Appendix

Proof of Theorem 1. Define

$$X_n(t) = \left[\frac{\hat{\Lambda}^*(D) - \hat{\Lambda}^*(t)}{D - t} - \frac{\hat{\Lambda}^*(t) - \hat{\Lambda}^*(0)}{t}\right] g(t(D-t)), \quad 0 < t < D$$

and

$$X_n^0(t) = \left[\frac{\hat{\Lambda}_0^*(D) - \hat{\Lambda}_0^*(t)}{D - t} - \frac{\hat{\Lambda}_0^*(t) - \hat{\Lambda}_0^*(0)}{t}\right] g(t(D-t)), \quad 0 < t < D,$$

where $\hat{\Lambda}^*(t) = -\log\{[\exp(-\hat{\Lambda}(t)) - 1 + \hat{p}]/\hat{p}\}$, $\hat{\Lambda}(t)$ is the Nelson-Aalen estimator of $\Lambda(t)$ defined in (3.8), $\hat{p}$ is defined in (3.4), and $\hat{\Lambda}_0^*(t) = -\log\{[\exp(-\hat{\Lambda}(t)) - 1 + p]/p\}$. Note that $\Lambda^*(0) = \hat{\Lambda}_0^*(0) = \Lambda(0) = 0$. Then

$$X(t) = t^{q-1}(D-t)^{q-1}\left\{t\Lambda^*(D) - D\Lambda^*(t)\right\}.$$

For any given small $\varepsilon > 0$, let $c_1 \in (0, \min\{X(\tau) - X(\tau - \varepsilon), X(\tau) - X(\tau + \varepsilon)\})$, which depends on $\varepsilon, \tau_1, \tau_2, \theta$ and q. Then $X(\tau) - X(t) > c_1$ whenever $|t - \tau| > \varepsilon$. Noticing that $X_n(t)$ attains its maximum at $\hat{\tau}_n$, for sufficiently large n, we have

$$\begin{aligned}
&\Pr\{\mid \hat{\tau}_n - \tau \mid > \varepsilon\} \leq \Pr\{X(\tau) - X(\hat{\tau}_n) > c_1\} \\
&\leq \Pr\{X(\tau) - X(\hat{\tau}_n) + X_n(\hat{\tau}_n) - X_n(\tau) > c_1\} \\
&\leq \Pr\{\mid X_n(\hat{\tau}_n) - X(\hat{\tau}_n) \mid + \mid X_n(\tau_n) - X(\tau) \mid > c_1\} \\
&\leq \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid X_n(t) - X(t) \mid > \frac{c_1}{2}\right\} \\
&\leq \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid X_n(t) - X_n^0(t) \mid > \frac{c_1}{4}\right\} + \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid X_n^0(t) - X(t) \mid > \frac{c_1}{4}\right\} \\
&\leq \Pr\left\{D\tau_1^{p-1}(D - \tau_2)^{p-1} \sup_{\tau_1 < t < \tau_2} \mid \tilde{U}_n^0(t) \mid + \tau_2^p (D - \tau_2)^{p-1} \mid \tilde{U}_n^0(D) \mid > \frac{c_1}{4}\right\} \\
&+ \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid X_n(t) - X_n^0(t) \mid > \frac{c_1}{4}\right\}, \qquad (A.1)
\end{aligned}$$

where $U_n^0(t) = \hat{\Lambda}_0^*(t) - \Lambda^*(t)$. Consequently, there exist $c_2 > 0$ and $c_3 > 0$, depending on c_1, τ_1, τ_2, D and q, such that

$$\begin{aligned}
\Pr\{\mid \hat{\tau}_n - \tau \mid > \varepsilon\} &\leq \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid U_n^0(t) \mid > c_2\right\} + \Pr\{\mid U_n^0(D) \mid > c_3\} \\
&+ \Pr\left\{\sup_{\tau_1 < t < \tau_2} \mid X_n(t) - X_n^0(t) \mid > \frac{c_1}{4}\right\} \\
&= I + II + III. \qquad (A.2)
\end{aligned}$$

