Supplementary Material for "Optimal Conditional Quantile Prediction via Model Averaging of Partially Linear Additive Models"

Jing Lv[1] and Shujie Ma[2]

[1]*School of Mathematics and Statistics, Southwest University, Chongqing, China*

[2]*Department of Statistics, University of California, Riverside, USA*

This document supplements the paper entitled "Optimal Conditional Quantile Prediction via Model Averaging of Partially Linear Additive Models". Section S1 provides additional numerical analysis and discussion. Section S2 contains additional numerical results of the article. Section S3 provides the technical proofs of three lemmas, Proposition 1 and Theorem 1.

S1 Additional numerical analysis and discussion

S1.1 The bootstrap procedure for the model weight estimation

We describe the bootstrap steps as follows. First, we use the original samples $\{Y_i, \boldsymbol{X}_i\}_{i=1}^n$ to build S semiparametric sub-models $\mathbb{M}_1, \cdots, \mathbb{M}_S$ by *Strategy A* or *Strategy B* in Section 4.2. Second, we randomly draw n samples $\{Y_i^\dagger, \boldsymbol{X}_i^\dagger\}_{i=1}^n$ from the original samples $\{Y_i, \boldsymbol{X}_i\}_{i=1}^n$ with replacement. Without loss of generality, we repeat the above sampling process for T times (i.e., T = 400). Based on the t-th bootstrap samples $\{Y_i^{\dagger[\mathrm{t}]}, \boldsymbol{X}_i^{\dagger[\mathrm{t}]}\}_{i=1}^n$, we use the same way of Section 2.1 to obtain the bootstrap weight estimator $\hat{\boldsymbol{w}}^{\dagger[\mathrm{t}]} = \left(\hat{w}_1^{\dagger[\mathrm{t}]}, \cdots, \hat{w}_S^{\dagger[\mathrm{t}]}\right)^\top$ for $1 \leq \mathrm{t} \leq \mathrm{T}$. Note that the model forms of candidate models have been established in the first step and thus are fixed for each bootstrap step when we obtain the bootstrap weight estimator. Third, for $s = 1, \dots, S$, we obtain the standard error of the sth model weight estimator $\hat{w}_s$ by calculating the sample standard deviation of $\hat{w}_s^{\dagger[1]}, \cdots, \hat{w}_s^{\dagger[\mathrm{T}]}$, and acquire the lower and upper bound of $100(1-\alpha)\%$ confidence interval for the sth model weight w_s through computing the $(\alpha/2)$th and $(1-\alpha/2)$th sample quantiles of $\hat{w}_s^{\dagger[1]}, \cdots, \hat{w}_s^{\dagger[\mathrm{T}]}$.

S1.2 Comprehensive comparison between SMAQP and SSMAQP

Next, we examine whether a combination of modeling quantile functions parametrically and integrated loss yields more accurate predictions than that of the direct application of standard quantile regression. Specifically, we compare the proposed SMAQP with the alternative model averaging method SSMAQP at the entire quantile process and a single quantile, where SSMAQP is based on the standard quantile loss function for a single quantile. We refer more details to Remark 3. Note that SMAQP and SSMAQP rely on strategies on building sub-models. Thus two versions for SMAQP and SSMAQP are considered here. Let SMAQP1 (SSMAQP1) and SMAQP2 (SSMAQP2) be the respective model averaging methods with candidate models constructed by *Strategy A* and *Strategy B*.

Figure S1 displays the average OAQPE, winning ratio and Loss to SMAQP2 of SMAQP1, SMAQP2, SSMAQP1 and SSMAQP2 for model II and $t = 3$ in Example

2. From Figure S1, we find that SMAQP1 (or SMAQP2) is superior to SSMAQP1 (or SSMAQP2) in terms of the OAQPE, showing that using the integrated loss can result in better predictions for the entire quantile process. Moreover, SMAQP2 performs best with the highest winning ratio.

In some application areas, one may be interested in a specific quantile rather than the entire quantile process. We stress that the proposed SMAQP (the formula (2.9) of the paper) can be used to predict the conditional quantile function $\mu(\boldsymbol{X}_i, \tau)$ at any interested single quantile as long as basis functions are specified in advance. Therefore, we also make a full comparison for SMAQP and SSMAQP at a single specific quantile. To measure out-of-sample quantile prediction errors at a specific quantile, we just take $K = 1$ in OAQPE defined in Section 5.1 and fix τ_k at any quantile level τ, then OAQPE reduces to $\mid \mathcal{I} \mid^{-1} \sum_{i \in \mathcal{I}} \rho_\tau (Y_i - \hat{\mu}(\boldsymbol{X}_i, \tau)) \triangleq \mathrm{FPE}_\tau$ that is commonly used in existing literature such as (Lu and Su, 2015; Lee and Shin, 2023), where $\mathcal{I}$ is the the testing set with the size $|\mathcal{I}| = 100$. Here the subscript τ in FPE_τ is used to stress the dependence on the quantile level τ. Thus, FPE_τ is adopted to measure the out-of-sample quantile prediction error at a specific quantile in our numerical studies.

Table S1 lists the results of Example 1 with $t = 3$ and $p = 10$ at $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$. Figure S2 plots standard deviations of the out-of-sample quantile prediction error (FPE_τ) of SSMAQP1, SSMAQP2, SMAQP1 and SMAQP2 for model II and $t = 3$ of Example 2. From Table S1, we can see that SMAQP outperforms SSMAQP in the term of FPE_τ. It is found from Figure S2 that SMAQP1 (or SMAQP2) enjoys more stable predictions than that of SSMAQP1 (or SSMAQP2) in general. Also, SMAQP2 dominates the other three approaches due to its smallest standard deviations at most of quantiles. The results for other cases are similar. Thus, we omit those results to save space. Overall, these numerical results indicate that SMAQP outperforms SSMAQP in general when we consider the entire quantile process or a single interested quantile.

S1.3 A sensitivity analysis of SMAQP for $\lambda_n = 1$ and $\log(n)/2$

To investigate whether there is a significant difference in the numerical results when $\lambda_n = 1$ or $\log(n)/2$ is used, we conduct a sensitivity analysis of the proposed SMAQP. For $\lambda_n = \log(n)/2$, recall that SMAQP1 and SMAQP2 are the respective optimal model averaging methods with candidate models constructed by *Strategy A* and *Strategy B*. Similarly, by taking $\lambda_n = 1$, we define that SMAQP1* and SMAQP2* are the proposed model averaging methods for *Strategy A* and *Strategy B* respectively.

Figure S3 plots the 2.5%, 50% and 97.5% quantiles of differences of FPE_τ between SMAQP1 and SMAQP1*(SMAQP2 and SMAQP2*) at various quantiles. We only display the results of Example 1 with $t = 3$ and $p = 10$ because of similar performance in other cases. To save space, we omit to report those results. Figure S3 shows that the differences of FPE_τ between SMAQP1 and SMAQP1* (SMAQP2 and SMAQP2*) are very close to zero for various quantiles, indicating that there is little difference between the prediction results when two types of λ_n are used. Thus, both $\lambda_n = 1$ and $\log(n)/2$ can be used in practice. We fix $\lambda_n = \log(n)/2$ in our numerical studies.

S2 Additional numerical results

Tables S2 and S3 present additional results of the simulation study reported in Subsection 5.1 of the paper. For the Boston housing price example in Subsection 5.2, Tables S4 and S5 report the estimated model weights, their standard errors, and the 95% confidence intervals of the model weights for different basis functions. The standard errors and confidence intervals are computed based on 400 bootstrap samples with the procedure described in Section S1.1. Figure S4 presents the boxplots of OAQPE with

Table S1: The simulation results of SSMAQP1, SSMAQP2, SMAQP1 and SMAQP2 at $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$ in Example 1 with $t = 3$ and $p = 10$.

error	τ	homoscedasticity (*case a*)				heteroscedasticity (*case b*)			
		SSMAQP1	SSMAQP2	SMAQP1	SMAQP2	SSMAQP1	SSMAQP2	SMAQP1	SMAQP2
					Average FPE_τ				
	0.1	0.215	0.211	0.213	0.206	0.219	0.216	0.216	0.210
	0.3	0.411	0.403	0.407	0.398	0.415	0.408	0.413	0.404
	0.5	0.469	0.461	0.466	0.456	0.475	0.467	0.472	0.463
	0.7	0.412	0.404	0.410	0.401	0.417	0.410	0.414	0.406
	0.9	0.217	0.211	0.214	0.208	0.220	0.214	0.216	0.209
					Winning Ratio				
	0.1	8.0%	26.5%	8.5%	57.0%	6.5%	17.5%	18.0%	58.0%
	0.3	5.0%	24.5%	5.0%	65.5%	6.0%	32.5%	6.0%	55.5%
case i	0.5	6.5%	23.0%	5.5%	65.0%	1.5%	32.0%	6.0%	60.5%
	0.7	5.5%	31.5%	8.0%	55.0%	2.5%	31.0%	6.5%	60.0%
	0.9	11.0%	22.5%	7.5%	59.0%	7.0%	30.5%	6.5%	56.0%
					Loss to SMAQP2				
	0.1	77.0%	66.0%	85.0%	NA	79.5%	74.0%	80.0%	NA
	0.3	86.5%	70.5%	92.0%	NA	84.0%	62.5%	87.0%	NA
	0.5	84.0%	71.5%	91.5%	NA	81%	64.0%	91.0%	NA
	0.7	83.5%	64.5%	86.5%	NA	85%	67.0%	92.0%	NA
	0.9	81.0%	67.5%	85.5%	NA	80.0%	63.0%	87.5%	NA
					Average FPE_τ				
	0.1	0.332	0.326	0.323	0.316	0.339	0.333	0.333	0.327
	0.3	0.563	0.556	0.559	0.550	0.571	0.564	0.566	0.558
	0.5	0.628	0.619	0.625	0.615	0.634	0.626	0.631	0.623
	0.7	0.559	0.551	0.556	0.548	0.566	0.559	0.563	0.555
	0.9	0.327	0.321	0.319	0.314	0.331	0.325	0.324	0.318
					Winning Ratio				
	0.1	5.0%	21.5%	15.5%	58.0%	6.0%	29.5%	16.5%	48.0%
	0.3	12.0%	25.0%	6.5%	56.5%	6.5%	31.5%	5.5%	56.5%
case ii	0.5	6.5%	25.0%	9.0%	59.5%	4.5%	34.0%	8.0%	53.5%
	0.7	9.0%	33.0%	7.5%	50.5%	5.5%	32.0%	8.0%	54.5%
	0.9	9.0%	21.5%	19.5%	50.0%	8.5%	22.5%	14.5%	54.5%
					Loss to SMAQP2				
	0.1	86.0%	73.5%	82.0%	NA	75.5%	65.0%	79.0%	NA
	0.3	76.0%	63.0%	83.5%	NA	77.5%	61.5%	86.5%	NA
	0.5	79.5%	65.0%	84.0%	NA	84.0%	61.5%	86.5%	NA
	0.7	70.0%	58.0%	82.5%	NA	76.5%	63.0%	85.0%	NA
	0.9	81.5%	71.5%	75.0%	NA	81.5%	72.5%	82.0%	NA
					Average FPE_τ				
	0.1	0.273	0.269	0.268	0.263	0.288	0.284	0.285	0.280
	0.3	0.483	0.477	0.479	0.472	0.498	0.491	0.493	0.486
	0.5	0.546	0.539	0.543	0.534	0.558	0.55	0.555	0.547
	0.7	0.488	0.48	0.485	0.477	0.499	0.492	0.495	0.487
	0.9	0.285	0.279	0.281	0.275	0.291	0.285	0.286	0.280
					Winning Ratio				
	0.1	11.5%	19.5%	13.0%	56.0%	10.5%	26.0%	15.5%	48.0%
	0.3	6.5%	21.5%	10.5%	61.5%	7.5%	28.5%	6.0%	58.0%
case iii	0.5	3.5%	28.0%	10.5%	58.0%	4.5%	35.5%	7.5%	52.5%
	0.7	4.0%	31.0%	8.5%	56.5%	6.0%	25.0%	7.0%	62.0%
	0.9	10.0%	26.5%	10.0%	53.5%	6.5%	25.0%	14.0%	54.5%
					Loss to SMAQP2				
	0.1	77.5%	71.5 %	79.0%	NA	71.0%	63.5%	80.0%	NA
	0.3	81.5%	71.5%	83.0%	NA	80.5%	64.5%	87.0%	NA
	0.5	85.0%	67.0%	84.0%	NA	81.0%	58.5%	88.5%	NA
	0.7	81.5%	63.5%	86.0%	NA	85.0%	67.5%	87.5%	NA
	0.9	78.0%	65.0%	81.5%	NA	81.0%	69.5%	81.5%	NA

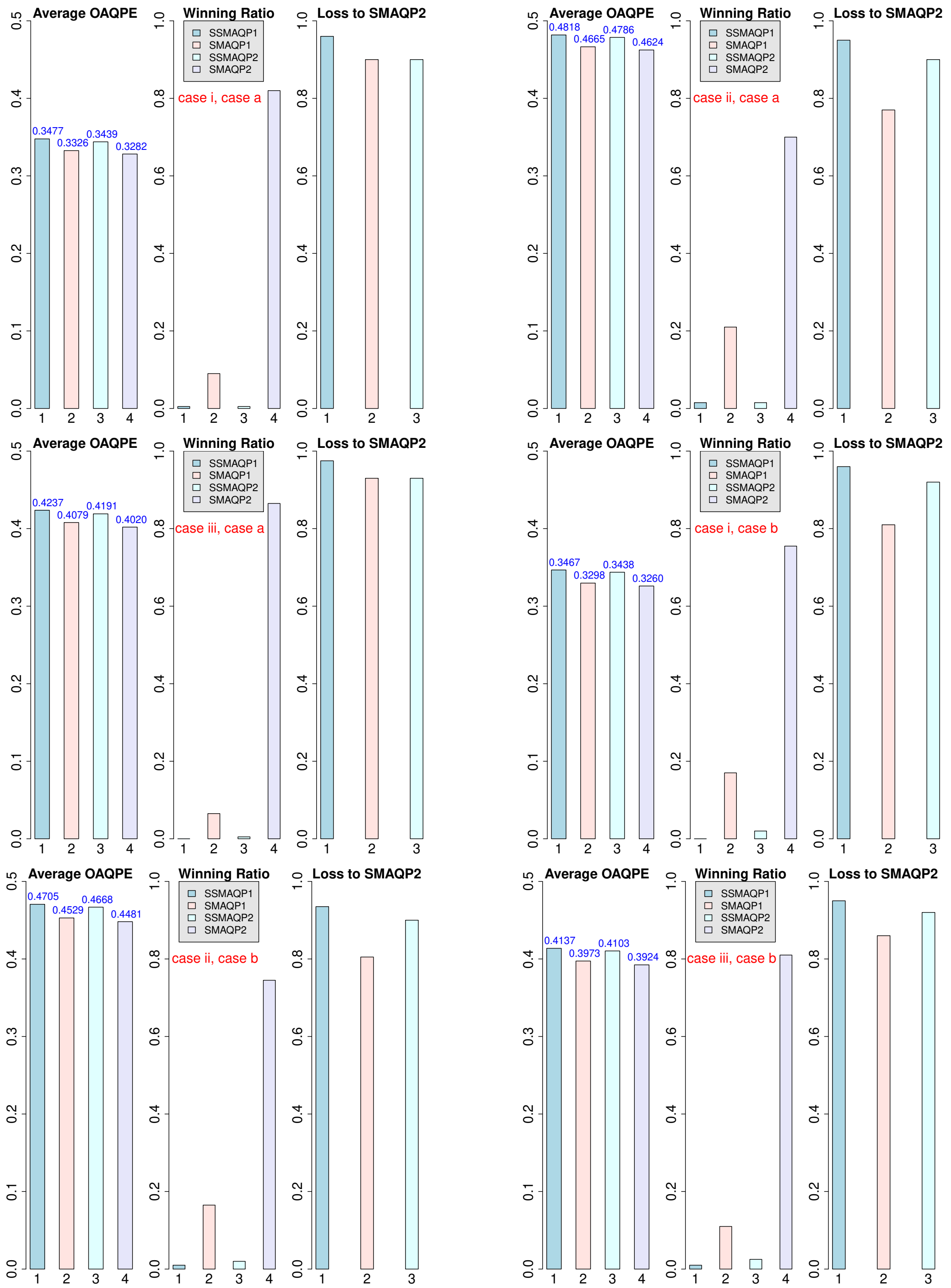

Figure S1: The average OAQPE, winning ratio and Loss to SMAQP2 for model II and $t = 3$ in Example 2.

