where $\boldsymbol{\gamma}_t = \left(\xi_t, \boldsymbol{\eta}_t^\top\right)^\top$ denotes the model parameters and $g_t\left(\boldsymbol{\eta}_t, \widetilde{\mathbf{x}}_{it}, \widetilde{\mathbf{y}}_{it}\right)$ some functions of $\left(\boldsymbol{\eta}_t, \widetilde{\mathbf{x}}_{it}, \widetilde{\mathbf{y}}_{it}\right)$. For example, if y_{it} and $\mathbf{x}_{it}$ are all continuous, we can set $g_t\left(\boldsymbol{\eta}_t, \widetilde{\mathbf{x}}_{it}, \widetilde{\mathbf{y}}_{it}\right) = \boldsymbol{\eta}_{xt}^\top \widetilde{\mathbf{x}}_{it} + \boldsymbol{\eta}_{yt}^\top \widetilde{\mathbf{y}}_{it}$ with $\boldsymbol{\eta}_t = \left(\boldsymbol{\eta}_{xt}^\top, \boldsymbol{\eta}_{yt}^\top\right)^\top$. More complex forms of $g_t\left(\boldsymbol{\eta}_t, \widetilde{\mathbf{x}}_{it}, \widetilde{\mathbf{y}}_{it}\right)$ as well as cases with other types of y_{it} and $\mathbf{x}_{it}$, such as binary, are similarly considered. Under MMDP, we have (see exercise):

$$\pi_{it}\left(\boldsymbol{\gamma}\right) = p_{it} \Pr\left(r_{i(t-1)} = 1 \mid H_{i(t-1)}\right) = \prod_{s=2}^{t} p_{is}\left(\boldsymbol{\gamma}_s\right), \qquad (4.88)$$
$$2 \le t \le m, \quad 1 \le i \le n$$

where $\boldsymbol{\gamma} = \left(\boldsymbol{\gamma}_1^\top, \ldots, \boldsymbol{\gamma}_m^\top\right)^\top$. Thus, we can estimate π_{it} from p_{it} in (4.87) using the above relationship.

Note that when modeled by (4.87) and (4.88), π_{it} in theory ranges between 0 and 1. However, $\mathbf{x}_{it}$ can be assumed to be bounded and the requirement $\pi_{it} \ge c > 0$ is satisfied for all practical purposes. Note also that when π_{it} is estimated by (4.88), the asymptotic variance in (4.83) or (4.85) is likely to underestimate the variability of $\widehat{\boldsymbol{\beta}}$. Unlike the parameter vector $\boldsymbol{\alpha}$ in G_i, the variability of $\widehat{\boldsymbol{\gamma}}$ cannot be ignored even for $\sqrt{n}$-consistent estimate. If $\widehat{\boldsymbol{\gamma}}$ is obtained from GEE or maximum likelihood method, we can readily find the asymptotic variance of $\widehat{\boldsymbol{\beta}}$ obtained from (4.82) with $\boldsymbol{\gamma}$ substituted by $\widehat{\boldsymbol{\gamma}}$.

Let $\widehat{\boldsymbol{\gamma}}$ be the solution to the equation: $\mathbf{u}_n\left(\boldsymbol{\gamma}\right) = \sum_{i=1}^n \mathbf{u}_{ni}\left(\boldsymbol{\gamma}\right) = \mathbf{0}$, where $\mathbf{u}_{ni}$ is the score (for maximum likelihood estimation) or score-like vector (for GEE estimation) for the ith subject ($1 \le i \le n$). As in the proof of Theorem 1 of Section 4.3.2, let $\mathbf{u}_n$ and $\mathbf{w}_n$ denote the normalized versions of $\mathbf{u}_n$ and $\mathbf{w}_n$ in (4.82), that is,

$$\mathbf{u}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{u}_{ni}, \quad \mathbf{w}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{w}_{ni} = \frac{1}{n}\sum_{i=1}^n G_i\left(\mathbf{x}_i\right) \Delta_i S_i$$

Then, by a Taylor series expansion, we have

$$\begin{aligned}\sqrt{n}\left(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right) &= \left(-\frac{\partial}{\partial \boldsymbol{\gamma}} \mathbf{u}_n\right)^{-\top} \sqrt{n}\mathbf{u}_n + \mathbf{o}_p\left(1\right) \\ &= -H^{-1}\frac{\sqrt{n}}{n}\sum_{i=1}^n \mathbf{u}_{ni} + \mathbf{o}_p\left(1\right)\end{aligned}$$