# Supplementary Information for

## Testing the Drift-Diffusion Model

**Drew Fudenberg, Whitney Newey, Philipp Strack, Tomasz Strzalecki**

**Corresponding Author Drew Fudenberg.**
**E-mail: drewf@mit.edu**

### This PDF file includes:

Supplementary text
SI References

## Supporting Information Text

## 1. Monte Carlo Examples

Table 1 reports Monte Carlo mean and standard deviation for the estimator $\hat{\delta}$ of the drift parameter $\delta_0$, where the true boundary is constant at $-1$ and 1, $\delta_0 \in \{.25, .5, 1.0\}$, $p^K(G) = (1, G, G^2, G^3)'$, 1000 Monte Carlo replications, and sample sizes $n = 100$ and $n = 1000$. The code for these results and for those in the Appendix is available at https://www.dropbox.com/sh/hopgdabw9dohiw4/AADtxHeGkwIyzSOFGsIWP7-oa?dl=0.

Table 1: Mean and Std Dev of $\hat{\delta}$

| | $\delta = .25$ | $\delta = .50$ | $\delta = 1.0$ |
|---|---|---|---|
| $n = 100$ | .31 | .55 | 1.07 |
| | (.10) | (.10) | (.14) |
| $n = 1000$ | .26 | .50 | 1.02 |
| | (.03) | (.03) | (.04) |

Here we see that $\hat{\delta}$ is slightly upward biased for $n = 100$ but the bias disappears for $n = 1000$. The drift parameter is quite precisely estimated for $n = 1000$. We expect that this small variance results from averaging over observed $\tau_i$ values. The delta method implies that averaging lowers the sample variance to be equal to the estimator of the unconditional log odds, which is smaller than the variance of the log odds for the regression.

Table 2 gives additional Monte Carlo results for $\hat{\delta}$ when $\hat{p}(t)$ is piecewise linear in $G(\tau)$, as in Table 3 of the Appendix, the true boundary is constant at $-1$ and 1, $\delta_0 = .5$, $n = 1000$, and there are 1000 Monte Carlo replications. It also reports the median (Med) and median absolute deviation (MED) of $\hat{\delta}$ to avoid problems from the possible nonexistence of the population mean and standard deviation; these give about the same results.

Table 2: Additional Monte Carlo Results for $\hat{\delta}$

| Boundary Estimate | Mean | Med | SD | MAD |
|---|---|---|---|---|
| Constant | .500 | .500 | 033 | .026 |
| Linear | .501 | .501 | .033 | .026 |
| 1 Slope Change | .502 | .502 | .033 | .026 |
| 2 Slope Changes | .504 | .503 | .033 | .026 |

The bias is slightly larger for richer $\hat{p}(t)$ specifications but still less than one percent of $\delta_0 = .5$, and overall the specification of $\hat{p}(t)$ has little effect on the properties of $\hat{\delta}$.

The large size of the quantile bands for the boundary estimator in Figure 2 are consistent with delta method calculations. When estimating a constant boundary the numerator and denominator of the boundary estimator $\hat{b}(t)$ are highly positively correlated leading to a precise boundary estimator. When the boundary is allowed to depend on $t$ the variance of the slope is much larger than the variance of a constant when $t$ is far from the middle of the distribution of $\tau$. Furthermore, the variance of the slope is magnified by the fact that the boundary depends on a log odds ratio. Note that $\partial \ln(p/[1-p])/\partial p = 1/[p(1-p)] \geq 4$ so that the standard deviation of a log odds ratio is at least 4 times the standard deviation of an estimator of $p$. If $\delta = .5$, $n = 1000$, the true probability is constant, is estimated by a linear regression of $\gamma_i$ on $G(\tau_i)$, and $G(\tau_i)$ is approximately uniformly distributed as in the simulation, then in the tails of the distribution of $\tau_i$ the boundary estimator has standard deviation of about $\sqrt{12/1000} \approx .11$, with a corresponding distance between upper and lower quantiles of about .44, consistent with Figure 2. Thus we see that both Monte Carlo results and delta method calculations deliver the conclusion that the boundary estimator is quite variable. We do not think this results from the choice of the least squares estimator of the probability, as other estimators would have similar variances. The high variance of the boundary seems to come instead from the fact it depends directly on a log odds ratio, which is quite variable.

## 2. Proofs from Section 4

**A. Proof of Lemma 1.** Dividing Eq. (1) in the paper by $\alpha$ and observing that $\inf\{t \geq 0 : |Z_t| \geq b(t)\} = \inf\{t \geq 0 : |\frac{Z_t}{\alpha}| \geq \frac{b(t)}{\alpha}\}$ yields that $p^*\Big(\delta(x,y), b, 0, \alpha\Big) = p^*\Big(\frac{1}{\alpha}\delta(x,y), \frac{b}{\alpha}, 0, 1\Big)$ and thus the result. **Q.E.D.**

**B. Proof of Theorem 1.** As stated in the text, we restrict attention to cases where $F^c(0) = 0$, $F^c$ admits a density and is strictly increasing with $\lim_{t\to\infty} F^c(t) = 10 < p^c(t) < 1$ for all $t$. We call $(p^c, F^c)$ a choice process.

Consider a continuous and eventually bounded boundary $b : \mathbb{R}_+ \to \mathbb{R}_+$, i.e. there exists $T, \bar{b}$ such that for all $t \geq T$ the boundary satisfies $b(t) \leq \bar{b}$. Let $\tau = \inf\{t \colon |Z_t| \geq b(t)\}$ and

$$Z_t = \delta t + B_t\,. \qquad [1]$$

We denote by $F(t) = F^*(t, \delta, b)$ the distribution of stopping times $\tau \sim F$ and assume throughout that $F$ admits a positive density $f > 0$. We suppose that there exists a regular conditional distribution $(\Pr \cdot \tau = t)_{t \geq 0}$. We denote $p(t) = \Pr Z_\tau = b(\tau)\tau = t$.

**Lemma 2** We have that $p(t) \in (0, 1)$ for all $t > 0$ and

$$2\,\delta\, b(t) = \log\left(\frac{p(t)}{1 - p(t)}\right) . \qquad [2]$$

**Proof**: Let $\mathbb{P}^\delta$ be the probability measure under which $Z$ is a Brownian motion with drift $\delta$. Girsanov's Theorem implies that $W_t = B_t + 2\delta t$ is a Brownian motion under the probability measure $\mathbb{P}^{-\delta}$ that has density $L_t = \exp(-2\delta Z_t)$ with respect to the original probability measure $\mathbb{P}^\delta$ under which $B$ is a Brownian motion (1, Theorem 5.1 in Chapter 3.5). We thus have that

$$\begin{aligned}
\mathbb{P}^\delta\left[\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = b(\tau)\right] &= \mathbb{E}^\delta\left[\mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = b(\tau)}\right] \\
&= \mathbb{E}^\delta\left[L_t\, e^{2\delta Z_t} \mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = b(\tau)}\right] \\
&= \mathbb{E}^{-\delta}\left[e^{2\delta Z_t} \mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = b(\tau)}\right] \\
&= \mathbb{E}^\delta\left[e^{2\delta Z_t} \mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = -b(\tau)}\right] ,
\end{aligned}$$

where the last step follows by considering the process $-Z_t$. As $p(t)$ is well defined in the support of $F$

$$\begin{aligned}
p(t) &= \lim_{\epsilon \to 0} \frac{\mathbb{P}^\delta\left[\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = b(\tau)\right]}{\mathbb{P}^\delta[\tau \in (t-\epsilon, t+\epsilon)]} = \lim_{\epsilon \to 0} \frac{\mathbb{E}^\delta\left[e^{2\delta Z_t} \mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = -b(\tau)}\right]}{\mathbb{P}^\delta[\tau \in (t-\epsilon, t+\epsilon)]} \\
&\leq \lim_{\epsilon \to 0} \frac{\left(\max_{s \in (t-\epsilon, t+\epsilon)} e^{2\delta b(s)}\right) \mathbb{E}^\delta\left[\mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = -b(\tau)}\right]}{\mathbb{P}^\delta[\tau \in (t-\epsilon, t+\epsilon)]} \\
&= e^{2\delta b(t)} \lim_{\epsilon \to 0} \frac{\mathbb{E}^\delta\left[\mathbf{1}_{\tau \in (t-\epsilon, t+\epsilon) \text{ and } Z_\tau = -b(\tau)}\right]}{\mathbb{P}^\delta[\tau \in (t-\epsilon, t+\epsilon)]} = e^{2\delta b(t)}(1 - p(t)) .
\end{aligned}$$

By a symmetric argument we have that $p(t) \geq e^{2\delta b(t)}(1 - p(t))$ and thus

$$p(t) = e^{2\delta b(t)}(1 - p(t)) .$$

Note that this equation can not be satisfied if $p(t) \in \{0, 1\}$. Dividing by $1 - p(t)$ and taking the logarithm yields the result. **Q.E.D.**