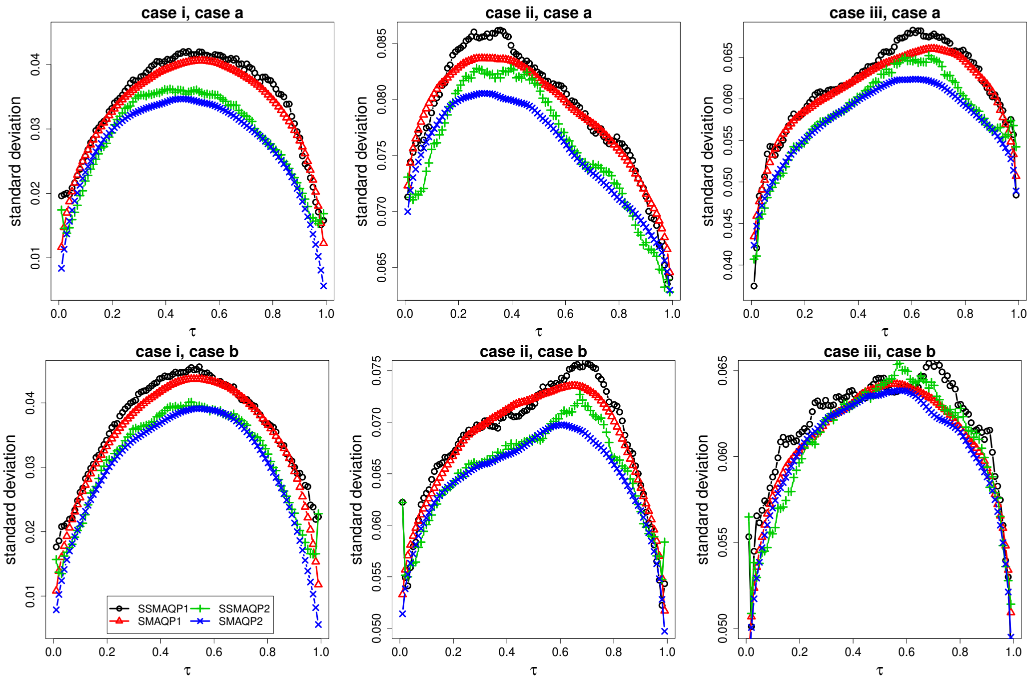

Figure S2: The standard deviations of FPE$_\tau$s for SSMAQP1, SSMAQP2, SMAQP1 and SMAQP2 at various quantiles for model II and $t = 3$ in Example 2.

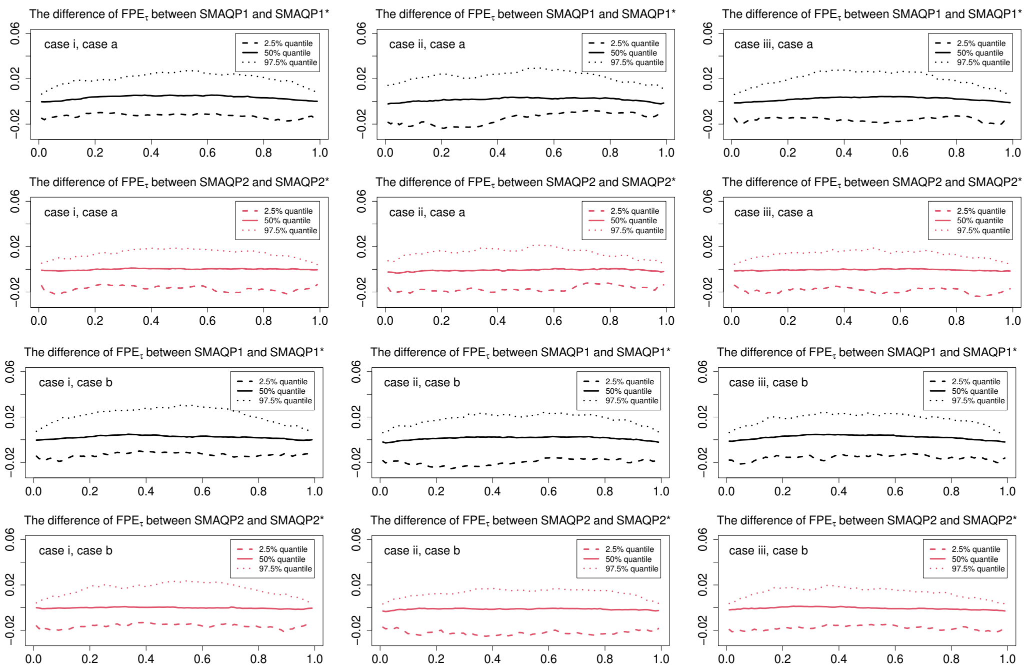

Figure S3: The 2.5%, 50% and 97.5% quantiles of differences of FPE$_\tau$ between SMAQP1 and SMAQP1*(SMAQP2 and SMAQP2*) at various quantiles in Example 1 with $p = 10$ and $t = 3$.

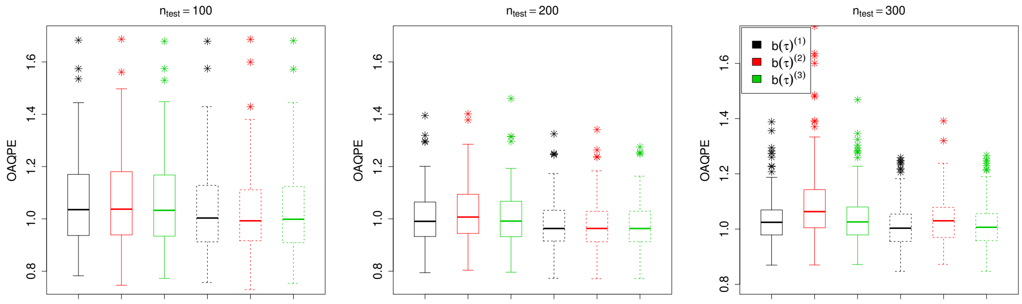

Figure S4: The OAQPEs of SMAQP1 (solid line) and SMAQP2 (dotted line) with three different bias functions $\boldsymbol{b}(\tau)^{(1)}$, $\boldsymbol{b}(\tau)^{(2)}$ and $\boldsymbol{b}(\tau)^{(3)}$ for the Boston housing price data.

different test sample sizes and basis functions. Furthermore, Figures S5 and S6 plot the estimates of the regression coefficients and the corresponding 95% pointwise confidence intervals of the coefficients for the 5th candidate model and the 28th candidate model, respectively. We can clearly see that almost all estimated coefficients are smooth over the quantile levels, indicating that modeling quantile regression coefficients as smooth functions of the quantile level is reasonable in practice. The plots of the regression coefficients for other candidate models show similar patterns, so we omit the results to save space.

Table S2: Simulation results of OAQPE and $\hat{\boldsymbol{w}} = (\hat{w}_1, \cdots, \hat{w}_{10})^\top$ for SMAQP1 with $t = 1$, $p = 10$ and homoscedasticity (*case a*) for different errors and basis functions.

Example	error	basis	OAQPE	$\hat{w}_1$	$\hat{w}_2$	$\hat{w}_3$	$\hat{w}_4$	$\hat{w}_5$	$\hat{w}_6$	$\hat{w}_7$	$\hat{w}_8$	$\hat{w}_9$	$\hat{w}_{10}$
1	case i	$\boldsymbol{b}(\tau)^{(1)}$	0.342	0.048	0.812	0.096	0.045	0	0	0	0	0	0
			(0.038)	(0.006)	(0.152)	(0.143)	(0.094)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.342	0.043	0.815	0.096	0.046	0	0	0	0	0	0
			(0.038)	(0.006)	(0.151)	(0.143)	(0.095)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.342	0.048	0.811	0.096	0.045	0	0	0	0	0	0
			(0.038)	(0.006)	(0.152)	(0.144)	(0.095)	(0)	(0)	(0)	(0)	(0)	(0)
	case ii	$\boldsymbol{b}(\tau)^{(1)}$	0.468	0.062	0.846	0.074	0.018	0	0	0	0	0	0
			(0.055)	(0.015)	(0.132)	(0.123)	(0.059)	(0.004)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.467	0.056	0.852	0.073	0.019	0	0	0	0	0	0
			(0.055)	(0.016)	(0.131)	(0.123)	(0.058)	(0.005)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.467	0.062	0.843	0.076	0.018	0	0	0	0	0	0
			(0.055)	(0.015)	(0.134)	(0.125)	(0.060)	(0.004)	(0)	(0)	(0)	(0)	(0)
	case iii	$\boldsymbol{b}(\tau)^{(1)}$	0.413	0.052	0.816	0.096	0.035	0	0	0	0	0	0
			(0.051)	(0.009)	(0.151)	(0.139)	(0.089)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.413	0.046	0.820	0.099	0.035	0	0	0	0	0	0
			(0.051)	(0.009)	(0.152)	(0.142)	(0.083)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.413	0.053	0.815	0.097	0.035	0	0	0	0	0	0
			(0.051)	(0.009)	(0.151)	(0.140)	(0.089)	(0)	(0)	(0)	(0)	(0)	(0)
2(Model I)	case i	$\boldsymbol{b}(\tau)^{(1)}$	0.439	0.621	0.256	0.093	0.023	0.006	0	0	0	0	0
			(0.074)	(0.249)	(0.231)	(0.151)	(0.072)	(0.041)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.439	0.614	0.258	0.098	0.023	0.006	0	0	0	0	0
			(0.074)	(0.247)	(0.228)	(0.156)	(0.072)	(0.041)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.439	0.621	0.256	0.094	0.023	0.006	0	0	0	0	0
			(0.074)	(0.249)	(0.231)	(0.151)	(0.072)	(0.042)	(0)	(0)	(0)	(0)	(0)
	case ii	$\boldsymbol{b}(\tau)^{(1)}$	0.541	0.691	0.217	0.069	0.023	0.001	0	0	0	0	0
			(0.077)	(0.256)	(0.238)	(0.144)	(0.076)	(0.006)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.541	0.686	0.220	0.072	0.022	0.001	0	0	0	0	0
			(0.077)	(0.260)	(0.236)	(0.144)	(0.073)	(0.012)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.541	0.690	0.217	0.069	0.024	0.001	0	0	0	0	0
			(0.077)	(0.256)	(0.238)	(0.145)	(0.077)	(0.005)	(0)	(0)	(0)	(0)	(0)
	case iii	$\boldsymbol{b}(\tau)^{(1)}$	0.509	0.628	0.238	0.104	0.026	0.004	0	0	0	0	0
			(0.077)	(0.257)	(0.246)	(0.160)	(0.072)	(0.024)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.508	0.629	0.240	0.099	0.028	0.004	0	0	0	0	0
			(0.077)	(0.256)	(0.245)	(0.152)	(0.073)	(0.024)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.508	0.627	0.239	0.104	0.025	0.005	0	0	0	0	0
			(0.077)	(0.257)	(0.246)	(0.160)	(0.070)	(0.024)	(0)	(0)	(0)	(0)	(0)
2(Model II)	case i	$\boldsymbol{b}(\tau)^{(1)}$	0.349	0.850	0.129	0.021	0	0	0	0	0	0	0
			(0.027)	(0.157)	(0.147)	(0.067)	(0.002)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.349	0.841	0.135	0.023	0	0	0	0	0	0	0
			(0.027)	(0.159)	(0.149)	(0.070)	(0.004)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.349	0.853	0.126	0.021	0	0	0	0	0	0	0
			(0.027)	(0.157)	(0.147)	(0.067)	(0.002)	(0)	(0)	(0)	(0)	(0)	(0)
	case ii	$\boldsymbol{b}(\tau)^{(1)}$	0.472	0.890	0.099	0.0110	0	0	0	0	0	0	0
			(0.053)	(0.148)	(0.143)	(0.044)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.471	0.884	0.102	0.0140	0	0	0	0	0	0	0
			(0.053)	(0.151)	(0.144)	(0.049)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.471	0.891	0.097	0.0110	0	0	0	0	0	0	0
			(0.053)	(0.148)	(0.142)	(0.043)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
	case iii	$\boldsymbol{b}(\tau)^{(1)}$	0.415	0.865	0.120	0.016	0	0	0	0	0	0	0
			(0.049)	(0.163)	(0.159)	(0.057)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(2)}$	0.414	0.858	0.126	0.016	0	0	0	0	0	0	0
			(0.049)	(0.167)	(0.162)	(0.056)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
		$\boldsymbol{b}(\tau)^{(3)}$	0.414	0.867	0.117	0.016	0	0	0	0	0	0	0
			(0.049)	(0.162)	(0.158)	(0.057)	(0)	(0)	(0)	(0)	(0)	(0)	(0)

Note: The standard errors of OAQPE and $\hat{\boldsymbol{w}} = (\hat{w}_1, \cdots, \hat{w}_{10})^\top$ are denoted inside the parentheses.

