x}\left(\frac{\partial^{s_1+s_2+r}}{\partial\theta_k^{s_1}\partial\theta_j^{s_2}\partial\eta^r}logf_\varepsilon(y-\alpha_\theta(x);\theta)f_\varepsilon(y-g_0(x);\theta)\right)\right|dy<C,$ with $s_1,s_2,r\geq 0$, where $(s_1,s_2,r)=(0,0,1),(1,0,1),(0,1,1),(1,1,1)$.

Consider the case $(s_1,s_2,r)=(0,0,1)$.

$$\begin{aligned}&\int\sup_{x\in G}\left|\frac{\partial}{\partial x}\left(\frac{\partial}{\partial\eta}logf_\varepsilon(y-\alpha_\theta(x);\theta)f_\varepsilon(y-g_0(x);\theta)\right)\right|dy\\ \leq\ &\int\sup_{x\in G}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y-\alpha_\theta(x);\theta)f_\varepsilon(y-g_0(x);\theta)\frac{\partial}{\partial x}\alpha_\theta(x)\right|dy\\ +\ &\int\sup_{x\in G}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y-\alpha_\theta(x);\theta)\frac{\partial}{\partial x}f_\varepsilon(y-g_0(x);\theta)\right|dy=I_1+I_2\end{aligned}$$

$I_1\leq\sup_{x\in G,\theta\in\Theta}|\frac{\partial}{\partial x}\alpha_\theta(x)|\underbrace{\int\sup_{x\in G}\left|\frac{\partial^2}{\partial\eta^2}logf_\varepsilon(y-\alpha_\theta(x);\theta)f_\varepsilon(y-g_0(x);\theta)\right|dy}_{I_{11}}$. Since $\sup_{x\in G,\theta\in\Theta}|\frac{\partial}{\partial x}\alpha_\theta(x)|<\infty$, and for s_0^2 and λ_0, $I_{11}\leq\int\sup_{\eta\in\mathcal{H}}\left|-\frac{1}{s^2}-\Phi^{-2}(-\frac{\lambda}{s}e)\phi^2(-\frac{\lambda}{s}e)(\frac{\lambda}{s})^2+\Phi^{-1}(-\frac{\lambda}{s}e)\phi(-\frac{\lambda}{s}e)e(\frac{\lambda}{s})^3\right|\sup_{\eta\in\mathcal{H}}\frac{2}{s_0}\phi(\frac{e}{s_0})\Phi(-\frac{\lambda_0}{s_0}e)dy$ $<\infty$, since all terms involve only σ_u^2 or σ_v^2 are bounded, $\Phi(\cdot)\leq 1$ and the term

$$\begin{aligned}&\int\sup_{\eta\in\mathcal{H}}\Phi^{-i}(-\tfrac{\lambda}{s}e)\phi^i(-\tfrac{\lambda}{s}e)e^j\sup_{\eta\in\mathcal{H}}\tfrac{2}{s_0}\phi(\tfrac{e}{s_0})\Phi(-\tfrac{\lambda_0}{s_0}e)dy\\ \leq\ &\int\sup_{\eta\in\mathcal{H}}c[1+\tfrac{\lambda}{s}|y-\eta|]^i|y-\eta|^j\sup_{\eta\in\mathcal{H}}\phi(\tfrac{e}{s_0})dy\leq\int\sup_{\eta\in\mathcal{H}}c[|y-\eta|^j+|y-\eta|^{i+j}]\sup_{\eta\in\mathcal{H}}\phi(\tfrac{e}{s_0})dy<\infty\end{aligned}$$