**Lemma 3** We have that

$$2\delta^2 = \frac{\int_0^\infty [2\,p(t) - 1] \log\left(\frac{p(t)}{1-p(t)}\right) \mathrm{d}F(t)}{\int_0^\infty t\ \mathrm{d}F(t)} .$$

**Proof**: By the definition of $\tau$ and the continuity of $b$ we have almost surely

$$Z_\tau = \operatorname{sgn}(Z_\tau) b(\tau) . \qquad [3]$$

By Eq. (1), we have that $Z_\tau = \delta\tau + B_\tau$. Combining these two equations and taking expectations, it follows from Doob's optional sampling theorem that for every $n \geq 0$

$$\mathbb{E}\left[Z_{\min\{\tau, n\}}\right] = \mathbb{E}\left[\delta \min\{\tau, n\} + B_{\min\{\tau, n\}}\right] = \mathbb{E}\left[\delta \min\{\tau, n\}\right] . \qquad [4]$$

Recall that we require that $b$ is eventually bounded, i.e. there exists $T, \bar{b}$ such that $b(t) \leq \bar{b}$ for all $t \geq T$. For $t \leq T$ we can bound $|Z_t|$ by

$$|Z_t| \leq |\delta| T + \max_{s \in [0,T]} |B_s| .$$

We can thus bound $|Z_\tau|$

$$|Z_\tau| \leq \max\{|\delta| T + \max_{s \in [0,T]} |B_s|, \bar{b}\} \leq |\delta| T + \max_{s \in [0,T]} |B_s| + \bar{b} =: C .$$

As the quadratic variation of the Brownian motion is given by $[B]_t = t$, the Burkholder-Davis-Gundi inequality (Theorem 4.1 in Chapter IV 2) implies that $\mathbb{E}\left[\max_{s \in [0,T]} |B_s|\right] \leq c\sqrt{T} < \infty$ and thus that the random variable $C$ has finite expectation

$$\mathbb{E}\left[|\delta| T + \max_{s \in [0,T]} |B_s| + \bar{b}\right] = |\delta| T + \bar{b} + \mathbb{E}\left[\max_{s \in [0,T]} |B_s|\right] < \infty .$$

We can thus apply the dominated convergence theorem to get that

$$\mathbb{E}\left[Z_\tau\right] = \mathbb{E}\left[\lim_{n\to\infty} Z_{\min\{\tau,n\}}\right] = \lim_{n\to\infty} \mathbb{E}\left[Z_{\min\{\tau,n\}}\right] = \lim_{n\to\infty} \mathbb{E}\left[\delta \min\{\tau, n\}\right] = \delta \lim_{n\to\infty} \mathbb{E}\left[\min\{\tau, n\}\right] . \qquad [5]$$

We note that

$$\begin{aligned} \mathbb{P}\left[\tau > t\right] &\leq \mathbb{P}\left[|Z_t| < b(t)\right] \leq \mathbb{P}\left[|Z_t| < \bar{b}\right] = \mathbb{P}\left[-\bar{b} < Z_t < \bar{b}\right] = \mathbb{P}\left[-\bar{b} < \delta t + B_t < \bar{b}\right] \\ &= \Phi\left(\frac{\bar{b} - \delta t}{\sqrt{t}}\right) - \Phi\left(\frac{-\bar{b} - \delta t}{\sqrt{t}}\right) . \end{aligned}$$

Taking the limit $t \to \infty$ yields $\lim_{t\to\infty} \mathbb{P}\left[\tau > t\right] = 0$ and $\tau < \infty$ almost surely. As $\tau < \infty$ a.s. we have that $\tau = \lim_{n\to\infty} \min\{\tau, n\}$ a.s. and the monotone convergence theorem implies that

$$\mathbb{E}[\tau] = \mathbb{E}\left[\lim_{n\to\infty} \min\{\tau, n\}\right] = \lim_{n\to\infty} \mathbb{E}[\min\{\tau, n\}] .$$

Combining the above equation with Eq. (3) and Eq. (5) yields

$$\delta\, \mathbb{E}\left[\tau\right] = \mathbb{E}\left[\operatorname{sgn}(Z_\tau) b(\tau)\right] \qquad [6]$$

We can plug Eq. (2) into Eq. (6) and get that

$$\delta\, \mathbb{E}\left[\tau\right] = \mathbb{E}\left[\operatorname{sgn}(Z_\tau) \frac{1}{2\delta} \log\left(\frac{p(\tau)}{1 - p(\tau)}\right)\right]$$

Dividing by $\mathbb{E}\left[\tau\right]$ and multiplying by $2\delta$ yields

$$\begin{aligned} 2\delta^2 &= \frac{\mathbb{E}\left[\operatorname{sgn}(Z_\tau) \log\left(\frac{p^x(\tau)}{1-p(\tau)}\right)\right]}{\mathbb{E}\left[\tau\right]} = \frac{\mathbb{E}\left[\left[\mathbf{1}_{Z_\tau > 0} - \mathbf{1}_{Z_\tau < 0}\right] \log\left(\frac{p^x(\tau)}{1-p(\tau)}\right)\right]}{\mathbb{E}\left[\tau\right]} \\ &= \frac{\mathbb{E}\left[\mathbb{E}\left[\left[\mathbf{1}_{Z_\tau > 0} - \mathbf{1}_{Z_\tau < 0}\right] \log\left(\frac{p(t)}{1-p(t)}\right) \mid \tau\right]\right]}{\int_0^\infty t \, \mathrm{d}F(t)} \\ &= \frac{\int_0^\infty \mathbb{E}\left[\left[\mathbf{1}_{Z_\tau > 0} - \mathbf{1}_{Z_\tau < 0}\right] \log\left(\frac{p(t)}{1-p(t)}\right) \mid \tau = t\right] \mathrm{d}F(t)}{\int_0^\infty t \, \mathrm{d}F(t)} \\ &= \frac{\int_0^\infty \left[p(t) - (1 - p)(t)\right] \log\left(\frac{p(t)}{1-p(t)}\right) \mathrm{d}F(t)}{\int_0^\infty t \, \mathrm{d}F(t)} = \frac{\int_0^\infty \left[2\, p(t) - 1\right] \log\left(\frac{p(t)}{1-p(t)}\right) \mathrm{d}F(t)}{\int_0^\infty t \, \mathrm{d}F(t)} . \end{aligned}$$

**Q.E.D.**

Recall that we call a function $b : \mathbb{R}_+ \to \mathbb{R}$ a valid boundary if $b(t) \geq 0$ for all $t$, $b$ is continuous, and $b$ it is eventually bounded. We defined the revealed drift $\tilde{\delta}^c$ for a choice process $(p^c, F^c)$ by

$$\tilde{\delta}^c = \sqrt{\frac{\bar{I}^c}{2\bar{T}^c}} = \sqrt{\frac{\int_0^\infty \left[2\, p(t) - 1\right] \log\left(\frac{p(t)}{1-p(t)}\right) \mathrm{d}F(t)}{2 \int_0^\infty t \, \mathrm{d}F(t)}} \qquad [7]$$

and the revealed boundary $\tilde{b}^c$ by

$$\tilde{b}^c(t) = \frac{\ln(p^c(t)) - \ln(1 - p^c(t))}{2\tilde{\delta}^c} . \qquad [8]$$

**Theorem 1** For $c$ with $\tilde{\delta}^c \neq 0$ the choice process $(p^c, F^c)$ admits a DDM representation if and only if $\tilde{b}^c$ is a valid boundary, and for all $t \geq 0$

$$F^c(t) = F^*(t, \tilde{\delta}^c, \tilde{b}^c).$$

Moreover, if such a representation exists, it is unique (up to the choice of $\alpha$) and given by $(\tilde{\delta}^c, \tilde{b}^c)$.