Table S3: Simulation results over different settings in Example 2.

error	t,model	homoscedasticity (*case a*)					heteroscedasticity (*case b*)				
		QRCM	PAQRM	JQLMA	SMAQP1	SMAQP2	QRCM	PAQRM	JQLMA	SMAQP1	SMAQP2
		Average OAQPE									
case i	0,I	0.395	0.407	0.393	0.398	0.392	0.399	0.408	0.397	0.402	0.396
	1,I	0.458	0.461	0.458	0.444	0.429	0.455	0.456	0.457	0.440	0.425
	3,I	0.703	0.472	0.701	0.464	0.441	0.697	0.450	0.695	0.438	0.414
	0,II	0.363	0.378	0.366	0.364	0.360	0.357	0.372	0.360	0.357	0.353
	1,II	0.356	0.356	0.356	0.350	0.338	0.353	0.357	0.353	0.349	0.339
	3,II	0.456	0.357	0.453	0.346	0.328	0.446	0.347	0.444	0.341	0.323
		Winning Ratio									
	0,I	13%	15.5%	28.5%	10.0%	33.0%	13.0%	19.0%	28.5%	14.5%	24.0%
	1,I	14.5%	13.0%	5.0%	6.5%	60.0%	14.5%	10.5%	6.5%	6.5%	61.0%
	3,I	0.5%	13.0%	0.5%	10.0%	74.5%	0%	10.0%	0%	12.0%	78.0%
	0,II	10.5%	2.5%	10.5%	16.0%	60.0%	12.5%	2.0%	4.0%	20.0%	60.5%
	1,II	4.5%	8.5%	6.0%	9.5%	71.5%	9.5%	5.0%	8.0%	11.5%	65.5%
	3,II	0%	4.0%	0.5%	6.0%	89.0%	0%	8.0%	0%	7.0%	84.0%
		Loss to SMAQP2									
	0,I	57.5%	80.0%	57.5%	82.0%	NA	56.5%	73.0%	54.0%	77.0%	NA
	1,I	80.0%	86.5%	82.0%	91.0%	NA	79.0%	85.0%	80.5%	91.0%	NA
	3,I	98.5%	85.0%	98.5%	86.5%	NA	100%	88.0%	100%	86.0%	NA
	0,II	77.0%	96.5%	86.0%	80.0%	NA	80.5%	97.0%	90.0%	77.0%	NA
	1,II	90.0%	90.5%	92.0%	89.0%	NA	86.0%	93.5%	85.0%	87.5%	NA
	3,II	99.0%	95.0%	90.0%	91.0%	NA	99.0%	92.0%	99.0%	92.0%	NA
		Average OAQPE									
case ii	0,I	0.519	0.521	0.516	0.516	0.510	0.515	0.523	0.512	0.519	0.513
	1,I	0.565	0.584	0.566	0.554	0.542	0.561	0.578	0.563	0.552	0.540
	3,I	0.780	0.569	0.777	0.560	0.537	0.775	0.580	0.774	0.563	0.539
	0,II	0.484	0.510	0.486	0.486	0.482	0.483	0.511	0.485	0.484	0.481
	1,II	0.488	0.495	0.488	0.482	0.472	0.486	0.495	0.486	0.479	0.472
	3,II	0.563	0.485	0.559	0.469	0.453	0.560	0.488	0.558	0.476	0.459
		Winning Ratio									
	0,I	12.5%	21.5%	26.5%	8.0%	31%	9.0%	16.5%	33.0%	12.0%	28.5%
	1,I	12.0%	4.5%	8.5%	10.5%	64%	12.0%	6.5%	8.0%	10.0%	63.0%
	3,I	0.5%	10.5%	0%	4.5%	84%	0.5%	6%	0.5%	10.5%	82.0%
	0,II	20.5%	2.5%	14.0%	18.5%	44%	18.5%	1.0%	10.5%	21.5%	47.5%
	1,II	5.5%	1.0%	8.0%	14.0%	71.0%	7.0%	1.0%	4.0%	16.5%	71.0%
	3,II	0%	1.0%	0.5%	10.5%	87.0%	0%	3.0%	0%	10.0%	86.5%
		Loss to SMAQP2									
	0,I	60.5%	70.5%	56.0%	80.5%	NA	52.5%	75.0%	47.0%	75.0%	NA
	1,I	80.0%	92.5%	80.5%	88.5%	NA	80.5%	92.0%	82.0%	88.5%	NA
	3,I	99.0%	88.5%	99.0%	92.0%	NA	99.0%	92.0%	98.5%	88.0%	NA
	0,II	62.0%	97.0%	74.0%	75.0%	NA	69.0%	97.0%	77.0%	73.5%	NA
	1,II	89.0%	97.0%	87.0%	84.5%	NA	88.5%	99.5%	89.5%	81.5%	NA
	3,II	99.0%	96.5%	99.0%	89.0%	NA	100%	96.0%	100%	89.0%	NA
		Average OAQPE									
case iii	0,I	0.468	0.477	0.466	0.471	0.466	0.453	0.461	0.451	0.452	0.448
	1,I	0.513	0.517	0.514	0.496	0.484	0.527	0.534	0.528	0.509	0.498
	3,I	0.748	0.529	0.745	0.517	0.493	0.758	0.523	0.755	0.513	0.491
	0,II	0.434	0.458	0.437	0.436	0.431	0.429	0.454	0.432	0.430	0.426
	1,II	0.432	0.439	0.432	0.423	0.415	0.421	0.431	0.422	0.419	0.408
	3,II	0.530	0.441	0.526	0.425	0.409	0.509	0.421	0.507	0.411	0.393
		Winning Ratio									
	0,I	9.5%	22.5%	31.5%	12.0%	24.0%	11.0%	16.0%	26.0%	15.0%	31.0%
	1,I	8.5%	6.5%	6.0%	13.0%	65.0%	11.0%	10.0%	5.0%	12.0%	61.0%
	3,I	0%	15.0%	0.5%	6.5%	77.0%	0%	14.5%	0.5%	9.0%	75.0%
	0,II	16.5%	1.0%	11.0%	15.0%	55.0%	18.0%	1.5%	9.0%	14.5%	56.5%
	1,II	7.0%	2.0%	6.5%	14.5%	69.5%	8.0%	3.5%	7.5%	12.5%	68.0%
	3,II	0%	3.0%	0%	10.0%	86.5%	0%	5.0%	1.0%	8.0%	85.0%
		Loss to SMAQP2									
	0,I	58.5%	73.0%	54.5%	81.0%	NA	60.0%	77.0%	58.5%	77.0%	NA
	1,I	85.0%	93.0%	84.5%	85.0%	NA	85.0%	90.0%	84.5%	86.0%	NA
	3,I	98.5%	83.0%	99.0%	91.0%	NA	99.0%	82.5%	99.0%	87.0%	NA
	0,II	74.0%	98.0%	79.0%	80.5%	NA	72.0%	97.0%	80.5%	81.5%	NA
	1,II	87.0%	97.0%	88.5%	81.0%	NA	84.5%	96.0%	87.5%	86.7%	NA
	3,II	100%	95.0%	100%	89.0%	NA	98.5%	94.0%	98.5%	92.0%	NA

Table S4: The estimated model weights are obtained from SMAQP1 and their standard errors (in round brackets) and 95% confidence intervals (in square brackets) based on the bootstrap resampling method for the Boston housing price data.

	$\boldsymbol{b}(\tau)^{(1)}$	$\boldsymbol{b}(\tau)^{(2)}$	$\boldsymbol{b}(\tau)^{(3)}$
$\hat{w}_1$	0.3033 (0.0957)	0.3503 (0.2435)	0.3129 (0.1037)
	[0.0923,0.4842]	[0,0.8916]	[0.0756,0.4534]
$\hat{w}_2$	0.1761 (0.1612)	0.1697 (0.2560)	0.2356 (0.1643)
	[0,0.5279]	[0, 0.9610]	[0, 0.5614]
$\hat{w}_3$	0 (0.1228)	0.0163 (0.1960)	0 (0.1311)
	[0,0.4762]	[0,0.6029]	[0, 0.5827]
$\hat{w}_4$	0 (0.0329)	0 (0.0902)	0 (0.0512)
	[0,0.3000]	[0,0.3732]	[0, 0.2008]
$\hat{w}_5$	0.5205 (0.2389)	0.2486 (0.1386)	0.4515 (0.2245)
	[0,0.6871]	[0, 0.4516]	[0, 0.7208]
$\hat{w}_6$	0 (0.0577)	0 (0.0866)	0 (0.0844)
	[0,0.3767]	[0, 0.3022]	[0, 0.4927]
$\hat{w}_7$	0 (0.0850)	0 (0.0789)	0 (0.1078)
	[0,0.2591]	[0, 0.2812]	[0, 0.3632]
$\hat{w}_8$	0 (0.1373)	0.2151 (0.0680)	0 (0.1405)
	[0,0.4374]	[0, 0.2054]	[0, 0.4722]
$\hat{w}_9$	0 (0.0398)	0 (0.0574)	0 (0.0480)
	[0,0.2318]	[0, 0.0993]	[0, 0.2411]
$\hat{w}_{10}$	0 (0.0450)	0 (0.0298)	0 (0.0587)
	[0,0.2153]	[0, 0.1256]	[0, 0.2041]

Table S5: The estimated model weights are obtained from SMAQP2 and their standard errors (in round brackets) and 95% confidence intervals (in square brackets) based on the bootstrap resampling method for the Boston housing price data.

	$\boldsymbol{b}(\tau)^{(1)}$	$\boldsymbol{b}(\tau)^{(2)}$	$\boldsymbol{b}(\tau)^{(3)}$		$\boldsymbol{b}(\tau)^{(1)}$	$\boldsymbol{b}(\tau)^{(2)}$	$\boldsymbol{b}(\tau)^{(3)}$
$\hat{w}_1$	0.2564 (0.1189)	0.0743 (0.1846)	0 (0.1306)	$\hat{w}_{16}$	0 (0.0808)	0.1253 (0.0866)	0 (0.1089)
	[0 ,0.4222]	[0,0.6381]	[0,0.4360]		[0, 0.3336]	[0,0.2985]	[0, 0.3610]
$\hat{w}_2$	0 (0.0533)	0.0953 (0.0950)	0 (0.0394)	$\hat{w}_{17}$	0 (0.0454)	0 (0.0445)	0 (0.0381)
	[0,0.1280]	[0,0.2201]	[0,0.1214]		[0, 0.1845]	[0,0.2701]	[0, 0.2549]
$\hat{w}_3$	0 (0)	0 (0.0116)	0 (0)	$\hat{w}_{18}$	0 (0.0647)	0 (0.0895)	0.0262 (0.0787)
	[0, 0]	[0,0.0135]	[0, 0]		[0,0.2601]	[0,0.2213]	[0, 0.3418]
$\hat{w}_4$	0 (0)	0 (0)	0 (0)	$\hat{w}_{19}$	0 (0.0246)	0 (0.0802)	0 (0.0276)
	[0, 0]	[0, 0]	[0, 0]		[0, 0]	[0,0.1802]	[0, 0.0001]
$\hat{w}_5$	0 (0)	0 (0.0030)	0 (0.0000)	$\hat{w}_{20}$	0 (0.1124)	0.0321 (0.0807)	0 (0.1129)
	[0, 0]	[0, 0]	[0, 0]		[0,0.4246]	[0,0.3764]	[0, 0.3773]
$\hat{w}_6$	0.0046 (0.1412)	0.1072 (0.1571)	0.2841 (0.1363)	$\hat{w}_{21}$	0 (0.1238)	0.4547 (0.1202)	0.2915 (0.1211)
	[0,0.5022]	[0,0.5345]	[0,0.5468]		[0, 0.3967]	[0,0.3235]	[0, 0.3616]
$\hat{w}_7$	0.2102 (0.1037)	0 (0.1078)	0.2204 (0.1099)	$\hat{w}_{22}$	0 (0)	0 (0.0478)	0 (0.0001)
	[0,0.4206]	[0,0.4117]	[0,0.4356]		[0, 0]	[0,0.1610]	[0,0]
$\hat{w}_8$	0 (0.0289)	0 (0.0716)	0 (0.0399)	$\hat{w}_{23}$	0 (0.0202)	0.0347 (0.0715)	0 (0.0238)
	[0,0.0664]	[0,0.2816]	[0,0.1070]		[0,0.1188]	[0,0.1393]	[0, 0.0914]
$\hat{w}_9$	0 (0.0621)	0 (0.0993)	0 (0.0806)	$\hat{w}_{24}$	0.1500 (0.0718)	0 (0.0874)	0.1778 (0.0759)
	[0,0.3142]	[0,0.3072]	[0,0.3257]		[0, 0.2296]	[0,0.2779]	[0, 0.2445]
$\hat{w}_{10}$	0 (0.0212)	0 (0.0324)	0 (0.0209)	$\hat{w}_{25}$	0 (0.0005)	0 (0.0176)	0 (0.0004)
	[0,0.0712]	[0,0.1782]	[0,0.0773]		[0, 0]	[0,0.0634]	[0,0]
$\hat{w}_{11}$	0 (0.0221)	0 (0.0581)	0 (0.0143)	$\hat{w}_{26}$	0 (0.0347)	0 (0.0389)	0 (0.0195)
	[0,0.1033]	[0, 0.1614]	[0,0.1023]		[0, 0]	[0,0.0895]	[0,0]
$\hat{w}_{12}$	0 (0.0124)	0 (0.0326)	0 (0.0136)	$\hat{w}_{27}$	0 (0.0620)	0 (0.0347)	0 (0.0469)
	[0 ,0.1169]	[0,0.2064]	[0,0.0673]		[0,0.1901]	[0,0.2446]	[0, 0.2402]
$\hat{w}_{13}$	0 (0)	0 (0.0024)	0 (0)	$\hat{w}_{28}$	0.3788 (0.1526)	0 (0.0994)	0 (0.1770)
	[0, 0]	[0, 0]	[0, 0]		[0, 0.5431]	[0, 0.2376]	[0, 0.5025]
$\hat{w}_{14}$	0 (0)	0 (0.0067)	0 (0.0005)	$\hat{w}_{29}$	0 (0.1314)	0 (0.0950)	0 (0.1355)
	[0, 0]	[0,0.0188]	[0, 0]		[0,0.4275]	[0,0.2356]	[0, 0.3651]
$\hat{w}_{15}$	0 (0)	0 (0.0168)	0 (0.0001)	$\hat{w}_{30}$	0 (0.0566)	0.0766 (0.0616)	0 (0.0595)
	[0, 0]	[0,0.0583]	[0, 0]		[0,0.1772]	[0,0.2234]	[0, 0.1772]
				$\hat{w}_{31}$	0 (0.1265)	0 (0.0607)	0 (0.1160)
					[0,0.3223]	[0, 0.2076]	[0, 0.3936]

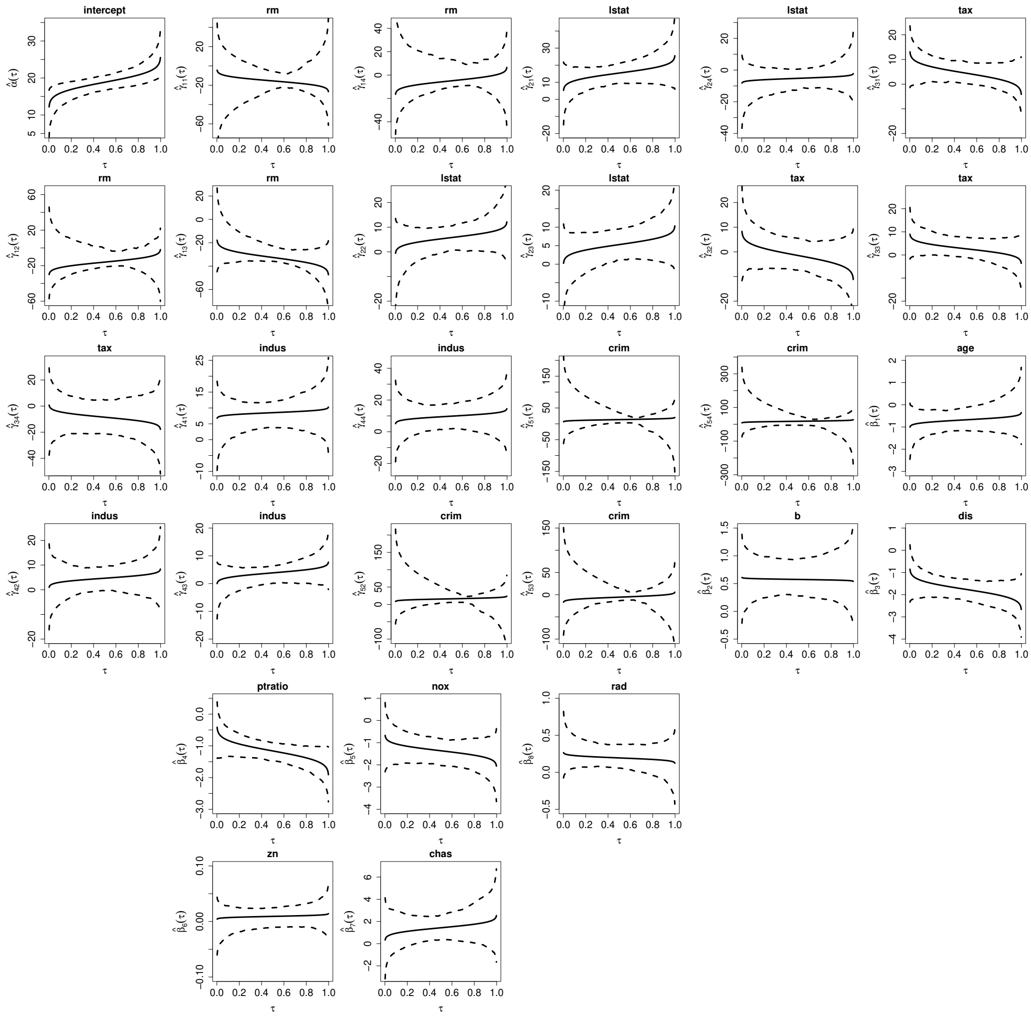

Figure S5: For the Boston housing price data. The estimated regression coefficients (solid line) and corresponding 95% pointwise confidence intervals (dashed line) of the 5th candidate model for SMAQP1 at various quantiles. In the 5th candidate model, *rm*, *lstat*, *tax*, *indus* and *crim* are taken as the nonlinear parts and the rest variables are taken as linear parts.

