## Online Supplement for "Conditional Score Approach to Errors-in-variable Current Status Data Under the Proportional Odds Model"

Chi-Chung Wen

*Department of Mathematics, Tamkang University*

Yi-Hau Chen

*Institute of Statistical Science, Academia Sinica*

# Appendix: Proof of theorems

*Proof of theorem 1.* We prove the case where $\sigma^2$ is known. The extension of the proof to the case where $\sigma^2$ is estimated by $\hat{\sigma}^2$ is straightforward given the consistency of $\hat{\sigma}^2$. Let $\hat{H}_{\boldsymbol{\eta}}$ be the maximizer of the log conditional likelihood, defined in equation (5) in the main text as a function $H$ with fixed $\boldsymbol{\eta}$. We apply the Wald's theorem (van der Vaart, 1998, Theorem 5.14) to prove that $\hat{H}_{\boldsymbol{\eta}_0}(t)$ is consistent for $H_0(t)$ for every $\tau_1 < t < \tau_2$. By arguments similar to those in Huang (1996), we consider $F \equiv e^H/(1+e^H)$ and $F_0 \equiv e^{H_0}/(1+e^{H_0})$. Write the conditional likelihood $L(\boldsymbol{\eta}, H, \sigma^2)$ defined in (2) in the main text as $L(\boldsymbol{\eta}, F, \sigma^2)$ and define

$$m(F)(\mathbf{Y}) = \log \frac{L(\boldsymbol{\eta}_0, F, \sigma_0^2)(\mathbf{Y}) + L(\boldsymbol{\eta}_0, F_0, \sigma_0^2)(\mathbf{Y})}{2L(\boldsymbol{\eta}_0, F_0, \sigma_0^2)(\mathbf{Y})}.$$

Then $m(F)$ is uniformly bounded. Also, the map $F \mapsto m(F)(\mathbf{Y})$ is continuous at $F$, relative to the weak topology, for every $\mathbf{Y} = (C, \Delta, \mathbf{Z}, \mathcal{W})$ such that $C$ is a continuous point of $F$. By assumption (A1) in Section 3.1, this includes almost every $\mathbf{Y}$, for every given $F$. Let $\hat{F}_{\boldsymbol{\eta}_0} = \exp(\hat{H}_{\boldsymbol{\eta}_0})/\{1 + \exp(\hat{H}_{\boldsymbol{\eta}_0})\}$. By the definition of $\hat{H}_{\boldsymbol{\eta}_0}$, $\hat{F}_{\boldsymbol{\eta}_0}$ maximizes $P_n \log L(\boldsymbol{\eta}_0, F, \sigma_0^2)$. Therefore

$$P_n m(\hat{F}_{\boldsymbol{\eta}_0}) \geq P_n\Big(\frac{1}{2}\log\frac{L(\boldsymbol{\eta}_0, \hat{F}_{\boldsymbol{\eta}_0}, \sigma_0^2)}{L(\boldsymbol{\eta}_0, F_0, \sigma_0^2)} + \frac{1}{2}\log 1\Big) \geq 0 = P_n m(F_0).$$

On the other hand, by the concavity of the mapping $u \mapsto \log((u+1)/2)$ and Jensen's inequality, we have $P_0 m(F) \leq 0$, and the equality holds only if $F = F_0$ on $(\tau_1, \tau_2)$. Therefore, it follows directly from Theorem 5.14 of van der Vaart (1998) that $\hat{F}_{\boldsymbol{\eta}_0}(t) \xrightarrow{p} F_0(t)$ for every $\tau_1 < t < \tau_2$. By the continuous mapping theorem (van der Vaart, 1998, Theorem 18.11), we obtain $\hat{H}_{\boldsymbol{\eta}_0}(t) \xrightarrow{p} H_0(t)$ for every $\tau_1 < t < \tau_2$.