**Proof**: Lemmas 2 and 3 established that $(\tilde{\delta}^c, \tilde{b}^c)$ is the unique candidate for a DDM representation. To show sufficiency, consider the DDM model with parameters $(\tilde{\delta}^c, \tilde{b}^c)$. By the assumption of the Theorem $\tilde{b}^c$ is, non-negative, eventually bounded, and continuous and hence a valid boundary. It follows from the assumption of the Theorem that $F^c$ equals the distribution over stopping times in the DDM model with boundary $\tilde{b}^c$ and drift $\delta^c$. Finally, we will show that this DDM model also generates the correct conditional stopping probabilities $p^c$. By Lemma 2 and Eq. (7) and Eq. (8), the conditional probability of stopping in the DDM model $p^*(t, \tilde{\delta}^c, \tilde{b}^c)$ satisfies for each $t \geq 0$

$$\frac{p^*(t, \tilde{\delta}^c, \tilde{b}^c)}{1 - p^*(t, \tilde{\delta}^c, \tilde{b}^c)} = \exp\left(2\tilde{\delta}^c\, \tilde{b}^c(t)\right) = \frac{p^c(t)}{1 - p^c(t)}\,.$$

This shows that $(\tilde{\delta}^c, \tilde{b}^c)$ is a DDM representation of $(p^c, F^c)$ and completes the proof. **Q.E.D.**

## 3. Construction of $\hat{V}$

To construct $\hat{V}$ we use the fact that there are three asymptotically independent sources of variation in $\bar{m} - \hat{m}$. These sources are the variation in $\tau_i$, the variation in $\hat{\beta}$, and the variation from simulation. The variation in $\tau_i$ affects both $\bar{m}$ and $\hat{\delta}$ and the variation in $\hat{\delta}$ has an effect through $\hat{m}$. Generally $\hat{m}$ will not be differentiable in $\hat{\delta}$ so we use a difference quotient to estimate the derivative of $\hat{m}$ with respect to $\delta$. To describe how this source of variation can be estimated let

$$\tau_s(\delta, \beta) = \inf\{t \geq 0 : |\delta t + B_i^s| \geq \frac{1}{\delta} \ln\left[\frac{q^K(G(t))'\beta}{1 - q^K(G(t))'\beta}\right]\},\ \hat{m}(\delta, \beta) = \frac{1}{S}\sum_{s=1}^{S} m_J(\tau_s(\delta, \beta)).$$

denote one simulation $\tau_s(\delta, \beta)$ of $\tau_s$ when $\delta$ is the true drift and $q_K(G(t))'\beta$ the true $p(t) = p^{xy}(t)$ and $\hat{m}(\delta, \beta)$ denote the average over $S$ simulations. Let

$$\hat{M}_\delta = \frac{\hat{m}(\hat{\delta} + \Delta, \hat{\beta}) - \hat{m}(\hat{\delta} - \Delta, \hat{\beta})}{2\Delta}$$

be the difference quotient that serves as an estimator of the derivative of the the expectation of the model moments with respect to the drift. Then

$$\hat{\psi}_{i1} = m_J(\tau_i) - \bar{m} - \hat{M}_\delta \frac{1}{2\hat{\delta}\bar{\tau}}[\hat{I}(\tau_i) - \bar{I} - \hat{\delta}^2\{\tau_i - \bar{\tau}\}]$$

will estimate the influence of $\tau_i$ on the difference of moments coming from the effect of $\tau_i$ on the sample moments as well as on $\hat{\delta}$. An estimator of the variance of the moment differences due to variation in $\tau_i$ is then

$$\hat{V}_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{\psi}_{i1}\hat{\psi}_{i1}'.$$

To estimate the component of the variance due to $\hat{\beta}$ we use

$$\hat{M}_k = \frac{\hat{m}(\hat{\delta}, \hat{\beta} + e_k\Delta) - \hat{m}(\hat{\delta}, \hat{\beta} - e_k\Delta)}{2\Delta}, \quad \hat{M}_\beta = [\hat{M}_1, ..., \hat{M}_K].$$

to estimate the derivative of $E[m_J(\tau_s(\delta, \beta))]$ with respect to $\beta$ at $\hat{\delta}$ and $\hat{\beta}$, where $e_k$ is the $k^{th}$ unit vector. Let $\hat{p}_i = \hat{p}(\tau_i)$ and $d(p) = d\ln[p/(1-p)]/dp = p^{-1}(1-p)^{-1}$. Accounting also for the effect of $\beta$ on $\hat{\delta}$, an estimator of the Jacobian of $E[m_J(\tau_s(\delta, \beta))]$ with respect to $\beta$ is

$$\hat{D}_\beta = \hat{M}_\delta \frac{1}{2\hat{\delta}\bar{\tau}n}\sum_{i=1}^{n} d(\hat{p}_i) q_i^{K\prime} + \hat{M}_\beta.$$

The variation in $\bar{m} - \hat{m}$ due to $\hat{\beta}$ can then be estimated by

$$\hat{V}_2 = \hat{D}_\beta \hat{\Sigma}^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} q_i^K q_i^{K\prime}(\gamma_i - \hat{p}_i)^2\right]\hat{\Sigma}^{-1}\hat{D}_\beta',\ \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} q_i^K q_i^{K\prime}.$$

This is a delta method estimator of the asymptotic variance of $E[m_J(\tau_s(\delta, \beta))]$ due to the $\hat{\beta}$ in the nonparametric estimator $\hat{p}(t)$. As in (3), it is formed by treating $\hat{m}$ as depending on the vector of parameters $\hat{\beta}$ and applying the delta method as if $K$ were fixed and not growing with the sample size.

The variation due to simulation is easy to estimate as $\hat{V}_3 = (n/S^2)\sum_{s=1}^{S}[m_J(\hat{\tau}_s) - \hat{m}][m_J(\hat{\tau}_s) - \hat{m}]'$. In the theory we assume that the number of simulations is large enough so that we can replace this $\hat{V}_3$ by zero without affecting the results. Computing $\hat{V}_3$ in practice may still be a good idea check whether the number of simulations is large enough to make $\hat{V}_3$ negligible.

The estimators of the variance from independent sources of variation can then be combined into an asymptotic variance estimator for $\sqrt{n}[\bar{m} - \hat{m}_S]$ as

$$\hat{V} = \hat{V}_1 + \hat{V}_2 + \hat{V}_3.$$

We give conditions in Theorem 3 sufficient for the chi-squared approximation to the distribution of $\hat{A}$ to be correct for $n$, $J$, and $S$ growing and $\Delta$ shrinking in specific ways.

## 4. Lemmas for Theorem 3

We will use two Lemmas on the asymptotic behavior of quadratic forms to prove the properties of the test statistic. For the first Lemma let $h_i$ be a $J \times 1$ vector of random variables with $E[h_i] = 0$ and $h_1, \ldots, h_n$ i.i.d. Let

$$\Omega = E\left[h_i h_i'\right], \ \bar{h} = \frac{1}{n}\sum_i h_i.$$

Consider $\hat{h}$ that is approximately equal to $\bar{h}$ in the sense that $\hat{h} - \bar{h}$ is small. Also consider an estimator $\hat{\Omega}$ of $\Omega$ and let $\|A\| = \sqrt{tr\,(A'A)}$ be the $L_2$ norm on matrices.

**Lemma 4**: If i) $\lambda_{\min}(\Omega) \geq c > 0$, ii) $J^{-1/2}\sqrt{n}tr\,(\Omega)^{1/2}\left\|\hat{h} - \bar{h}\right\| \xrightarrow{p} 0$, iii) $J^{-1/2}tr\,(\Omega)\left\|\hat{\Omega} - \Omega\right\| \xrightarrow{p} 0$, and iv) $E\left[(h_i'h_i)^2\right]/nJ \longrightarrow 0$ then for the $1-\alpha$ quantile $c\,(\alpha, J)$ of a chi-square distribution with $J$ degrees of freedom

$$\Pr\left(n\hat{h}'\hat{\Omega}^{-1}\hat{h} \geq c\,(\alpha, J)\right) \longrightarrow \alpha.$$

**Proof**: By i) we have $\lambda_{\min}(\Omega) \geq c$, so that $J^{-1/2}tr\,(\Omega)^{1/2} \geq c$. Then iii) implies $\left\|\hat{\Omega} - \Omega\right\| \xrightarrow{p} 0$ and hence w.p.a.1,

$$\lambda_{\min}\left(\hat{\Omega}\right) \geq c.$$

Since this event occurs w.p.a.1 we can assume it is true henceforth. Define

$$T_1 = n'\hat{h}\left(\hat{\Omega}^{-1} - \Omega^{-1}\right)\hat{h}, \ T_2 = n\left[\hat{h}'\Omega^{-1}\hat{h} - \bar{h}'\Omega^{-1}\bar{h}\right]$$

Note that $E[n\left\|\bar{h}\right\|^2] = nE[\bar{h}'\bar{h}] = tr(\Omega)$.Then by the Markov inequality we have

$$\sqrt{n}\left\|\bar{h}\right\| = O_p(tr(\Omega)^{1/2}).$$

Also by ii) $\sqrt{n}\left\|\hat{h} - \bar{h}\right\| \leq CJ^{-1/2}tr(\Omega)^{1/2}\sqrt{n}\left\|\hat{h} - \bar{h}\right\| \xrightarrow{p} 0$. Then by the triangle inequality

$$\sqrt{n}\left\|\hat{h}\right\| \leq \sqrt{n}\left\|\bar{h}\right\| + \sqrt{n}\left\|\hat{h} - \bar{h}\right\| = O_p(tr(\Omega)^{1/2}).$$