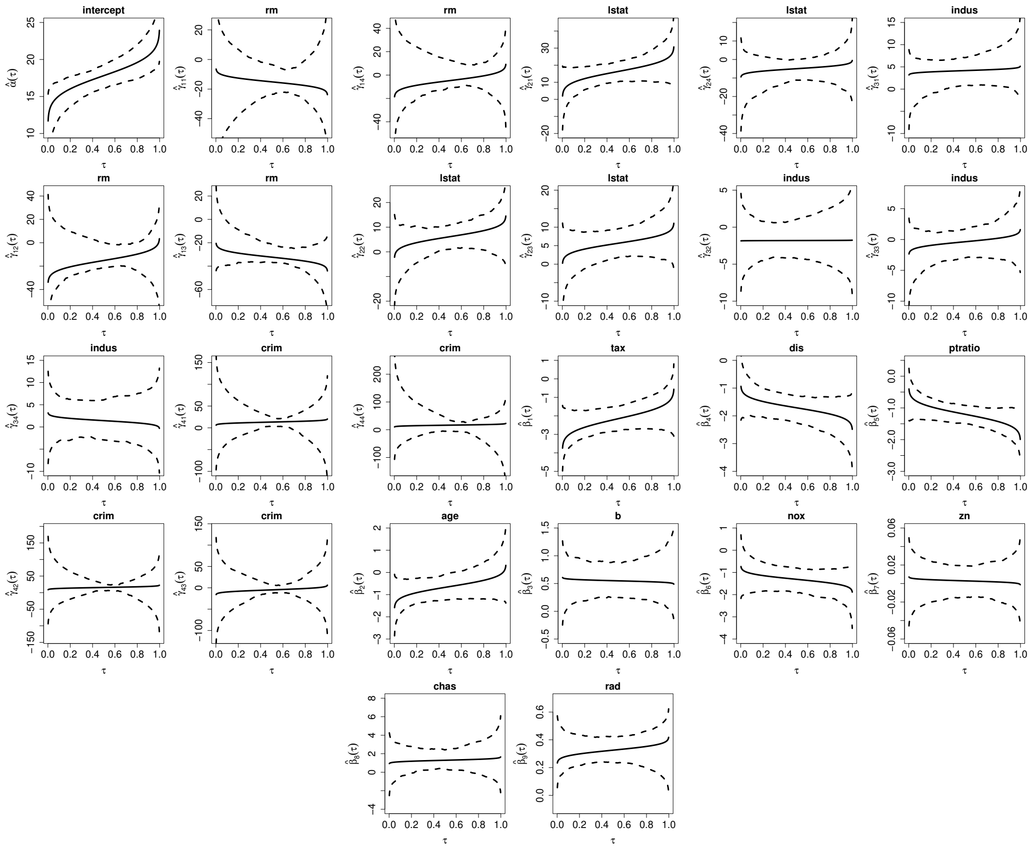

Figure S6: For the Boston housing price data. The estimated regression coefficients (solid line) and corresponding 95% pointwise confidence intervals (dashed line) of the 28th candidate model for SMAQP2 at various quantiles. In the 28th candidate model, *rm*, *lstat*, *indus* and *crim* are taken as the nonlinear parts and the rest variables are taken as linear parts.

S3 The technical proofs of three lemmas, Proposition 1 and Theorem 1

We use C, C_0, C_1, C_2 to denote generic positive constants whose values may vary from line to line.

Lemma 1. *(Bernstein's inequality, Lemma 2.2.9, Van der Vaart and Wellner (1996)) For independent random variables $Y_1, ..., Y_n$ with 0 means and bounded ranges $[-M, M]$, then*

$$P\left\{|Y_1 + \cdots + Y_n| > x\right\} \leq 2\exp\left\{-\frac{x^2}{2\left(v + Mx/3\right)}\right\}$$

for $v \geq Var\left(Y_1 + \cdots + Y_n\right)$.

Lemma 2. *Suppose conditions (C1)–(C2) and (C5) (i) hold. Then, we have*

$$\max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| d\tau = O_p\left(p^{1/2}\right).$$

Proof of Lemma 2 By applying the triangle inequality, we have

$$\begin{aligned}
&\max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| d\tau \\
\leq &\max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 E\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| d\tau + \max_{1\leq s\leq S}\left|\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left[\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| - E\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|\right] d\tau\right| \\
\triangleq &\Delta_1 + \Delta_2.
\end{aligned}$$

We first consider Δ_1. Based on the definitions of $\boldsymbol{D}_i^{(s)}(\tau)$ and $\boldsymbol{Z}_i^{(s)}$, we have $\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{D}_i^{(s)}(\tau) = \sum_{k=1}^{K} b_k^2(\tau)\boldsymbol{Z}_i^{(s)\top}\boldsymbol{Z}_i^{(s)}$ with $\boldsymbol{Z}_i^{(s)\top}\boldsymbol{Z}_i^{(s)} = \sum_{j=1}^{p_s}\boldsymbol{\psi}_j^\top\left(\bar{X}_{ij}^{(s)}\right)\boldsymbol{\psi}_j\left(\bar{X}_{ij}^{(s)}\right) + \frac{1}{J_n}\left(1 + \boldsymbol{X}_{i\mathcal{A}_s^c}^\top\boldsymbol{X}_{i\mathcal{A}_s^c}\right)$. Then we have

$$\begin{aligned}
\Delta_1 &= \max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 E\left\{\left[\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{D}_i^{(s)}(\tau)\right]^{1/2}\right\} d\tau \\
&\leq \max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\{E\left[\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{D}_i^{(s)}(\tau)\right]\right\}^{1/2} d\tau \\
&= \max_{1\leq s\leq S} \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\{E\left[\sum_{k=1}^{K} b_k^2(\tau)\boldsymbol{Z}_i^{(s)\top}\boldsymbol{Z}_i^{(s)}\right]\right\}^{1/2} d\tau \\
&= \max_{1\leq s\leq S} \int_0^1 \left\{\left[\sum_{k=1}^{K} b_k^2(\tau) E\left(\boldsymbol{Z}_i^{(s)\top}\boldsymbol{Z}_i^{(s)}\right)\right]\right\}^{1/2} d\tau \\
&= \max_{1\leq s\leq S} \int_0^1 \left[\sum_{k=1}^{K} b_k^2(\tau)\left\{\sum_{j=1}^{p_s} E\left[\boldsymbol{\psi}_j^\top\left(\bar{X}_{ij}^{(s)}\right)\boldsymbol{\psi}_j\left(\bar{X}_{ij}^{(s)}\right)\right] + \frac{1}{J_n}E\left(1 + \boldsymbol{X}_{i\mathcal{A}_s^c}^\top\boldsymbol{X}_{i\mathcal{A}_s^c}\right)\right\}\right]^{1/2} d\tau \\
&\leq \max_{1\leq s\leq S} C\left[C_0 p_s + \frac{1}{J_n}C_2 p\right]^{1/2} \\
&= O\left(p^{1/2}\right),
\end{aligned}$$

where the last inequality uses the following $E\left(1 + \boldsymbol{X}_{i\mathcal{A}_s^c}^\top\boldsymbol{X}_{i\mathcal{A}_s^c}\right) = tr\left(E\left[\left(1, \boldsymbol{X}_{\mathcal{A}_s^c}^\top\right)^\top\left(1, \boldsymbol{X}_{\mathcal{A}_s^c}^\top\right)\right]\right) \leq$

$C_2 p$ by the condition (C1) (iii), $E\left\{\boldsymbol{\psi}_j^\top\left(\bar{X}_{ij}^{(s)}\right)\boldsymbol{\psi}_j\left(\bar{X}_{ij}^{(s)}\right)\right\} = tr\left[E\left\{\boldsymbol{\psi}_j\left(\bar{X}_{ij}^{(s)}\right)\boldsymbol{\psi}_j^\top\left(\bar{X}_{ij}^{(s)}\right)\right\}\right] \leq C_0$ by the condition (C2), and $\int_0^1\left\{\sum_{k=1}^K b_k^2(\tau)\right\}^{1/2} d\tau \leq C$ by the fact that some $b_k(\tau)$ may be unbounded at $\tau = 0$ or $\tau = 1$, but $\int_0^1 b_k(\tau)\, d\tau$ is finite (Frumento and Bottai , 2016).

Next, we consider Δ_2. Let $v_i^{(s)} = \int_0^1\left[\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| - E\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|\right]d\tau$. Note that $E\left(v_i^{(s)}\right) = 0$, $Var\left(v_i^{(s)}\right) \leq C_1 p$ and $\left|v_i^{(s)}\right| \leq C_2 p^{1/2}$ for all $s = 1, \cdots, S$. Then for any $\delta > 0$, we have

$$
\begin{aligned}
P\{\Delta_2 \geq \delta\} =& P\left\{\max_{1\leq s\leq S}\left|\frac{1}{n}\sum_{i=1}^n v_i^{(s)}\right| \geq \delta\right\} \\
\leq& \sum_{s=1}^S P\left\{\left|\frac{1}{n}\sum_{i=1}^n v_i^{(s)}\right| \geq \delta\right\} \\
\leq& S \max_{1\leq s\leq S} P\left\{\left|\frac{1}{n}\sum_{i=1}^n v_i^{(s)}\right| \geq \delta\right\} \\
\leq& 2S\exp\left\{-\frac{n\delta^2}{2C_1 p + 2\delta C_2 p^{1/2}/3}\right\} \\
=& 2\exp\left\{-\frac{n\delta^2}{2C_1 p + 2\delta C_2 p^{1/2}/3} + \log S\right\} \\
=& o(1),
\end{aligned}
$$

where the first inequality uses Boole's inequality, the third inequality uses the Bernstein inequality and the last equality holds because of $p\log S/n = o(p\bar{\phi}J_n\log n/n) = o(1)$ by the condition (C5) (i). Thus we complete the proof. □

Lemma 3. *Suppose conditions (C1)–(C5)(i) hold. Let $\boldsymbol{G}_n^{(s)}$ be an $l_s \times K\phi_s$ matrix with a fixed positive integer l_s, and $\boldsymbol{G}^{(s)} = \lim_{n\to\infty}\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}$ exists and is positive definite. Then, we have*

(i) $\left\|\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right\| = O_p\left(\sqrt{J_n\phi_s/n}\right)$;

(ii) $\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left[\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right] \xrightarrow{d} N\left(\boldsymbol{0}, \boldsymbol{G}^{(s)}\right)$,

where $\boldsymbol{V}_n^{(s)} = \left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{C}_n^{(s)}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}$, $\boldsymbol{A}_n^{(s)} = \sum_{i=1}^n\left[\int_0^1 f\left(-u_i^{(s)}|\boldsymbol{X}_i\right)\boldsymbol{D}_i^{(s)}(\tau)\boldsymbol{D}_i^{(s)\top}(\tau)\,d\tau\right]$ *and* $\boldsymbol{C}_n^{(s)} = \sum_{i=1}^n Cov\left[\int_0^1\psi_\tau\left(\varepsilon_i + u_i^{(s)}\right)\boldsymbol{D}_i^{(s)}(\tau)\,d\tau\right]$.

Lemma 3 indicates that the estimator $\hat{\boldsymbol{\zeta}}^{(s)}$ converges to the pseudo-true parameter $\boldsymbol{\zeta}_0^{(s)}$ and has an asymptotic normal distribution even if the sth candidate model $\mathbb{M}_s$ may be misspecified. We need the results of Lemma 3 to establish the uniform convergence rate of $\hat{\boldsymbol{\zeta}}^{(s)}$ in Proposition 1.

Proof of Lemma 3 (i) Let $a_n = \sqrt{J_n\phi_s/n}$ and $\boldsymbol{v}^{(s)} = (\boldsymbol{v}_1^{(s)\top}, \cdots, \boldsymbol{v}_K^{(s)\top})^\top \in \mathbb{R}^{K\phi_s}$. It suffices to show that for any given $\delta > 0$, there exists a large constant C such that

$$
P\left\{\inf_{\|\boldsymbol{v}^{(s)}\|=C}\bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)} + a_n\boldsymbol{v}^{(s)}\right) > \bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}\right)\right\} \geq 1 - \delta. \tag{S.1}
$$

This implies that with probability at least $1-\delta$ that there exists a local minimum $\hat{\boldsymbol{\zeta}}^{(s)}$ in the ball $\left\{\boldsymbol{\zeta}_0^{(s)} + a_n\boldsymbol{v}^{(s)} : \left\|\boldsymbol{v}^{(s)}\right\| \leq C\right\}$ such that $\left\|\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right\| = O_p(a_n)$. Then, using the identity in Knight (1998)

$$\rho_\tau\left(u+v\right)-\rho_\tau\left(u\right)=v\psi_\tau\left(u\right)+\int_0^{-v}\left\{I\left(u\le t\right)-I\left(u\le 0\right)\right\}dt, \tag{S.2}$$

we have

$$\begin{aligned}
&\bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}+a_n\boldsymbol{v}^{(s)}\right)-\bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}\right)\\
=&\int_0^1\sum_{i=1}^n\left[\rho_\tau\left\{\varepsilon_i+u_i^{(s)}-a_n\boldsymbol{D}_i^{(s)\top}\left(\tau\right)\boldsymbol{v}^{(s)}\right\}-\rho_\tau\left\{\varepsilon_i+u_i^{(s)}\right\}\right]d\tau\\
=&-\int_0^1\sum_{i=1}^n a_n\boldsymbol{D}_i^{(s)\top}\left(\tau\right)\boldsymbol{v}^{(s)}\psi_\tau\left(\varepsilon_i+u_i^{(s)}\right)d\tau+\int_0^1\left(\sum_{i=1}^n\int_0^{a_n\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}}\eta_i^{(s)}\left(t\right)dt\right)d\tau\\
\triangleq&F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)+F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right),
\end{aligned} \tag{S.3}$$

where $\eta_i^{(s)}\left(t\right)=I\left(\varepsilon_i+u_i^{(s)}\le t\right)-I\left(\varepsilon_i+u_i^{(s)}\le 0\right)$.

