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 $\hat{H}_{\eta}$ be the maximizer of the log conditional likelihood, defined in equation (5) in the main text as a function $H$ with fixed $\eta$. We apply the Wald’s theorem (van der Vaart, 1998, Theorem 5.14) to prove that $\hat{H}_{\eta}$ is consistent for $H_0(t)$ for every $\tau_1 < t < \tau_2$. By arguments similar to those in Huang (1996), we consider $F = \exp(\hat{H}_\eta)/(1 + \exp(\hat{H}_\eta))$ and $F_0 = \exp(\hat{H}_0)/(1 + \exp(\hat{H}_0))$. Write the conditional likelihood $L(\eta, H, \sigma^2)$ defined in (2) in the main text as $L(\eta, F, \sigma^2)$ and define

$$m(F)(Y) = \log \frac{L(\eta_0, F, \sigma_0^2)(Y)}{2L(\eta_0, F_0, \sigma_0^2)(Y)}.$$

Then $m(F)$ is uniformly bounded. Also, the map $F \mapsto m(F)(Y)$ is continuous at $F$, relative to the weak topology, for every $Y = (C, \Delta, Z, W)$ such that $C$ is a continuous point of $F$. By assumption (A1) in Section 3.1, this includes almost every $Y$, for every given $F$. Let $\hat{F}_{\eta_0} = \exp(\hat{H}_{\eta_0})/(1 + \exp(\hat{H}_{\eta_0}))$. By the definition of $\hat{H}_{\eta_0}$, $\hat{F}_{\eta_0}$ maximizes $P_n \log L(\eta_0, F, \sigma_0^2)$. Therefore

$$P_n m(\hat{F}_{\eta_0}) \geq P_n \left( \frac{1}{2} \log \frac{L(\eta_0, \hat{F}_{\eta_0}, \sigma_0^2)}{L(\eta_0, F_0, \sigma_0^2)} + \frac{1}{2} \log 1 \right) \geq 0 = P_n m(F_0).$$

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 \( \hat{F}_{\eta_0}(t) \stackrel{p}{\to} F_0(t) \) for every \( \tau_1 < t < \tau_2 \). By the continuous mapping theorem (van der Vaart, 1998, Theorem 18.11), we obtain \( \hat{H}_{\eta_0}(t) \stackrel{p}{\to} H_0(t) \) for every \( \tau_1 < t < \tau_2 \).

With the consistency of \( \hat{H}_{\eta_0} \), we now restrict \( \eta \) to \( N_{\eta_0} \), a neighborhood of \( \eta_0 \), and restrict \( H \) to \( H_0 = \{ H \in \mathcal{H} \mid -M \leq H(\tau_1) \leq H(\tau_2) \leq M \} \). 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 \( \{ \ell_{\eta}((\eta, H, \sigma^2_0)) : (\eta, H) \in \mathcal{N}_{\eta_0} \times \mathcal{H}_0 \} \) is Donsker and hence Glivenko-Cantelli. By the fact that \( P_0 \ell_{\eta}(\zeta_0) = 0 \) and the consistency of \( \hat{H}_{\eta_0} \) shown above, it is easy to see that \( P_n \ell_{\eta}(\eta_0, \hat{H}_{\eta_0}, \sigma^2_0) = o_p(1) \). This together with assumption (A4) implies the existence of a consistent solution of \( \eta \) to the conditional score estimating equation \( 0 = P_n \ell_{\eta}(\eta, \hat{H}_{\eta}, \sigma^2_0) \).

**Proof of theorem 2.** We shall claim that

\[
\|\hat{H}_{\eta} - H_0\|_Q = O_p(\|\eta - \eta_0\| + n^{-1/3})
\]

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 \( P_0 \{ \log L(\zeta_0) - \log L(\eta, H_0, \sigma^2_0) \} \preceq \|\eta - \eta_0\|^2 \), which, in conjunction with assumption (A3), can verify \( P_0 \{ \log L(\eta, H, \sigma^2_0) - \log L(\eta, H_0, \sigma^2_0) \} \preceq -\|H - H_0\|_Q^2 + \|\eta - \eta_0\|^2 \), the condition (3.5) of Murphy and van der Vaart (1999).

Let \( \Psi = \{ \log L(\eta, H, \sigma^2_0) : (\eta, H) \in \mathcal{N}_{\eta_0} \times \mathcal{H}_0 \} \). It is easy to see that the element in \( \Psi \) is uniformly bounded and satisfies \( P_0 \{ \log L(\eta, H, \sigma^2_0) - \log L(\eta, H_0, \sigma^2_0) \} \preceq \|\eta - \eta_0\|^2 + \|H - H_0\|_Q^2 \). Furthermore, we can check that the bracketing number \( N_{\eta_0}(\epsilon, \Psi, L_2(P)) \) (see van der Vaart, 1998, p. 412) is \( O(1/\epsilon) \), and hence the bracketing integral \( J(\delta, \Psi, L_2(P)) = \int_0^\delta \{ 1 + \log N_{\eta_0}(\epsilon, \Psi, L_2(P)) \}^{1/2} d\epsilon \) 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