It therefore follows that

$$\begin{aligned} |T_1| &= \left|n\hat{h}'\hat{\Omega}^{-1}\left(\Omega - \hat{\Omega}\right)\Omega^{-1}\hat{h}\right| \leq \left\|\sqrt{n}\hat{h}'\hat{\Omega}^{-1}\right\|\left\|\hat{\Omega} - \Omega\right\|\left\|\sqrt{n}\hat{h}'\Omega^{-1}\right\| \leq cn\left\|\hat{h}\right\|^2\left\|\hat{\Omega} - \Omega\right\| \\ &= O_p(tr(\Omega))\left\|\hat{\Omega} - \Omega\right\| = o_p(J^{1/2}). \end{aligned}$$

Similarly we have

$$\begin{aligned} |T_2| &= n\left|\left(\hat{h} - \bar{h}\right)'\Omega^{-1}\hat{h} + \bar{h}'\Omega^{-1}\left(\hat{h} - \bar{h}\right)\right| \leq n(\left\|\hat{h} - \bar{h}\right\|(\left\|\hat{h}\right\| + \left\|\bar{h}\right\|)) \\ &= O_p(tr(\Omega)^{1/2}\sqrt{n}\left\|\hat{h} - \bar{h}\right\|) = o_p(J^{1/2}). \end{aligned}$$

It then follows by the triangle inequality that

$$n'\hat{h}\hat{\Omega}^{-1}\hat{h} - n\bar{h}\Omega^{-1}\bar{h} = T_1 + T_2 = o_p(J^{1/2}).$$

In addition, by iv) and Lemma A.15 of (4),

$$\frac{n\bar{h}'\Omega^{-1}\bar{h} - J}{\sqrt{2J}} \xrightarrow{d} N(0,1).$$

Also, by standard results for the chi-squared distribution, as $J \to \infty$ we have $(c(\alpha, J) - J)/\sqrt{2J}$ converges to the $1-\alpha$ quantile of a $N(0,1)$. Hence

$$\Pr\left(n\bar{h}'\Omega^{-1}\bar{h} \geq c(\alpha, J)\right) = \Pr\left(\frac{n\bar{h}'\Omega^{-1}\bar{h} - J}{\sqrt{2J}} \geq \frac{c(\alpha, J) - J}{\sqrt{2J}}\right) \longrightarrow \alpha.$$

The conclusion then follows by the Slutzky Lemma. **Q.E.D.**

The next Lemma gives a rate of growth for the number of simulation draws to ensure that the limiting distribution of the test statistic based on $\hat{m}_S$ is the same as that based on $\hat{m} = \int m\left(\tau_s\left(\hat{\delta}, \hat{b}\right)\right) dF(s)$.

Let $h_s$ be simulated moments. Then we have:

**Lemma 5**: If $\max_{1\leq j\leq J} \sup_{\tau>0} |m_{jJ}(\tau)| \leq C\sqrt{J}$ and $nJtr(\Omega)/S \longrightarrow 0$ then

$$J^{-1/2}\sqrt{n}tr(\Omega)^{1/2} \|\hat{m}_S - \hat{m}\| \xrightarrow{p} 0,$$

**Proof:** Let $Z = ((\gamma_1, \tau_1), \ldots, (\gamma_n, \tau_n))$ denote the data. Note that by definition, $E[\hat{m}_S|Z] = \hat{m}$. Then for any constant $\ell$

$$\lim Prob(\|\hat{m}_S - \hat{m}\| > \ell) = E\left[\Pr(\|\hat{m}_S - \hat{m}\| > \ell \mid Z)\right].$$

By the Markov inequality

$$\Pr(\|\hat{m}_S - \hat{m}\| > \ell \mid Z) = \Pr\left(\|\hat{m}_S - \hat{m}\|^2 > \ell^2 \mid Z\right) \leq E\left[\sum_{j=1}^{J}(\hat{m}_{Sj} - \hat{m}_j)^2 \mid Z\right]/\ell^2$$

$$\leq \frac{1}{S}\sum_{j=1}^{J} E\left[\hat{m}_j\left(\tau_s\left(\hat{\delta}, \hat{\beta}\right)\right)^2 \mid Z\right]/\ell^2 \leq \frac{C^2J^2}{S\ell^2}.$$

By iterated expectations we then have

$$\Pr(\|\hat{m}_S - \hat{m}\| > \ell) \leq \frac{C^2J^2}{S\ell^2}.$$

Let $\ell = J^{1/2}tr(\Omega)^{-1/2} n^{-1/2}\varepsilon$. Then

$$\Pr\left(J^{-1/2}tr(\Omega)^{1/2}\sqrt{n}\|\hat{m}_s - \hat{m}\| \geq \varepsilon\right) = \Pr(\|\hat{m}_s - \hat{m}\| \geq \ell) \leq C^2J^2\left[SJtr(\Omega)^{-1} n^{-1}\varepsilon^2\right]^{-1}$$
$$= \frac{J^2tr(\Omega)n}{SJ\varepsilon^2} = \frac{nJtr(\Omega)}{S}\frac{1}{\varepsilon^2} \longrightarrow 0.$$

**Q.E.D.**

We next give a uniform convergence rate for $\hat{p}(t)$. For notational simplicity we let $p(t) := p^{xy}(t)$.

**Lemma 6**: If Assumptions 2 and 3 are satisfied then

$$\sup_t |\hat{p}(t) - p(t)| = O_p(\sqrt{\frac{K\ln(K)}{n}} + K^{-s}).$$

**Proof:** Follows from (5), Theorem 4.3 and Comments 4.5 and 4.6. **Q.E.D.**

We next give an asymptotic expansion for $\hat{\delta}$. Define

$$I(p) = p\ln\left(\frac{p}{1-p}\right) + (1-p)\ln\left(\frac{1-p}{p}\right) = (1-2p)\ln\left(\frac{1-p}{p}\right),$$

$$\psi_i^\delta = \frac{1}{2E[\tau_i]\delta}\left\{I(p_i) - I_0 + I_p(p_i)(\gamma_i - p_i) - \delta^2(\tau_i - E[\tau_i])\right\}.$$

**Lemma 7**: If Assumptions 2 and 3 are satisfied and $\sqrt{n}\varepsilon_{pn}^2 \longrightarrow 0$ then

$$\hat{\delta} - \delta = \frac{1}{n}\sum_i \psi_i^\delta + O_p(\varepsilon_{pn}^2) = \frac{1}{n}\sum_i \psi_i^\delta + o_p(1/\sqrt{n}) = O_p(1/\sqrt{n}).$$

**Proof:** Equation (4) and Assumption 3 imply that $p(t)$ is bounded away from zero and one. It then follows from Lemma 6 that with probability approaching one (w.p.a.1) there is $\varepsilon > 0$ with $\varepsilon \leq \hat{p}(t) \leq 1-\varepsilon$. It is straightforward to check that $I(p)$ is twice continuously differentiable in $p \in (0,1)$ with first and second deriatives that are bounded when $p$ is bounded away from zero and one. It then follows by an expansion and Lemma 6 that

$$I\left(\hat{p}_i\right) = I\left(p_i\right) + I_p(p_i)\left(\hat{p}_i - p_i\right) + \hat{R}_i, \left|\hat{R}_i\right| \leq C|\hat{p}_i - p_i|^2.$$

Therefore we have

$$\hat{I} = \frac{1}{n}\sum_i I\left(\hat{p}_i\right) = \frac{1}{n}\sum_i [I\left(p_i\right) + I_p\left(p_i\right)\left(\hat{p}_i - p_i\right)] + \hat{R},\ \hat{R} = O_p(\varepsilon_{pn}^2).$$

Define

$$\begin{aligned}
\Gamma &= (\gamma_1, ..., \gamma_n)',\ P = (p_1, ..., p_n)',\ Q = [q^K(G_1), ..., q^K(G_n)]',\ I_p = (I_p(p_1), ..., I_p(p_n)),\\
H &= I - Q(Q'Q)^{-}Q.
\end{aligned}$$

Note that derivatives of $I_p(p)$ to any order are bounded on $[\varepsilon, 1-\varepsilon]$, so that by the fact that the approximation rate of a general $s$ differentiable function by a b-spline of at least order $s-1$ is $K^{-s}$ we have

$$\frac{1}{n}P'HP = O(K^{-2s}),\ \frac{1}{n}I_p'HI_p = O(K^{-2s}).$$

Note also that

$$\frac{1}{n}\sum_i I_p\left(p_i\right)\left(\hat{p}_i - p_i\right) - \frac{1}{n}\sum_i I_p\left(p_i\right)\left(\gamma_i - p_i\right) = -\frac{1}{n}I_p'H\Gamma$$

Furthermore,

$$E[-\frac{1}{n}I_p'H\Gamma|\tau_1, ..., \tau_n] = -\frac{1}{n}I_p'HP = O(K^{-2s}),\ Var(-\frac{1}{n}I_p'H\Gamma|\tau_1, ..., \tau_n) \leq \frac{1}{n^2}I_p'HI_p = O(\frac{K^{-2s}}{n}).$$