Note that $\boldsymbol{\zeta}_0^{(s)}$ is the minimum of the population objective (3.1), which implies that

$$E\left(\int_0^1\boldsymbol{D}_i^{(s)}\left(\tau\right)\psi_\tau\left(\varepsilon_i+u_i^{(s)}\right)d\tau\right)=\mathbf{0}. \tag{S.4}$$

Thus we can obtain $E\left[F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right]=\mathbf{0}$ by using (S.4). Based on Condition (C4) (ii) and the above result, we have

$$\begin{aligned}
Var\Big[F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\Big]=&a_n^2\sum_{i=1}^n E\left[\int_0^1\boldsymbol{D}_i^{(s)\top}\left(\tau\right)\boldsymbol{v}^{(s)}\psi_\tau\left(\varepsilon_i+u_i^{(s)}\right)d\tau\right]^2\\
=&a_n^2\Big[\boldsymbol{v}^{(s)}\Big]^\top\sum_{i=1}^n Cov\left[\int_0^1\boldsymbol{D}_i^{(s)}\left(\tau\right)\psi_\tau\left(\varepsilon_i+u_i^{(s)}\right)d\tau\right]\boldsymbol{v}^{(s)}\\
\le&na_n^2\Big[\boldsymbol{v}^{(s)}\Big]^\top\lambda_{\max}\left(\boldsymbol{C}^{(s)}\right)\boldsymbol{v}^{(s)}\\
\le&(na_n^2\bar{c}_{\boldsymbol{C}^{(s)}}/J_n)\left\|\boldsymbol{v}^{(s)}\right\|^2,
\end{aligned}$$

and therefore, we have

$$\begin{aligned}
F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)=&E\left[F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right]+O_p\left\{\sqrt{Var\left[F_{n1}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right]}\right\}\\
=&O_p(a_n\sqrt{n/J_n}\bar{c}_{\boldsymbol{C}^{(s)}}^{1/2})\left\|\boldsymbol{v}^{(s)}\right\|=O_p(\sqrt{\bar{c}_{\boldsymbol{C}^{(s)}}/\phi_s na_n^2/J_n})\left\|\boldsymbol{v}^{(s)}\right\|.
\end{aligned} \tag{S.5}$$

Next, we consider $F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right)$. Let $\mathcal{X}=(X_{11},\cdots,X_{1p},\cdots,X_{n1},\cdots,X_{np})^\top$. Note that

$$
\begin{aligned}
& E\left[F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right] \\
=& \sum_{i=1}^{n} \int_0^1 E\left\{E\left[\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \eta_i^{(s)}(t)\, dt \,|\mathcal{X}\right]\right\} d\tau \\
=& \sum_{i=1}^{n} \int_0^1 E\left\{\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \left[F\left(-u_i^{(s)} + t\,|\boldsymbol{X}_i\right) - F\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)\right] dt\right\} d\tau \\
=& \sum_{i=1}^{n} \int_0^1 E\left\{\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \left[f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)t + \frac{1}{2}f'\left(-u_i^{(s)*}\,|\boldsymbol{X}_i\right)t^2\right] dt\right\} d\tau \\
=& \sum_{i=1}^{n} \int_0^1 E\left\{\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \left[f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)t\left(1 + \frac{1}{2f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)}f'\left(-u_i^{(s)*}\,|\boldsymbol{X}_i\right)t\right)\right] dt\right\} d\tau \\
=& \sum_{i=1}^{n} \int_0^1 E\left\{\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \left[f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)t\{1+o_p(1)\}\right] dt\right\} d\tau \\
=& \frac{1}{2}a_n^2\left[\boldsymbol{v}^{(s)}\right]^\top \sum_{i=1}^{n} E\left[\int_0^1 f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)\boldsymbol{D}_i^{(s)}(\tau)\,\boldsymbol{D}_i^{(s)\top}(\tau)\,d\tau\right]\boldsymbol{v}^{(s)}\{1+o(1)\} \\
\geq& \frac{1}{2}na_n^2\left[\boldsymbol{v}^{(s)}\right]^\top \lambda_{\min}\left(\boldsymbol{A}^{(s)}\right)\boldsymbol{v}^{(s)}\{1+o(1)\} \\
\geq& \underline{c}_{\boldsymbol{A}^{(s)}} na_n^2\left\|\boldsymbol{v}^{(s)}\right\|^2/(4J_n),
\end{aligned} \tag{S.6}
$$

where $u_i^{(s)*}$ lies between $u_i^{(s)}$ and $u_i^{(s)} + t$. Here the first equality is true by the law of iterated expectations. The third equality uses Taylor's expansion. The fifth equality holds because of $0 < C_1 < f(\cdot\,|\boldsymbol{X}_i) \leq C$ and $|f'(\cdot\,|\boldsymbol{X}_i)| \leq C_2$ by the condition (C3) and $a_n\left|\boldsymbol{D}_i^{(s)\top}(\tau)\,\boldsymbol{v}^{(s)}\right| \leq a_n\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|\left\|\boldsymbol{v}^{(s)}\right\| \leq C\sqrt{pJ_n\phi_s/n} = o(1)$ by the condition (C5) (i). The last inequality holds due to the condition (C4) (i). Similarly, by the condition (C4)(i) and the fact $|\eta_i^{(s)}(t)| \leq 1$, we have

$$
\begin{aligned}
Var\left[F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right] \leq& \sum_{i=1}^{n} E\left[\int_0^1\left(\int_0^{a_n \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{v}^{(s)}} \eta_i^{(s)}(t)\,dt\right)d\tau\right]^2 \\
\leq& \sum_{i=1}^{n} E\left[\int_0^1\left(a_n\boldsymbol{D}_i^{(s)\top}(\tau)\,\boldsymbol{v}^{(s)}\right)^2 d\tau\right] \\
=& na_n^2\left[\boldsymbol{v}^{(s)}\right]^\top E\left[\int_0^1 \boldsymbol{D}_i^{(s)}(\tau)\,\boldsymbol{D}_i^{(s)\top}(\tau)\,d\tau\right]\boldsymbol{v}^{(s)} \\
\leq& \bar{c}_{\boldsymbol{A}^{(s)}} na_n^2\left\|\boldsymbol{v}^{(s)}\right\|^2/(C_0J_n),
\end{aligned} \tag{S.7}
$$

and thus

$$
\begin{aligned}
F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right) =& E\left[F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right] + O_p\left\{\sqrt{Var\left[F_{n2}^{(s)}\left(\boldsymbol{v}^{(s)}\right)\right]}\right\} \\
\geq& \underline{c}_{\boldsymbol{A}^{(s)}} na_n^2\left\|\boldsymbol{v}^{(s)}\right\|^2/(4J_n) + O_p(a_n\sqrt{n/J_n}\,\bar{c}_{\boldsymbol{A}^{(s)}}^{1/2})\left\|\boldsymbol{v}^{(s)}\right\| \\
=& \underline{c}_{\boldsymbol{A}^{(s)}} na_n^2\left\|\boldsymbol{v}^{(s)}\right\|^2/(4J_n) + O_p(\sqrt{\bar{c}_{\boldsymbol{A}^{(s)}}/\phi_s na_n^2/J_n})\left\|\boldsymbol{v}^{(s)}\right\|.
\end{aligned} \tag{S.8}
$$

By using (S.3), (S.5) and (S.8), we have

$$\begin{aligned}&\bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}+a_n\boldsymbol{v}^{(s)}\right)-\bar{\mathcal{L}}_n^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}\right)\\ \geq&\underline{c}_{\boldsymbol{A}^{(s)}}na_n^2\left\|\boldsymbol{v}^{(s)}\right\|^2/(4J_n)+O_p((\sqrt{\bar{c}_{\boldsymbol{C}^{(s)}}/\phi_s}+\sqrt{\bar{c}_{\boldsymbol{A}^{(s)}}/\phi_s})na_n^2/J_n)\left\|\boldsymbol{v}^{(s)}\right\|. \qquad \text{(S.9)}\end{aligned}$$

By the condition $(\bar{c}_{\boldsymbol{A}^{(s)}}+\bar{c}_{\boldsymbol{C}^{(s)}})/\phi_s=O(\underline{c}^2_{\boldsymbol{A}^{(s)}})$ and choosing a sufficiently large C, the first term dominates the second term uniformly in $\left\|\boldsymbol{v}^{(s)}\right\|=C$ on the right-hand side of (S.9), which is positive. Hence, by choosing a sufficiently large C, (S.1) holds. This completes the proof of the Lemma 3 (i).

(ii) Similar to the proof of Theorem 3.1 (ii) in Lu and Su (2015), we have

$$\begin{aligned}\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left[\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right]=&\sum_{i=1}^n\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{\mathcal{H}}_i^{(s)}+o_p(1)\\ \triangleq&\sum_{i=1}^n\boldsymbol{T}_{ni}^{(s)}+o_p(1),\end{aligned}$$

where $\boldsymbol{T}_{ni}^{(s)}=\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{\mathcal{H}}_i^{(s)}$ and $\boldsymbol{\mathcal{H}}_i^{(s)}=\int_0^1\psi_\tau\left(\varepsilon_i+u_i^{(s)}\right)\boldsymbol{D}_i^{(s)}(\tau)\,d\tau$. By (S.4), we have $E(\boldsymbol{T}_{ni}^{(s)})=\boldsymbol{0}$ and

$$\begin{aligned}&Var\left(\sum_{i=1}^n\boldsymbol{T}_{ni}^{(s)}\right)\\ =&\sum_{i=1}^nE\left(\boldsymbol{T}_{ni}^{(s)}\boldsymbol{T}_{ni}^{(s)\top}\right)\\ =&\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\sum_{i=1}^nCov\left(\boldsymbol{\mathcal{H}}_i^{(s)}\right)\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\boldsymbol{G}_n^{(s)\top}\\ =&\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{C}_n^{(s)}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\boldsymbol{G}_n^{(s)\top}\\ =&\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\to\boldsymbol{G}^{(s)}.\end{aligned}$$

To verify asymptotic normality, we should check the Lindeberg-Feller condition, that is, $\forall\delta>0$, $\sum_{i=1}^nE\left[\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^2I\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|\geq\delta\right)\right]\to0$. For any $\delta>0$,

$$\begin{aligned}\sum_{i=1}^nE\left[\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^2I\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|\geq\delta\right)\right]=&nE\left[\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^2I\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|\geq\delta\right)\right]\\ \leq&n\left\{E\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^4\right)\right\}^{\frac{1}{2}}\left\{P\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|\geq\delta\right)\right\}^{\frac{1}{2}}\\ \leq&n\delta^{-2}E\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^4\right). \qquad \text{(S.10)}\end{aligned}$$

Next, we need to derive the bound of $E\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^4\right)$. Based on the fact that $tr(\boldsymbol{AB})\leq\lambda_{\max}(\boldsymbol{A})tr(\boldsymbol{B})$ for any symmetric matrix $\boldsymbol{A}$ and positive semidefinite matrix $\boldsymbol{B}$ and

$\lambda_{\max}(\boldsymbol{G}_n^{(s)\top}\boldsymbol{G}_n^{(s)}) = \lambda_{\max}(\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}) \to c$ for some finite positive constant c, we have

$$
\begin{aligned}
&E\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^4\right)\\
=&E\left\{tr\left[\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{\mathcal{H}}_i^{(s)}\boldsymbol{\mathcal{H}}_i^{(s)\top}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\boldsymbol{G}_n^{(s)\top}\right]\right\}^2\\
=&E\left\{tr\left[\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\boldsymbol{G}_n^{(s)\top}\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\boldsymbol{\mathcal{H}}_i^{(s)}\boldsymbol{\mathcal{H}}_i^{(s)\top}\right]\right\}^2\\
\le&E\left\{\lambda_{\max}^2\left\{\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\boldsymbol{G}_n^{(s)\top}\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\right\}\left[tr(\boldsymbol{\mathcal{H}}_i^{(s)}\boldsymbol{\mathcal{H}}_i^{(s)\top})\right]^2\right\}\\
\le&E\left\|\boldsymbol{\mathcal{H}}_i^{(s)}\right\|^4\left[\lambda_{\max}(\boldsymbol{G}_n^{(s)\top}\boldsymbol{G}_n^{(s)})\right]^2\left[\lambda_{\max}\left\{\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\left(\boldsymbol{V}_n^{(s)}\right)^{-1}\left(\boldsymbol{A}_n^{(s)}\right)^{-1}\right\}\right]^2\\
=&E\left\|\boldsymbol{\mathcal{H}}_i^{(s)}\right\|^4\left[\lambda_{\max}(\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top})\right]^2\left[\lambda_{\max}\left\{\left(\boldsymbol{C}_n^{(s)}\right)^{-1}\right\}\right]^2\\
=&O\left(n^{-2}p^2J_n^2\underline{c}_{\boldsymbol{C}^{(s)}}^{-2}\right).
\end{aligned}
$$

This together with (S.10) and condition (C5) (i) leads to

$$
\begin{aligned}
\sum_{i=1}^n E\left[\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^2 I\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|\ge\delta\right)\right] \le& n\delta^{-2}E\left(\left\|\boldsymbol{T}_{ni}^{(s)}\right\|^4\right)\\
=&O\left(n^{-1}p^2J_n^2\underline{c}_{\boldsymbol{C}^{(s)}}^{-2}\right)\\
=&o\left(1\right).
\end{aligned}
$$

Thus, $\left\{\boldsymbol{T}_{ni}^{(s)}\right\}$ satisfies the conditions of the Lindeberg-Feller central limit theorem (see van der Vaart (1998)). This completes the proof. □

Proof of Proposition 1. Let $\delta_n = L\sqrt{J_n\bar{\phi}\log n/n}$ for some large positive constant $L < \infty$. Note that the objective function $\bar{Q}^{(s)}\left(\boldsymbol{\zeta}^{(s)}\right)$ is defined in (3.1). Now we define

$$
G(\delta_n) = \inf_{1\le s\le S}\inf_{\left\|\boldsymbol{\zeta}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\|>\delta_n}\left[\bar{Q}^{(s)}\left(\boldsymbol{\zeta}^{(s)}\right)-\bar{Q}^{(s)}\left(\boldsymbol{\zeta}_0^{(s)}\right)\right] \tag{S.11}
$$

and

$$
\varsigma^{(s)}\left(\delta_n\right) = \left\{\boldsymbol{\zeta}^{(s)} : \left\|\boldsymbol{\zeta}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\| > \delta_n, \sqrt{p}\left\|\boldsymbol{\zeta}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\| = o(1)\right\}.
$$

For $\boldsymbol{\zeta}^{(s)} \in \varsigma^{(s)}\left(\delta_n\right)$, we apply the Knight's identity (S.2), $u_i^{(s)} = \mu_i - \boldsymbol{D}_i^{(s)\top}\left(\tau\right)\boldsymbol{\zeta}_0^{(s)}$, (S.4) and condition (C4) (i) to obtain that

$$
\begin{aligned}
&\bar{Q}^{(s)}\left(\zeta^{(s)}\right)-\bar{Q}^{(s)}\left(\zeta_0^{(s)}\right)\\
=&E\left[\int_0^1\left\{\rho_\tau\left(\varepsilon_i+u_i^{(s)}-\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]\right)-\rho_\tau\left(\varepsilon_i+u_i^{(s)}\right)\right\}d\tau\right]\\
=&E\left[\int_0^1\left\{\int_0^{\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]}\left[I\left(\varepsilon_i+u_i^{(s)}\le t\right)-I\left(\varepsilon_i+u_i^{(s)}\le 0\right)\right]dt\right\}d\tau\right]\\
=&E\left[\int_0^1\left\{\int_0^{\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]}\left[F\left(-u_i^{(s)}+t\,|\boldsymbol{X}_i\right)-F\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)\right]dt\right\}d\tau\right]\\
=&E\left[\int_0^1\left\{\int_0^{\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]}\left[f\left(-u_i^{(s)}\,|X_i\right)t+\frac{1}{2}f'\left(-\tilde{u}_i^{(s)}\,|X_i\right)t^2\right]dt\right\}d\tau\right]\\
=&E\left[\int_0^1\left\{\int_0^{\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]}f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)tdt\right\}d\tau\right]\{1+o(1)\}\\
=&\frac{1}{2}\left[\zeta^{(s)}-\zeta_0^{(s)}\right]^\top E\left[\int_0^1 f\left(-u_i^{(s)}\,|\boldsymbol{X}_i\right)\boldsymbol{D}_i^{(s)}(\tau)\,\boldsymbol{D}_i^{(s)\top}(\tau)d\tau\right]\left[\zeta^{(s)}-\zeta_0^{(s)}\right]\{1+o(1)\}\\
\geq&\frac{1}{4}\underline{c}_{\boldsymbol{A}^{(s)}}\delta_n^2/J_n,
\end{aligned}
$$

where $\tilde{u}_i^{(s)}$ lies between $u_i^{(s)}$ and $u_i^{(s)}+t$. Here the fourth equality uses Taylor's expansion. The fifth equality holds because of condition (C3) and $\left|\boldsymbol{D}_i^{(s)\top}(\tau)\left[\zeta^{(s)}-\zeta_0^{(s)}\right]\right|\le\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|\left\|\zeta^{(s)}-\zeta_0^{(s)}\right\|\le C\sqrt{p}\left\|\zeta^{(s)}-\zeta_0^{(s)}\right\|=o(1)$.