We divide the area of integration for y into three cases: $I=\{y\leq\eta_L\}$, so $\sup_{\eta\in\mathcal{H}}\phi(\frac{e}{s_0})=\phi(\frac{y-\eta_L}{s_0})$ on I. $II=\{\eta_L<y\leq\eta_U\}$, so $\sup_{\eta\in\mathcal{H}}\phi(\frac{e}{s_0})\leq\frac{1}{\sqrt{2\pi}}$ on II. $III=\{\eta_U<y\}$, so $\sup_{\eta\in\mathcal{H}}\phi(\frac{e}{s_0})=\phi(\frac{y-\eta_U}{s_0})$ on III. $\int\sup_{\eta\in\mathcal{H}}C|y-\eta|^j\sup_{\eta\in\mathcal{H}}\phi(\frac{e}{s_0})dy\leq C\int_I[|y|^j+|\eta|^j]\phi(\frac{y-\eta_L}{s_0})dy+C\int_{II}[|y|^j+|\eta|^j]\frac{1}{\sqrt{2\pi}}dy+C\int_{III}[|y|^j+|\eta|^j]\phi(\frac{y-\eta_U}{s_0})dy<\infty$. Similarly $\int\sup_{\eta\in\mathcal{H}}C|y-\eta|^{i+j}\sup_{\eta\in\mathcal{H}}\phi(\frac{e}{s_0})dy\leq C$, hence $I_1<\infty$.

$I_2\leq\underbrace{\sup_{x\in G}|g_0'(x)|}_{<\infty}\underbrace{\int\sup_{\eta\in\mathcal{H}}\left|\frac{\partial}{\partial\eta}logf_\varepsilon(y-\eta;\theta)|\sup_{\eta\in\mathcal{H}}|\frac{\partial}{\partial\eta}f_\varepsilon(y-\eta;\theta)\right|dy}_{I_{21}}<\infty$, as

$I_{21}\leq\int\sup_{\eta\in\mathcal{H}}|\frac{e}{s^2}+\Phi^{-1}(-\frac{\lambda}{s}e)\phi(-\frac{\lambda}{s}e)\frac{\lambda}{s}|\sup_{\eta\in\mathcal{H}}|\frac{2}{s_0}\phi^{-1}(-\frac{e}{s_0})\Phi(-\frac{\lambda_0}{s_0}e)\frac{|e|}{s_0^2}+\frac{2}{s_0}\phi^{-1}(-\frac{e}{s_0})\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}|dy<\infty$, since all terms involving only σ_u^2 or σ_v^2 are bounded, $\Phi(\cdot)\leq 1$ and $\phi(\cdot)\leq\frac{1}{\sqrt{2\pi}}$. Using arguments similar to those used for I_1, the terms

$$S_1(y)=\int\sup_{\eta\in\mathcal{H}}\Phi^{-i}(-\frac{\lambda}{s}e)\phi^i(-\frac{\lambda}{s}e)e^j\sup_{\eta\in\mathcal{H}}\frac{2}{s_0}\phi(\frac{e}{s_0})\Phi(-\frac{\lambda_0}{s_0}e)\frac{|e|}{s_0^2}dy<\infty \text{ and}$$

$$S_2(y)=\int\sup_{\eta\in\mathcal{H}}\Phi^{-i}(-\frac{\lambda}{s}e)\phi^i(-\frac{\lambda}{s}e)e^j\sup_{\eta\in\mathcal{H}}\frac{2}{s_0}\phi(\frac{e}{s_0})\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}dy<\infty.$$

Verification of PC when $(s_1,s_2,r)=(1,0,1),(0,1,1),(1,1,1)$ folows in an analogous manner.

PD : $\sup_{x\in G,\theta\in\Theta,\eta\in\mathcal{H}}\left|\int\frac{\partial^{s_1+s_2+r}}{\partial\theta_k^{s_1}\partial\theta_j^{s_2}\partial\eta^r}logf_\varepsilon(y-\eta;\theta)\frac{\partial^j}{\partial x^j}f_\varepsilon(y-g_0(x);\theta_0)dy\right|<C$ with $s_1,s_2,r\geq 0$, for j=0, $s_1+s_2+r\leq 4$, and (s_1,s_2,r) = (0,0,5); for j=1, (s_1,s_2,r) = (0,0,1), (1,0,2), (0,1,2), (1,0,3), (0,1,3),

(0,0,4), (0,0,3), (1,1,2), (1,1,3), (0,1,4), (1,0,4), (0,0,5) and for j=2, (s_1, s_2, r) = (0,0,1), (0,0,2), (0,0,3), (1,0,1), (0,1,1), (1,0,2), (0,1,2), (1,0,3), (0,1,3), (0,0,4), (1,1,1), (1,1,2), (1,1,3), (0,1,4), (1,0,4), (0,0,5).