From the definition of $\Lambda^*(t)$ we find that

$$\begin{aligned} \mid U_n^0(t) \mid &= \left| \log\left\{ \frac{1}{p}\left[\exp(-\hat{\Lambda}(t)) - 1 + p\right]\right\} - \log\left\{\frac{1}{p}\left[\exp(-\Lambda(t)) - 1 + p\right]\right\}\right| \\ &= \left| \frac{\exp(-\eta(t))}{\exp(-\eta(t)) - 1 + p}\right| \cdot |\hat{\Lambda}(t) - \Lambda(t)|, \end{aligned} \qquad (A.3)$$

where $\eta(t)$ is a number between $\hat{\Lambda}(t)$ and $\Lambda(t)$. Thus $\exp(-\eta(t))$ lies on the segment between $\hat{S}(t) = 1 - \hat{F}_n(t)$ and $S(t) = 1 - F(t) = 1 - pF_0(t)$. Under the i.i.d. censoring model, according to Wang (1987), $\sup_{t\in[0,\tau_{F_0}]} |\hat{F}_n(t) - F(t)| \longrightarrow 0$ almost surely for $\tau_{F_0} \le \tau_G$. Thus for any $\alpha < 1 - pF_0(D)$ and sufficiently large n,

$$\exp(-\eta(t)) > [1 - F(D)] - \alpha = [1 - pF_0(D)] - \alpha = \phi(D)$$

provided that $\tau_{F_0} > D$. It follows from $(A.3)$ that

$$\mid U_n^0(t) \mid \le \frac{1}{\phi(D) - 1 + p} \mid \hat{\Lambda}(t) - \Lambda(t) \mid = \frac{1}{\phi(D) - 1 + p} \mid U_n(t) \mid,$$

where $U_n(t) = \hat{\Lambda}(t) - \Lambda(t)$ is a martingale. By $(A.2)$ and $(A.3)$ there exists $c_4 > 0$ depending on $c_1, c_2, \tau_1, \tau_2, D$, q, p and F_0, such that

$$\begin{aligned} I &\le \Pr\left\{ \sup_{\tau_1 < t < \tau_2} \mid U_n(t) \mid > c_4, \tau_2 \le Y_{(n)} \right\} + \Pr\{Y_{(n)} < \tau_2\} \\ &\le \Pr\left\{ \sup_{\tau_1 < t < \tau_2} \mid U_n(t \wedge Y_{(n)}) \mid > c_4 \right\} + \prod_{i=1}^{n} \Pr\{Y_i < \tau_2\} \\ &\le c_4^{-2} E\left[\sup_{\tau_1 < t < \tau_2} \mid U_n(t \wedge Y_{(n)}) \mid \right]^2 + (\Pr\{Y_1 < \tau_2\})^2. \end{aligned} \qquad (A.4)$$

We know that

$$\hat{\Lambda}(t \wedge Y_{(n)}) - \Lambda(t \wedge Y_{(n)}) = \int_0^{t \wedge Y_{(n)}} \left\{ \sum_{i=1}^{n} H_i(s) \right\}^{-1} dM_n(s) \qquad (A.5)$$

is a mean zero, square integrable martingale, and

$$E[\hat{\Lambda}(t \wedge Y_{(n)}) - \Lambda(t \wedge Y_{(n)})]^2 = E \int_0^{t \wedge Y_{(n)}} \left\{ \sum_{i=1}^{n} H_i(s) \right\}^{-1} \lambda(s) ds, \qquad (A.6)$$

where $M_n(t) = \sum_{i=1}^n N_i(t) - \int_0^t \sum_{i=1}^n H_i(s)\lambda(s)ds$ is the basic martingale.

In view of $(A.5)$ and $(A.6)$ and the martingale inequality (Hall and Heyde, 1980, p.15), the first term on the right hand side of $(A.4)$, denoted by I_1, satisfies

$$I_1 \leq 2c_4^{-2} E[U_n(\tau_2 \wedge Y_{(n)})]^2 = 2c_4^{-2} E \int_0^{\tau_2 \wedge Y_{(n)}} I(s \leq Y_{(n)}) H_n(s) \lambda(s) ds.$$

Thus I_1 converges to zero as $n \to \infty$ since $\int_0^{\tau_2 \wedge Y_{(n)}} H_n(s)\lambda(s)ds \to 0$ almost surely.

Next, by $(A.2)$, $II \leq \Pr\{| U_n(D) |\} > c_3, D \leq Y_{(n)}\} + \Pr\{Y_{(n)} < D\}$. Similarly, II converges to zero as $n \to \infty$.