Then by $2K^{-s}/\sqrt{n} \leq 1/n + K^{-2s} \leq \varepsilon_{pn}^2$ it follows that

$$\frac{1}{n}\sum_i I_p\left(p_i\right)\left(\hat{p}_i - p_i\right) - \frac{1}{n}\sum_i I_p\left(p_i\right)\left(\gamma_i - p_i\right) = O_p(\frac{K^{-s}}{\sqrt{n}} + K^{-2s}) = O_p(\varepsilon_{pn}^2).$$

Then by the triangle inequality

$$\hat{I} = \frac{1}{n}\sum_i I\left(\hat{p}_i\right) = \frac{1}{n}\sum_i [I\left(p_i\right) + I_p\left(p_i\right)\left(\gamma_i - p_i\right)] + O_p(\varepsilon_{pn}^2).$$

Note that for $\delta(I,\tau) = \sqrt{I/\tau}$,

$$\frac{\partial\delta(I,\tau)}{\partial I} = \frac{1}{2\delta(I,\tau)\tau},\ \frac{\partial\delta(I,\tau)}{\partial\tau} = -\frac{\delta(I,\tau)}{2\tau}.$$

The conclusion then follows by the usual delta method argument. **Q.E.D.**

Next for any $\alpha(\tau)$ define

$$\psi_i^\alpha = -\delta^{-1}\{E[\alpha(\tau_i)b(\tau_i)]\psi_i^\delta + \frac{\alpha(\tau_i)}{p(\tau_i)[1-p(\tau_i)]}(\gamma_i - p_i)\}.$$

The next result gives a rate of convergence for the boundary estimator $\hat{b}(t)$ and a uniform expansion for a mean square continuous linear functional of $\hat{b}(t)$

**Lemma 8**: If there is a constant $C$ such that $\alpha(G^{-1}(g))$ is continuously differentiable of order $s$ with $\left|d\alpha(G^{-1}(g))/dg\right| \leq C$ on $[0,1]$, then $\sup_t|\hat{b}(t) - b(t)| = O_p(\varepsilon_{pn})$ and

$$\int \alpha(\tau)\{\hat{b}(\tau) - b(\tau)\}F_0(d\tau) = \frac{1}{n}\sum_i \psi_i^\alpha + O_p(\varepsilon_{np}^2),$$

uniformly in $\alpha$.

**Proof:** Note that for $b(\delta,p)=\delta^{-1}\ln(p/[1-p])$,

$$\frac{\partial b(\delta,p)}{\partial\delta}=\frac{-b(\delta,p)}{\delta},\ \ \frac{\partial b(\delta,p)}{\partial p}=\frac{1}{\delta p(1-p)}.$$

Then by Lemma 7, a delta method argument similar to that used in the proof of Lemma 7, and $\hat{\delta}=\delta+O_p(1/\sqrt{n})$ we have

$$\hat{b}\,(t)=b(t)-b(t)\frac{[\hat{\delta}-\delta]}{\delta}+\frac{1}{\delta p(t)[1-p(t)]}[\hat{p}(t)-p(t)]+\hat{R}(t),\ \ \sup_t\left|\hat{R}(t)\right|=O_p(\varepsilon_{pn}^2).$$

The first conclusion then follows by $b(t)$ bounded, which implies $p(t)$ is bounded away from zero and one, and by Lemma 7. To show the second conclusion note that for any bounded $a(t)$ it follows by the proof of Corollary 10 of (6) that

$$\int a(\tau)[\hat{p}(\tau)-p(\tau)]F_0(d\tau)=\frac{1}{n}\sum_i a(\tau_i)[\gamma_i-p_i]+O_p(\varepsilon_{pn}^2),$$

uniformly in $a(\tau)$ with uniformly bounded derivatives to order $s$. Let $a(\tau)=\alpha(\tau)/\{\delta p(t)[1-p(t)]\}$. By plugging in the above expansion for $\hat{b}(t)$ and using boundedness of $\alpha(\tau)$ we obtain

$$\begin{aligned}
&\int\alpha(\tau)\{\hat{b}(\tau)-b(\tau)\}F_0(d\tau)\\
&= -\delta^{-1}\{E[\alpha(\tau_i)b(\tau_i)](\hat{\delta}-\delta)+\int a(\tau)[\hat{p}(\tau)-p(\tau)]F_0(d\tau)+\int\alpha(\tau)\hat{R}(\tau)F_0(d\tau).\\
&= \frac{1}{n}\sum_i\psi_i^{\alpha}+O_p(\varepsilon_{np}^2)+\int\alpha(\tau)\hat{R}(\tau)F_0(d\tau)=\frac{1}{n}\sum_i\psi_i^{\alpha}+O_p(\varepsilon_{np}^2).\textbf{Q.E.D.}
\end{aligned}$$

## 5. Proof of Theorem 3

We first show that conditions i)-iv) of Lemma 4 are satisfied. Let

$$\begin{aligned}
h_{ji} &= m_{ji}-E[m_{ji}]+M_{\delta j}\psi_i^{\tau}+\alpha_{j0}(\tau_i)(\gamma_i-p_i),\\
\psi_i^{\tau} &= \frac{1}{2\delta E[\tau_i]}\{I(p_i)-I_0-\delta^2(\tau_i-E[\tau_i])\},\\
M_{\delta j} &= \sqrt{J}(D_{0\tau_{j+1}}^{\delta}-D_{0\tau_j}^{\delta}-\delta^{-1}E[\{\alpha_{0,\tau_{j+1}}(\tau_i)-\alpha_{0,\tau_j}(\tau_i)\}b(\tau_i)])\\
\alpha_{j0}(\tau_i) &= M_{\delta j}\frac{1}{2E[\tau_i]\delta}I_p(p_i)+\frac{\sqrt{J}[\alpha_{0,\tau_{j+1}}(\tau_i)-\alpha_{0,\tau_j}(\tau_i)]}{\delta p_i[1-p_i]}.
\end{aligned}$$

Also let

$$\begin{aligned}
h_i &= (h_{i1},...,h_{iJ})'=m_i-E[m_i]+M_\delta\psi_i^{\tau}+\alpha_0(\tau_i)(\gamma_i-p_i),\\
M_\delta &= (M_{\delta 1},...,M_{\delta J})',\ \ \alpha_0(\tau)=(\alpha_{10}(\tau),...,\alpha_{J0}(\tau))',\\
\Omega &= E[h_ih_i'],\ \ V_1=Var(m_i+M_\delta\psi_i^{\tau}),\ \ V_2=E[\alpha_0(\tau_i)\alpha_0(\tau_i)'Var(\gamma_i|\tau_i)].
\end{aligned}$$

Note that $\Omega=V_1+V_2$ by $E[\gamma_i|\tau_i]=p(\tau_i)$.

To show condition i) of Lemma 4 it suffices to show that $\lambda_{\min}(V_1)\geq C$, which we now proceed to show. Let

$$\tilde{m}_i=(\sqrt{J+1}\psi_i^{\tau},m_i')'.$$

It follows in a straightforward way from Assumption 5 d) that

$$\lambda_{\min}(E[\tilde{m}_i\tilde{m}_i'])\geq C.$$

Also, for $B=[M_\delta,I]$ we have

$$V_1=BE[\tilde{m}_i\tilde{m}_i']B'.$$

Therefore for any conformable vector $\lambda$ with $\lambda'\lambda=1$,

$$\lambda'V_1\lambda=\frac{\lambda'BE[\tilde{m}_i\tilde{m}_i']B'\lambda}{\lambda'BB'\lambda}\lambda'BB'\lambda\geq C\lambda'BB'\lambda\geq C\lambda_{\min}(BB')\geq C\lambda_{\min}(I)=C.$$

We next show that condition ii) of the Lemma 4 is satisfied. Recall that

$$m_{jJ}(t) = \sqrt{J}1(\tau_{j,J} \leq t < \tau_{j+1,J}),\ (j = 1, ..., J).$$

Then taking expectations over the simulation,

$$\begin{aligned} E[m_{jS}(\delta, b)] &= \bar{m}_j(\delta, b) = \int m_{jJ}(\tau_s(\delta, b))F_s(ds) \\ &= \sqrt{J}[F(\tau_{j+1,J}|\delta, b) - F(\tau_{j,J}|\delta, b)],\ (j = 1, ..., J). \end{aligned}$$

From Assumption 5 let

$$\hat{D}_j(\tilde{\delta}, \tilde{b}) = D(\tilde{\delta}, \tilde{b}; \hat{\delta}, \hat{b}, \tau_j),\ \ D_j(\tilde{\delta}, \tilde{b}) = D(\tilde{\delta}, \tilde{b}; \delta, b, \tau_j).$$