For $j = 0$ we need to show $\sup_{x\in G,\theta\in\Theta,\eta\in\mathcal{H}} \left| E\left(\frac{\partial^{s_1+s_2+r}}{\partial\theta_k^{s_1}\partial\theta_j^{s_2}\partial\eta^r} log f_\varepsilon(y-\eta;\theta)|x\right)\right| < C$. Consider specifically the case where $s_1 = 2$, $s_2 = r = 0$.

$$\begin{aligned} E(\tfrac{\partial^2}{\partial(\sigma_u^2)^2} log f_\varepsilon(y-\eta;\theta)|x) &= \tfrac{1}{2s^4} - \tfrac{E(e^2|x)}{s^6} - w^2 E(\Phi^{-2}(-\tfrac{\lambda}{s}e)\phi^2(-\tfrac{\lambda}{s}e)e^2|x) \\ &+ \tfrac{w}{2s^4}E(\Phi^{-1}(-\tfrac{\lambda}{s}e)\phi(-\tfrac{\lambda}{s}e)e^3|x) - \tfrac{\partial w}{\partial\sigma_u^2}E(\Phi^{-1}(-\tfrac{\lambda}{s}e)\phi(-\tfrac{\lambda}{s}e)e|x) \end{aligned}$$

Again, all terms involving only σ_u^2 or σ_v^2 are bounded. The typical term in the remaining terms are for $i + j \le 2$,

$$\begin{aligned} & E(\Phi^{-i}(-\tfrac{\lambda}{s}e)\phi^i(-\tfrac{\lambda}{s}e)e^j|x) \\ \le\ & CE([1+\tfrac{\lambda}{s}|y-\eta|]^i|y-\eta|^j|x) \le E([|y-\eta|^j + C|y-\eta|]^{i+j}|x) \le cE(|y-\eta|^{i+j}|x) \\ \le\ & CE|y|^{i+j}|x) + c \text{ since } \eta\in\mathcal{H} \\ \le\ & CE(|g_0(x)+\epsilon|^{i+j}|x) \le C + CE|\epsilon|^{i+j} < \infty \text{ since } g_0(x)\in\mathcal{H} \text{ and } \epsilon \text{ is independent of } x. \end{aligned}$$

Hence, $\sup_{x\in G,\theta\in\Theta,\eta\in\mathcal{H}} |E(\frac{\partial^2}{\partial(\sigma_u^2)^2} log f_\varepsilon(y-\eta;\theta)|x)| < \infty$. The other terms can be shown to be bounded in a similar fashion. Now consider the case where $j = 1$ and $(s_1, s_2, r) = (0, 0, 2)$. Then,

$$\begin{aligned} & \sup_{x\in G,\theta\in\Theta,\eta\in\mathcal{H}} \left| \int \tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta) \tfrac{\partial}{\partial x} f_\varepsilon(y-g_0(x);\theta_0)dy \right| \\ & \le \sup_{x\in G}|g_0'(x)| \sup_{\theta\in\Theta,\eta\in\mathcal{H}} \int |\tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta)| \sup_{\eta\in\mathcal{H}} |\tfrac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta_0)|dy \\ & \le C\sup_{\theta\in\Theta} \int \sup_{\eta\in\mathcal{H}} |\tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta)| \sup_{\eta\in\mathcal{H}} |\tfrac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta_0)|dy < \infty \text{ as in } S_1(y) \text{ and } S_2(y) \text{ in the verification of PC.} \end{aligned}$$