With the consistency of $\hat{H}_{\boldsymbol{\eta}_0}$, we now restrict $\boldsymbol{\eta}$ to $\mathcal{N}_{\boldsymbol{\eta}_0}$, a neighborhood of $\boldsymbol{\eta}_0$, and restrict $H$ to $\mathcal{H}_0 = \{H \in \mathcal{H} | -M \leq H(\tau_1) \leq H(\tau_2) \leq M\}$. Since the class of monotone and uniformly bounded functions is a Donsker class, by Theorem 2.10.6 of van der Vaart and Wellner (1996), we can show that the class $\{\ell_{\boldsymbol{\eta}}(\boldsymbol{\eta}, H, \sigma_0^2) | (\boldsymbol{\eta}, H) \in \mathcal{N}_{\boldsymbol{\eta}_0} \times \mathcal{H}_0\}$ is Donsker and hence Glivenko-Cantelli. By the fact that $P_0 \ell_{\boldsymbol{\eta}}(\boldsymbol{\zeta}_0) = 0$ and the consistency of $\hat{H}_{\boldsymbol{\eta}_0}$ shown above, it is easy to see that $P_n \ell_{\boldsymbol{\eta}}(\boldsymbol{\eta}_0, \hat{H}_{\boldsymbol{\eta}_0}, \sigma_0^2) = o_p(1)$. This together with assumption $(A4)$ implies the existence of a consistent solution of $\boldsymbol{\eta}$ to the conditional score estimating equation $0 = P_n \ell_{\boldsymbol{\eta}}(\boldsymbol{\eta}, \hat{H}_{\boldsymbol{\eta}}, \sigma_0^2)$.

*Proof of theorem 2.* We shall claim that

$$\|\hat{H}_{\boldsymbol{\eta}} - H_0\|_Q = O_p(\|\boldsymbol{\eta} - \boldsymbol{\eta}_0\| + n^{-1/3})$$

by verifying the conditions (3.5) and (3.6) in Theorem 3.2 of Murphy and van der Vaart (1999). Let the symbol $\preceq$ mean smaller than, up to a constant. A Taylor series argument gives $P_0\{\log L(\boldsymbol{\zeta}_0) - \log L(\boldsymbol{\eta}, H_0, \sigma_0^2)\} \preceq \|\boldsymbol{\eta} - \boldsymbol{\eta}_0\|^2$, which, in conjunction with assumption $(A3)$, can verify $P_0\{\log L(\boldsymbol{\eta}, H, \sigma_0^2) - \log L(\boldsymbol{\eta}, H_0, \sigma_0^2)\} \preceq -\|H - H_0\|_Q^2 + \|\boldsymbol{\eta} - \boldsymbol{\eta}_0\|^2$, the condition (3.5) of Murphy and van der Vaart (1999).

Let $\Psi = \{\log L(\boldsymbol{\eta}, H, \sigma_0^2) | (\boldsymbol{\eta}, H) \in \mathcal{N}_{\boldsymbol{\eta}_0} \times \mathcal{H}_0\}$. It is easy to see that the element in $\Psi$ is uniformly bounded and satisfies $P_0\{\log L(\boldsymbol{\eta}, H, \sigma_0^2) - \log L(\boldsymbol{\eta}, H_0, \sigma_0^2)\}^2 \preceq \|\boldsymbol{\eta} - \boldsymbol{\eta}_0\|^2 + \|H - H_0\|_Q^2$. Furthermore, we can check that the bracketing number $N_{[\,]}(\varepsilon, \Psi, L_2(P))$ (see van der Vaart, 1998, p. 412) is $O(1/\varepsilon)$, and hence the bracketing integral $J(\delta, \Psi, L_2(P)) \equiv \int_0^\delta \{1 + \log N_{[\,]}(\varepsilon, \Psi, L_2(P))\}^{1/2} d\varepsilon$ is $O(\delta^{1/2})$. It then follows from Lemma 3.3 of Murphy and van der Vaart (1999) that their condition (3.6) is satisfied for $\phi_n(\delta) = \delta^{1/2}$. This completes the proof.

*Proof theorem 3.* Define

$$\check{\ell}(\boldsymbol{\zeta}) = \ell_{\boldsymbol{\eta}}(\boldsymbol{\zeta}) - \ell_H(\boldsymbol{\zeta})[\tilde{\mathbf{g}}_0],$$

where $\tilde{\mathbf{g}}_0 = \tilde{\mathbf{g}}(\boldsymbol{\zeta}_0)$. We first verify

$$\sqrt{n} P_0 \check{l}(\boldsymbol{\eta}_0, \hat{H}, \sigma_0^2) = o_p(1), \tag{1}$$