In order to prove $III \to 0$, we rewrite $X_n(t)$ and $X_n^0(t)$ as

$$X_n(t) = t^{q-1}(D-t)^{q-1} \left\{ t \left[\hat{\Lambda}^*(D) - \hat{\Lambda}^*(t) \right] - (D-t)\hat{\Lambda}^*(t) \right\} \tag{A.7}$$

and

$$X_n^0(t) = t^{q-1}(D-t)^{q-1} \left\{ t \left[\hat{\Lambda}_0^*(D) - \hat{\Lambda}_0^*(t) \right] - (D-t)\hat{\Lambda}_0^*(t) \right\}. \tag{A.8}$$

By $(A.1)$, $(A.7)$ and $(A.8)$,

$$\begin{aligned} III &\leq \Pr\left\{ \sup_{\tau_1 < t < \tau_2} | X_n(t) - X_n^0(t) | > \frac{c_1}{4}, D < Y_{(n)} \right\} + \Pr\{Y_{(n)} \leq D\} \\ &\leq \Pr\left\{ 2 \sup_{0<t<D} \left| \hat{\Lambda}^*(t) - \hat{\Lambda}_0^*(t) \right| \cdot \tau_2^q (D-\tau_2)^{q-1} > \frac{c_4}{8} \right\} \\ &\quad + \Pr\left\{ \sup_{\tau_1 < t < \tau_2} \left| \hat{\Lambda}_0^*(t) - \Lambda^*(t) \right| \cdot \tau_2^{q-1}(D-\tau_2)^q > \frac{c_4}{8} \right\} + \Pr\{Y_1 \leq D\}^n \\ &= I_2 + I_3 + (\Pr\{Y_1 \leq D\})^n, \quad \text{say.} \end{aligned}$$

It is easy to see that

$$I_2 \leq \Pr\left\{ |\log \hat{p} - \log p| + \sup_{0<t<D} \left| \log \frac{\exp(-\hat{\Lambda}(t)) - 1 + \hat{p}}{\exp(-\hat{\Lambda}(t)) - 1 + p} \right| > \frac{c_4}{8} \right\}. \tag{A.9}$$

Since $\hat{p}$ converges to p in probability under conditions of Maller and Zhou (1996, p.67), $\log x$ is a continuous function for $x > 0$, and $\sup_{0<t<D} |\hat{\Lambda}(t) - \Lambda(t)| \to 0$ almost surely (cf. Anderson *et al.*, 1993, p.193), $(A.9)$ shows that I_2 converges to zero as $n \to \infty$. Similarly, $I_3 \to 0$ as $n \to \infty$. This completes the proof of Theorem 1. ∎

In order to show the asymptotic properties in Section 4, we need some Lemmas. We first state some conditions from Hu (1998), which correspond to Conditions of 3, 1, 4, 2, 5 of Huang (1996), respectively. Note that $o_{p^*}(1)$ in the following representations indicates convergence to zero in outer probability in case that the term involved is not Borel measurable.

Condition 1. (*Stochastic Equicontinuity Condition*)

$$\frac{|\sqrt{n}(P_n - P_0)\dot{l}_\mu(\hat{\mu},\hat{\nu}) - \sqrt{n}(P_n - P_0)\dot{l}_\mu(\mu_0,\nu_0)|}{1+\sqrt{n}|\hat{\mu}-\mu_0|} = o_{p^*}(1),$$

where $|\hat{\mu}-\mu_0| = o_{p^*}(1)$ and $|\hat{\nu}-\nu_0| = o_{p^*}(1)$.

Condition 2. $\sqrt{n}P_n\dot{l}_\mu(\mu_0,\nu_0) = O_{p^*}(1)$.

For i.i.d. observations, Condition 2 holds automatically if $P_0\dot{l}_\mu^2(\mu_0,\nu_0) < \infty$ by the central limit theorem.

Condition 3. (*Smoothness Condition*) For $(\mu,\nu) \in D_n$,

$$\begin{aligned}&|P_0\dot{l}_\mu(\mu,\nu) - P_0\dot{l}_\mu(\mu_0,\nu_0) - P_0\ddot{l}_{\mu\mu}(\mu_0,\nu_0)(\mu-\mu_0) - P_0\ddot{l}_{\mu\nu}(\mu_0,\nu_0)(\nu-\nu_0)|\\ &\quad = o(|\mu-\mu_0|) + o(|\nu-\nu_0|),\end{aligned}$$

where $D_n = \left\{(\mu,\nu) : |\mu-\mu_0| \le \eta_n \downarrow 0, |\nu-\nu_0| \le cn^{-1/2}\right\}$ for some constant c.