The other cases for $j = 1$ can be shown to be bounded in a similar fashion. For $j = 2$, consider specifically the case $(s_1, s_2, r) = (0, 0, 2)$. Then

$$\begin{aligned} & \sup_{x\in G,\theta\in\Theta,\eta\in\mathcal{H}} \left| \int \tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta) \tfrac{\partial^2}{\partial x^2} f_\varepsilon(y-g_0(x);\theta_0)dy \right| \\ & \le \sup_{\theta\in\Theta,\eta\in\mathcal{H}} \int |\tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta)| [\sup_{\eta\in\mathcal{H}} |\tfrac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta_0)] \sup_{x\in G}|g_0^{(2)}(x)| + [\sup_{\eta\in\mathcal{H}} |\tfrac{\partial^2}{\partial\eta^2} f_\varepsilon(y-\eta;\theta_0)] \sup_{x\in G}|g_0'(x)|^2|dy \\ & \le \sup_{\theta\in\Theta} \int \sup_{\eta\in\mathcal{H}} |\tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta)| \sup_{\eta\in\mathcal{H}} |\tfrac{\partial}{\partial\eta} f_\varepsilon(y-\eta;\theta_0)|dy \sup_{x\in G}|g_0^{(2)}(x)| \\ & + \sup_{\theta\in\Theta} \int \sup_{\eta\in\mathcal{H}} |\tfrac{\partial^2}{\partial\eta^2} log f_\varepsilon(y-\eta;\theta)| \sup_{\eta\in\mathcal{H}} |\tfrac{\partial^2}{\partial\eta^2} f_\varepsilon(y-\eta;\theta_0)|dy \sup_{x\in G}|g_0'(x)|^2 = I_1 + I_2. \end{aligned}$$

$I_1 < \infty$ is shown when $j = 1$.

$$\begin{aligned} I_2 < & C\sup_{\theta\in\Theta} \int \sup_{\eta\in\mathcal{H}} |-\tfrac{1}{s^2} - \Phi^{-2}(-\tfrac{\lambda}{s}e)\phi^2(-\tfrac{\lambda}{s}e)(\tfrac{\lambda}{s})^2 + \Phi^{-1}(-\tfrac{\lambda}{s}e)\phi(-\tfrac{\lambda}{s}e)e(\tfrac{\lambda}{s})^3| \\ & \times \sup_{\eta\in\mathcal{H}} |\tfrac{2}{s_0}\phi(\tfrac{e}{s_0})\tfrac{(y-g_0(x))^2}{s_0^4}\Phi(-\tfrac{\lambda_0}{s_0}(y-g_0(x))) + \tfrac{2}{s_0}\phi(\tfrac{e}{s_0})\phi(-\tfrac{\lambda_0}{s_0}(y-g_0(x)))(\tfrac{\lambda_0}{s_0})^2 \\ & + \tfrac{2}{s_0}\phi(\tfrac{e}{s_0})\phi(-\tfrac{\lambda_0}{s_0}(y-g_0(x)))(\tfrac{\lambda_0}{s_0})\tfrac{y-g_0(x)}{s_0^2} + \tfrac{2}{s_0}\phi(\tfrac{e}{s_0})\phi(-\tfrac{\lambda_0}{s_0}(y-g_0(x)))(\tfrac{\lambda_0}{s_0})^2\tfrac{y-g_0(x)}{s_0^2}|dy \end{aligned}$$

Using arguments similar to those in verifying PC we obtain $I_2 < \infty$. The other cases for $j = 2$ can be shown in a similar fashion.