By directly using the Boole's inequality, (S.11) and the fact

$$
\begin{aligned}
\bar{Q}^{(s)}\left(\hat{\zeta}^{(s)}\right)-\bar{Q}^{(s)}\left(\zeta_0^{(s)}\right)=&\frac{1}{2}\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]^\top\boldsymbol{A}^{(s)}\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]\{1+o(1)\},\\
\le&\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]^\top\boldsymbol{A}^{(s)}\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right],
\end{aligned}
$$

we can obtain

$$
\begin{aligned}
P\left\{\max_{1\le s\le S}\left\|\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right\|\ge\delta_n\right\}\le&S\max_{1\le s\le S}P\left\{\left\|\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right\|\ge\delta_n\right\}\\
\le&S\max_{1\le s\le S}P\left\{\bar{Q}^{(s)}\left(\hat{\zeta}^{(s)}\right)-\bar{Q}^{(s)}\left(\zeta_0^{(s)}\right)\ge G(\delta_n)\right\}\\
\le&S\max_{1\le s\le S}P\left\{\mathcal{W}^{(s)}\ge nG(\delta_n)\right\},
\end{aligned}
$$

where $\mathcal{W}^{(s)}=n\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]^\top\boldsymbol{A}^{(s)}\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]$.

To prove the Proposition 1, motivated by the proof of Lu and Su (2015), we need to show the bound $P\left\{\mathcal{W}^{(s)}\ge nG(\delta_n)\right\}$. We can obtain $\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left[\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right]\xrightarrow{d}N\left(\boldsymbol{0},\boldsymbol{G}^{(s)}\right)$ by applying the proof of Lemma 3 (ii). Let $\hat{\boldsymbol{\psi}}^{(s)}\triangleq\left(\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right)^{-1/2}\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\left(\hat{\zeta}^{(s)}-\zeta_0^{(s)}\right)$. Then we have $\hat{\boldsymbol{\psi}}^{(s)}\xrightarrow{d}N\left(\boldsymbol{0},\boldsymbol{I}_{l_s}\right)$. So we obtain $\hat{\zeta}^{(s)}-\zeta_0^{(s)}=\left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2}\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\right\}^+\hat{\boldsymbol{\psi}}^{(s)}$, where $\boldsymbol{A}^+$ stands for the Moore-Penrose generalized inverse of $\boldsymbol{A}$. Furthermore, using the fact $\boldsymbol{A}^\top\boldsymbol{B}\boldsymbol{A}\le\lambda_{\max}(\boldsymbol{B})\boldsymbol{A}^\top\boldsymbol{A}$ for any real symmetric matrix $\boldsymbol{B}$ and conformable matrix $\boldsymbol{A}$ and $(\boldsymbol{A}^{+})^{\top}\boldsymbol{A}^{+} = (\boldsymbol{A}\boldsymbol{A}^{\top})^{+}$ yields that

$$\begin{aligned}
\mathcal{W}^{(s)} =& n\left[\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right]^{\top} \boldsymbol{A}^{(s)} \left[\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right] \\
=& n\hat{\boldsymbol{\psi}}^{(s)\top} \left[\left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2} \boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\right\}^{+}\right]^{\top} \boldsymbol{A}^{(s)} \\
&\times \left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2} \boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\right\}^{+} \hat{\boldsymbol{\psi}}^{(s)} \\
\leq& n\lambda_{\max}\left(\boldsymbol{A}^{(s)}\right)\hat{\boldsymbol{\psi}}^{(s)\top} \left[\left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2} \boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\right\}^{+}\right]^{\top} \\
&\times \left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2} \boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1/2}\right\}^{+} \hat{\boldsymbol{\psi}}^{(s)} \\
=& n\lambda_{\max}\left(\boldsymbol{A}^{(s)}\right)\hat{\boldsymbol{\psi}}^{(s)\top} \left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2} \left(\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1}\boldsymbol{G}_n^{(s)\top}\right) \left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{-1/2}\right\}^{-1} \hat{\boldsymbol{\psi}}^{(s)} \\
=& n\lambda_{\max}\left(\boldsymbol{A}^{(s)}\right)\hat{\boldsymbol{\psi}}^{(s)\top} \left\{\left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{1/2} \left(\boldsymbol{G}_n^{(s)}\left(\boldsymbol{V}_n^{(s)}\right)^{-1}\boldsymbol{G}_n^{(s)\top}\right)^{-1} \left[\boldsymbol{G}_n^{(s)}\boldsymbol{G}_n^{(s)\top}\right]^{1/2}\right\} \hat{\boldsymbol{\psi}}^{(s)} \\
\leq& n\lambda_{\max}\left(\boldsymbol{A}^{(s)}\right)\lambda_{\max}\left(\boldsymbol{V}_n^{(s)}\right)\hat{\boldsymbol{\psi}}^{(s)\top}\hat{\boldsymbol{\psi}}^{(s)} \\
\leq& (\bar{c}_{\boldsymbol{A}}\bar{c}_{\boldsymbol{C}}/\underline{c}_{\boldsymbol{A}}^2)\left\|\hat{\boldsymbol{\psi}}^{(s)}\right\|^2.
\end{aligned}$$

Define $c_{\boldsymbol{AC}} = \bar{c}_{\boldsymbol{A}}\bar{c}_{\boldsymbol{C}}/\underline{c}_{\boldsymbol{A}}^2$ and $\bar{l} = \max_{1\leq s\leq S} l_s$. Then, by Lemma 2.1 of Shibata (1981), we have

$$\begin{aligned}
& S\max_{1\leq s\leq S} P\left\{\mathcal{W}^{(s)} \geq nG\left(\delta_n\right)\right\} \\
\leq& S\max_{1\leq s\leq S} P\left\{\left\|\hat{\boldsymbol{\psi}}^{(s)}\right\|^2 \geq nG\left(\delta_n\right)/c_{\boldsymbol{AC}}\right\} \\
\leq& \limsup_{n\to\infty} S\max_{1\leq s\leq S} P\left\{\chi^2\left(l_s\right) \geq nG\left(\delta_n\right)/c_{\boldsymbol{AC}}\right\} \\
\leq& \limsup_{n\to\infty} SP\left\{\chi^2\left(\bar{l}\right) \geq nG\left(\delta_n\right)/c_{\boldsymbol{AC}}\right\} \\
\leq& \limsup_{n\to\infty} SP\left\{\chi^2\left(\bar{l}\right) \geq \bar{l} + \left[n\delta_n^2\underline{c}_{\boldsymbol{A}}/(J_n c_{\boldsymbol{AC}}) - \bar{l}\right]\right\} \\
\leq& \limsup_{n\to\infty} S\exp\left\{-\frac{\left[n\delta_n^2\underline{c}_{\boldsymbol{A}}/(J_n c_{\boldsymbol{AC}}) - \bar{l}\right]}{4}\right\} \\
=& 0.
\end{aligned}$$

Here the last equality holds because of $Sn^{-L^2\bar{\phi}\underline{c}_{\boldsymbol{A}}^3/(4\bar{c}_{\boldsymbol{A}}\bar{c}_{\boldsymbol{C}})} = o(1)$ by the condition (C5) (ii) and taking some large L. □

Proof of Theorem 1. Following the proof of Theorem 3.3 in Lu and Su (2015), we should show, as $n \to \infty$, that

$$\sup_{\boldsymbol{w}\in\mathbb{W}} \left|\frac{\mathcal{Q}_n\left(\boldsymbol{w}\right) - \mathrm{OAQPE}_n\left(\boldsymbol{w}\right)}{\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)}\right| = o_p(1). \tag{S.12}$$

By applying the Knight's identity, we have

$$
\begin{aligned}
&\mathcal{Q}_n(\boldsymbol{w})-\mathrm{OAQPE}_n(\boldsymbol{w})\\
=&\left\{\int_0^1 \frac{1}{n}\sum_{i=1}^{n}\left[\rho_\tau\left(Y_i-\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}\right)-\rho_\tau(\varepsilon_i)\right]d\tau\right\}+\frac{1}{n}\lambda_n\sum_{s=1}^{S} w_s\phi_s\\
&-\left\{\mathrm{OAQPE}_n(\boldsymbol{w})-E\left[\int_0^1 \rho_\tau(\varepsilon)d\tau\right]\right\}+\frac{1}{n}\sum_{i=1}^{n}\left\{\int_0^1 \rho_\tau(\varepsilon_i)\,d\tau-E\left[\int_0^1 \rho_\tau(\varepsilon)\,d\tau\right]\right\}\\
=&\int_0^1 \frac{1}{n}\sum_{i=1}^{n}\left[\mu_i-\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}\right]\psi_\tau(\varepsilon_i)d\tau\\
&+\int_0^1 \frac{1}{n}\sum_{i=1}^{n}\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}-\mu_i}\left[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)\right]dtd\tau\\
&-E\left[\int_0^1\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}-\mu}\left[I(\varepsilon\le t)-I(\varepsilon\le 0)\right]dtd\tau\,|\mathcal{D}_n\right]\\
&+\frac{1}{n}\lambda_n\sum_{s=1}^{S} w_s\phi_s+\frac{1}{n}\sum_{i=1}^{n}\left\{\int_0^1 \rho_\tau(\varepsilon_i)\,d\tau-E\left[\int_0^1 \rho_\tau(\varepsilon_i)\,d\tau\right]\right\}\\
\triangleq&\Lambda_{n1}(\boldsymbol{w})+\Lambda_{n2}(\boldsymbol{w})+\Lambda_{n3}(\boldsymbol{w})+\Lambda_{n4}(\boldsymbol{w})+\Lambda_{n5}(\boldsymbol{w})+\Lambda_{n6}(\boldsymbol{w})+\Lambda_{n7},
\end{aligned}
$$

where $\mu\triangleq\mu(\mathbf{x},\tau)$,

$$
\begin{aligned}
\Lambda_{n1}(\boldsymbol{w})=&\frac{1}{n}\sum_{i=1}^{n}\int_0^1\left[\mu_i-\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}\right]\psi_\tau(\varepsilon_i)d\tau,\\
\Lambda_{n2}(\boldsymbol{w})=&\frac{1}{n}\sum_{i=1}^{n}\int_0^1\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(\boldsymbol{s})}-\mu_i}\left\{\left[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)\right]\right.\\
&\left.-\left[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right]\right\}dtd\tau,
\end{aligned}
$$

$$\Lambda_{n3}(\boldsymbol{w}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\{ \int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i} \left[F(t\,|\boldsymbol{X}_i) - F(0\,|\boldsymbol{X}_i)\right]dt \right.$$
$$\left. - E_{\boldsymbol{X}_i}\left[\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i} \left[F(t\,|\boldsymbol{X}_i) - F(0\,|\boldsymbol{X}_i)\right]dt\right]\right\} d\tau,$$

$$\Lambda_{n4}(\boldsymbol{w}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left\{ E_{\boldsymbol{X}_i}\left[\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i} \left[F(t\,|\boldsymbol{X}_i) - F(0\,|\boldsymbol{X}_i)\right]dt\right] \right.$$
$$\left. - E_{\boldsymbol{X}_i}\left[\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i} \left[F(t\,|\boldsymbol{X}_i) - F(0\,|\boldsymbol{X}_i)\right]dt\right]\right\} d\tau,$$

$$\Lambda_{n5}(\boldsymbol{w}) = -\int_0^1 E_{\mathbf{x}}\left[\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[F(t\,|\mathbf{x}) - F(0\,|\mathbf{x})\right]dt\right] d\tau$$
$$+ \int_0^1 E_{\mathbf{x}}\left[\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu} \left[F(t\,|\mathbf{x}) - F(0\,|\mathbf{x})\right]dt\right] d\tau,$$

$$\Lambda_{n6}(\boldsymbol{w}) = \frac{1}{n}\lambda_n \sum_{s=1}^{S} w_s\phi_s \text{ and } \Lambda_{n7} = \frac{1}{n}\sum_{i=1}^{n}\left\{\int_0^1 \rho_\tau(\varepsilon_i)\,d\tau - E\left[\int_0^1 \rho_\tau(\varepsilon_i)\,d\tau\right]\right\}.$$

It is easy to derive all terms except $\Lambda_{n4}(\boldsymbol{w})$ and $\Lambda_{n5}(\boldsymbol{w})$. Let $E_{\mathbf{x}}$ be an expectation with respect to a random variable $\mathbf{x}$. To get $\Lambda_{n4}(\boldsymbol{w})$ and $\Lambda_{n5}(\boldsymbol{w})$, we need the following two results

$$E_{\boldsymbol{X}_i}\left\{\int_0^{\sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}\boldsymbol{\zeta}_0^{(s)}-\mu_i} \left[F(t\,|\boldsymbol{X}_i) - F(0\,|\boldsymbol{X}_i)\right]dt\right\}$$
$$= E_{\mathbf{x}}\left\{\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}\boldsymbol{\zeta}_0^{(s)}-\mu} \left[F(t\,|\mathbf{x}) - F(0\,|\mathbf{x})\right]dt\right\}, \tag{S.13}$$

and

$$E\left\{\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[I\{\varepsilon \le t\} - I\{\varepsilon \le 0\}\right]dt\,|\mathcal{D}_n\right\}$$
$$= E_{\mathbf{x}}\left\{\int_0^{\sum_{s=1}^{S} w_s \mathbb{D}^{(s)\top}\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[F(t\,|\mathbf{x}) - F(0\,|\mathbf{x})\right]dt\right\}. \tag{S.14}$$

The identity (S.13) holds because $\boldsymbol{\zeta}_0^{(s)}$ does not depend on the ith observation, and the result (S.14) comes from

$$
\begin{aligned}
& E\left\{\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[I\left\{\varepsilon\leq t\right\}-I\left\{\varepsilon\leq 0\right\}\right]dt\left|\mathcal{D}_n\right.\right\} \\
=& \int_{(\mathbf{x},\varepsilon)}\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[I\left\{\varepsilon\leq t\right\}-I\left\{\varepsilon\leq 0\right\}\right]dtf\left(\mathbf{x},\varepsilon\left|\mathcal{D}_n\right.\right)d\mathbf{x}d\varepsilon \\
=& \int_{(\mathbf{x},\varepsilon)}\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[I\left\{\varepsilon\leq t\right\}-I\left\{\varepsilon\leq 0\right\}\right]dtf\left(\mathbf{x},\varepsilon\right)d\mathbf{x}d\varepsilon \\
=& \int_{\mathbf{x}}\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \int_{\varepsilon}\left[I\left\{\varepsilon\leq t\right\}-I\left\{\varepsilon\leq 0\right\}\right]f\left(\varepsilon\left|\mathbf{x}\right.\right)f\left(\mathbf{x}\right)d\varepsilon d\mathbf{x}dt \\
=& \int_{\mathbf{x}}\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[F\left(t\left|\mathbf{x}\right.\right)-F\left(0\left|\mathbf{x}\right.\right)\right]f\left(\mathbf{x}\right)d\mathbf{x}dt \\
=& E_{\mathbf{x}}\left\{\int_0^{\sum\limits_{s=1}^{S} w_s\mathbb{D}^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu} \left[F\left(t\left|\mathbf{x}\right.\right)-F\left(0\left|\mathbf{x}\right.\right)\right]dt\right\}.
\end{aligned}
$$