similar to the 'no bias condition' in Murphy and van der Vaart (2000). It is easy to see that $P_{\boldsymbol{\zeta}}\check{\ell}(\boldsymbol{\zeta}) = 0$. Therefore

$$P_0\check{\ell}(\boldsymbol{\eta}_0, H, \sigma_0^2) = (P_0 - P_{\boldsymbol{\eta}_0,H,\sigma_0^2})\{\check{\ell}(\boldsymbol{\eta}_0, H, \sigma_0^2) - \check{\ell}(\boldsymbol{\zeta}_0)\} - P_{\boldsymbol{\eta}_0,H,\sigma_0^2}\check{\ell}(\boldsymbol{\zeta}_0). \quad (2)$$

Both the two terms on the right side of (2) are bounded by a multiple of $\|H - H_0\|_Q^2$; the bound for the first term is obtained by using the mean value theorem twice, and the bound for the second term is obtained if we claim $|E_{(\boldsymbol{\eta}_0,H,\sigma_0^2)}\{\check{\ell}(\boldsymbol{\zeta}_0)|C, S(\boldsymbol{\eta}_0, \sigma_0^2)\}| \preceq (H - H_0)^2(C)$. Note that

$$\begin{aligned} &|E_{(\boldsymbol{\eta}_0,H,\sigma_0^2)}\{\check{\ell}(\boldsymbol{\zeta}_0)|C, S(\boldsymbol{\eta}_0, \sigma_0^2)\}| \\ &= |E_{(\boldsymbol{\eta}_0,H,\sigma_0^2)}\{\check{\ell}(\boldsymbol{\zeta}_0)|C, S(\boldsymbol{\eta}_0, \sigma_0^2)\} - E_{(\boldsymbol{\zeta}_0)}\{\check{\ell}(\boldsymbol{\zeta}_0)|C, S(\boldsymbol{\eta}_0, \sigma_0^2)\}|. \quad (3) \end{aligned}$$

Hence the claim is obtained by first applying Taylor expansion for $H(C) \mapsto L(\boldsymbol{\eta}_0, H, \sigma_0^2)$ around $H_0(C)$, i.e., $|L(\boldsymbol{\eta}_0, H, \sigma_0^2) - L(\boldsymbol{\zeta}_0) - \ell_H(\boldsymbol{\zeta}_0)[H - H_0]L(\boldsymbol{\zeta}_0)| \preceq (H - H_0)^2(C)$, and then employing the fact that $\check{\ell}(\boldsymbol{\zeta}_0)$ is the efficient conditional score $\tilde{\ell}_{\boldsymbol{\eta},0}$. Applying the rate of convergence on $\hat{H}$ to (2), we have $P_0\check{\ell}(\boldsymbol{\eta}_0, \hat{H}, \sigma_0^2) = O_P(\|\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0\|^2 + n^{-2/3})$, which implies (1).

It is known that the class of uniformly bounded functions of bounded variations is a Donsker class. Applying assumption (A5) and Theorem 2.10.6 of van der Vaart and Wellner (1996), it can be verified that $\{\check{\ell}(\boldsymbol{\zeta})|\boldsymbol{\zeta} \in \mathcal{N}_{\boldsymbol{\eta}_0} \times \mathcal{H}_0 \times \mathcal{Q}\}$ and $\{\varphi(\sigma^2)|\sigma^2 \in \mathcal{Q}\}$, where $\mathcal{Q}$ denotes the parameter space for $\sigma^2$, are uniformly bounded Donsker classes; the proof of which is technical and hence omitted here. Combining this with the consistency of $\hat{\boldsymbol{\zeta}} \equiv (\hat{\boldsymbol{\eta}}, \hat{H}, \hat{\sigma}^2)$ leads to

$$\sqrt{n}(P_n - P_0)\left[\begin{array}{c} \check{\ell}(\hat{\boldsymbol{\zeta}}) - \check{\ell}(\boldsymbol{\zeta}_0) \\ \varphi(\hat{\sigma}^2) - \varphi(\sigma_0^2) \end{array}\right] = o_p(1).$$