5 Asymptotic variance for $\tilde{\theta}$

We derive the exact form of $\mathcal{I}_0$ as the asymptotic variance for $\tilde{\theta}$. In this case the $\mathcal{I}_0$ matrix has (i,j) elements with $i,j=1,2$ given by

$$\begin{aligned}
\mathcal{I}_{\theta_0}(1,1) &= \frac{1}{2s_0^4} + w^2 I_1 + \alpha'^2_{\sigma_u^2}(\frac{1}{s_0^2} + (\lambda_0/s_0)^2 I_2) + \frac{1}{s_0^4}\left(\alpha'_{\sigma_u^2}\gamma(\theta_0) + \alpha'_{\sigma_u^2}\frac{\lambda_0}{s_0}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1}\right) \\
&- \frac{1}{s_0^2}\alpha'_{\sigma_u^2}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\left(\frac{\lambda_0}{s_0} + 2w\frac{s_0^2}{\lambda_0^2+1}\right) - 2\alpha'_{\sigma_u^2}\frac{\lambda_0}{s_0}wI \\
&+ \frac{1}{s_0^6}\alpha'_{\sigma_u^2}\left(-3\sigma_{v0}^2\sqrt{2/\pi\sigma_{u0}^2} - \frac{(2\sigma_{u0}^2)^{3/2}}{\sqrt{\pi}}\right)
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_{\theta_0}(1,2) = \mathcal{I}_{\theta_0}(2,1) &= \frac{1}{2s_0^4} + ww_1 I_1 + \alpha'_{\sigma_u^2}\alpha'_{\sigma_v^2}(\frac{1}{s_0^2} + (\lambda_0/s_0)^2 I_2) + \frac{1}{2s_0^4}\Big((\alpha'_{\sigma_u^2} + \alpha'_{\sigma_v^2})(\gamma(\theta_0) \\
&+ \frac{\lambda_0}{s_0}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1})\Big) \\
&- \frac{1}{s_0^2}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\left(\frac{\lambda_0}{2s_0}(\alpha'_{\sigma_u^2} + \alpha'_{\sigma_v^2}) + (w\alpha'_{\sigma_v^2} + w_1\alpha'_{\sigma_u^2})\frac{s_0^2}{\lambda_0^2+1}\right) - (w\alpha'_{\sigma_v^2} + w_1\alpha'_{\sigma_u^2})\frac{\lambda_0}{s_0}I \\
&+ \frac{1}{2s_0^6}(\alpha'_{\sigma_u^2} + \alpha'_{\sigma_v^2})\left(-3\sigma_{v0}^2\sqrt{2/\pi\sigma_{u0}^2} - \frac{(2\sigma_{u0}^2)^{3/2}}{\sqrt{\pi}}\right)
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_{\theta_0}(2,2) &= \frac{1}{2s_0^4} + w_1^2 I_1 + \alpha'^2_{\sigma_v^2}(\frac{1}{s_0^2} + (\lambda_0/s_0)^2 I_2) + \frac{1}{s_0^4}\left(\alpha'_{\sigma_v^2}\gamma(\theta_0) + \alpha'_{\sigma_v^2}\frac{\lambda_0}{s_0}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1}\right) \\
&- \frac{1}{s_0^2}\alpha'_{\sigma_v^2}\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\left(\frac{\lambda_0}{s_0} + 2w_1\frac{s_0^2}{\lambda_0^2+1}\right) - 2\alpha'_{\sigma_v^2}\frac{\lambda_0}{s_0}w_1 I \\
&+ \frac{1}{s_0^6}\alpha'_{\sigma_v^2}\left(-3\sigma_{v0}^2\sqrt{2/\pi\sigma_{u0}^2} - \frac{(2\sigma_{u0}^2)^{3/2}}{\sqrt{\pi}}\right)
\end{aligned}$$