The second equality is true because the sample $\{\boldsymbol{X}_i,\varepsilon_i\}$ is independent of the random variable $\{\mathbf{x},\varepsilon\}$. Hence

$$
\begin{aligned}
& \sup_{\boldsymbol{w}\in\mathbb{W}}\left|\frac{\mathcal{Q}_n\left(\boldsymbol{w}\right)-\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)}{\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)}\right| \\
=& \sup_{\boldsymbol{w}\in\mathbb{W}}\left|\frac{\Lambda_{n1}\left(\boldsymbol{w}\right)+\Lambda_{n2}\left(\boldsymbol{w}\right)+\Lambda_{n3}\left(\boldsymbol{w}\right)+\Lambda_{n4}\left(\boldsymbol{w}\right)+\Lambda_{n5}\left(\boldsymbol{w}\right)+\Lambda_{n6}\left(\boldsymbol{w}\right)+\Lambda_{n7}}{\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)}\right| \\
\leq& \frac{\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n1}\left(\boldsymbol{w}\right)+\Lambda_{n2}\left(\boldsymbol{w}\right)+\Lambda_{n3}\left(\boldsymbol{w}\right)+\Lambda_{n4}\left(\boldsymbol{w}\right)+\Lambda_{n5}\left(\boldsymbol{w}\right)+\Lambda_{n6}\left(\boldsymbol{w}\right)+\Lambda_{n7}\right|}{\inf\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)\right|} \\
\leq& \left\{\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n1}\left(\boldsymbol{w}\right)\right|+\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n2}\left(\boldsymbol{w}\right)\right|+\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3}\left(\boldsymbol{w}\right)\right|+\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n4}\left(\boldsymbol{w}\right)\right|\right. \\
& \left.+\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n5}\left(\boldsymbol{w}\right)\right|+\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n6}\left(\boldsymbol{w}\right)\right|+\left|\Lambda_{n7}\right|\right\}/\inf_{\boldsymbol{w}\in\mathbb{W}}\left|\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)\right|. \qquad \text{(S.15)}
\end{aligned}
$$

To prove (S.12), we need the following results: (i) $\inf\limits_{\boldsymbol{w}\in\mathbb{W}}\mathrm{OAQPE}_n\left(\boldsymbol{w}\right)\geq E\left\{\int_0^1\rho_\tau\left(\varepsilon\right)d\tau\right\}-o_p(1)$; (ii) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n1}\left(\boldsymbol{w}\right)\right|=o_p(1)$; (iii) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n2}\left(\boldsymbol{w}\right)\right|=o_p(1)$; (iv) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3}\left(\boldsymbol{w}\right)\right|=o_p(1)$; (v) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n4}\left(\boldsymbol{w}\right)\right|=o_p(1)$; (vi) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n5}\left(\boldsymbol{w}\right)\right|=o_p(1)$; (vii) $\sup\limits_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n6}\left(\boldsymbol{w}\right)\right|=o(1)$ and $(viii)\Lambda_{n7}=o_p(1)$. Now we need to prove (i)–(vii), and $(viii)$ holds by the weak law of large numbers.

(i) Let $u(\boldsymbol{w})=\mu-\sum_{s=1}^{S}w_s\mathbb{D}^{(s)\top}(\tau)\boldsymbol{\zeta}_{\mathbf{0}}^{(\boldsymbol{s})}$. By the Knight's identity, Taylor's expansion, Jensen's inequality, $\sum_i|a_ib_i|\leq(\max_i|a_i|)\sum_i|b_i|$, conditions (C4)–(C5) and

Proposition 1, we have

$$
\begin{aligned}
&\text{OAQPE}_n(\boldsymbol{w}) - E\left[\int_0^1 \rho_\tau(\varepsilon + u(\boldsymbol{w}))\,d\tau\right]\\
=&E\left\{\int_0^1 \left[\rho_\tau\left(\varepsilon + u(\boldsymbol{w}) - \sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\right) - \rho_\tau(\varepsilon + u(\boldsymbol{w}))\right] d\tau\,|\mathcal{D}_n\right\}\\
=&E\left\{\int_0^1 \int_0^{\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)} \left[I\{\varepsilon + u(\boldsymbol{w}) \le t\} - I\{\varepsilon + u(\boldsymbol{w}) \le 0\}\right] dt d\tau\,|\mathcal{D}_n\right\}\\
=&E_{\mathbf{x}}\left\{\int_0^1 \int_0^{\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)} \left[F(t - u(\boldsymbol{w})\,|\mathbf{x}) - F(-u(\boldsymbol{w})\,|\mathbf{x})\right] dt d\tau\right\}\\
=&E_{\mathbf{x}}\left\{\int_0^1 \int_0^{\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)} \left\{f(-u(\boldsymbol{w})\,|\mathbf{x})\,t + \frac{1}{2} f'(-u^*(\boldsymbol{w})\,|\mathbf{x})\,t^2\right\} dt d\tau\right\}\\
=&E_{\mathbf{x}}\left\{\int_0^1 \int_0^{\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)} f(-u(\boldsymbol{w})\,|\mathbf{x})\,t\left(1 + \frac{1}{2f(-u(\boldsymbol{w})\,|\mathbf{x})} f'(-u^*(\boldsymbol{w})\,|\mathbf{x})\,t\right) dt d\tau\right\}\\
=&E_{\mathbf{x}}\left\{\int_0^1 \int_0^{\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)} f(-u(\boldsymbol{w})\,|\mathbf{x})\,t dt d\tau\right\}\{1 + o_p(1)\}\\
=&2^{-1}E_{\mathbf{x}}\left\{\int_0^1 f(-u(\boldsymbol{w})\,|\mathbf{x})\left[\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\right]^2 d\tau\right\}\{1 + o_p(1)\}\\
\le&2^{-1}CE_{\mathbf{x}}\left\{\int_0^1 \sum_{s=1}^S w_s \left[\mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\right]^2 d\tau\right\}\{1 + o_p(1)\}\\
=&2^{-1}C\sum_{s=1}^S w_s \left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)^\top E_{\mathbf{x}}\left\{\int_0^1 \mathbb{D}^{(s)}(\tau)\mathbb{D}^{(s)\top}(\tau) d\tau\right\}\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\{1 + o_p(1)\}\\
\le&\frac{C\bar{c}_A}{2J_n C_0} \max_{1\le s\le S}\left\|\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right\|^2 \{1 + o_p(1)\}\\
=&o_p(1),
\end{aligned} \tag{S.16}
$$

where $u^*(\boldsymbol{w})$ lies between $u(\boldsymbol{w})$ and $u(\boldsymbol{w}) + t$. Here the fourth equality uses Taylor's expansion, the sixth equality holds because of $0 < C_1 \le f(\cdot\,|\boldsymbol{X}_i) \le C$ and $|f'(\cdot\,|\boldsymbol{X}_i)| \le C_2$ for some positive constants C, C_1 and C_2 by the condition (C3), and the fact

$$
\begin{aligned}
\left|\sum_{s=1}^S w_s \mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\right| \le& \sum_{s=1}^S w_s \left|\mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right)\right|\\
\le& \max_{1\le s\le S}\left\|\mathbb{D}^{(s)}(\tau)\right\|\left\|\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right\|\sum_{s=1}^S w_s\\
=& \max_{1\le s\le S}\left\|\mathbb{D}^{(s)}(\tau)\right\|\left\|\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)}\right\|\\
=&O\left(\sqrt{p\bar{\phi}J_n\log n/n}\right) = o(1)
\end{aligned}
$$

by the condition (C5) (i).

Note that $\inf_{\boldsymbol{w}\in\mathbb{W}} E\left[\int_0^1 \rho_\tau(\varepsilon + u(\boldsymbol{w}))\,d\tau\right] \ge E\left[\int_0^1 \rho_\tau(\varepsilon) d\tau\right]$. Consequently, combine with (S.16), we have $\inf_{\boldsymbol{w}\in\mathbb{W}} \mathrm{OAQPE}_n(\boldsymbol{w}) = \inf_{\boldsymbol{w}\in\mathbb{W}} E\left[\int_0^1 \rho_\tau(\varepsilon + u(\boldsymbol{w}))\,d\tau\right] - o_p(1) \geq E\left[\int_0^1 \rho_\tau(\varepsilon)d\tau\right] - o_p(1)$.

(ii) We decompose $\Lambda_{n1}(\boldsymbol{w}) \triangleq \Lambda_{n1,1}(\boldsymbol{w}) + \Lambda_{n1,2}(\boldsymbol{w})$, where

$$\Lambda_{n1,1}(\boldsymbol{w}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^1 [\mu_i - \sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}]\psi_\tau(\varepsilon_i)d\tau,$$

and

$$\Lambda_{n1,2}(\boldsymbol{w}) = -\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \sum_{s=1}^{S} w_s \boldsymbol{D}_i^{(s)\top}(\tau)(\hat{\boldsymbol{\zeta}}^{(s)} - \boldsymbol{\zeta}_0^{(s)})\psi_\tau(\varepsilon_i)d\tau.$$

Let $h_n = 1/(p^{1/2}\bar{\phi}^{1/2}\log n)$, $\|\boldsymbol{w} - \bar{\boldsymbol{w}}\|_1 = \sum_{s=1}^{S}|w_s - \bar{w}_s|$ for any $\boldsymbol{w} = (w_1, \cdots, w_S)^\top \in \mathbb{W}$ and $\bar{\boldsymbol{w}} = (\bar{w}_1, \cdots, \bar{w}_S)^\top \in \mathbb{W}$, and consider grids based on regions of the form $W_j = \{\boldsymbol{w} : \|\boldsymbol{w} - \boldsymbol{w}_j\|_1 \leq h_n\}$. By assigning $\boldsymbol{w}_j = (w_{j1}, \cdots, w_{jS})^\top$ to lie on a grid, $\mathbb{W}$ can be covered with $N = O(1/h_n^{S-1})$ regions $W_j, j = 1, \cdots, N$. By the Lemma 2, we have

$$\begin{aligned}
&\sup_{\boldsymbol{w}\in W_j}|\Lambda_{n1,1}(\boldsymbol{w}) - \Lambda_{n1,1}(\boldsymbol{w}_j)| \\
=&\sup_{\boldsymbol{w}\in W_j}\left|\sum_{s=1}^{S}(w_s - w_{js})\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}\psi_\tau(\varepsilon_i)d\tau\right| \\
\leq&\max_{1\leq s\leq S}\left\|\boldsymbol{\zeta}_0^{(s)}\right\|\left[\max_{1\leq s\leq S}\frac{1}{n}\sum_{i=1}^{n}\int_0^1\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|d\tau\right]\sup_{\boldsymbol{w}\in W_j}\sum_{s=1}^{S}|w_s - w_{js}| \\
=&O\left(\bar{\phi}^{1/2}\right)O_p\left(p^{1/2}\right)h_n \\
=&o_p(1)
\end{aligned}$$

uniformly for j. Thus, combined with the triangle inequality, we have

$$\begin{aligned}
\sup_{\boldsymbol{w}\in\mathbb{W}}|\Lambda_{n1,1}(\boldsymbol{w})| &= \max_{1\leq j\leq N}\sup_{\boldsymbol{w}\in W_j}|\Lambda_{n1,1}(\boldsymbol{w})| \\
&\leq \max_{1\leq j\leq N}|\Lambda_{n1,1}(\boldsymbol{w}_j)| + \max_{1\leq j\leq N}\sup_{\boldsymbol{w}\in W_j}|\Lambda_{n1,1}(\boldsymbol{w}) - \Lambda_{n1,1}(\boldsymbol{w}_j)| \\
&\leq \max_{1\leq j\leq N}|\Lambda_{n1,1}(\boldsymbol{w}_j)| + o_p(1). \qquad \text{(S.17)}
\end{aligned}$$

Next, we need to prove $\max_{1\leq j\leq N}|\Lambda_{n1,1}(\boldsymbol{w}_j)| = o_p(1)$. Let $u_{ij}(\tau) = \mu_i - \sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}$. Note that

$$\begin{aligned}
&E\left[\int_0^1 u_{ij}(\tau)\psi_\tau(\varepsilon_i)\,d\tau\right] \\
=&E\left\{E\left[\int_0^1 u_{ij}(\tau)\psi_\tau(\varepsilon_i)\,d\tau\,|\boldsymbol{X}_i\right]\right\} \\
=&E\left\{\int_0^1 u_{ij}(\tau)E\left[\psi_\tau(\varepsilon_i)\,|\boldsymbol{X}_i\right]d\tau\right\} \\
=&0,
\end{aligned}$$

where the last equality holds because $E[\psi_\tau(\varepsilon_i)\,|\boldsymbol{X}_i] = 0$. Due to the boundedness of $|\psi_\tau(\varepsilon_i)| \leq 1$ and $|u_{ij}(\tau)| \leq |\mu_i| + \max_{1\leq s\leq S}\left|\boldsymbol{D}_i^{(s)T}(\tau)\boldsymbol{\zeta}_0^{(s)}\right| \leq |\mu_i| + \max_{1\leq s\leq S}\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|\left\|\boldsymbol{\zeta}_0^{(s)}\right\|$, we have

$$\left|\int_0^1 u_{ij}(\tau)\psi_\tau\left(\varepsilon_i\right)d\tau\right| \leq \int_0^1 \left|u_{ij}(\tau)\right| d\tau$$
$$\leq |\mu_i| + \max_{1\leq s\leq S}\int_0^1 \left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| d\tau \max_{1\leq s\leq S}\left\|\boldsymbol{\zeta}_0^{(s)}\right\|$$
$$\leq C_1 p^{1/2}\bar{\phi}^{1/2},$$

and

$$\begin{aligned}
& Var\left[\int_0^1 u_{ij}(\tau)\psi_\tau\left(\varepsilon_i\right)d\tau\right] \\
=& E\left[\int_0^1 u_{ij}(\tau)\psi_\tau\left(\varepsilon_i\right)d\tau\right]^2 \\
\leq& E\left(\int_0^1 u_{ij}^2(\tau)d\tau\int_0^1 \psi_\tau^2\left(\varepsilon_i\right)d\tau\right) \\
\leq& E\left(\int_0^1 u_{ij}^2(\tau)d\tau\right) \\
\leq& 2E\left(\int_0^1\left[|\mu_i|^2 + \max_{1\leq s\leq S}\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|^2\left\|\boldsymbol{\zeta}_0^{(s)}\right\|^2\right]d\tau\right) \\
\leq& 2\left\{E\left(\mu_i^2\right) + \max_{1\leq s\leq S} E\int_0^1\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|^2 d\tau \max_{1\leq s\leq S}\left\|\boldsymbol{\zeta}_0^{(s)}\right\|^2\right\} \\
\leq& C_2 p\bar{\phi}.
\end{aligned}$$

Then, by the Boole's and Bernstein's inequalities, for any $\delta > 0$, we have

$$\begin{aligned}
& P\left\{\max_{1\leq j\leq N}\left|\Lambda_{n1,1}\left(\boldsymbol{w}_j\right)\right| \geq \delta\right\} \\
=& P\left\{\max_{1\leq j\leq N}\left|\frac{1}{n}\sum_{i=1}^n\int_0^1 u_{ij}\left(\tau\right)\psi_\tau\left(\varepsilon_i\right)d\tau\right| \geq \delta\right\} \\
\leq& N\max_{1\leq j\leq N} P\left\{\frac{1}{n}\left|\sum_{i=1}^n\int_0^1 u_{ij}\left(\tau\right)\psi_\tau\left(\varepsilon_i\right)d\tau\right| \geq \delta\right\} \\
\leq& 2N\max_{1\leq j\leq N}\exp\left\{-\frac{(n\delta)^2}{2\left(nC_2 p\bar{\phi} + n\delta C_1 p^{1/2}\bar{\phi}^{1/2}/3\right)}\right\} \\
=& 2\exp\left\{\log\left(N\right) - \frac{n\delta^2}{2C_2 p\bar{\phi} + 2\delta C_1 p^{1/2}\bar{\phi}^{1/2}/3}\right\} \\
=& 2\exp\left\{S\log(p^{1/2}\bar{\phi}^{1/2}\log n) - \frac{n\delta^2}{2C_2 p\bar{\phi} + 2\delta C_1 p^{1/2}\bar{\phi}^{1/2}/3}\right\} \\
=& o(1)
\end{aligned} \tag{S.18}$$

by the condition $p^{3/2}\bar{\phi}^{3/2}S\log n/n \to 0$, as $n \to \infty$. Thus $\max_{1\leq j\leq N}\left|\Lambda_{n1,1}\left(w_j\right)\right| = o_p\left(1\right)$. Combine with (S.17), we can obtain $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n1,1}\left(\boldsymbol{w}\right)\right| = o_p(1)$. Next, we consider $\Lambda_{n1,2}\left(\boldsymbol{w}\right)$.