Adding (1) to the first row of preceding display and using the facts that $P_0\check{\ell}(\boldsymbol{\zeta}_0) = P_0\varphi(\sigma_0^2) = 0$, $P_n\check{\ell}(\hat{\boldsymbol{\zeta}}) = P_n\varphi(\hat{\sigma}^2) = 0$, and $\check{\ell}(\boldsymbol{\zeta}_0) = \tilde{\ell}_{\boldsymbol{\eta},0}$, it is seen that

$$-\sqrt{n}P_0\left[\begin{array}{c} \check{\ell}(\hat{\boldsymbol{\zeta}}) - \check{\ell}(\boldsymbol{\eta}_0, \hat{H}, \sigma_0^2) \\ \varphi(\hat{\sigma}^2) - \varphi(\sigma_0^2) \end{array}\right] = \sqrt{n}P_n\left[\begin{array}{c} \tilde{\ell}_{\boldsymbol{\eta},0} \\ \varphi_0 \end{array}\right] + o_p(1).$$

By the mean value theorem, there exists $(\tilde{\boldsymbol{\eta}}, \tilde{\sigma}^2)$ lying between $(\hat{\boldsymbol{\eta}}, \hat{\sigma}^2)$ and $(\boldsymbol{\eta}_0, \sigma_0^2)$ such that

$$-\sqrt{n}\left\{P_0\left[\begin{array}{cc} \frac{\partial}{\partial\boldsymbol{\eta}}\check{\ell}(\tilde{\boldsymbol{\eta}}, \hat{H}, \tilde{\sigma}^2) & \frac{\partial}{\partial\sigma^2}\check{\ell}(\tilde{\boldsymbol{\eta}}, , \hat{H}, \tilde{\sigma}^2) \\ 0 & \frac{\partial}{\partial\sigma^2}\varphi(\tilde{\sigma}^2) \end{array}\right]\right\}\left[\begin{array}{c} \hat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0 \\ \hat{\sigma}^2 - \sigma_0^2 \end{array}\right] = \sqrt{n}P_n\left[\begin{array}{c} \tilde{\ell}_{\boldsymbol{\eta},0} \\ \varphi_0 \end{array}\right] + o_p(1).$$

By the consistency of $\hat{\boldsymbol{\zeta}}$, we have

$$\sqrt{n}\begin{bmatrix} \hat{\boldsymbol{\eta}}-\boldsymbol{\eta}_0 \\ \hat{\sigma}^2-\sigma_0^2 \end{bmatrix} = \mathcal{I}_*^{-1}\sqrt{n}P_n\begin{bmatrix} \tilde{\ell}_{\boldsymbol{\eta},0} \\ \varphi_0 \end{bmatrix} + o_P(1) \xrightarrow{d} N(0,\mathcal{I}_*^{-1}\mathcal{I}(\mathcal{I}_*^{-1})'),$$

where $\mathcal{I} = P_0\{(\tilde{\ell}'_{\boldsymbol{\eta},0},\varphi_0)'(\tilde{\ell}_{\boldsymbol{\eta},0},\varphi_0)\}$ and

$$\begin{aligned}\mathcal{I}_* &= P_0\begin{bmatrix} -\frac{\partial}{\partial\boldsymbol{\eta}}\check{\ell}(\boldsymbol{\zeta}_0) & -\frac{\partial}{\partial\sigma^2}\check{\ell}(\boldsymbol{\zeta}_0) \\ 0 & -\frac{\partial}{\partial\sigma^2}\varphi(\sigma_0^2) \end{bmatrix} \\ &= P_0\begin{bmatrix} -\frac{\partial}{\partial\boldsymbol{\eta}}\tilde{\ell}_{\boldsymbol{\eta}}(\boldsymbol{\zeta}_0) & -\frac{\partial}{\partial\sigma^2}\tilde{\ell}_{\boldsymbol{\eta}}(\boldsymbol{\zeta}_0) \\ 0 & -\frac{\partial}{\partial\sigma^2}\varphi(\sigma_0^2) \end{bmatrix}.\end{aligned}$$

# References


Huang, J. (1996). Efficient estimation for the Cox model with interval censoring. *Ann. Statist.* **24,** 540-568.

Murphy, S. A. & van der Vaart, A. W. (1999). Observed information in semiparametric models. *Bernoulli* **5,** 381-412.

Murphy, S. A. & van der Vaart, A. W. (2000). On profile likelihood. *J. Amer. Statist. Assoc.* **95,** 449-465.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

van der Vaart, A. W. & Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.