By Assumption 5 a) and Lemma 7,

$$\begin{aligned} \bar{m}_j(\hat{\delta}, \hat{b}) - \bar{m}_j(\delta, b) &= \sqrt{J}[D_{j+1}(\hat{\delta} - \delta, \hat{b} - b) - D_j(\hat{\delta} - \delta, \hat{b} - b)] + \hat{R}_j, \\ \left|\hat{R}_j\right| &\leq \sqrt{J}2C[(\hat{\delta} - \delta)^2 + \sup_t |\hat{b}(t) - b(t)|^2] = O_p(\sqrt{J}\varepsilon_{pn}^2), \end{aligned}$$

uniformly in $j$. By Assumption 5 b) and Lemmas 7 and 8,

$$\begin{aligned} &\sqrt{J}[D_{j+1}(\hat{\delta} - \delta, \hat{b} - b) - D_j(\hat{\delta} - \delta, \hat{b} - b)] \\ =\ &\sqrt{J}[(D^{\delta}_{0\tau_{j+1}} - D^{\delta}_{0\tau_j})(\hat{\delta} - \delta) + \int \{\alpha_{0,\tau_{j+1}}(\tau) - \alpha_{0,\tau_j}(\tau)\}\{\hat{b}(\tau) - b(\tau)\}F_0(d\tau)] \\ =\ &\sqrt{J}[(D^{\delta}_{0\tau_{j+1}} - D^{\delta}_{0\tau_j})\{\frac{1}{n}\sum_i \psi_i^{\delta} + O_p(\varepsilon_{pn}^2)\}] \\ &-\sqrt{J}\delta^{-1}E[\{\alpha_{0,\tau_{j+1}}(\tau_i) - \alpha_{0,\tau_j}(\tau_i)\}b(\tau_i)]) \left(\frac{1}{n}\sum_i \psi_i^{\delta}\right) \\ &+\sqrt{J}\frac{1}{n}\sum_i \frac{[\alpha_{0,\tau_{j+1}}(\tau_i) - \alpha_{0,\tau_j}(\tau_i)]}{\delta p_i[1 - p_i]}(\gamma_i - p_i) + \sqrt{J}O_p(\varepsilon_{pn}^2) \\ =\ &\frac{1}{n}\sum_i h_{ji} + O_p(\sqrt{J}\varepsilon_{pn}^2) \end{aligned}$$

Then by $tr(\Omega)^{1/2} = O(J)$ we have

$$J^{-1/2}\sqrt{n}tr\,(\Omega)^{1/2}\left\|\hat{h} - \bar{h}\right\| \leq CJ^{1/2}\sqrt{n}\left\|\hat{h} - \bar{h}\right\| \leq C\sqrt{n}\sqrt{J}O_p(\sqrt{J}\varepsilon_{pn}^2).$$

Hypothesis ii) of Lemma 4 then follows by $\sqrt{n}J\varepsilon_{pn}^2 \longrightarrow 0$, and by Lemma 5 and $nJ^3/S \longrightarrow 0$.

Next we verify hypothesis iii) of Lemma 4. Note that

$$\hat{M}_{\delta j} = \frac{\hat{m}_j(\hat{\delta} + \Delta, \hat{\beta}) - \hat{m}_j(\hat{\delta} - \Delta, \hat{\beta})}{2\Delta}$$

Let $\bar{m}_j\,(\delta, \beta) = \int m_j\,(\tau_s\,(\delta, \beta))\,F\,(ds)$ and

$$\bar{M}_{\delta j} = \frac{\bar{m}_j\left(\hat{\delta} + \Delta, \hat{\beta}\right) - \bar{m}_j\left(\hat{\delta} - \Delta, \hat{\beta}\right)}{2\Delta}.$$

By the simulations i.i.d. given $\hat{\delta}, \hat{\beta}$ and $m_{jJ}(\tau) \leq C\sqrt{J}$,

$$E\left[\left(\hat{M}_{\delta j} - \bar{M}_{\delta j}\right)^2 \mid \hat{\delta}, \hat{\beta}\right] \leq \frac{CJ}{S\Delta^2}.$$

Then for $\bar{M}_{\delta} = (\bar{M}_{\delta 1}, ..., \bar{M}_{\delta J})'$ the Markov inequality gives

$$E\left[\left\|\hat{M}_{\delta} - \bar{M}_{\delta}\right\|^2\right] \leq \frac{CJ^2}{S\Delta^2},\ \ \left\|\hat{M}_{\delta} - \bar{M}_{\delta}\right\| = O_p\left(\frac{J}{\sqrt{S}\Delta}\right).$$

Note that replacing $\hat{\delta}$ with $\hat{\delta}+\Delta$ in the boundary estimator $\hat{b}$ gives $[\hat{\delta}/(\hat{\delta}+\Delta)]\hat{b}$ and replacing $\hat{\delta}$ with $\hat{\delta}-\Delta$ gives $[\hat{\delta}/(\hat{\delta}-\Delta)]\hat{b}$. Also,

$$\frac{\hat{\delta}}{\hat{\delta}+\Delta}-1=\frac{-\Delta}{\hat{\delta}+\Delta},\ \frac{\hat{\delta}}{\hat{\delta}-\Delta}-1=\frac{\Delta}{\hat{\delta}-\Delta}$$

Let $\hat{D}_j(\delta,b)=D(\delta,b;\hat{\delta},\hat{b},j)$ and $D_j(\delta,b)=D(\delta,b;\delta_0,b_0,j)$ for true values $\delta_0$ and $b_0$. Then by Assumption 5 a),

$$\begin{aligned}\bar{M}_{\delta j} &= \frac{\bar{m}_j\left(\hat{\delta}+\Delta,\hat{\beta}\right)-\bar{m}_j(\hat{\delta},\hat{\beta})-[\bar{m}_j\left(\hat{\delta}-\Delta,\hat{\beta}\right)-\bar{m}_j(\hat{\delta},\hat{\beta})]}{2\Delta}\\ &= \frac{\sqrt{J}[\hat{D}_{j+1}(\Delta,\frac{-\Delta}{\hat{\delta}+\Delta}\hat{b})-\hat{D}_{j+1}(-\Delta,\frac{\Delta}{\hat{\delta}-\Delta}\hat{b};)]}{2\Delta}-\frac{\sqrt{J}[\hat{D}_j(\Delta,\frac{-\Delta}{\hat{\delta}+\Delta}\hat{b})-\hat{D}_j(-\Delta,\frac{\Delta}{\hat{\delta}-\Delta}\hat{b})]}{2\Delta}+\hat{R}_j\\ \left|\hat{R}_j\right| &\leq C\sqrt{J}\Delta^{-1}(\Delta^2+\left|\frac{\Delta}{\hat{\delta}+\Delta}\hat{b}\right|^2+\left|\frac{\Delta}{\hat{\delta}-\Delta}\hat{b}\right|^2)\leq C\sqrt{J}\Delta(1+\left|\hat{b}\right|^2).\end{aligned}$$

We also have

$$\begin{aligned}\sqrt{J}\frac{1}{\Delta}\hat{D}_{j+1}(\Delta,\frac{-\Delta}{\hat{\delta}+\Delta}\hat{b}) &= \sqrt{J}\hat{D}_{j+1}(1,\frac{-1}{\hat{\delta}+\Delta}\hat{b}),\\ \sqrt{J}|\hat{D}_{j+1}(1,\frac{-1}{\hat{\delta}+\Delta}\hat{b})-D_{j+1}(1,\frac{-1}{\hat{\delta}+\Delta}\hat{b})| &\leq C\sqrt{J}\left|\frac{\hat{b}}{\hat{\delta}+\Delta}\right|(|\hat{\delta}-\delta|+|\hat{b}-b|)\leq C\sqrt{J}O_p(\varepsilon_{pn}).\end{aligned}$$

Also,

$$\begin{aligned}&\sqrt{J}\left|D_{j+1}(1,\frac{-1}{\hat{\delta}+\Delta}\hat{b})-D^{\delta}_{0\tau_{j+1}}+\frac{1}{\delta}\int\alpha_{0,\tau_{j+1}}(\tau)b(\tau)F_0(d\tau)\right|\\ \leq\ & C\sqrt{J}(|\hat{\delta}-\delta|+|\hat{b}-b|)=\sqrt{J}O_p(\varepsilon_{pn}).\end{aligned}$$

Applying an analogous set of inequalities to other terms and collecting remainders gives

$$\left|\bar{M}_{\delta j}-M_{\delta j}\right|\leq C\sqrt{J}(\Delta+O_p(\varepsilon_{pn})).$$

Combining results and stacking over $j$ then give

$$\left\|\hat{M}_{\delta}-M_{\delta}\right\|=O_p(J(\frac{1}{\sqrt{S}\Delta}+\Delta+\varepsilon_{pn})).$$

Next, for $\hat{\psi}^{\tau}_i=\left(2\hat{\delta}\bar{\tau}\right)^{-1}[\hat{I}(\tau_i)-\bar{I}-\hat{\delta}^2\{\tau_i-\bar{\tau}\}]$ it follows straighforwardly that

$$\frac{1}{n}\sum_{i=1}^{n}\left(\hat{\psi}^{\tau}_i-\psi^{\tau}_i\right)^2=O_p(\varepsilon^2_{pn}).$$