Due to the boundedness of $|\psi_\tau(\varepsilon_i)| \le 1$, we have

$$\begin{aligned}\sup_{\boldsymbol{w}\in\mathbb{W}} |\Lambda_{n1,2}(\boldsymbol{w})| \le& \sup_{\boldsymbol{w}\in\mathbb{W}} \sum_{s=1}^{S} w_s \frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left|\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\psi_\tau(\varepsilon_i)\right| d\tau\\
\le& \max_{1\le s\le S}\frac{1}{n}\sum_{i=1}^{n}\int_0^1 \left|\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right| d\tau\\
\le& \left\{\max_{1\le s\le S}\left\|\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\|\right\}\left\{\max_{1\le s\le S}\frac{1}{n}\sum_{i=1}^{n}\int_0^1\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\| d\tau\right\}\\
=&O_p\left(\sqrt{J_n\bar{\phi}\log n/n}\right)O_p\left(p^{1/2}\right)=o_p(1)\end{aligned}$$

by the condition (C5) (i). Thus $\sup_{\boldsymbol{w}\in\mathbb{W}}|\Lambda_{n1,2}(\boldsymbol{w})| = o_p(1)$. Combine with $\sup_{\boldsymbol{w}\in\mathbb{W}}|\Lambda_{n1,1}(\boldsymbol{w})| = o_p(1)$, we have $\sup_{\boldsymbol{w}\in\mathbb{W}}|\Lambda_{n1}(\boldsymbol{w})| = o_p(1)$.

Thirdly, we establish (iii). Decompose

$$\begin{aligned}\Lambda_{n2}(\boldsymbol{w}) =& \frac{1}{n}\sum_{i=1}^{n}\int_0^1\int_0^{\sum_{s=1}^{S} w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\Big\{[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)]-[F(t|\boldsymbol{X}_i)-F(0|\boldsymbol{X}_i)]\Big\}dtd\tau\\
&+\frac{1}{n}\sum_{i=1}^{n}\int_0^1\int_{\sum_{s=1}^{S} w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}^{\sum_{s=1}^{S} w_s\boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i}\Big\{[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)]-[F(t|\boldsymbol{X}_i)-F(0|\boldsymbol{X}_i)]\Big\}dtd\tau\\
\triangleq&\Lambda_{n2,1}(\boldsymbol{w})+\Lambda_{n2,2}(\boldsymbol{w}).\end{aligned}$$

Similar to the proof of $\Lambda_{n1,1}(\boldsymbol{w})$, we have

$$\sup_{\boldsymbol{w}\in\mathbb{W}}|\Lambda_{n2,1}(\boldsymbol{w})| \le \max_{1\le j\le N}|\Lambda_{n2,1}(\boldsymbol{w}_j)| + o_p(1). \tag{S.19}$$

Therefore, we need to prove $\max_{1\le j\le N}|\Lambda_{n2,1}(\boldsymbol{w}_j)| = o_p(1)$. Note that

$$\begin{aligned}&E\Big\{\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}([I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)]-[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)])\,dtd\tau\Big\}\\
=&E\Big\{E\Big[\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}([I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)]\,[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)])\,dtd\tau\,|\boldsymbol{X}_i\Big]\Big\}\\
=&E\Big\{\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}E\left\{(I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)\,-\,F(t\,|\boldsymbol{X}_i)+F(0\,|\boldsymbol{X}_i))\,|\boldsymbol{X}_i\right\}dtd\tau\Big\}\\
=&0,\end{aligned}$$

where the last equality is true by the fact that

$$E\left\{(I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)-F(t\,|\boldsymbol{X}_i)+F(0\,|\boldsymbol{X}_i))\,|\boldsymbol{X}_i\right\}=0.$$

In view of the fact that $|[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)]-[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)]| \le 2$, we have

$$\begin{aligned}&\left|\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}[I(\varepsilon_i\le t)-I(\varepsilon_i\le 0)-F(t\,|\boldsymbol{X}_i)+F(0\,|\boldsymbol{X}_i)]dtd\tau\right|\\
\le&2|\mu_i|+2\max_{1\le s\le S}\int_0^1\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|d\tau\max_{1\le s\le S}\left\|\boldsymbol{\zeta}_0^{(s)}\right\|\le C_1p^{1/2}\bar{\phi}^{1/2},\end{aligned}$$

and

$$
\begin{aligned}
& Var\left\{\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left(\left[I\left(\varepsilon_i\le t\right)-I\left(\varepsilon_i\le 0\right)\right]-\left[F\left(t\left|\boldsymbol{X}_i\right.\right)-F\left(0\left|\boldsymbol{X}_i\right.\right)\right]\right)dtd\tau\right\} \\
=& E\left\{\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left(\left[I\left(\varepsilon_i\le t\right)-I\left(\varepsilon_i\le 0\right)\right]-\left[F\left(t\left|\boldsymbol{X}_i\right.\right)-F\left(0\left|\boldsymbol{X}_i\right.\right)\right]\right)dtd\tau\right\}^2 \\
\le& 4E\left[\int_0^1\left(|\mu_i|+\sum_{s=1}^{S} w_{js}\left|\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}\right|\right)d\tau\right]^2 \\
\le& 4E\left[\int_0^1\left(|\mu_i|+\sum_{s=1}^{S} w_{js}\left|\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}\right|\right)^2 d\tau\right] \\
\le& 4E\left[\int_0^1\left(|\mu_i|+\max_{1\le s\le S}\left|\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}\right|\right)^2 d\tau\right] \\
\le& 8E\left[\int_0^1\left(\mu_i^2+\max_{1\le s\le S}\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|^2\left\|\boldsymbol{\zeta}_0^{(s)}\right\|^2\right)d\tau\right] \\
\le& 8\left\{E\left(\mu_i^2\right)+\max_{1\le s\le S}E\int_0^1\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|^2 d\tau\max_{1\le s\le S}\left\|\boldsymbol{\zeta}_0^{(s)}\right\|^2\right\} \\
\le& C_2 p\bar{\phi}.
\end{aligned}
$$

Using the Boole's and Bernstein's inequalities, for any $\delta>0$, we have

$$
\begin{aligned}
& P\left\{\max_{1\le j\le N}\left|\Lambda_{n2,1}\left(\boldsymbol{w}_j\right)\right|\ge\delta\right\} \\
=& P\Big\{\max_{1\le j\le N}\Big|\frac{1}{n}\sum_{i=1}^{n}\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left(\left[I\left(\varepsilon_i\le t\right)-I\left(\varepsilon_i\le 0\right)\right]\right. \\
& \left.-\left[F\left(t|\boldsymbol{X}_i\right)-F\left(0|\boldsymbol{X}_i\right)\right]\right)dtd\tau\Big|\ge\delta\Big\} \\
\le& \sum_{j=1}^{N}\max_{1\le j\le N}P\Big\{\Big|\sum_{i=1}^{n}\int_0^1\int_0^{\sum_{s=1}^{S} w_{js}\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left(\left[I\left(\varepsilon_i\le t\right)-I\left(\varepsilon_i\le 0\right)\right]\right. \\
& \left.-\left[F\left(t\left|\boldsymbol{X}_i\right.\right)-F\left(0\left|\boldsymbol{X}_i\right.\right)\right]\right)dtd\tau\Big|\ge n\delta\Big\} \\
\le& 2N\exp\left\{-\frac{n^2\delta^2}{2nC_2p\bar{\phi}+2n\delta C_1p^{1/2}\bar{\phi}^{1/2}/3}\right\} \\
\le& 2\exp\left\{-\frac{n\delta^2}{2C_2p\bar{\phi}+2\delta C_1(p\bar{\phi})^{1/2}/3}+S\log(p^{1/2}\bar{\phi}^{1/2}\log n)\right\} \\
=& o(1) \qquad \text{(S.20)}
\end{aligned}
$$

by the condition $p^{3/2}\bar{\phi}^{3/2}S\log n/n\to 0$, as $n\to\infty$. Thus $\max_{1\le j\le N}\left|\Lambda_{n2,1}\left(\boldsymbol{w}_j\right)\right|=o_p(1)$. Combine with (S.19), we have $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n2,1}\left(\boldsymbol{w}\right)\right|=o_p(1)$.

By $\left|\left[I\left(\varepsilon_i\le t\right)-I\left(\varepsilon_i\le 0\right)\right]-\left[F\left(t\left|\boldsymbol{X}_i\right.\right)-F\left(0\left|\boldsymbol{X}_i\right.\right)\right]\right|\le 2$, Proposition 1 and condition (C5) (i), we have

$$\begin{aligned}
&\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n2,2}(\boldsymbol{w})\right| \\
\leq&\frac{2}{n}\sum_{i=1}^{n}\int_{0}^{1}\left|\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|d\tau \\
\leq&2\left\{\max_{1\leq s\leq S}\left\|\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\|\right\}\left\{\max_{1\leq s\leq S}\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{1}\left\|\boldsymbol{D}_i^{(s)}(\tau)\right\|d\tau\right\} \\
=&O_p\left(\sqrt{J_n\bar{\phi}\log n/n}\right)O_p\left(p^{1/2}\right) \\
=&o_p(1).
\end{aligned}$$

(iv) Observe that

$$\begin{aligned}
\Lambda_{n3}(\boldsymbol{w})=&\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{1}\left\{\int_{0}^{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right]dt\right. \\
&\left.-E_{\boldsymbol{X}_i}\left[\int_{0}^{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}\left[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right]dt\right]\right\}d\tau \\
&+\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{1}\left\{\int_{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}^{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i}\left[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right]dt\right. \\
&\left.-E_{\boldsymbol{X}_i}\left[\int_{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\boldsymbol{\zeta}_0^{(s)}-\mu_i}^{\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\hat{\boldsymbol{\zeta}}^{(s)}-\mu_i}\left[F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right]dt\right]\right\}d\tau \\
\triangleq&\Lambda_{n3,1}(\boldsymbol{w})+\Lambda_{n3,2}(\boldsymbol{w}).
\end{aligned}$$

Note that $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3,1}(\boldsymbol{w})\right|\leq\max_{1\leq j\leq N}\left|\Lambda_{n3,1}(\boldsymbol{w}_j)\right|+o_p(1)$. Similar to the proof of (S.20), we have $\max_{1\leq j\leq N}\left|\Lambda_{n3,1}(\boldsymbol{w}_j)\right|=o_p(1)$, and thus $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3,1}(\boldsymbol{w})\right|=o_p(1)$.

In view of the fact that $\left|F(t\,|\boldsymbol{X}_i)-F(0\,|\boldsymbol{X}_i)\right|\leq 1$, we have

$$\begin{aligned}
\left|\Lambda_{n3,2}(\boldsymbol{w})\right|\leq&\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{1}\left|\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|d\tau \\
&+\frac{1}{n}\sum_{i=1}^{n}E_{\boldsymbol{X}_i}\left\{\int_{0}^{1}\left|\sum_{s=1}^{S}w_s\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|d\tau\right\} \\
\triangleq&\Lambda_{n3,21}(\boldsymbol{w})+\Lambda_{n3,22}(\boldsymbol{w}).
\end{aligned}$$

The proof of $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3,21}(\boldsymbol{w})\right|=o_p(1)$ is similar to that of $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n2,2}(\boldsymbol{w})\right|=o_p(1)$ in (iii) and is also omitted. By the triangle inequality and Cauchy-Schwarz inequality, Proposition 1, condition (C5) (i), and $\boldsymbol{A}^\top\boldsymbol{B}\boldsymbol{A}\leq\lambda_{\max}(\boldsymbol{B})\boldsymbol{A}^\top\boldsymbol{A}$ for any symmetric matrix

$\boldsymbol{B}$, we have

$$
\begin{aligned}
&\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n3,22}\left(\boldsymbol{w}\right)\right| \\
\leq &\sup_{\boldsymbol{w}\in\mathbb{W}}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{S}w_s E_{\boldsymbol{X}_i}\left\{\int_0^1\left|\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|d\tau\right\} \\
\leq &\sup_{\boldsymbol{w}\in\mathbb{W}}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{S}w_s\left\{\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)^{\top}E\left[\int_0^1\boldsymbol{D}_i^{(s)}(\tau)\boldsymbol{D}_i^{(s)\top}(\tau)d\tau\right]\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right\}^{1/2} \\
\leq &\max_{1\leq s\leq S}\left[\lambda_{\max}\left(\boldsymbol{\Sigma}^{(s)}\right)\right]^{1/2}\max_{1\leq s\leq S}\left\|\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right\| \\
= &o\left(1\right).
\end{aligned}
$$

$(v)-(vi)$ By the same arguments used for $\Lambda_{n3,22}\left(\boldsymbol{w}\right)$ and $\left|F\left(t\left|\boldsymbol{X}_i\right.\right)-F\left(0\left|\boldsymbol{X}_i\right.\right)\right|\leq 1$, we have $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n4}\left(\boldsymbol{w}\right)\right|\leq\sup_{\boldsymbol{w}\in\mathbb{W}}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{S}w_s E_{\boldsymbol{X}_i}\left\{\int_0^1\left|\boldsymbol{D}_i^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|d\tau\right\}=o_p(1)$ and $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n5}\left(\boldsymbol{w}\right)\right|\leq\sup_{\boldsymbol{w}\in\mathbb{W}}E_{\mathbf{x}}\left|\int_0^1\sum_{s=1}^{S}w_s\mathbb{D}^{(s)\top}(\tau)\left(\hat{\boldsymbol{\zeta}}^{(s)}-\boldsymbol{\zeta}_0^{(s)}\right)\right|=o_p\left(1\right)$.

(vii) Please observe that $\sup_{\boldsymbol{w}\in\mathbb{W}}\left|\Lambda_{n6}\left(\boldsymbol{w}\right)\right|=\sup_{\boldsymbol{w}\in\mathbb{W}}n^{-1}\lambda_n\sum_{s=1}^{S}w_s\phi_s\leq n^{-1}\lambda_n\max_{1\leq s\leq S}\phi_s=O(n^{-1}\lambda_n\bar{\phi})=o\left(1\right)$ by the condition $\lambda_n\bar{\phi}=o(n)$. This completes the proof of the theorem.

References

Frumento, P., Bottai, M. (2016). Parametric modeling of quantile regression coefficient functions. Biometrics, 72, 74–84.

Knight, K. (1998). Limiting distribution for L_1 regression estimators under general conditions. Annals of Statistics, 26, 755–770.

Lee, J. and Shin, Y. (2023). Complete subset averaging for quantile regressions. Econometric Theory, 39, 146–188.

Lu, X., Su, L. (2015). Jackknife model averaging for quantile regressions. Journal of Econometrics, 188, 40–58.

Shibata, R. (1981). An optimal selection of regression variables. Biometrika 68, 45–54.

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

Van der Vaart, A. W., Wellner, J.A. (1996). Weak convergence and empirical processes: with applications to statistics. Springer, New York.