\[
\bar{\ell}(\zeta) = \ell_{\eta}(\zeta) - \ell_H(\zeta)[\bar{g}_0],
\]

where \( \bar{g}_0 = \bar{g}(\zeta_0) \). We first verify

\[
\sqrt{n} P_0 \bar{\ell}(\eta_0, \hat{H}, \sigma^2_0) = o_p(1), \quad (1)
\]
similar to the ‘no bias condition’ in Murphy and van der Vaart (2000). It is easy to see that 
\[ P_\zeta \hat{\ell}(\zeta) = 0. \] 
Therefore
\[ P_0 \hat{\ell}(\eta_0, H, \sigma_0^2) = (P_0 - P_0 H, \sigma_0^2) \{ \hat{\ell}(\eta_0, H, \sigma_0^2) - \hat{\ell}(\zeta_0) \} - P_0 H, \sigma_0^2 \hat{\ell}(\zeta_0). \]  
(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 \( |E_0 H, \sigma_0^2 \{ \hat{\ell}(\zeta_0)|C, S(\eta_0, \sigma_0^2) \} | \leq (H - H_0)^2 (C) \). 

Note that
\[ \eta_0 P \] 
leads to
\[ \eta_0 P \]
which implies
\[ \eta_0 P \]
and then
\[ \eta_0 P \]
Hence the claim is obtained by first applying Taylor expansion for \( H(C) \rightleftharpoons L(\eta_0, H, \sigma_0^2) \) around \( H_0(C) \), i.e., \( |L(\eta_0, H, \sigma_0^2) - L(\zeta_0) - \ell_H(\zeta_0)[H - H_0]L(\zeta_0) | \leq (H - H_0)^2 (C) \), and then employing the fact that \( \hat{\ell}(\zeta_0) \) is the efficient conditional score \( \hat{\zeta}_0 \). Applying the rate of convergence on \( \hat{H} \) to (2), we have \( P_0 \hat{\ell}(\eta_0, H, \sigma_0^2) = O_P(\| \eta - \eta_0 \|_2^2 + n^{-2/3}) \), 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 \( \{ \hat{\ell}(\zeta) | \zeta \in \mathcal{N}(\eta_0, H_0, Q) \} \) and \( \{ \varphi(\sigma^2) | \sigma^2 \in Q \} \), where \( Q \) 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 \( \hat{\zeta} \equiv (\hat{\eta}, \hat{H}, \hat{\sigma}^2) \) leads to
\[ \sqrt{n}(P_n - P_0) \left[ \hat{\ell}(\zeta) - \hat{\ell}(\zeta_0) \right] = o_P(1). \]

Adding (1) to the first row of preceding display and using the facts that \( P_0 \hat{\ell}(\zeta_0) = P_0 \varphi(\sigma_0^2) = 0, P_n \hat{\ell}(\zeta) = P_n \varphi(\hat{\sigma}^2) = 0, \) and \( \hat{\ell}(\zeta_0) = \hat{\ell}(\eta_0) \), it is seen that
\[ -\sqrt{n} P_0 \left[ \hat{\ell}(\zeta) - \hat{\ell}(\eta_0, H, \sigma_0^2) \right] = \sqrt{n} P_n \left[ \hat{\eta}_0 \right] + o_P(1). \]

By the mean value theorem, there exists \( (\hat{\eta}, \hat{\sigma}^2) \) lying between \( (\eta_0, \sigma_0^2) \) and \( (\eta_0, \sigma_0^2) \) such that
\[ -\sqrt{n} \left[ P_0 \left[ \frac{\partial}{\partial \eta} \hat{\ell}(\eta, H, \sigma^2) \begin{bmatrix} \hat{\eta} - \eta_0 \ 
\hat{\sigma}^2 - \sigma_0^2 \end{bmatrix} \right] - \sqrt{n} P_n \left[ \hat{\eta}_0 \right] + o_P(1). \]
By the consistency of $\hat{\zeta}$, we have
\[
\sqrt{n} \left[ \frac{\hat{\eta} - \eta_0}{\hat{\sigma}^2 - \sigma_0^2} \right] = \mathcal{I}_*^{-1} \sqrt{n} P_n \left[ \tilde{\eta}_{0,0} \right] + o_P(1) \rightarrow N(0, \mathcal{I}_*^{-1} \mathcal{I}(\mathcal{I}_*^{-1})'),
\]
where $\mathcal{I} = P_0\{(\tilde{\eta}_{0,0}, \varphi_0)\}'(\tilde{\eta}_{0,0}, \varphi_0)\}$ and
\[
\mathcal{I}_* = P_0 \left[ \begin{array}{cc}
-\frac{\partial}{\partial \eta} \tilde{\eta}(\zeta_0) & -\frac{\partial}{\partial \sigma^2} \tilde{\eta}(\zeta_0)
\end{array} \right] = P_0 \left[ \begin{array}{cc}
-\frac{\partial}{\partial \eta} \tilde{\eta}(\zeta_0) & -\frac{\partial}{\partial \sigma^2} \tilde{\eta}(\zeta_0)
\end{array} \right].
\]

References