When $\theta = \theta_0$, let's define $w = \frac{1}{2\lambda_0 s_0^3}$, $w_1 = -\frac{1}{2\lambda_0}\frac{(\sigma_{u0}^2 + 2\sigma_{v0}^2)\sigma_{u0}^2}{(\sigma_{v0}^2)^2 s_0^3}$, $I = \int \frac{e\frac{\sqrt{2}}{s_0\pi^{3/2}}}{1-erf(e\frac{\lambda_0}{s_0\sqrt{2}})}exp(-e^2(\frac{\lambda_0^2}{s_0^2} + \frac{1}{2s_0^2}))de$, $I_1 = \int \frac{e^2\frac{\sqrt{2}}{s_0\pi^{3/2}}}{1-erf(e\frac{\lambda_0}{s_0\sqrt{2}})}exp(-e^2(\frac{\lambda_0^2}{s_0^2} + \frac{1}{2s_0^2}))de$, $I_2 = \int \frac{\frac{\sqrt{2}}{s_0\pi^{3/2}}}{1-erf(e\frac{\lambda_0}{s_0\sqrt{2}})}exp(-e^2(\frac{\lambda_0^2}{s_0^2} + \frac{1}{2s_0^2}))de$, $C_1 = \frac{\gamma(\theta_0)}{s_0^4} + \frac{\lambda_0}{s_0}wI - (\frac{\lambda_0}{s_0})^2 w\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1} + w\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}$, $C_2 = \frac{\gamma(\theta_0)}{s_0^4} + \frac{\lambda_0}{s_0}w_1 I - (\frac{\lambda_0}{s_0})^2 w_1\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1} + w_1\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}$ where erf is the Gaussian error function. Then $\alpha'_{\sigma_u^2} = C_1(\frac{1}{s_0^2} + (\frac{\lambda_0}{s_0})^2 I_2)^{-1}$, $\alpha'_{\sigma_v^2} = C_2(\frac{1}{s_0^2} + (\frac{\lambda_0}{s_0})^2 I_2)^{-1}$.

$$\begin{aligned}
&\mathcal{I}_{\theta_0}(1,1) = E\left(\frac{\partial}{\partial\sigma_u^2}f_\varepsilon(y - g_0(x);\theta_0) + \frac{d_F}{dg}f_\varepsilon(y - g_0(x);\theta_0)\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)}\right)^2 \\
= \;& E(-\frac{1}{s_0^2} + \frac{e^2}{2s_0^4} - \Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)ew + [\frac{e}{s_0^2} + \Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}]\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)})^2
\end{aligned}$$

$$\begin{aligned}
&\alpha'_{\sigma_u^2} = \frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)} = -\frac{E(\frac{\partial^2}{\partial\sigma_u^2\partial\eta}f_\varepsilon(y-g_0(x);\theta_0)|x)}{E(\frac{\partial^2}{\partial\eta^2}f_\varepsilon(y-g_0(x);\theta_0)|x)} \\
= \;& -[-\frac{1}{s_0^2} - (\frac{\lambda_0}{s_0})^2 E(\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)|x) + E(\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)(\frac{\lambda_0}{s_0})^3 e|x)]^{-1}[E(-\frac{e}{s_0^4}|x) \\
+ \;& (\frac{\lambda_0}{s_0})wE\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)e|x) - (\frac{\lambda_0}{s_0})^2 wE\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)e^2|x) + wE\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)|x)] \\
= \;& [\frac{1}{s_0^2} + (\frac{\lambda_0}{s_0})^2 I_2]^{-1}C_1
\end{aligned}$$

since $E(\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)(\frac{\lambda_0}{s_0})^3e|x)=0$, $E(\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)|x)=I_2$, $E(-\frac{e}{s_0^4}|x)=\frac{\gamma\theta_0}{s_0^4}$, $E\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)e|x)=I$, $E\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)e^2|x)=\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}\frac{s_0^2}{\lambda_0^2+1}$, $E\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)|x)=\sqrt{\frac{2}{\pi(\lambda_0^2+1)}}$. The expression for $\mathcal{I}_{\theta_0}(1,1)$ is obtained with the expression for $\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)}$ plugged in, and realize $Ee^2=s_0^2$, $Ee^4=3s_0^4$, $E\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)e^i|x)=0$ for $i=1,3$.