Condition 4. $\sqrt{n}P_0\ddot{l}_{\mu\nu}(\mu_0,\nu_0)|\hat{\nu}-\nu_0| = O_p(1)$. When $\hat{\nu}$ is a $\sqrt{n}$-consistent, this condition holds automatically.

Condition 5. under the true probability P_0,

$$\sqrt{n}\begin{bmatrix}(P_n - P_0)\dot{l}_\mu(\mu_0,\nu_0)\\ \hat{\nu}-\nu_0\end{bmatrix} \xrightarrow{d} \Lambda = \begin{bmatrix}\Lambda_1\\ \Lambda_2\end{bmatrix}, \qquad (A.10)$$

where $\Lambda \sim N_4(0,\Sigma)$ with Σ being a 4×4 positive definite matrix.

The following Lemmas 1–4 are due to Hu (1998), which also correspond to Theorem 6.1 in Huang (1996) for the semiparametric model with a infinite-dimensional parameter space.

Lemma 1. (Consistency) *Suppose that μ_0 is the unique solution to $P_0\dot{l}_\mu(\mu,\nu_0) = 0$ and $\hat{\nu}$ is an estimator of ν_0 such that $|\hat{\nu}-\nu_0| = o_{p^*}(1)$. If*

$$\sup_{\mu\in\Theta_1, |\nu-\nu_0|\le\eta_n} \frac{|P_n\dot{l}_\mu(\mu,\nu) - P_0\dot{l}_\mu(\mu,\nu_0)|}{1+|P_n\dot{l}_\mu(\mu,\nu)| + |P_0\dot{l}_\mu(\mu,\nu_0)|} = o_{p^*}(1)$$

for every sequence $\{\eta_n\} \downarrow 0$, then the $\hat{\mu}$ almost surely solving $P_n\dot{l}_\mu(\hat{\mu},\hat{\nu}) = o_{p^*}(1)$ converges in outer probability to μ_0.

Proof. See Theorem 3.1.1 of Hu (1998). ∎

Lemma 2. *Suppose that the class of functions* $\{\psi(\mu,\nu) : |\mu-\mu_0| < \gamma, |\nu-\nu_0| < \gamma\}$ *is* P_0*-Donsker for some* $\gamma > 0$*, and that* $P_0|\psi(\mu,\nu|X) - \psi(\mu_0,\nu_0|X)|^2 \to 0$*, as* $|\mu-\mu_0| \to 0$ *and* $|\nu-\nu_0| \to 0$*. If* $\hat{\mu} \xrightarrow{p^*} \mu_0$ *and* $|\hat{\nu}-\nu_0| \xrightarrow{p^*} 0$*, then*

$$|\sqrt{n}(P_n - P_0)\left(\psi(\hat{\mu},\hat{\nu}) - \psi(\mu_0,\nu_0)\right)| = o_{p^*}(1).$$

Proof. See Lemma 3.1.1 of Hu (1998). ∎

We should note that the conditions of Lemma 2 imply Condition 1. But they give a set of simple sufficient conditions for Condition 1, so we will verify the conditions of Lemma 2 in the proof of Theorem 4 below.

Lemma 3. (Rate of Convergence) *Suppose that* $\hat{\mu}$ *satisfies* $P_n\dot{l}_\mu(\hat{\mu},\hat{\nu}) = o_{p^*}(n^{-1/2})$ *and is a consistent estimator of* μ*, which is the unique point for which* $P_0\dot{l}_\mu(\mu,\nu_0) = 0$*, and* $\hat{\nu}$ *is an estimator of* ν_0 *satisfying* $|\hat{\nu}-\nu_0| = O_{p^*}(n^{-1/2})$*. Then under Conditions 1-4,* $\sqrt{n}(\hat{\mu}-\mu_0) = O_{p^*}(1)$.