Let $\tilde{V}_1=n^{-1}\sum_{i=1}^{n}\psi_{1i}\psi_{1i}'$ and $\psi_{1i}=m_i-E[m_i]+M_\delta\psi^{\tau}_i$. Note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\psi}_{1i}-\psi_{1i}\right\|^2 &\leq \|\bar{m}-E[m_i]\|^2+\left\|\hat{M}_{\delta}-M_{\delta}\right\|^2\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\psi}^{\tau}_i\right\|^2+\|M_\delta\|^2\frac{1}{n}\sum_{i=1}^{n}(\hat{\psi}_{1i}-\psi_{1i})^2\\ &= O_p(\frac{J^2}{n})+O_p(J^2(\frac{1}{\sqrt{S}\Delta}+\Delta+\varepsilon_{pn})^2)+O_p(J^2\varepsilon^2_{pn})\\ &= O_p(J^2(\frac{1}{\sqrt{S}\Delta}+\Delta+\varepsilon_{pn})^2).\end{aligned}$$

Then by the Cauchy-Schwartz and triangle inequalities,

$$\begin{aligned}\left\|\hat{V}_1-\tilde{V}_1\right\| &\leq \frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\psi}_{1i}-\psi_{1i}\right\|^2+\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{\psi}_{1i}-\psi_{1i}\right\|^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\|\psi_{1i}\|^2}\\ &= O_p(J^2(\frac{1}{\sqrt{S}\Delta}+\Delta+\varepsilon_{pn})).\end{aligned}$$

It follows similarly that $\left\|\tilde{V}_1 - V_1\right\| = O_p(J^{3/2}/\sqrt{n})$, so by the triangle inequality,

$$\left\|\hat{V}_1 - V_1\right\| = O_p(J^2(\frac{1}{\sqrt{S}\Delta} + \Delta + \varepsilon_{pn})).$$

Next we derive a convergence rate for $\left\|\hat{V}_2 - V_2\right\|$. Let

$$\begin{aligned} D_\beta &= E[\alpha_0(\tau_i)q_i^{K\prime}],\ \Sigma = E[q_i^K q_i^{K\prime}],\ \alpha_K(\tau_i) = D_\beta \Sigma^{-1} q_i^K, \\ \Lambda &= E[q_i^K q_i^{K\prime}(\gamma_i - p_i)^2],\ \bar{V}_2 = D_\beta \Sigma^{-1}\Lambda\Sigma^{-1}D_\beta' = E[\alpha_K(\tau_i)\alpha_K(\tau_i)'(\gamma_i - p_i)^2]. \end{aligned}$$

Note that by Assumption 5 b) and standard approximation properties of splines

$$E[\{(\alpha_{0j}(\tau_i) - \alpha_{Kj}(\tau_i))(\gamma_i - p_i)\}^2] \leq CE[\{\alpha_{0j}(\tau_i) - \alpha_{Kj}(\tau_i)\}^2] \leq CK^{-2s_\alpha},$$

for a constant $C$ that does not epend on $j$. Then we have

$$\begin{aligned} \left\|\bar{V}_2 - V_2\right\|^2 &= \sum_{j,\ell=1}^{J}\{E[\alpha_{Kj}(\tau_i)\alpha_{K\ell}(\tau_i)(\gamma_i - p_i)^2] - E[\alpha_{0j}(\tau_i)\alpha_{0\ell}(\tau_i)(\gamma_i - p_i)^2]\}^2 \\ &= \sum_{j,\ell=1}^{J}\{E[\{\alpha_{Kj}(\tau_i) - \alpha_{0j}(\tau_i)\}\alpha_{K\ell}(\tau_i)(\gamma_i - p_i)^2] + E[\alpha_{0j}(\tau_i)\{\alpha_{K\ell}(\tau_i) - \alpha_{0\ell}(\tau_i)\}(\gamma_i - p_i)^2]\}^2 \\ &\leq C\sum_{j,\ell=1}^{J}\{\sqrt{E[\{\alpha_{Kj}(\tau_i) - \alpha_{0j}(\tau_i)\}^2]}\sqrt{E[\alpha_{K\ell}(\tau_i)^2]} \\ &\quad + \sqrt{E[\{\alpha_{K\ell}(\tau_i) - \alpha_{0\ell}(\tau_i)\}^2]}\sqrt{E[\alpha_{0j}(\tau_i)^2]}\}^2 \\ &\leq C\left(\sum_{j=1}^{J}E[\{\alpha_{Kj}(\tau_i) - \alpha_{0j}(\tau_i)\}^2]\right)\left(\sum_{j=1}^{J}\{E[\alpha_{0j}(\tau_i)^2] + E[\alpha_{K\ell}(\tau_i)^2]\}\right) \leq CJ^2K^{-2s_\alpha}. \end{aligned}$$

Taking square roots we have

$$\left\|\bar{V}_2 - V_2\right\| \leq CJK^{-s_\alpha}.$$

Define

$$\bar{M}_{\beta jk} = \frac{\bar{m}_j\left(\hat{\delta}, \hat{\beta} + e_k\Delta\right) - \bar{m}_j\left(\hat{\delta}, \hat{\beta} - e_k\Delta\right)}{2\Delta}.$$

It follows similarly to $\left\|\hat{M}_\delta - \bar{M}_\delta\right\| = \left\|\hat{M}_\delta - \bar{M}_\delta\right\| = O_p\left(J/\sqrt{S}\Delta\right)$ that

$$\left\|\hat{M}_\beta - \bar{M}_\beta\right\| = O_p\left(J\sqrt{K}/\sqrt{S}\Delta\right).$$

Next, let $\hat{p}_{\Delta k}(t) = \hat{p}(t) + \Delta q_{kK}(G(t))$ and $\hat{b}_{\Delta k}(t) = \hat{\delta}^{-1}\ln(\hat{p}_{\Delta k}(t)/[1 - \hat{p}_{\Delta k}(t)])$. By $\Delta\sqrt{K} \longrightarrow 0$ and $\sup_{G\in[0,1]}|q_{kK}(G)| \leq C\sqrt{K}$ it follows that $\sup_t \Delta q_{kK}(G(t)) \longrightarrow 0$. Then w.p.a.1 we have

$$\hat{b}_{\Delta k}(t) = \hat{b}(t) + \frac{\Delta q_{kK}(G(t))}{\hat{\delta}\hat{p}(t)[1 - \hat{p}(t)]} + \hat{R}_k(t, \Delta),\ \ \left|\hat{R}_k(t, \Delta)\right| \leq C\Delta^2 K.$$

Then we have

$$\begin{aligned} \bar{M}_{\beta jk} &= \frac{\bar{m}_j\left(\hat{\delta}, \hat{\beta} + e_k\Delta\right) - \bar{m}_j(\hat{\delta}, \hat{\beta}) - [\bar{m}_j\left(\hat{\delta}, \hat{\beta} - e_k\Delta\right) - \bar{m}_j(\hat{\delta}, \hat{\beta})]}{2\Delta} \\ &= \frac{\sqrt{J}[\hat{D}_{j+1}(0, \hat{b}_{\Delta k} - \hat{b}) - \hat{D}_{j+1}(0, \hat{b}_{-\Delta k} - \hat{b})]}{2\Delta} \\ &\quad - \frac{\sqrt{J}[\hat{D}_j(0, \hat{b}_{\Delta k} - \hat{b}) - \hat{D}_j(0, \hat{b}_{-\Delta k} - \hat{b})]}{2\Delta} + \hat{R}_{jk} \\ \left|\hat{R}_{jk}\right| &\leq C\sqrt{J}\Delta^{-1}(\left|\hat{b}_{\Delta k} - \hat{b}\right|^2 + \left|\hat{b}_{-\Delta,k} - \hat{b}\right|^2) \leq C\sqrt{J}\Delta K. \end{aligned}$$

We also have

$$\begin{aligned}
\sqrt{J}\frac{1}{\Delta}\hat{D}_{j+1}(0,\hat{b}_{\Delta k}-\hat{b}) &= \sqrt{J}\hat{D}_{j+1}(0,\frac{\hat{b}_{\Delta k}-\hat{b}}{\Delta}),\\
\sqrt{J}|\hat{D}_{j+1}(0,\frac{\hat{b}_{\Delta k}-\hat{b}}{\Delta})-D_{j+1}(0,\frac{\hat{b}_{\Delta k}-\hat{b}}{\Delta})| &\leq C\sqrt{J}\left|\frac{\hat{b}_{\Delta k}-\hat{b}}{\Delta}\right|(|\hat{\delta}-\delta|+|\hat{b}-b|)\leq C\sqrt{J}\sqrt{K}O_p(\varepsilon_{pn}).
\end{aligned}$$