$$
\begin{aligned}
&\mathcal{I}_{\theta_0}(2,2)=E\left(\frac{\partial}{\partial\sigma_v^2}f_\varepsilon(y-g_0(x);\theta_0)+\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)}\right)^2\\
=\ &E(-\frac{1}{s_0^2}+\frac{e^2}{2s_0^4}-\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)ew_1+[\frac{e}{s_0^2}+\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}]\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)})^2
\end{aligned}
$$

$$
\begin{aligned}
&\alpha'_{\sigma_v^2}=\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)}=-\frac{E(\frac{\partial^2}{\partial\sigma_v^2\partial\eta}f_\varepsilon(y-g_0(x);\theta_0)|x)}{E(\frac{\partial^2}{\partial\eta^2}f_\varepsilon(y-g_0(x);\theta_0)|x)}\\
=\ &-[-\frac{1}{s_0^2}-(\frac{\lambda_0}{s_0})^2E(\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)|x)+E(\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)(\frac{\lambda_0}{s_0})^3e|x)]^{-1}[E(-\frac{e}{s_0^4}|x)\\
+\ &(\frac{\lambda_0}{s_0})w_1E\Phi^{-2}(-\frac{\lambda_0}{s_0}e)\phi^2(-\frac{\lambda_0}{s_0}e)e|x)-(\frac{\lambda_0}{s_0})^2w_1E\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)e^2|x)+w_1E\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)|x)]\\
=\ &[\frac{1}{s_0^2}+(\frac{\lambda_0}{s_0})^2I_2]^{-1}C_2
\end{aligned}
$$

The expression for $\mathcal{I}_{\theta_0}(2,2)$ is obtained with the expression for $\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)}$ plugged in and follow similar arguments as in $\mathcal{I}_{\theta_0}(1,1)$.

$$
\begin{aligned}
&\mathcal{I}_{\theta_0}(1,2)=\mathcal{I}_{\theta_0}(2,1)\\
=\ &E\left(\frac{\partial}{\partial\sigma_u^2}f_\varepsilon(y-g_0(x);\theta_0)+\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)}\right)\\
\times\ &\left(\frac{\partial}{\partial\sigma_v^2}f_\varepsilon(y-g_0(x);\theta_0)+\frac{d_F}{dg}f_\varepsilon(y-g_0(x);\theta_0)\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)}\right)\\
=\ &E(-\frac{1}{s_0^2}+\frac{e^2}{2s_0^4}-\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)ew+[\frac{e}{s_0^2}+\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}]\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)})\\
\times\ &(-\frac{1}{s_0^2}+\frac{e^2}{2s_0^4}-\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)ew_1+[\frac{e}{s_0^2}+\Phi^{-1}(-\frac{\lambda_0}{s_0}e)\phi(-\frac{\lambda_0}{s_0}e)\frac{\lambda_0}{s_0}]\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)})
\end{aligned}
$$

The expression for $\mathcal{I}_{\theta_0}(1,2)$ is obtained with the expression for $\frac{\partial}{\partial\sigma_u^2}\alpha_{\theta_0(x)}$ and $\frac{\partial}{\partial\sigma_v^2}\alpha_{\theta_0(x)}$ plugged in and follow similar arguments as in $\mathcal{I}_{\theta_0}(1,1)$ and $\mathcal{I}_{\theta_0}(2,2)$.

References

[1] Azzalini, A., 1985, A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12, 171-178.

[2] Burridge, J., 1981, A Note on maximum likelihood estimation for regression models using grouped data. Journal of the Royal Statistical Society, Series B, 43, 41-45.

[3] Graves, L. M., 1927, Riemann integration and Taylor's Theorem in general analysis. Transactions of the American Mathematical Society, 29,163-177.

[4] Martins-Filho, C. and F. Yao, 2006, A Note on the use of V and U statistics in nonparametric models of regression. Annals of the Institute of Statistical Mathematics, 58, 389-406

[5] Martins-Filho, C. and F. Yao, 2007, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319.

[6] Martins-Filho, C. and F. Yao, 2011, Semiparametric stochastic frontier estimation via profile likelihood. Unpublished manuscript. Available at http://spot.colorado.edu/∼martinsc/Research.html

[7] Newey, W. and D. McFadden, 1994, Large sample estimation and hypothesis testing. In: R. F. Engle and D. L. McFadden, (Eds.), Handbook of Econometrics, Volume 4. Elsevier Science B.V., Amsterdam.