Proof. See Theorem 3.1.3 of Hu (1998). ∎

Lemma 4. (Normality) *Suppose that* μ_0 *is the unique solution to* $P_0\dot{l}_\mu(\mu,\nu_0) = 0$ *and* $\hat{\nu}$ *is an estimator of* ν_0 *satisfying* $|\hat{\nu}-\nu_0| = O_{p^*}(1)$*. Then under Conditions 1 and 3-5,* $\sqrt{n}(\hat{\mu}-\mu_0) \xrightarrow{d} \left(-P_0\ddot{l}_{\mu\mu}(\mu_0,\nu_0)\right)^{-1} N_4\left(0,V\right)$*, where* $V = Var(\Lambda_1 + P_0\ddot{l}_{\mu\nu}(\mu_0,\nu_0)\Lambda_2)$*.*

Proof. See Corollary 3.1.2 of Hu (1998). ∎

Lemma 5. *For* $\dot{l}_\beta(\mu,\nu|X)$ *and* $\dot{l}_\theta(\mu,\nu|X)$ *defined in (4.3) and (4.4), if* $|\mu-\mu_0| \le \eta_n \downarrow 0$ *and* $|\nu-\nu_0| \le cn^{-1/2}$*, then* $P_0|\dot{l}_\mu(\mu,\nu)) - \dot{l}_\mu(\mu_0,\nu_0))|^2 = o_p(1)$*.*

Proof. We only show that

$$P_0|\dot{l}_\beta(\mu,\nu)) - \dot{l}_\beta(\mu_0,\nu_0))|^2 = o_p(1),$$

when $|\mu-\mu_0| \le \eta_n \downarrow 0$ and $|\nu-\nu_0| \le cn^{-1/2}$, as the proof for $\dot{l}_\theta$ is similar. Denote

$$A(\mu,\nu,y) = \frac{(1-\delta)(1-p)y}{1-p+p\exp(-\beta y))}$$

and

$$B(\mu,\nu,y) = -\frac{\delta\theta}{\beta(\beta+\theta)} + \frac{(1-\delta)(1-p)y}{1-p+p\exp(-\beta y-\theta(y-\tau))}.$$

Then

$$\dot{l}_\beta(\mu,\nu)) - \dot{l}_\beta(\mu_0,\nu_0)) = [A(\mu,\nu,y)I(y \le \tau) - A(\mu_0,\nu_0,y)I(y \le \tau_0)] \\ + [B(\mu,\nu,y)I(y > \tau) - B(\mu_0,\nu_0,y)I(y > \tau_0)] + [\delta/\beta - \delta/\beta_0] \,.$$

Thus it suffices to show

$$P_0|A(\mu,\nu,y)I(y \le \tau) - A(\mu_0,\nu_0,y)I(y \le \tau_0)|^2 = o_p(1) \qquad (A.11)$$

and

$$P_0|B(\mu,\nu,y)I(y > \tau) - B(\mu_0,\nu_0,y)I(y > \tau_0)|^2 = o_p(1). \qquad (A.12)$$

Note that $A(\mu,\nu,y)$ is continuous for $(\mu,\nu) \in C_0 \times C_\eta$,

$$\begin{aligned} &|A(\mu,\nu,y)I(y \le \tau) - A(\mu_0,\nu_0,y)I(y \le \tau_0)|^2 \\ &\quad = |[A(\mu,\nu,y)I(y \le \tau) - A(\mu,\nu,y)I(y \le \tau_0)] \\ &\qquad + [A(\mu,\nu,y)I(y \le \tau_0) - A(\mu_0,\nu_0,y)I(y \le \tau_0)]|^2 \\ &\quad = A^2(\mu,\nu,y)I^2(\tau_0 < y \le \tau) + [A(\mu,\nu,y) - A(\mu_0,\nu_0,y)]^2 I^2(y \le \tau_0), \end{aligned}$$

and $P_0(I^2(\tau_0 < y \le \tau)) = p[F_0(\tau) - F_0(\tau_0)] \to 0$ as $\tau \to \tau_0$. Thus $(A.11)$ is proved. The proof of $(A.12)$ is similar. ∎