In addition

$$\begin{aligned}
\sqrt{J}D_{j+1}(0,\frac{\hat{b}_{\Delta k}-\hat{b}}{\Delta};\delta,b,\tau_{j+1}) &= \sqrt{J}D(0,\frac{q_{kK}(G(\cdot))}{\hat{\delta}\hat{p}(\cdot)[1-\hat{p}(\cdot)]};\delta,b,\tau_{j+1})+\sqrt{J}\Delta D(0,\hat{R}_k(\cdot,\Delta);\delta,b,\tau_{j+1})\\
&= \sqrt{J}D(0,\frac{q_{kK}(G(\cdot))}{\delta p(\cdot)[1-p(\cdot)]};\delta,b,\tau_{j+1})+\hat{R}_{jk},\\
\left|\hat{R}_{jk}\right| &\leq \sqrt{J}\sqrt{K}O_p(\varepsilon_{pn})+\sqrt{J}K\Delta.
\end{aligned}$$

Combining terms we have

$$\left\|\hat{M}_\beta-M_\beta\right\|=O_p(J\sqrt{K}/\sqrt{S}\Delta+JK\varepsilon_{pn}+JK^{3/2}\Delta)$$

Next, we have

$$\begin{aligned}
&\left\|\hat{M}_\delta\frac{1}{2\hat{\delta}\bar{\tau}n}\sum_{i=1}^n I_p(\hat{p}_i)q_i^{K\prime}-M_\delta\frac{1}{2\delta E[\tau_i]}E[I_p(p_i)q_i^{K\prime}]\right\|\\
\leq\ &\left\|\hat{M}_\delta-M_\delta\right\|\frac{1}{2\hat{\delta}\bar{\tau}}\left(\frac{1}{n}\sum_{i=1}^n I_p(\hat{p}_i)^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^n q_i^{K\prime}q_i^K\right)^{1/2}\\
&+\|M_\delta\|\left\|\frac{1}{2\hat{\delta}\bar{\tau}n}\sum_{i=1}^n I_p(\hat{p}_i)q_i^{K\prime}-\frac{1}{2\delta E[\tau_i]}E[I_p(p_i)q_i^{K\prime}]\right\|\\
=\ &O_p(J\sqrt{K}(\frac{1}{\sqrt{S}\Delta}+\Delta+\varepsilon_{pn}))+O_p(JK\varepsilon_{pn})=O_p(J\sqrt{K}(\frac{1}{\sqrt{S}\Delta}+\Delta+\sqrt{K}\varepsilon_{pn})).
\end{aligned}$$

Combining terms we then have

$$\left\|\hat{D}_\beta-D_\beta\right\|=O_p(J\sqrt{K}/\sqrt{S}\Delta+JK\varepsilon_{pn}+JK^{3/2}\Delta).$$

Next, for $\hat{\pi}=\hat{\Sigma}^{-1}\hat{D}_\beta$ and $\pi=\Sigma^{-1}D_\beta$ note that $\hat{V}_2=\hat{\pi}'\hat{\Lambda}\hat{\pi}$ and $\bar{V}_2=\pi'\Lambda\pi$. Also we have

$$\hat{V}_2-\bar{V}_2=(\hat{\pi}-\pi)'\hat{\Lambda}(\hat{\pi}-\pi)+2\pi'\hat{\Lambda}(\hat{\pi}-\pi)+\pi'(\hat{\Lambda}-\Lambda)\pi.$$

By the law of large number for symmetric matrices, $\left\|\hat{\Sigma}-\Sigma\right\|_{op}=O_p\left(\sqrt{n^{-1}K\ln K}\right)=o_p(1)$, where $\|\cdot\|_{op}$ denotes the operator norm on symmetric matrices. Then by the eigenvalues of $\Sigma$ bounded and bounded away from zero, $\lambda_{\max}(\hat{\Sigma})=O_p(1)$ and $1/\lambda_{\min}(\hat{\Sigma})=O_p(1)$. Let $\tilde{\Lambda}=\frac{1}{n}\sum_i q_i^K q_i^{K\prime}(\gamma_i-p_i)^2$. Note that

$$\begin{aligned}
\hat{\Lambda}-\tilde{\Lambda} &= \frac{1}{n}\sum_i q_i^K q_i^{K\prime}\left[(\gamma_i-\hat{p}_i)^2-(\gamma_i-p_i)^2\right]\leq\frac{1}{n}\sum_i q_i^K q_i^{K\prime}\left|(\gamma_i-\hat{p}_i)^2-(\gamma_i-p_i)^2\right|\\
&\leq C\hat{\Sigma}\max_i|\hat{p}_i-p_i|=\hat{\Sigma}O_p(\varepsilon_{pn}),\ \hat{\Lambda}-\tilde{\Lambda}\geq -C\hat{\Sigma}O_p(\varepsilon_{pn}).
\end{aligned}$$

Also by the law of large numbers for symmetric matrices $\left\|\tilde{\Lambda}-\Lambda\right\|_{op}=O_p\left(\sqrt{n^{-1}K\ln K}\right)$. Therefore by the triangle inequality,

$$\left\|\hat{\Lambda}-\Lambda\right\|_{op}=O_p(\varepsilon_{pn}).$$

It follows that $\lambda_{\max}(\hat{\Lambda})=O_p(1)$, $1/\lambda_{\min}(\hat{\Lambda})=O_p(1)$, and for $\hat{\Upsilon}=\hat{\Lambda}-\Lambda$,

$$\left\|\hat{\Upsilon}\right\|=\sqrt{tr(\hat{\Upsilon}^2)}\leq C\sqrt{J}\left\|\hat{\Lambda}-\Lambda\right\|_{op}=O_p(\sqrt{J}\varepsilon_{pn}).$$

Similarly we have $\left\|\hat{\Sigma}-\Sigma\right\|=O_p(K\sqrt{\ln(K)/n})$. We also have $\|D_\beta\|\leq CJ\sqrt{K}$.Then it follows that for $\varepsilon_{Dn}=J\sqrt{K}/\sqrt{S}\Delta+JK\varepsilon_{pn}+JK^{3/2}\Delta$

$$\|\hat{\pi}-\pi\|\leq\left\|(\hat{D}_\beta-D_\beta)'\hat{\Sigma}^{-1}\right\|+\left\|D_\beta{}'\hat{\Sigma}^{-1}(\Sigma-\hat{\Sigma})\Sigma^{-1}\right\|\leq O_p(\varepsilon_{Dn})+O_p(JK\sqrt{\ln(K)/n})=O_p(\varepsilon_{Dn}).$$

It then follows by the triangle inequality that

$$\begin{aligned} \left\|\hat{V}_2 - \bar{V}_2\right\| &\leq O_p(1)(\|\hat{\pi} - \pi\|^2 + \|\pi\| \|\hat{\pi} - \pi\| + \|\pi\|^2 \left\|\hat{\Lambda} - \Lambda\right\|) \\ &= O_p(J\sqrt{K}\varepsilon_{Dn} + J^2K^2\sqrt{\ln(K)/n}) = O_p(J^2K/\sqrt{S}\Delta + J^2K^{3/2}\varepsilon_{pn} + J^2K\Delta). \end{aligned}$$

By the triangle inequality we then have

$$\left\|\hat{\Omega} - \Omega\right\| = O_p(J^2K/\sqrt{S}\Delta + J^2K\Delta + J^2K^{3/2}\varepsilon_{pn} + JK^{-s_\alpha})$$

It then follows that Assumption iii) is satisfied by Assumption 5 e).

Finally, for Assumption iv) of Lemma 4, note that

$$\left(h_i'h_i\right)^2 = \left(\sum_{j=1}^{J} h_{ij}^2\right)^2 = \sum_{j=1}^{J}\sum_{k=1}^{K} h_{ij}^2 h_{ik}^2 \leq CJ\sum_{j=1}^{J} h_{ij}^4 \leq CJ^4,$$

so that

$$E\left[\left(h_i'h_i\right)^2\right]/nJ \leq CJ^3/n \longrightarrow 0.$$

Therefore condition iv) is satisfied. **Q.E.D.**

## References


1. I Karatzas, SE Shreve, *Brownian Motion and Stochastic Calculus.* (Springer), (1998).
2. D Revuz, M Yor, *Continuous martingales and Brownian motion.* (Springer) Vol. 293, (1999).
3. WK Newey, The asymptotic variance of semiparametric estimators. *Econometrica* **62**, 1349–1382 (1994).
4. WK Newey, F Windmeijer, Generalized method of moments with many weak moment conditions. *Econometrica* **77**, 687–719 (2009).
5. A Belloni, V Chernozhukov, D Chetverikov, K Kato, Some new asymptotic theory for least squares series: Pointwise and uniform results. *J. Econom.* **186**, 345–366 (2015).
6. H Ichimura, WK Newey, The influence function of semiparametric estimators. *mimeo* (2018).