Proof of Theorem 2. To prove the consistency of the pseudo estimator $\hat{\mu}$, we mainly need

$$\sup_{\mu \in C_0, |\nu-\nu_0| \le \eta_n} |P_n \dot{l}_\mu(\mu,\nu) - P_0 \dot{l}_\mu(\mu,\nu_0)| = o_{p^*}(1)$$

for every sequence $\{\eta_n\} \downarrow 0$. Then the consistency of $\hat{\mu}$ follows from Lemma 1. Since

$$|P_n \dot{l}_\mu(\mu,\nu) - P_0 \dot{l}_\mu(\mu,\nu_0)| \le |(P_n - P_0)\dot{l}_\mu(\mu,\nu)| + |P_0(\dot{l}_\mu(\mu,\nu) - \dot{l}_\mu(\mu,\nu_0))|,$$

and by (4.4) the second term obviously tends to zero when $|\nu - \nu_0| \le \eta_n \downarrow 0$, it suffices to show that the class of functions $F_\eta \equiv \{\dot{l}_\mu(\mu,\nu) : \mu \in C_0 \subset R^2, |\nu - \nu_0| \le \eta\}$ is a VC-class for some $\eta > 0$, where C_0 is defined in (4.1). This implies that the uniform strong law of large numbers holds, i.e., $\sup_{f \in F_\eta}(P_n - P_0)f \xrightarrow{p} 0$ (see Van der Vaart and Wellner, 1996, Chap. 2.6–2.7, for details).

Let $F_{1\eta} = \{I_{(-\infty,\tau]}(y) : |\tau - \tau_0| \le \eta_1\}$. Then the VC-index of the class of functions $F_{1\eta}$ is 2 by Example 2.6.1 of Van der Vaart and Wellner (1996). Thus the class of functions

$$\left\{ \frac{(1-\delta)(1-p)yI(y \le \tau)}{1 - p + p\exp(-\beta y)} : \beta > A_1, v \in C_\eta \right\}$$

is Donsker by Lemma 2.6.18 and Example 2.10.8 of Van der Vaart and Wellner (1996), because $(1-\delta)(1-p)/(1-p+p\exp(-\beta y))$ is bounded.

Let $F_{2\eta} = \{I_{(\tau,\infty)}(y) : |\tau - \tau_0| \le \eta_1\}$, we apply Lemma 2.6.18 of Van der Vaart and Wellner (1996) to show that $F_{2\eta}$ is VC-class. Thus the class of functions

$$\left\{\frac{\delta\theta}{\beta(\beta+\theta)} I(y > \tau) : \mu \in C_0, |\tau - \tau_0| \le \eta_1\right\}$$

is Donsker since $\delta\theta/\beta(\beta+\theta)$ is bounded. It is similar to show that the other classes of functions are also Donsker. Thus the class of functions of F_η is VC-class by applying Example 2.10.7 and Theorem 2.10.6 of Van der Vaart and Wellner (1996). Finally, by Lemma 1, $\hat{\mu}$ is consistent. ∎

Proof of Theorem 3. We first verify the stochastic equicontinuity condition:

$$|\sqrt{n}(P_n - P_0)[\dot{l}_\mu(\hat{\mu},\hat{\nu}) - \dot{l}_\mu(\mu_0,\nu_0)]| = o_{p^*}(1). \qquad (A.13)$$

Let $F_\gamma = \{\dot{l}_\mu(\mu,\nu) - \dot{l}_\mu(\mu_0,\nu_0) : |\mu - \mu_0| \le \gamma, |\nu - \nu_0| \le \gamma\}$. Simular to the proof of Theorem 1 we can show that F_γ is a VC-class. Thus $(A.13)$ follows from Lemma 2 together with Lemma 5.

Next, the smoothness Condition 3 holds by (4.5) and Lemma 5, and $P_n\dot{l}_\mu(\mu_0,\nu_0)$ converges in distribution to a normal random variable by the central limit theorem. Thus $\sqrt{n}|\hat{\mu} - \mu| = O_{p^*}(1)$ by Lemma 3. ∎

Proof of Theorem 4. By the consistency of $\hat{p}$ and $\hat{\tau}$ together with Slutsky's theorem and the central limit theorem, we can show that $(A.10)$ holds with normally distributed Λ_1 with mean zero and positive variance. Hence by Lemma 4, $\sqrt{n}(\hat{\mu} - \mu_0)$ is asymptotically normal with mean 0 and variance $\{P_0\ddot{l}_{\mu\mu}(\mu_0,\nu_0)\}^{-2}V$. ∎